

2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2023)

**Vancouver, British Columbia, Canada
18-22 June 2023**

Pages 1-371



**IEEE Catalog Number: CFP23003-POD
ISBN: 979-8-3503-0130-4**

**Copyright © 2023 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP23003-POD
ISBN (Print-On-Demand):	979-8-3503-0130-4
ISBN (Online):	979-8-3503-0129-8
ISSN:	1063-6919

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) CVPR 2023

Table of Contents

Message from the 2023 General and Program Chairs	ccci
2023 Organizing Committee	cccii
2023 Outstanding Reviewers	ccciv
Sponsors	cccv

Poster-Tue-AM

Megahertz Light Steering Without Moving Parts	1
<i>Adithya Pediredla (Dartmouth College; Dartmouth College), Srinivasa G. Narasimhan (Carnegie Mellon University), Maysamreza Chamanzar (Carnegie Mellon University), and Ioannis Gkioulekas (Carnegie Mellon University)</i>	
Robust Dynamic Radiance Fields	13
<i>Yu-Lun Liu (National Taiwan University), Chen Gao (Meta), Andréas Meuleman (KAIST), Hung-Yu Tseng (Meta), Ayush Saraf (Meta), Changil Kim (Meta), Yung-Yu Chuang (National Taiwan University), Johannes Kopf (Meta), and Jia-Bin Huang (Meta; University of Maryland, College Park)</i>	
DBARF: Deep Bundle-Adjusting Generalizable Neural Radiance Fields	24
<i>Yu Chen (National University of Singapore) and Gim Hee Lee (National University of Singapore)</i>	
VDN-NeRF: Resolving Shape-Radiance Ambiguity via View-Dependence Normalization	35
<i>Bingfan Zhu (Zhejiang University), Yanchao Yang (The University of Hong Kong; Stanford University), Xulong Wang (Zhejiang University), Youyi Zheng (Zhejiang University), and Leonidas Guibas (Stanford University)</i>	
AligNeRF: High-Fidelity Neural Radiance Fields via Alignment-Aware Training	46
<i>Yifan Jiang (University of Texas at Austin), Peter Hedman (Google Research), Ben Mildenhall (University of Texas at Austin; Google Research), DeJia Xu (Google Research), Jonathan T. Barron (The Chinese University of Hong Kong), Zhangyang Wang (University of Texas at Austin), and Tianfan Xue (The Chinese University of Hong Kong)</i>	

SeaThru-NeRF: Neural Radiance Fields in Scattering Media	56
<i>Deborah Levy (University of Haifa), Amit Peleg (University of Haifa), Naama Pearl (University of Haifa), Dan Rosenbaum (University of Haifa), Derya Akkaynak (University of Haifa; The Inter-University Institute for Marine Sciences in Eilat), Simon Korman (University of Haifa), and Tali Treibitz (University of Haifa)</i>	
Exact-NeRF: An Exploration of a Precise Volumetric Parameterization for Neural Radiance Fields	66
<i>Brian K. S. Isaac-Medina (Department of Computer Science, Durham University, UK), Chris G. Willcocks (Department of Computer Science, Durham University, UK), and Toby P. Breckon (Department of Computer Science, Durham University, UK; Department of Engineering, Durham University, UK)</i>	
Neural Residual Radiance Fields for Streamably Free-Viewpoint Videos	76
<i>Liao Wang (ShanghaiTech University; NeuDim), Qiang Hu (ShanghaiTech University), Qihan He (ShanghaiTech University; DGene), Ziyu Wang (ShanghaiTech University), Jingyi Yu (ShanghaiTech University), Tinne Tuytelaars (KU Leuven), Lan Xu (ShanghaiTech University), and Minye Wu (KU Leuven)</i>	
PlenVDB: Memory Efficient VDB-Based Radiance Fields for Fast Training and Rendering	88
<i>Han Yan (Al Institute, Shanghai Jiao Tong University), Celong Liu (ByteDance Inc.), Chao Ma (Al Institute, Shanghai Jiao Tong University), and Xing Mei (ByteDance Inc.)</i>	
Local Implicit Ray Function for Generalizable Radiance Field Representation	97
<i>Xin Huang (School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China), Qi Zhang (Tencent AI Lab), Ying Feng (Tencent AI Lab), Xiaoyu Li (Tencent AI Lab), Xuan Wang (Tencent AI Lab), and Qing Wang (School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China)</i>	
SurfelNeRF: Neural Surfel Radiance Fields for Online Photorealistic Reconstruction of Indoor Scenes	108
<i>Yiming Gao (ARC Lab, Tencent PCG), Yan-Pei Cao (ARC Lab, Tencent PCG), and Ying Shan (ARC Lab, Tencent PCG)</i>	
Frequency-Modulated Point Cloud Rendering With Easy Editing	119
<i>Yi Zhang (Shanghai Jiao Tong University, Shanghai 200240, China), Xiaoyang Huang (Shanghai Jiao Tong University, Shanghai 200240, China), Bingbing Ni (Shanghai Jiao Tong University, Shanghai 200240, China), Teng Li (Anhui University), and Wenjun Zhang (Shanghai Jiao Tong University, Shanghai 200240, China)</i>	
HexPlane: A Fast Representation for Dynamic Scenes	130
<i>Ang Cao (University of Michigan, Ann Arbor) and Justin Johnson (University of Michigan, Ann Arbor)</i>	
Differentiable Shadow Mapping for Efficient Inverse Graphics	142
<i>Markus Worchel (TU Berlin) and Marc Alexa (TU Berlin)</i>	
Hybrid Neural Rendering for Large-Scale Scenes With Motion Blur	154
<i>Peng Dai (The University of Hong Kong), Yinda Zhang (Google), Xin Yu (The University of Hong Kong), Xiaoyang Lyu (The University of Hong Kong), and Xiaojuan Qi (The University of Hong Kong)</i>	

TensorIR: Tensorial Inverse Rendering	165
<i>Haian Jin (Zhejiang University), Isabella Liu (UC San Diego), Peijia Xu (Kingstar Technology Inc.), Xiaoshuai Zhang (UC San Diego), Songfang Han (UC San Diego), Sai Bi (Adobe Research), Xiaowei Zhou (Zhejiang University), Zexiang Xu (Adobe Research), and Hao Su (UC San Diego)</i>	
ShadowNeuS: Neural SDF Reconstruction by Shadow Ray Supervision	175
<i>Jingwang Ling (Tsinghua University), Zhibo Wang (SenseTime Research), and Feng Xu (Tsinghua University)</i>	
Realistic Saliency Guided Image Enhancement	186
<i>S. H. Miangoleh (Simon Fraser University), Zoya Bylinskii (Adobe Research), Eric Kee (Adobe Research), Eli Shechtman (Adobe Research), and Yağiz Aksoy (Simon Fraser University)</i>	
LightPainter: Interactive Portrait Relighting With Freehand Scribble	195
<i>Yiqun Mei (Johns Hopkins University), He Zhang (Adobe Inc.), Xuaner Zhang (Adobe Inc.), Jianming Zhang (Adobe Inc.), Zhixin Shu (Adobe Inc.), Yilin Wang (Adobe Inc.), Zijun Wei (Adobe Inc.), Shi Yan (Adobe Inc.), HyunJoon Jung (Adobe Inc.), and Vishal M. Patel (Johns Hopkins University)</i>	
A Unified Spatial-Angular Structured Light for Single-View Acquisition of Shape and Reflectance	206
<i>Xianmin Xu (Zhejiang University), Yuxin Lin (Zhejiang University), Haoyang Zhou (Zhejiang University), Chong Zeng (Zhejiang University), Yaxin Yu (Zhejiang University), Kun Zhou (Zhejiang University; ZJU-FaceUnity Joint Lab of Intelligent Graphics), and Hongzhi Wu (Zhejiang University)</i>	
Learning Visibility Field for Detailed 3D Human Reconstruction and Relighting	216
<i>Ruichen Zheng (Tsinghua University, China; Weilan Tech, China), Peng Li (Tsinghua University, China), Haoqian Wang (Tsinghua University, China), and Tao Yu (Tsinghua University, China)</i>	
Unsupervised Contour Tracking of Live Cells by Mechanical and Cycle Consistency Losses	227
<i>Junbong Jang (KAIST), Kwonmoo Lee (Boston Children's Hospital Harvard Medical School), and Tae-Kyun Kim (KAIST)</i>	
NeUDF: Leaning Neural Unsigned Distance Fields With Volume Rendering	237
<i>Yu-Tao Liu (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Li Wang (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Jie Yang (Institute of Computing Technology, Chinese Academy of Sciences), Weikai Chen (Digital Content Technology Center, Tencent Games), Xiaoxu Meng (Digital Content Technology Center, Tencent Games), Bo Yang (Digital Content Technology Center, Tencent Games), and Lin Gao (University of Chinese Academy of Sciences)</i>	
NeAT: Learning Neural Implicit Surfaces With Arbitrary Topologies From Multi-View Images ...	248
<i>Xiaoxu Meng (Digital Content Technology Center), Weikai Chen (Digital Content Technology Center), and Bo Yang (Digital Content Technology Center)</i>	

ALTO: Alternating Latent Topologies for Implicit 3D Reconstruction	259
<i>Zhen Wang (University of California, Los Angeles), Shijie Zhou (University of California, Los Angeles), Jeong Joon Park (Stanford University), Despoina Paschalidou (Stanford University), Suyu You (DEVCOM Anny Research Laboratory), Gordon Wetzstein (Stanford University), Leonidas Guibas (Stanford University), and Achuta Kadambi (University of California, Los Angeles)</i>	
Controllable Mesh Generation Through Sparse Latent Point Diffusion Models	271
<i>Zhaoyang Lyu (Shanghai AI Laboratory), Jinyi Wang (Shanghai AI Laboratory; Shanghai Jiao Tong University), Yuwei An (Shanghai AI Laboratory; Tsinghua University), Ya Zhang (Shanghai AI Laboratory; Shanghai Jiao Tong University), Dahua Lin (Shanghai AI Laboratory; The Chinese University of Hong Kong), and Bo Dai (Shanghai AI Laboratory)</i>	
Power Bundle Adjustment for Large-Scale 3D Reconstruction	281
<i>Simon Weber (Technical University of Munich; Munich Center for Machine Learning), Nikolaus Demmel (Technical University of Munich; Munich Center for Machine Learning), Tin Chon Chan (Technical University of Munich; Munich Center for Machine Learning), and Daniel Cremers (Technical University of Munich; Munich Center for Machine Learning; University of Oxford)</i>	
Neural Pixel Composition for 3D-4D View Synthesis From Multi-Views	290
<i>Aayush Bansal (Reality Labs Research, Pittsburgh, USA) and Michael Zollhöfer (Reality Labs Research, Pittsburgh, USA)</i>	
Magic3D: High-Resolution Text-to-3D Content Creation	300
<i>Chen-Hsuan Lin (NVIDIA Corporation), Jun Gao (NVIDIA Corporation), Luming Tang (NVIDIA Corporation), Towaki Takikawa (NVIDIA Corporation), Xiaohui Zeng (NVIDIA Corporation), Xun Huang (NVIDIA Corporation), Karsten Kreis (NVIDIA Corporation), Sanja Fidler (NVIDIA Corporation), Ming-Yu Liu (NVIDIA Corporation), and Tsung-Yi Lin (NVIDIA Corporation)</i>	
3D Video Loops From Asynchronous Input	310
<i>Li Ma (The Hong Kong University of Science and Technology), Xiaoyu Li (Tencent AI Lab), Jing Liao (City University of Hong Kong), and Pedro V. Sander (The Hong Kong University of Science and Technology)</i>	
High-Fidelity 3D GAN Inversion by Pseudo-Multi-View Optimization	321
<i>Jiaxin Xie (HKUST), Hao Ouyang (HKUST), Jingtian Piao (CUHK), Chenyang Lei (HKISI-CAS), and Qifeng Chen (HKUST)</i>	
Lift3D: Synthesize 3D Training Data by Lifting 2D GAN to 3D Generative Radiance Field	332
<i>Leheng Li (HKUST(GZ)), Qing Lian (HKUST), Luozhou Wang (HKUST(GZ)), Ningning Ma (NIO Autonomous Driving), and Ying-Cong Chen (HKUST(GZ); HKUST)</i>	
3D GAN Inversion With Facial Symmetry Prior	342
<i>Fei Yin (Tsinghua University), Yong Zhang (Tencent AI Lab), Xuan Wang (Ant Group), Tengfei Wang (HKUST), Xiaoyu Li (Tencent AI Lab), Yuan Gong (Tsinghua University), Yanbo Fan (Tencent AI Lab), Xiaodong Cun (Tencent AI Lab), Ying Shan (Tencent AI Lab), Cengiz Oztireli (University of Cambridge), and Yujiu Yang (Tsinghua University)</i>	

StyleIPSB: Identity-Preserving Semantic Basis of StyleGAN for High Fidelity Face Swapping	352
<i>Diqiong Jiang (Zhejiang University, China), Dan Song (Tianjin University, China), Ruofeng Tong (Zhejiang University, China), and Min Tang (Zhejiang University, China)</i>	
FFHQ-UV: Normalized Facial UV-Texture Dataset for 3D Face Reconstruction	362
<i>Haoran Bai (Nanjing University of Science and Technology), Di Kang (Tencent AI Lab), Haoxian Zhang (Tencent AI Lab), Jinshan Pan (Nanjing University of Science and Technology), and Linchao Bao (Tencent AI Lab)</i>	
Robust Model-Based Face Reconstruction Through Weakly-Supervised Outlier Segmentation .	372
<i>Chunlu Li (Southeast University; University of Basel), Andreas Morel-Forster (University of Basel), Thomas Vetter (University of Basel), Bernhard Egger (Friedrich-Alexander-Universitat Erlangen-Nürnberg), and Adam Kortylewski (University of Freiburg; Max Planck Institute for Informatics)</i>	
Learning Neural Proto-Face Field for Disentangled 3D Face Modeling in the Wild	382
<i>Zhenyu Zhang (Tencent Youtu Lab, Shanghai, China), Renwang Chen (Tencent Youtu Lab, Shanghai, China), Weijian Cao (Tencent Youtu Lab, Shanghai, China), Ying Tai (Tencent Youtu Lab, Shanghai, China), and Chengjie Wang (Tencent Youtu Lab, China; Shanghai Jiao Tong University, Shanghai, China)</i>	
A Hierarchical Representation Network for Accurate and Detailed Face Reconstruction From In-the-Wild Images	394
<i>Biwen Lei (DAMO Academy, Alibaba Group), Jianqiang Ren (DAMO Academy, Alibaba Group), Mengyang Feng (DAMO Academy, Alibaba Group), Miaomiao Cui (DAMO Academy, Alibaba Group), and Xuansong Xie (DAMO Academy, Alibaba Group)</i>	
BlendFields: Few-Shot Example-Driven Facial Modeling	404
<i>Kacper Kania (Warsaw University of Technology; University of British Columbia), Stephan J. Garbin (Microsoft), Andrea Tagliasacchi (Simon Fraser University; Google Brain), Virginia Estellers (Microsoft), Kwang Moo Yi (University of British Columbia), Julien Valentin (Microsoft), Tomasz Trzcinski (Warsaw University of Technology; IDEAS NCBR; Tooploox), and Marek Kowalski (Microsoft)</i>	
Implicit Neural Head Synthesis via Controllable Local Deformation Fields	416
<i>Chuhan Chen (Carnegie Mellon University), Matthew O'Toole (Carnegie Mellon University), Gaurav Bharaj (Flawless AI), and Pablo Garrido (Flawless AI)</i>	
DPE: Disentanglement of Pose and Expression for General Video Portrait Editing	427
<i>Youxin Pang (MAIS & NLPR, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences; Tencent AI Lab, China), Yong Zhang (Tencent AI Lab, China), Weize Quan (MAIS & NLPR, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences), Yanbo Fan (Tencent AI Lab, China), Xiaodong Cun (Tencent AI Lab, China), Ying Shan (Tencent AI Lab, China), and Dong-Ming Yan (MAIS & NLPR, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences)</i>	

GANHead: Towards Generative Animatable Neural Head Avatars	437
<i>Sijing Wu (Shanghai Jiao Tong University), Yichao Yan (Shanghai Jiao Tong University), Yunhao Li (Shanghai Jiao Tong University), Yuhao Cheng (Shanghai Jiao Tong University), Wenhan Zhu (Shanghai Jiao Tong University), Ke Gao (Alibaba Group), Xiaobo Li (Alibaba Group), and Guangtao Zhai (Shanghai Jiao Tong University; Shanghai Jiao Tong University)</i>	
EDGE: Editable Dance Generation From Music	448
<i>Jonathan Tseng (Stanford University), Rodrigo Castellon (Stanford University), and Karen Liu (Stanford University)</i>	
Blowing in the Wind: CycleNet for Human Cinemagraphs From Still Images	459
<i>Hugo Bertiche (Universitat de Barcelona; Computer Vision Center), Niloy J. Mitra (University College London; Adobe Research), Kuldeep Kulkarni (Adobe Research), Chun-Hao P. Huang (Adobe Research), Tuanfeng Y. Wang (Adobe Research), Meysam Madadi (Universitat de Barcelona; Computer Vision Center), Sergio Escalera (Universitat de Barcelona; Computer Vision Center), and Duygu Ceylan (Adobe Research)</i>	
Generating Holistic 3D Human Motion From Speech	469
<i>Hongwei Yi (Max Planck Institute for Intelligent Systems, Germany), Hualin Liang (South China University of Technology), Yifei Liu (South China University of Technology), Qiong Cao (JD Explore Academy), Yandong Wen (Max Planck Institute for Intelligent Systems, Germany), Timo Bolkart (Max Planck Institute for Intelligent Systems, Germany), Dacheng Tao (JD Explore Academy), and Michael J. Black (Max Planck Institute for Intelligent Systems, Germany)</i>	
Avatars Grow Legs: Generating Smooth Human Motion From Sparse Tracking Inputs With Diffusion Model	481
<i>Yuming Du (Meta AI), Robin Kips (Meta AI), Albert Pumarola (Meta AI), Sebastian Starke (Meta AI), Ali Thabet (Meta AI), and Artsiom Sanakoyeu (Meta AI)</i>	
Learning Anchor Transformations for 3D Garment Animation	491
<i>Fang Zhao (Tencent AI Lab), Zekun Li (Tencent AI Lab), Shaoli Huang (Tencent AI Lab), Junwu Weng (Tencent AI Lab), Tianfei Zhou (ETH Zurich), Guo-Sen Xie (Nanjing University of Science and Technology), Jue Wang (Tencent AI Lab), and Ying Shan (Tencent AI Lab)</i>	
CloSET: Modeling Clothed Humans on Continuous Surface With Explicit Template Decomposition.	501
<i>Hongwen Zhang (Tsinghua University), Siyou Lin (Tsinghua University), Ruizhi Shao (Tsinghua University), Yuxiang Zhang (Tsinghua University), Zerong Zheng (Tsinghua University), Han Huang (OPPO Research Institute), Yandong Guo (OPPO Research Institute), and Yebin Liu (Tsinghua University)</i>	
ECON: Explicit Clothed Humans Optimized via Normal Integration	512
<i>Yuliang Xiu (Max Planck Institute for Intelligent Systems, Germany), Jinlong Yang (Max Planck Institute for Intelligent Systems, Germany), Xu Cao (Osaka University, Japan), Dimitrios Tzionas (University of Amsterdam, the Netherlands), and Michael J. Black (Max Planck Institute for Intelligent Systems, Germany)</i>	

PersonNeRF: Personalized Reconstruction From Photo Collections	524
<i>Chung-Yi Weng (University of Washington), Pratul P. Srinivasan (Google Research), Brian Curless (University of Washington; Google Research), and Ira Kemelmacher-Shlizerman (University of Washington; Google Research)</i>	
3D Human Mesh Estimation From Virtual Markers	534
<i>Xiaoxuan Ma (Peking University), Jiajun Su (Peking University), Chunyu Wang (Microsoft Research Asia), Wentao Zhu (Peking University), and Yizhou Wang (Peking University; Nat'l Eng. Research Center of Visual Technology)</i>	
Overcoming the Trade-Off Between Accuracy and Plausibility in 3D Hand Shape Reconstruction....	544
<i>Ziwei Yu (National University of Singapore), Chen Li (National University of Singapore), Linlin Yang (National University of Singapore), Xiaoxu Zheng (National University of Singapore), Michael Bi Mi (National University of Singapore), Gim Hee Lee (National University of Singapore), and Angela Yao (National University of Singapore)</i>	
Recovering 3D Hand Mesh Sequence From a Single Blurry Image: A New Dataset and Temporal Unfolding	554
<i>Yeonguk Oh (Dept. of ECE&ASRI), JoonKyu Park (Dept. of ECE&ASRI), Jaeha Kim (Dept. of ECE&ASRI), Gyeongsik Moon (Meta Reality Labs Research), and Kyoung Mu Lee (Dept. of ECE&ASRI; IPAI, Seoul National University, Korea)</i>	
MeMaHand: Exploiting Mesh-Mano Interaction for Single Image Two-Hand Reconstruction	564
<i>Congyi Wang (ByteDance), Feida Zhu (ByteDance), and Shilei Wen (ByteDance)</i>	
PLIKS: A Pseudo-Linear Inverse Kinematic Solver for 3D Human Body Estimation	574
<i>Karthik Shetty (FAU Erlangen-NUrnberg, Germany; Siemens Healthineers AG, Germany), Annette Birkhold (Siemens Healthineers AG, Germany), Srikrishna Jaganathan (FAU Erlangen-NUrnberg, Germany; Siemens Healthineers AG, Germany), Norbert Strobel (Siemens Healthineers AG, Germany; University of Applied Sciences WUrzburg-Schweinfurt, Germany), Markus Kowarschik (Siemens Healthineers AG, Germany), Andreas Maier (FAU Erlangen-NUrnberg, Germany), and Bernhard Egger (FAU Erlangen-NUrnberg, Germany)</i>	
CAMS: CANonicalized Manipulation Spaces for Category-Level Functional Hand-Object Manipulation Synthesis	585
<i>Juntian Zheng (Tsinghua University; Shanghai Qi Zhi Institute), Qingyuan Zheng (Shanghai Qi Zhi Institute), Lixing Fang (Tsinghua University; Shanghai Qi Zhi Institute), Yun Liu (Tsinghua University), and Li Yi (Tsinghua University; Shanghai Artificial Intelligence Laboratory; Shanghai Qi Zhi Institute)</i>	
Instant-NVR: Instant Neural Volumetric Rendering for Human-Object Interactions From Monocular RGBD Stream	595
<i>Yuheng Jiang (ShanghaiTech University; NeuDim), Kaixin Yao (ShanghaiTech University; NeuDim), Zhuo Su (Pico IDL, ByteDance), Zhehao Shen (ShanghaiTech University), Haimin Luo (ShanghaiTech University), and Lan Xu (ShanghaiTech University)</i>	

BundleSDF: Neural 6-DoF Tracking and 3D Reconstruction of Unknown Objects	606
<i>Bowen Wen (NVIDIA), Jonathan Tremblay (NVIDIA), Valts Blukis (NVIDIA), Stephen Tyree (NVIDIA), Thomas Müller (NVIDIA), Alex Evans (NVIDIA), Dieter Fox (NVIDIA), Jan Kautz (NVIDIA), and Stan Birchfield (NVIDIA)</i>	
Human-Art: A Versatile Human-Centric Dataset Bridging Natural and Artificial Scenes	618
<i>Xuan Ju (International Digital Economy Academy; The Chinese University of Hong Kong), Ailing Zeng (International Digital Economy Academy), Jianan Wang (International Digital Economy Academy), Qiang Xu (The Chinese University of Hong Kong), and Lei Zhang (International Digital Economy Academy)</i>	
Omnimatte3D: Associating Objects and Their Effects in Unconstrained Monocular Video	630
<i>Mohammed Suhail (University of British Columbia; Vector Institute for AI), Erika Lu (Google), Zhengqi Li (Google), Noah Snavely (Google), Leonid Sigal (University of British Columbia; Vector Institute for AI; Canada CIFAR AI Chair), and Forrester Cole (Google)</i>	
On the Benefits of 3D Pose and Tracking for Human Action Recognition	640
<i>Jathushan Rajasegaran (UC Berkeley; Meta AI, FAIR), Georgios Pavlakos (UC Berkeley), Angjoo Kanazawa (UC Berkeley), Christoph Feichtenhofer (Meta AI, FAIR), and Jitendra Malik (UC Berkeley)</i>	
Towards Stable Human Pose Estimation via Cross-View Fusion and Foot Stabilization	650
<i>Li'an Li'an Zhuo (Alibaba Group), Jian Cao (Alibaba Group), Qi Wang (Alibaba Group), Bang Zhang (Alibaba Group), and Liefeng Bo (Alibaba Group)</i>	
Human Pose As Compositional Tokens	660
<i>Zigang Geng (University of Science and Technology of China; Microsoft Research Asia), Chunyu Wang (Microsoft Research Asia), Yixuan Wei (Tsinghua University; Microsoft Research Asia), Ze Liu (University of Science and Technology of China; Microsoft Research Asia), Houqiang Li (University of Science and Technology of China), and Han Hu (Microsoft Research Asia)</i>	
PoseExaminer: Automated Testing of Out-of-Distribution Robustness in Human Pose and Shape Estimation	672
<i>Qihao Liu (Johns Hopkins University), Adam Kortylewski (Max Planck Institute for Informatics; University of Freiburg), and Alan L. Yuille (Johns Hopkins University)</i>	
SLOPER4D: A Scene-Aware Dataset for Global 4D Human Pose Estimation in Urban Environments	682
<i>Yudi Dai (Xiamen University; Xiamen University), Yitai Lin (Xiamen University; Xiamen University), Xiping Lin (Xiamen University; Xiamen University), Chenglu Wen (Xiamen University; Xiamen University), Lan Xu (ShanghaiTech University, China), Hongwei Yi (Max Planck Institute for Intelligent Systems, Germany), Siqi Shen (Xiamen University; Xiamen University), Yuexin Ma (ShanghaiTech University, China), and Cheng Wang (Xiamen University; Xiamen University)</i>	

Semi-Supervised 2D Human Pose Estimation Driven by Position Inconsistency Pseudo Label Correction Module	693
<i>Linzhi Huang (Beijing University of Posts and Telecommunications; Beike), Yulong Li (Beike), Hongbo Tian (Beijing University of Posts and Telecommunications; Beike), Yue Yang (Beike), Xiangang Li (Beike), Weihong Deng (Beijing University of Posts and Telecommunications), and Jieping Ye (Beike)</i>	
Human Pose Estimation in Extremely Low-Light Conditions	704
<i>Sohyun Lee (Graduate School of AI), Jaesung Rim (Graduate School of AI), Boseung Jeong (Graduate School of AI), Geonu Kim (Graduate School of AI), Byungju Woo (Graduate School of AI), Haechan Lee (Graduate School of AI), Sunghyun Cho (Graduate School of AI), and Suha Kwak (Graduate School of AI)</i>	
Flexible-Cm GAN: Towards Precise 3D Dose Prediction in Radiotherapy	715
<i>Riqiang Gao (Digital Technology and Innovation, Siemens Healthineers, NJ), Bin Lou (Digital Technology and Innovation, Siemens Healthineers, NJ), Zhoubing Xu (Digital Technology and Innovation, Siemens Healthineers, NJ), Dorin Comaniciu (Digital Technology and Innovation, Siemens Healthineers, NJ), and Ali Kamen (Digital Technology and Innovation, Siemens Healthineers, NJ)</i>	
DualRefine: Self-Supervised Depth and Pose Estimation Through Iterative Epipolar Sampling and Refinement Toward Equilibrium	726
<i>Antyanta Bangunharcana (Mechatronics, Systems, and Control Laboratory, Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea), Ahmed Magd (Vehicular Systems Design and Control Lab, Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea), and Kyung-Soo Kim (Mechatronics, Systems, and Control Laboratory, Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea)</i>	
A Rotation-Translation-Decoupled Solution for Robust and Efficient Visual-Inertial Initialization	739
<i>Yijia He (Chinese Academy of Sciences Wuhan University), Bo Xu (Chinese Academy of Sciences Wuhan University), Zhanpeng Ouyang (ShanghaiTech University), and Hongdong Li (Australian National University)</i>	
Semidefinite Relaxations for Robust Multiview Triangulation	749
<i>Linus Härenstam-Nielsen (Technical University of Munich; Munich Center for Machine Learning), Niclas Zeller (Karlsruhe University of Applied Sciences), and Daniel Cremers (Technical University of Munich; Munich Center for Machine Learning; University of Oxford)</i>	
A Probabilistic Attention Model With Occlusion-Aware Texture Regression for 3D Hand Reconstruction From a Single RGB Image	758
<i>Zheng Jiang (Lancaster University), Hossein Rahmani (Lancaster University), Sue Black (University of Oxford), and Bryan M. Williams (Lancaster University)</i>	
Instant Multi-View Head Capture Through Learnable Registration	768
<i>Timo Bolkart (MPI for Intelligent Systems, Tübingen), Tianye Li (University of Southern California), and Michael J. Black (MPI for Intelligent Systems, Tübingen)</i>	

On the Importance of Accurate Geometry Data for Dense 3D Vision Tasks	780
<i>HyunJun Jung (Technical University of Munich), Patrick Ruhkamp (Technical University of Munich; 3Dwe.ai), Guangyao Zhai (Technical University of Munich), Nikolas Brasch (Technical University of Munich), Yitong Li (Technical University of Munich), Yannick Verdie (Technical University of Munich; Huawei Noah's Ark Lab), Jifei Song (Huawei Noah's Ark Lab), Yiren Zhou (Huawei Noah's Ark Lab), Anil Armagan (Huawei Noah's Ark Lab), Slobodan Ilic (Technical University of Munich; Siemens AG), Aleš Leonardis (Huawei Noah's Ark Lab), Nassir Navab (Technical University of Munich), and Benjamin Busam (Technical University of Munich; 3Dwe.ai)</i>	
Learning 3D Scene Priors With 2D Supervision	792
<i>Yinyu Nie (Technical University of Munich), Angela Dai (Technical University of Munich), Xiaoguang Han (The Chinese University of Hong Kong (Shenzhen)), and Matthias Nießner (Technical University of Munich)</i>	
OmniObject3D: Large-Vocabulary 3D Object Dataset for Realistic Perception, Reconstruction and Generation	803
<i>Tong Wu (Shanghai Artificial Intelligence Laboratory; The Chinese University of Hong Kong), Jiarui Zhang (Shanghai Artificial Intelligence Laboratory; SenseTime Research), Xiao Fu (Shanghai Artificial Intelligence Laboratory), Yuxin Wang (Shanghai Artificial Intelligence Laboratory; Hong Kong University of Science and Technology), Jiawei Ren (Nanyang Technological University), Liang Pan (Nanyang Technological University), Wayne Wu (Shanghai Artificial Intelligence Laboratory), Lei Yang (Shanghai Artificial Intelligence Laboratory; SenseTime Research), Jiaqi Wang (Shanghai Artificial Intelligence Laboratory), Chen Qian (Shanghai Artificial Intelligence Laboratory), Dahua Lin (Shanghai Artificial Intelligence Laboratory; The Chinese University of Hong Kong), and Ziwei Liu (Nanyang Technological University)</i>	
OpenScene: 3D Scene Understanding With Open Vocabularies	815
<i>Songyou Peng (Google Research; ETH Zurich; MPI for Intelligent Systems, Tübingen), Kyle Genova (Google Research), Chiyu Max Jiang (Waymo LLC), Andrea Tagliasacchi (Google Research; Simon Fraser University), Marc Pollefeys (ETH Zurich), and Thomas Funkhouser (Google Research)</i>	
Multi-View Azimuth Stereo via Tangent Space Consistency	825
<i>Xu Cao (Osaka University), Hiroaki Santo (Osaka University), Fumio Okura (Osaka University), and Yasuyuki Matsushita (Osaka University)</i>	
Progressive Transformation Learning for Leveraging Virtual Images in Training	835
<i>Yi-Ting Shen (University of Maryland, College Park), Hyungtae Lee (DEVCOM Army Research Laboratory), Heesung Kwon (DEVCOM Army Research Laboratory), and Shuvra S. Bhattacharyya (University of Maryland, College Park)</i>	
Connecting the Dots: Floorplan Reconstruction Using Two-Level Queries	845
<i>Yuanwen Yue (Photogrammetry and Remote Sensing, ETH Zurich), Theodora Kontogianni (ETH AI Center, ETH Zurich), Konrad Schindler (Photogrammetry and Remote Sensing, ETH Zurich; ETH AI Center, ETH Zurich), and Francis Engelmann (ETH AI Center, ETH Zurich)</i>	

NeRF-Supervised Deep Stereo	855
<i>Fabio Tosi (University of Bologna), Alessio Tonioni (University of Bologna), Daniele De Gregorio (University of Bologna), and Matteo Poggi (University of Bologna)</i>	
Semantic Scene Completion With Cleaner Self	867
<i>Fengyun Wang (Nanjing University of Science & Technology), Dong Zhang (The Hong Kong University of Science & Technology), Hanwang Zhang (Nanyang Technological University), Jinhui Tang (Nanjing University of Science & Technology), and Qianru Sun (Singapore Management University)</i>	
PanelNet: Understanding 360 Indoor Environment via Panel Representation	878
<i>Haozheng Yu (Tencent America), Lu He (Tencent America), Bing Jian (Tencent America), Weiwei Feng (Tencent America), and Shan Liu (Tencent America)</i>	
Implicit View-Time Interpolation of Stereo Videos Using Multi-Plane Disparities and Non-Uniform Coordinates	888
<i>Avinash Paliwal (Texas A&M University), Andrii Tsarov (Leia Inc.), and Nima Khademi Kalantari (Texas A&M University)</i>	
Depth Estimation From Indoor Panoramas With Neural Scene Representation	899
<i>Wenjie Chang (University of Science and Technology of China), Yueyi Zhang (University of Science and Technology of China), and Zhiwei Xiong (University of Science and Technology of China)</i>	
NeuralPCI: Spatio-Temporal Neural Field for 3D Point Cloud Multi-Frame Non-Linear Interpolation	909
<i>Zehan Zheng (Tongji University), Danni Wu (Tongji University), Ruisi Lu (Tongji University), Fan Lu (Tongji University), Guang Chen (Tongji University), and Changjun Jiang (Tongji University)</i>	
RIAV-MVS: Recurrent-Indexing an Asymmetric Volume for Multi-View Stereo	919
<i>Changjiang Cai (OPPO US Research Center, InnoPeak Technology, Inc), Pan Ji (OPPO US Research Center, InnoPeak Technology, Inc), Qingan Yan (OPPO US Research Center, InnoPeak Technology, Inc), and Yi Xu (OPPO US Research Center, InnoPeak Technology, Inc)</i>	
NeuMap: Neural Coordinate Mapping by Auto-Transdecoder for Camera Localization	929
<i>Shitao Tang (Simon Fraser University), Sicong Tang (Simon Fraser University), Andrea Tagliasacchi (Simon Fraser University), Ping Tan (Simon Fraser University), and Yasutaka Furukawa (Simon Fraser University)</i>	
MACARONS: Mapping and Coverage Anticipation With RGB Online Self-Supervision	940
<i>Antoine Guédon (LIGM, Ecole des Ponts, Univ Gustave Eiffel, CNRS, France), Tom Monnier (LIGM, Ecole des Ponts, Univ Gustave Eiffel, CNRS, France), Pascal Monasse (LIGM, Ecole des Ponts, Univ Gustave Eiffel, CNRS, France), and Vincent Lepetit (LIGM, Ecole des Ponts, Univ Gustave Eiffel, CNRS, France)</i>	
vMAP: Vectorised Object Mapping for Neural Field SLAM	952
<i>Xin Kong (Dyson Robotics Lab, Imperial College London), Shikun Liu (Dyson Robotics Lab, Imperial College London), Marwan Taher (Dyson Robotics Lab, Imperial College London), and Andrew J. Davison (Dyson Robotics Lab, Imperial College London)</i>	

Seeing a Rose in Five Thousand Ways	962
<i>Yunzhi Zhang (Stanford University), Shangzhe Wu (University of Oxford), Noah Snavely (Cornell Tech), and Jiajun Wu (Stanford University)</i>	
Propagate and Calibrate: Real-Time Passive Non-Line-of-Sight Tracking	972
<i>Yihao Wang (Shanghai AI Laboratory), Zhigang Wang (Shanghai AI Laboratory), Bin Zhao (Shanghai AI Laboratory; Northwestern Polytechnical University), Dong Wang (Shanghai AI Laboratory), Mulin Chen (Shanghai AI Laboratory; Northwestern Polytechnical University), and Xuelong Li (Shanghai AI Laboratory; Northwestern Polytechnical University)</i>	
Seeing With Sound: Long-range Acoustic Beamforming for Multimodal Scene Understanding .	982
<i>Praneeth Chakravarthula (Princeton University), Jim Aldon D'Souza (Algolux), Ethan Tseng (Princeton University), Joe Bartusek (Princeton University), and Felix Heide (Princeton University; Algolux)</i>	
Distilling Focal Knowledge From Imperfect Expert for 3D Object Detection	992
<i>Jia Zeng (OpenDriveLab, Shanghai AI Lab; Shanghai Jiao Tong University), Li Chen (OpenDriveLab, Shanghai AI Lab), Hanming Deng (SenseTime Research), Lewei Lu (SenseTime Research), Junchi Yan (Shanghai Jiao Tong University; OpenDriveLab, Shanghai AI Lab), Yu Qiao (OpenDriveLab, Shanghai AI Lab), and Hongyang Li (OpenDriveLab, Shanghai AI Lab)</i>	
BEV-LaneDet: An Efficient 3D Lane Detection Based on Virtual Camera via Key-Points	1002
<i>Ruihao Wang (HAOMO.AI Technology Co., Ltd.; Xi'an Jiaotong University), Jian Qin (HAOMO.AI Technology Co., Ltd.), Kaiying Li (HAOMO.AI Technology Co., Ltd.), Yaochen Li (Xi'an Jiaotong University), Dong Cao (HAOMO.AI Technology Co., Ltd.), and Jintao Xu (HAOMO.AI Technology Co., Ltd.)</i>	
AShapeFormer: Semantics-Guided Object-Level Active Shape Encoding for 3D Object Detection via Transformers	1012
<i>Zechuan Li (Hunan University), Hongshan Yu (Hunan University; National University of Defense Technology; Quanzhou Institute of Industrial Design and Machine Intelligence Innovation, Hunan University), Zhengeng Yang (Hunan University), Tongjia Chen (Hunan University), and Naveed Akhtar (The University of Western Australia)</i>	
Benchmarking Robustness of 3D Object Detection to Common Corruptions	1022
<i>Yinpeng Dong (Tsinghua University; RealAI), Caixin Kang (Beihang University), Jinlai Zhang (Guangxi University), Zijian Zhu (Shanghai Jiao Tong University), Yikai Wang (Tsinghua University), Xiao Yang (Tsinghua University), Hang Su (Tsinghua University; Peng Cheng Laboratory, Pazhou Laboratory (Huangpu), Guangzhou, China), Xingxing Wei (Beihang University), and Jun Zhu (Tsinghua University; RealAI; Peng Cheng Laboratory, Pazhou Laboratory (Huangpu), Guangzhou, China)</i>	

Gaussian Label Distribution Learning for Spherical Image Object Detection	1033
<i>Hang Xu (Hangzhou Dianzi University, China; Chinese Academy of Sciences, Beijing, China), Xinyuan Liu (Chinese Academy of Sciences, Beijing, China), Qiang Zhao (Chinese Academy of Sciences, Beijing, China), Yike Ma (Chinese Academy of Sciences, Beijing, China), Chenggang Yan (Hangzhou Dianzi University, Hangzhou, China), and Feng Dai (Chinese Academy of Sciences, Beijing, China)</i>	
Deep Depth Estimation From Thermal Image	1043
<i>Ukcheol Shin (KAIST), Jinsun Park (Pusan National University), and In So Kweon (KAIST)</i>	
LidarGait: Benchmarking 3D Gait Recognition With Point Clouds	1054
<i>Chuanfu Shen (The University of Hong Kong; Southern University of Science and Technology), Chao Fan (Southern University of Science and Technology; Southern University of Science and Technology), Wei Wu (Southern University of Science and Technology), Rui Wang (Southern University of Science and Technology), George Q. Huang (The Hong Kong Polytechnic University), and Shiqi Yu (Southern University of Science and Technology; Southern University of Science and Technology)</i>	
Generalized UAV Object Detection via Frequency Domain Disentanglement	1064
<i>Kunyu Wang (University of Science and Technology of China, China), Xueyang Fu (University of Science and Technology of China, China), Yukun Huang (University of Science and Technology of China, China), Chengzhi Cao (University of Science and Technology of China, China), Gege Shi (University of Science and Technology of China, China), and Zheng-Jun Zha (University of Science and Technology of China, China)</i>	
Learning Compact Representations for LiDAR Completion and Generation	1074
<i>Yuwen Xiong (Waabi; University of Toronto), Wei-Chiu Ma (Waabi; Massachusetts Institute of Technology), Jingkan Wang (Waabi; University of Toronto), and Raquel Urtasun (Waabi; University of Toronto)</i>	
CXTrack: Improving 3D Point Cloud Tracking With Contextual Information	1084
<i>Tian-Xing Xu (Tsinghua University, China), Yuan-Chen Guo (Tsinghua University, China), Yu-Kun Lai (Cardiff University, United Kingdom), and Song-Hai Zhang (Tsinghua University, China)</i>	
Multispectral Video Semantic Segmentation: A Benchmark Dataset and Baseline	1094
<i>Wei Ji (University of Alberta), Jingjing Li (University of Alberta), Cheng Bian (ByteDance), Zongwei Zhou (Johns Hopkins University), Jiaying Zhao (ByteDance), Alan L. Yuille (Johns Hopkins University), and Li Cheng (University of Alberta)</i>	
LinK: Linear Kernel for LiDAR-Based 3D Perception	1105
<i>Tao Lu (Nanjing University), Xiang Ding (Nanjing University), Haisong Liu (Nanjing University), Gangshan Wu (Nanjing University), and Limin Wang (Nanjing University; Shanghai AI Lab)</i>	
Point Cloud Forecasting as a Proxy for 4D Occupancy Forecasting	1116
<i>Tarasha Khurana (Carnegie Mellon University), Peiyun Hu (Carnegie Mellon University), David Held (Carnegie Mellon University), and Deva Ramanan (Carnegie Mellon University)</i>	

Curricular Object Manipulation in LiDAR-Based Object Detection	1125
<i>Ziyue Zhu (Nankai University), Qiang Meng (Didi Chuxing), Xiao Wang (Didi Chuxing), Ke Wang (Didi Chuxing), Liujiang Yan (Didi Chuxing), and Jian Yang (Nankai University)</i>	
Delivering Arbitrary-Modal Semantic Segmentation	1136
<i>Jiaming Zhang (Karlsruhe Institute of Technology), Ruiping Liu (Karlsruhe Institute of Technology), Hao Shi (Zhejiang University), Kailun Yang (Hunan University), Simon Reiß (Karlsruhe Institute of Technology), Kunyu Peng (Karlsruhe Institute of Technology), Haodong Fu (Beihang University), Kaiwei Wang (Zhejiang University), and Rainer Stiefelhagen (Karlsruhe Institute of Technology)</i>	
Robust Outlier Rejection for 3D Registration With Variational Bayes	1148
<i>Haobo Jiang (Nanjing University of Science and Technology, China), Zheng Dang (EPFL, Switzerland), Zhen Wei (EPFL, Switzerland), Jin Xie (Nanjing University of Science and Technology, China), Jian Yang (Nanjing University of Science and Technology, China), and Mathieu Salzmann (EPFL, Switzerland)</i>	
3D Human Keypoints Estimation From Point Clouds in the Wild Without Human Labels	1158
<i>Zhenzhen Weng (Stanford University), Alexander S. Gorban (Waymo), Jingwei Ji (Waymo), Mahyar Najibi (Waymo), Yin Zhou (Waymo), and Dragomir Anguelov (Waymo)</i>	
Self-Supervised Pre-Training With Masked Shape Prediction for 3D Scene Understanding	1168
<i>Li Jiang (Max Planck Institute for Informatics, Saarland Informatics Campus), Zetong Yang (CUHK), Shaoshuai Shi (Max Planck Institute for Informatics, Saarland Informatics Campus), Vladislav Golyanik (Max Planck Institute for Informatics, Saarland Informatics Campus), Dengxin Dai (Max Planck Institute for Informatics, Saarland Informatics Campus), and Bernt Schiele (Max Planck Institute for Informatics, Saarland Informatics Campus)</i>	
ULIP: Learning a Unified Representation of Language, Images, and Point Clouds for 3D Understanding	1179
<i>Le Xue (Salesforce Research, Palo Alto, USA), Mingfei Gao (Salesforce Research, Palo Alto, USA), Chen Xing (Salesforce Research, Palo Alto, USA), Roberto Martín-Martín (Salesforce Research, Palo Alto, USA), Jiajun Wu (Salesforce Research, Palo Alto, USA; UT Austin, Texas, USA), Caiming Xiong (Stanford University, Stanford, USA), Ran Xu (Salesforce Research, Palo Alto, USA), Juan Carlos Niebles (Salesforce Research, Palo Alto, USA), and Silvio Savarese (Salesforce Research, Palo Alto, USA)</i>	
Open-Vocabulary Point-Cloud Object Detection Without 3D Annotation	1190
<i>Yuheng Lu (Peking University), Chenfeng Xu (University of California Berkeley), Xiaobao Wei (Peking University), Xiaodong Xie (Peking University), Masayoshi Tomizuka (University of California Berkeley), Kurt Keutzer (University of California Berkeley), and Shanghang Zhang (Peking University)</i>	
FlatFormer: Flattened Window Attention for Efficient Point Cloud Transformer	1200
<i>Zhijian Liu (MIT), Xinyu Yang (MIT; Shanghai Jiao Tong University), Haotian Tang (MIT), Shang Yang (MIT; Tsinghua University), and Song Han (MIT)</i>	

PointCMP: Contrastive Mask Prediction for Self-Supervised Learning on Point Cloud Videos .	1212
<i>Zhiqiang Shen (Shanghai Jiao Tong University; CloudWalk), Xiaoxiao Sheng (Shanghai Jiao Tong University), Longguang Wang (Aviation University of Air Force), Yulan Guo (Sun Yat-sen University), Qiong Liu (CloudWalk), and Xi Zhou (CloudWalk)</i>	
E2PN: Efficient SE(3)-Equivariant Point Network	1223
<i>Minghan Zhu (University of Michigan), Maani Ghaffari (University of Michigan), William A. Clark (Cornell University), and Huei Peng (University of Michigan)</i>	
Poly-PC: A Polyhedral Network for Multiple Point Cloud Tasks at Once	1233
<i>Tao Xie (Harbin Institute of Technology; SenseTime Research), Shiguang Wang (University of Electronic Science and Technology of China; SenseTime Research), Ke Wang (Harbin Institute of Technology), Linqi Yang (Harbin Institute of Technology), Zhiqiang Jiang (Harbin Institute of Technology), Xingcheng Zhang (SenseTime Research), Kun Dai (Harbin Institute of Technology), Ruifeng Li (Harbin Institute of Technology), and Jian Cheng (University of Electronic Science and Technology of China)</i>	
Improving Graph Representation for Point Cloud Segmentation via Attentive Filtering	1244
<i>Nan Zhang (Peking University), Zhiyi Pan (Peking University), Thomas H. Li (Peking University), Wei Gao (Peking University), and Ge Li (Peking University)</i>	
BUFFER: Balancing Accuracy, Efficiency, and Generalizability in Point Cloud Registration	1255
<i>Sheng Ao (Sun Yat-sen University), Qingyong Hu (University of Oxford), Hanyun Wang (Information Engineering University), Kai Xu (National University of Defense Technology), and Yulan Guo (Sun Yat-sen University)</i>	
TopDiG: Class-Agnostic Topological Directional Graph Extraction From Remote Sensing Images.....	1265
<i>Bingnan Yang (Wuhan University, China), Mi Zhang (Wuhan University, China), Zhan Zhang (Wuhan University, China), Zhili Zhang (Wuhan University, China), and Xiangyun Hu (Wuhan University, China)</i>	
Recognizing Rigid Patterns of Unlabeled Point Clouds by Complete and Continuous Isometry Invariants With No False Negatives and No False Positives	1275
<i>Daniel Widdowson (University of Liverpool, UK) and Vitaliy Kurlin (University of Liverpool, UK)</i>	
Both Style and Distortion Matter: Dual-Path Unsupervised Domain Adaptation for Panoramic Semantic Segmentation	1285
<i>Xu Zheng (Northeastern University), Jinjing Zhu (AI Thrust, HKUST(GZ)), Yexin Liu (AI Thrust, HKUST(GZ)), Zidong Cao (AI Thrust, HKUST(GZ)), Chong Fu (Northeastern University; Key Laboratory of Intelligent Computing in Medical Image, Ministry of Education, China), and Lin Wang (AI Thrust, HKUST(GZ); Dept. of CSE, HKUST)</i>	
CCuantuMM: Cycle-Consistent Quantum-Hybrid Matching of Multiple Shapes	1296
<i>Harshil Bhatia (Indian Institute of Technology, Jodhpur; SIC), Edith Tretschk (SIC), Zorah Löhner (Universität Siegen), Marcel Seelbach Benkner (Universität Siegen), Michael Moeller (Universität Siegen), Christian Theobalt (SIC), and Vladislav Golyanik (SIC)</i>	

Enhancing Deformable Local Features by Jointly Learning To Detect and Describe Keypoints	1306
<i>Guilherme Potje (Universidade Federal de Minas Gerais), Felipe Cadar (Universidade Federal de Minas Gerais), André Araujo (Google Research), Renato Martins (Universite de Bourgogne; Universite de Lorraine, LORIA, Inria), and Erickson R. Nascimento (Universidade Federal de Minas Gerais; Microsoft)</i>	
Understanding and Improving Features Learned in Deep Functional Maps	1316
<i>Souhaib Attaiki (LIX, École Polytechnique, IP Paris) and Maks Ovsjanikov (LIX, École Polytechnique, IP Paris)</i>	
High-Frequency Stereo Matching Network	1327
<i>Haoliang Zhao (Guizhou University, China; Ghost-Valley AI Technology, China), Huizhou Zhou (Guangdong University of Technology, China; Ghost-Valley AI Technology, China), Yongjun Zhang (Guizhou University, China), Jie Chen (Peking University, China), Yitong Yang (Guizhou University, China), and Yong Zhao (Peking University, China; Ghost-Valley AI Technology, China)</i>	
Rethinking Optical Flow From Geometric Matching Consistent Perspective	1337
<i>Qiaole Dong (Fudan University), Chenjie Cao (Fudan University), and Yanwei Fu (Fudan University)</i>	
Efficient Robust Principal Component Analysis via Block Krylov Iteration and CUR Decomposition	1348
<i>Shun Fang (Wuhan University of Science and Technology, China; Wuhan University of Science and Technology, China), Zhengqin Xu (Shanghai Jiao Tong University, China), Shiqian Wu (Wuhan University of Science and Technology, China; Hubei Province Key Laboratory of Intelligent Information Processing and Real-Time Industrial Systems), and Shoulie Xie (Signal Processing, RF & Optical Dept. Institute for Infocomm Research A*STAR, Singapore)</i>	
VectorFloorSeg: Two-Stream Graph Attention Network for Vectorized Roughcast Floorplan Segmentation	1358
<i>Bingchen Yang (University of Chinese Academy of Sciences), Haiyong Jiang (University of Chinese Academy of Sciences), Hao Pan (Microsoft Research Asia), and Jun Xiao (University of Chinese Academy of Sciences)</i>	
TBP-Former: Learning Temporal Bird's-Eye-View Pyramid for Joint Perception and Prediction in Vision-Centric Autonomous Driving	1368
<i>Shaoheng Fang (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University), Zi Wang (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University), Yiqi Zhong (Department of Computer Science, University of Southern California), Junhao Ge (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University), and Siheng Chen (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University; Shanghai AI Laboratory)</i>	
Implicit Occupancy Flow Fields for Perception and Prediction in Self-Driving	1379
<i>Ben Agro (Waabi, University of Toronto), Quinlan Sykora (Waabi, University of Toronto), Sergio Casas (Waabi, University of Toronto), and Raquel Urtasun (Waabi, University of Toronto)</i>	

UniSim: A Neural Closed-Loop Sensor Simulator	1389
<i>Ze Yang (Waabi; University of Toronto), Yun Chen (Waabi; University of Toronto), Jingkang Wang (Waabi; University of Toronto), Sivabalan Manivasagam (Waabi; University of Toronto), Wei-Chiu Ma (Waabi; Massachusetts Institute of Technology), Anqi Joyce Yang (Waabi; University of Toronto), and Raquel Urtasun (Waabi; University of Toronto)</i>	
FEND: A Future Enhanced Distribution-Aware Contrastive Learning Framework for Long-Tail Trajectory Prediction	1400
<i>Yuning Wang (Xi'an Jiaotong University, China), Pu Zhang (DiDi Chuxing, China), Lei Bai (Shanghai AI Laboratory, China), and Jianru Xue (Xi'an Jiaotong University, China)</i>	
EqMotion: Equivariant Multi-Agent Motion Prediction With Invariant Interaction Reasoning ..	1410
<i>Chenxin Xu (Shanghai Jiao Tong University; National University of Singapore), Robby T. Tan (National University of Singapore), Yuhong Tan (Shanghai Jiao Tong University), Siheng Chen (Shanghai Jiao Tong University; Shanghai AI Laboratory), Yu Guang Wang (Shanghai Jiao Tong University), Xinchao Wang (National University of Singapore), and Yanfeng Wang (Shanghai AI Laboratory; Shanghai Jiao Tong University)</i>	
Lookahead Diffusion Probabilistic Models for Refining Mean Estimation	1421
<i>Guoqiang Zhang (University of Technology Sydney), Kenta Niwa (NTT Communication Science Laboratories), and W. Bastiaan Kleijn (Victoria University of Wellington)</i>	
Neural Volumetric Memory for Visual Locomotion Control	1430
<i>Ruihan Yang (UC San Diego), Ge Yang (Institute of AI and Fundamental Interactions; MIT CSAIL), and Xiaolong Wang (UC San Diego)</i>	
Gazeformer: Scalable, Effective and Fast Prediction of Goal-Directed Human Attention	1441
<i>Sounak Mondal (Stony Brook University), Zhibo Yang (Stony Brook University; Waymo LLC), Seoyoung Ahn (Stony Brook University), Dimitris Samaras (Stony Brook University), Gregory Zelinsky (Stony Brook University), and Minh Hoai (Stony Brook University; VinAI Research)</i>	
DrapeNet: Garment Generation and Self-Supervised Draping	1451
<i>Luca De Luigi (University of Bologna), Ren Li (CYLab, EPFL), Benoît Guillard (CYLab, EPFL), Mathieu Salzmann (CYLab, EPFL), and Pascal Fua (CYLab, EPFL)</i>	
Tracking Multiple Deformable Objects in Egocentric Videos	1461
<i>Mingzhen Huang (State University of New York at Buffalo; Meta Reality Labs), Xiaoxing Li (Meta Reality Labs), Jun Hu (Meta Reality Labs), Honghong Peng (Meta Reality Labs), and Siwei Lyu (Meta Reality Labs)</i>	

Good Is Bad: Causality Inspired Cloth-Debiasing for Cloth-Changing Person Re-Identification	1472
<i>Zhengwei Yang (National Engineering Research Center for Multimedia Software, Institute of Artificial Intelligence; Wuhan University; Hubei Key Laboratory of Multimedia and Network Communication Engineering), Meng Lin (National Engineering Research Center for Multimedia Software, Institute of Artificial Intelligence; Wuhan University; Hubei Key Laboratory of Multimedia and Network Communication Engineering), Xian Zhong (Wuhan University of Technology), Yu Wu (National Engineering Research Center for Multimedia Software, Institute of Artificial Intelligence; Wuhan University), and Zheng Wang (National Engineering Research Center for Multimedia Software, Institute of Artificial Intelligence; Wuhan University; Hubei Key Laboratory of Multimedia and Network Communication Engineering)</i>	
Micron-BERT: BERT-Based Facial Micro-Expression Recognition	1482
<i>Xuan-Bac Nguyen (University of Arkansas, USA), Chi Nhan Duong (Concordia University, Canada), Xin Li (West Virginia University, USA), Susan Gauch (University of Arkansas, USA), Han-Seok Seo (University of Arkansas, USA), and Khoa Luu (University of Arkansas, USA)</i>	
MARLIN: Masked Autoencoder for Facial Video Representation Learning	1493
<i>Zhixi Cai (Monash University), Shreya Ghosh (Monash University; Curtin University), Kalin Stefanov (Monash University), Abhinav Dhall (Monash University; Indian Institute of Technology Ropar), Jianfei Cai (Monash University), Hamid Rezatofighi (Monash University), Reza Haffari (Monash University), and Munawar Hayat (Monash University)</i>	
StyleSync: High-Fidelity Generalized and Personalized Lip Sync in Style-Based Generator	1505
<i>Jiazhi Guan (Department of Computer Vision Technology (VIS), Baidu Inc.; Tsinghua University), Zhanwang Zhang (Department of Computer Vision Technology (VIS), Baidu Inc.), Hang Zhou (Department of Computer Vision Technology (VIS), Baidu Inc.), Tianshu Hu (Department of Computer Vision Technology (VIS), Baidu Inc.), Kaisiyuan Wang (The University of Sydney), Dongliang He (Department of Computer Vision Technology (VIS), Baidu Inc.), Haocheng Feng (Department of Computer Vision Technology (VIS), Baidu Inc.), Jingtuo Liu (Department of Computer Vision Technology (VIS), Baidu Inc.), Errui Ding (Department of Computer Vision Technology (VIS), Baidu Inc.), Ziwei Liu (Nanyang Technological University), and Jingdong Wang (Department of Computer Vision Technology (VIS), Baidu Inc.)</i>	
ReallImpact: A Dataset of Impact Sound Fields for Real Objects	1516
<i>Samuel Clarke (Stanford University), Ruohan Gao (Stanford University), Mason Wang (Stanford University), Mark Rau (Stanford University), Julia Xu (Stanford University), Jui-Hsien Wang (Adobe Research), Doug L. James (Stanford University), and Jiajun Wu (Stanford University)</i>	
STMT: A Spatial-Temporal Mesh Transformer for MoCap-Based Action Recognition	1526
<i>Xiaoyu Zhu (Carnegie Mellon University), Po-Yao Huang (FAIR, Meta AI), Junwei Liang (HKUST), Celso M. de Melo (DEVCOM Army Research Laboratory), and Alexander G. Hauptmann (Carnegie Mellon University)</i>	

Progressive Spatio-Temporal Alignment for Efficient Event-Based Motion Estimation	1537
<i>Xueyan Huang (University of Science and Technology of China), Yueyi Zhang (University of Science and Technology of China), and Zhiwei Xiong (University of Science and Technology of China)</i>	
Event-Based Shape From Polarization	1547
<i>Manasi Muglikar (University of Zurich, Switzerland), Leonard Bauersfeld (University of Zurich, Switzerland), Diederik Paul Moeys (Advanced Sensors and Modelling Group, SONY R&D Center Europe, SL1), and Davide Scaramuzza (University of Zurich, Switzerland)</i>	
Learning Spatial-Temporal Implicit Neural Representations for Event-Guided Video Super-Resolution	1557
<i>Yunfan Lu (HKUST(GZ)), Zipeng Wang (HKUST(GZ)), Minjie Liu (HKUST(GZ)), Hongjian Wang (Tsinghua University), and Lin Wang (HKUST(GZ); HKUST)</i>	
BiFormer: Learning Bilateral Motion Estimation via Bilateral Transformer for 4K Video Frame Interpolation	1568
<i>Junheum Park (Korea University), Jintae Kim (Korea University), and Chang-Su Kim (Korea University)</i>	
A Unified Pyramid Recurrent Network for Video Frame Interpolation	1578
<i>Xin Jin (Samsung Electronics (China) R&D Center), Longhai Wu (Samsung Electronics (China) R&D Center), Jie Chen (Samsung Electronics (China) R&D Center), Youxin Chen (Samsung Electronics (China) R&D Center), Jayoon Koo (Samsung Electronics, South Korea), and Cheul-hee Hahm (Samsung Electronics, South Korea)</i>	
Event-Based Blurry Frame Interpolation Under Blind Exposure	1588
<i>Wenming Weng (University of Science and Technology of China, China), Yueyi Zhang (University of Science and Technology of China, China), and Zhiwei Xiong (University of Science and Technology of China, China)</i>	
FlowFormer++: Masked Cost Volume Autoencoding for Pretraining Optical Flow Estimation .	1599
<i>Xiaoyu Shi (CUHK MMLab; NVIDIA AI Technology Center), Zhaoyang Huang (CUHK MMLab; NVIDIA AI Technology Center), Dasong Li (CUHK MMLab), Manyuan Zhang (CUHK MMLab), Ka Chun Cheung (NVIDIA AI Technology Center), Simon See (NVIDIA AI Technology Center), Hongwei Qin (SenseTime Research), Jifeng Dai (Tsinghua University), and Hongsheng Li (CUHK MMLab; CII under InnoHK)</i>	
POTTER: Pooling Attention Transformer for Efficient Human Mesh Recovery	1611
<i>Ce Zheng (University of Central Florida), Xianpeng Liu (North Carolina State University), Guo-Jun Qi (OPPO Seattle Research Center, USA; Westlake University), and Chen Chen (University of Central Florida)</i>	
Adaptive Patch Deformation for Textureless-Resilient Multi-View Stereo	1621
<i>Yuesong Wang (Huazhong University of Science & Technology), Zhaojie Zeng (Huazhong University of Science & Technology), Tao Guan (Huazhong University of Science & Technology), Wei Yang (Huazhong University of Science & Technology), Zhuo Chen (Huazhong University of Science & Technology), Wenkai Liu (Huazhong University of Science & Technology), Luoyuan Xu (Huazhong University of Science & Technology), and Yawei Luo (Huazhong University of Science & Technology)</i>	

On the Difficulty of Unpaired Infrared-to-Visible Video Translation: Fine-Grained Content-Rich Patches Transfer	1631
<i>Zhenjie Yu (Beijing Institute of Technology), Shuang Li (Beijing Institute of Technology), Yirui Shen (Beijing Institute of Technology), Chi Harold Liu (Beijing Institute of Technology), and Shuigen Wang (Yantai IRay Technologies Lt. Co.)</i>	
Thermal Spread Functions (TSF): Physics-Guided Material Classification	1641
<i>Aniket Dashpute (Electrical and Computer Engineering; Rice University), Vishwanath Saragadam (Rice University), Emma Alexander (Northwestern University), Florian Willomitzer (University of Arizona), Aggelos Katsaggelos (Electrical and Computer Engineering), Ashok Veeraraghavan (Rice University), and Oliver Cossairt (Electrical and Computer Engineering; Northwestern University)</i>	
Better “CMOS” Produces Clearer Images: Learning Space-Variant Blur Estimation for Blind Image Super-Resolution	1651
<i>Xuhai Chen (Zhejiang University), Jiangning Zhang (Youtu Lab, Tencent), Chao Xu (Zhejiang University), Yabiao Wang (Youtu Lab, Tencent), Chengjie Wang (Youtu Lab, Tencent), and Yong Liu (Zhejiang University)</i>	
Learning Semantic-Aware Knowledge Guidance for Low-Light Image Enhancement	1662
<i>Yuhui Wu (University of Electronic Science and Technology of China, China), Chen Pan (University of Electronic Science and Technology of China, China), Guoqing Wang (University of Electronic Science and Technology of China, China), Yang Yang (University of Electronic Science and Technology of China, China), Jiwei Wei (University of Electronic Science and Technology of China, China), Chongyi Li (Nanyang Technological University, Singapore), and Heng Tao Shen (University of Electronic Science and Technology of China, China)</i>	
CutMIB: Boosting Light Field Super-Resolution via Multi-View Image Blending	1672
<i>Zeyu Xiao (University of Science and Technology of China), Yutong Liu (University of Science and Technology of China), Ruisheng Gao (University of Science and Technology of China), and Zhiwei Xiong (University of Science and Technology of China)</i>	
sRGB Real Noise Synthesizing With Neighboring Correlation-Aware Noise Model	1683
<i>Zixuan Fu (Nanyang Technological University, Singapore), Lanqing Guo (Nanyang Technological University, Singapore), and Bihan Wen (Nanyang Technological University, Singapore)</i>	
Masked Image Training for Generalizable Deep Image Denoising	1692
<i>Haoyu Chen (The Hong Kong University of Science and Technology, Guangzhou), Jinjin Gu (Shanghai AI Lab; The University of Sydney), Yihao Liu (Shanghai AI Lab; Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Salma Abdel Magid (The Hong Kong University of Science and Technology), Chao Dong (Shanghai AI Lab; Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), Qiong Wang (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), Hanspeter Pfister (Harvard University), and Lei Zhu (Shanghai AI Lab; The Hong Kong University of Science and Technology)</i>	

DR2: Diffusion-Based Robust Degradation Remover for Blind Face Restoration	1704
<i>Zhixin Wang (CMIC, Shanghai Jiao Tong University), Ziyang Zhang (CMIC, Shanghai Jiao Tong University), Xiaoyun Zhang (CMIC, Shanghai Jiao Tong University), Huangjie Zheng (The University of Texas at Austin), Mingyuan Zhou (The University of Texas at Austin), Ya Zhang (Shanghai AI Lab), and Yanfeng Wang (CMIC, Shanghai Jiao Tong University; Shanghai AI Lab)</i>	
Learning Distortion Invariant Representation for Image Restoration From a Causality Perspective	1714
<i>Xin Li (University of Science and Technology of China), Bingchen Li (University of Science and Technology of China), Xin Jin (Eastern Institute for Advanced Study), Cuiling Lan (Microsoft Research Asia), and Zhibo Chen (University of Science and Technology of China)</i>	
Perception-Oriented Single Image Super-Resolution Using Optimal Objective Estimation	1725
<i>Seung Ho Park (Seoul National University, Korea; Samsung Electronics, Korea), Young Su Moon (Samsung Electronics, Korea), and Nam Ik Cho (Seoul National University, Korea)</i>	
Catch Missing Details: Image Reconstruction With Frequency Augmented Variational Autoencoder	1736
<i>Xinmiao Lin (Rochester Institute of Technology), Yikang Li (OPPO US Research), Jenhao Hsiao (OPPO US Research), Chiuman Ho (OPPO US Research), and Yu Kong (Michigan State University)</i>	
MD-VQA: Multi-Dimensional Quality Assessment for UGC Live Videos	1746
<i>Zicheng Zhang (Shanghai Jiao Tong University), Wei Wu (Alibaba Group), Wei Sun (Shanghai Jiao Tong University), Danyang Tu (Shanghai Jiao Tong University), Wei Lu (Shanghai Jiao Tong University), Xiongkuo Min (Shanghai Jiao Tong University), Ying Chen (Alibaba Group), and Guangtao Zhai (Shanghai Jiao Tong University; Shanghai Jiao Tong University)</i>	
CABM: Content-Aware Bit Mapping for Single Image Super-Resolution Network With Large Input..	1756
<i>Senmao Tian (Beijing Jiaotong University; OPPO Research Institute), Ming Lu (Intel Labs China), Jiaming Liu (OPPO Research Institute; Peking University), Yandong Guo (Beijing University of Posts and Telecommunications), Yurong Chen (Intel Labs China), and Shunli Zhang (Beijing Jiaotong University)</i>	
Initialization Noise in Image Gradients and Saliency Maps	1766
<i>Ann-Christin Woerl (Johannes Gutenberg University Mainz, Germany), Jan Disselhoff (Johannes Gutenberg University Mainz, Germany), and Michael Wand (Johannes Gutenberg University Mainz, Germany)</i>	
Local Implicit Normalizing Flow for Arbitrary-Scale Image Super-Resolution	1776
<i>Jie-En Yao (National Tsing Hua University), Li-Yuan Tsao (National Tsing Hua University), Yi-Chen Lo (MediaTek Inc.), Roy Tseng (MediaTek Inc.), Chia-Che Chang (MediaTek Inc.), and Chun-Yi Lee (National Tsing Hua University)</i>	

Deep Arbitrary-Scale Image Super-Resolution via Scale-Equivariance Pursuit	1786
<i>Xiaohang Wang (Shanghai Jiao Tong University, China), Xuanhong Chen (Shanghai Jiao Tong University, China), Bingbing Ni (Shanghai Jiao Tong University, China), Hang Wang (Huawei), Zhengyan Tong (Shanghai Jiao Tong University, China), and Yutian Liu (Shanghai Jiao Tong University, China)</i>	
CiaoSR: Continuous Implicit Attention-in-Attention Network for Arbitrary-Scale Image Super-Resolution	1796
<i>Jiezhong Cao (ETHZürich), Qin Wang (ETHZürich), Yongqin Xian (ETHZürich), Yawei Li (ETHZürich), Bingbing Ni (Huawei Inc.), Zhiming Pi (Huawei Inc.), Kai Zhang (ETHZürich), Yulun Zhang (ETHZürich), Radu Timofte (ETHZürich; University of Würzburg), and Luc Van Gool (ETHZürich; KU Leuven)</i>	
Multiplicative Fourier Level of Detail	1808
<i>Yishun Dou (Shanghai Jiao Tong University, China), Zhong Zheng (Shanghai Jiao Tong University, China), Qiaoqiao Jin (Huawei), and Bingbing Ni (Shanghai Jiao Tong University, China; Huawei)</i>	
Document Image Shadow Removal Guided by Color-Aware Background	1818
<i>Ling Zhang (Wuhan University of Science and Technology), Yinghao He (Wuhan University of Science and Technology), Qing Zhang (Sun Yat-Sen University), Zheng Liu (China University of Geosciences), Xiaolong Zhang (Wuhan University of Science and Technology), and Chunxia Xiao (Wuhan University)</i>	
StyleRes: Transforming the Residuals for Real Image Editing With StyleGAN	1828
<i>Hamza Pehlivan (Bilkent University), Yusuf Dalva (Bilkent University), and Aysegül Dundar (Bilkent University)</i>	
TopNet: Transformer-Based Object Placement Network for Image Compositing	1838
<i>Sijie Zhu (University of Central Florida), Zhe Lin (Adobe Research), Scott Cohen (Adobe Research), Jason Kuen (Adobe Research), Zhifei Zhang (Adobe Research), and Chen Chen (University of Central Florida)</i>	
VecFontSDF: Learning To Reconstruct and Synthesize High-Quality Vector Fonts via Signed Distance Functions	1848
<i>Zeqing Xia (Peking University, China; Baidu Inc.), Bojun Xiong (Peking University, China), and Zhouhui Lian (Peking University, China)</i>	
CF-Font: Content Fusion for Few-Shot Font Generation	1858
<i>Chi Wang (Zhejiang University; Alibaba Group), Min Zhou (Alibaba Group), Tiezheng Ge (Alibaba Group), Yuning Jiang (Alibaba Group), Hujun Bao (Zhejiang University), and Weiwei Xu (Zhejiang University)</i>	
SIEDOB: Semantic Image Editing by Disentangling Object and Background	1868
<i>Wuyang Luo (Fudan University), Su Yang (Fudan University), Xinjian Zhang (Fudan University), and Weishan Zhang (China University of Petroleum)</i>	
MaskSketch: Unpaired Structure-Guided Masked Image Generation	1879
<i>Dina Bashkirova (Boston University), José Lezama (Google Research), Kihyuk Sohn (Google Research), Kate Saenko (Boston University; MIT-IBM Watson AI Lab), and Irfan Essa (Google Research; Georgia Institute of Technology)</i>	

Text2Scene: Text-Driven Indoor Scene Stylization With Part-Aware Details	1890
<i>Inwoo Hwang (Seoul National University), Hyeonwoo Kim (Seoul National University), and Young Min Kim (Seoul National University; Seoul National University)</i>	
Uncovering the Disentanglement Capability in Text-to-Image Diffusion Models	1900
<i>Qiucheng Wu (UC, Santa Barbara), Yujian Liu (UC, Santa Barbara), Handong Zhao (Adobe Research), Ajinkya Kale (Adobe Research), Trung Bui (Adobe Research), Tong Yu (Adobe Research), Zhe Lin (Adobe Research), Yang Zhang (MIT-IBM Watson AI Lab), and Shiyu Chang (UC, Santa Barbara)</i>	
VectorFusion: Text-to-SVG by Abstracting Pixel-Based Diffusion Models	1911
<i>Ajay Jain (UC Berkeley), Amber Xie (UC Berkeley), and Pieter Abbeel (UC Berkeley)</i>	
Plug-and-Play Diffusion Features for Text-Driven Image-to-Image Translation	1921
<i>Narek Tumanyan (Weizmann Institute of Science), Michal Geyer (Weizmann Institute of Science), Shai Bagon (Weizmann Institute of Science), and Tali Dekel (Weizmann Institute of Science)</i>	
Multi-Concept Customization of Text-to-Image Diffusion	1931
<i>Nupur Kumari (Carnegie Mellon University), Bingliang Zhang (Tsinghua University), Richard Zhang (Adobe Research), Eli Shechtman (Adobe Research), and Jun-Yan Zhu (Carnegie Mellon University)</i>	
Unifying Layout Generation With a Decoupled Diffusion Model	1942
<i>Mude Hui (Xi'an Jiaotong University), Zhizheng Zhang (Microsoft Research Asia), Xiaoyi Zhang (Microsoft Research Asia), Wenxuan Xie (Microsoft Research Asia), Yuwang Wang (Tsinghua University), and Yan Lu (Microsoft Research Asia)</i>	
BBDM: Image-to-Image Translation With Brownian Bridge Diffusion Models	1952
<i>Bo Li (Nanchang Hangkong University, China), Kaitao Xue (Nanchang Hangkong University, China), Bin Liu (Nanchang Hangkong University, China), and Yu-Kun Lai (Cardiff University, UK)</i>	
Towards Practical Plug-and-Play Diffusion Models	1962
<i>Hyojun Go (Riiid AI Research), Yunsung Lee (Riiid AI Research), Jin-Young Kim (Riiid AI Research), Seunghyun Lee (Riiid AI Research), Myeongho Jeong (Riiid AI Research), Hyun Seung Lee (Riiid AI Research), and Seungtaek Choi (Riiid AI Research)</i>	
Post-Training Quantization on Diffusion Models	1972
<i>Yuzhang Shang (Illinois Institute of Technology; Cisco Research), Zhihang Yuan (Houmo AI), Bin Xie (Illinois Institute of Technology), Bingzhe Wu (Tencent AI Lab), and Yan Yan (Illinois Institute of Technology)</i>	
DiffTalk: Crafting Diffusion Models for Generalized Audio-Driven Portraits Animation	1982
<i>Shuai Shen (Tsinghua University), Wenliang Zhao (Tsinghua University), Zibin Meng (Tsinghua University), Wanhua Li (Tsinghua University), Zheng Zhu (PhiGent Robotics), Jie Zhou (Tsinghua University), and Jiwen Lu (Tsinghua University)</i>	
Mask-Guided Matting in the Wild	1992
<i>Kwanyong Park (KAIST), Sanghyun Woo (KAIST), Seoung Wug Oh (Adobe Research), In So Kweon (KAIST), and Joon-Young Lee (Adobe Research)</i>	

Not All Image Regions Matter: Masked Vector Quantization for Autoregressive Image Generation	2002
<i>Mengqi Huang (University of Science and Technology of China, China), Zhendong Mao (University of Science and Technology of China, China; Institute of Artificial intelligence, Hefei Comprehensive National Science Center, China), Quan Wang (Beijing University of Posts and Telecommunications, China), and Yongdong Zhang (University of Science and Technology of China, China; Institute of Artificial intelligence, Hefei Comprehensive National Science Center, China)</i>	
Compression-Aware Video Super-Resolution	2012
<i>Yingwei Wang (Dalian University of Technology), Takashi Isobe (Kuaishou Technology), Xu Jia (Xiaohongshu Inc.), Xin Tao (Dalian University of Technology), Huchuan Lu (Dalian University of Technology; Peng Cheng Laboratory), and Yu-Wing Tai (The Hong Kong University of Science and Technology)</i>	
Neural Rate Estimator and Unsupervised Learning for Efficient Distributed Image Analytics in Split-DNN Models	2022
<i>Nilesh Ahuja (Intel Labs), Parual Datta (Intel Labs), Bhavya Kanzariya (Intel Labs; IIT Hyderabad), V. Srinivasa Somayazulu (Intel Labs), and Omesh Tickoo (Intel Labs)</i>	
DNeRV: Modeling Inherent Dynamics via Difference Neural Representation for Videos	2031
<i>Qi Zhao (Nanjing University), M. Salman Asif (Nanjing University), and Zhan Ma (Nanjing University)</i>	
Polynomial Implicit Neural Representations for Large Diverse Datasets	2041
<i>Rajhans Singh (Geometric Media Lab, Arizona State University), Ankita Shukla (Geometric Media Lab, Arizona State University), and Pavan Turaga (Geometric Media Lab, Arizona State University)</i>	
Learning Decorrelated Representations Efficiently Using Fast Fourier Transform	2052
<i>Yutaro Shigeto (Chiba Institute of Technology, Japan), Masashi Shimbo (Chiba Institute of Technology, Japan), Yuya Yoshikawa (Chiba Institute of Technology, Japan), and Akikazu Takeuchi (Chiba Institute of Technology, Japan)</i>	
SparseViT: Revisiting Activation Sparsity for Efficient High-Resolution Vision Transformer	2061
<i>Xuanyao Chen (Shanghai Qi Zhi Institute; Fudan University), Zhijian Liu (MIT), Haotian Tang (MIT), Li Yi (Shanghai Qi Zhi Institute; Tsinghua University), Hang Zhao (Shanghai Qi Zhi Institute; Tsinghua University), and Song Han (MIT)</i>	
N-Gram in Swin Transformers for Efficient Lightweight Image Super-Resolution	2071
<i>Haram Choi (Sogang University), Jeongmin Lee (LG Innotek), and Jihoon Yang (Sogang University)</i>	
Slide-Transformer: Hierarchical Vision Transformer With Local Self-Attention	2082
<i>Xuran Pan (BNRist, Tsinghua University), Tianzhu Ye (BNRist, Tsinghua University), Zhuofan Xia (BNRist, Tsinghua University), Shiji Song (BNRist, Tsinghua University), and Gao Huang (BNRist, Tsinghua University)</i>	

Joint Token Pruning and Squeezing Towards More Aggressive Compression of Vision Transformers	2092
<i>Siyuan Wei (MEGVII Technology), Tianzhu Ye (Tsinghua University), Shen Zhang (MEGVII Technology), Yao Tang (MEGVII Technology), and Jiajun Liang (MEGVII Technology)</i>	
Top-Down Visual Attention From Analysis by Synthesis	2102
<i>Baifeng Shi (UC Berkeley), Trevor Darrell (UC Berkeley), and Xin Wang (Microsoft Research)</i>	
Probing Neural Representations of Scene Perception in a Hippocampally Dependent Task Using Artificial Neural Networks	2113
<i>Markus Frey (NTNU, Norway; Max-Planck-Institute for Human Cognitive and Brain Sciences, Germany), Christian F. Doeller (NTNU, Norway; Max-Planck-Institute for Human Cognitive and Brain Sciences, Germany), and Caswell Barry (UCL, United Kingdom)</i>	
Masked Image Modeling With Local Multi-Scale Reconstruction	2122
<i>Haoqing Wang (Peking University), Yehui Tang (Peking University; Huawei Noah's Ark Lab), Yunhe Wang (Huawei Noah's Ark Lab), Jianyuan Guo (Huawei Noah's Ark Lab), Zhi-Hong Deng (Peking University), and Kai Han (Huawei Noah's Ark Lab)</i>	
Siamese Image Modeling for Self-Supervised Vision Representation Learning	2132
<i>Chenxin Tao (Tsinghua University), Xizhou Zhu (SenseTime Research; Shanghai Artificial Intelligence Laboratory), Weijie Su (University of Science and Technology of China), Gao Huang (Tsinghua University), Bin Li (University of Science and Technology of China), Jie Zhou (Tsinghua University), Yu Qiao (Shanghai Artificial Intelligence Laboratory), Xiaogang Wang (The Chinese University of Hong Kong), and Jifeng Dai (Tsinghua University; Shanghai Artificial Intelligence Laboratory)</i>	
MAGE: MAsked Generative Encoder To Unify Representation Learning and Image Synthesis .	2142
<i>Tianhong Li (MIT CSAIL), Huiwen Chang (Google Research), Shlok Mishra (University of Maryland), Han Zhang (Google Research), Dina Katabi (MIT CSAIL), and Dilip Krishnan (Google Research)</i>	
Diverse Embedding Expansion Network and Low-Light Cross-Modality Benchmark for Visible-Infrared Person Re-Identification	2153
<i>Yukang Zhang (Xiamen University, China) and Hanzi Wang (Xiamen University, China; Shanghai Artificial Intelligence Laboratory, China)</i>	
DistilPose: Tokenized Pose Regression With Heatmap Distillation	2163
<i>Suhang Ye (Xiamen University), Yingyi Zhang (Tencent Youtu Lab), Jie Hu (Xiamen University), Liujuan Cao (Xiamen University), Shengchuan Zhang (Xiamen University), Lei Shen (Tencent Youtu Lab), Jun Wang (Tencent Youtu Lab), Shouhong Ding (Tencent WeChat Pay Lab), and Rongrong Ji (Xiamen University)</i>	
Graph Transformer GANs for Graph-Constrained House Generation	2173
<i>Hao Tang (CVL, ETH Zurich), Zhenyu Zhang (Tencent Youtu Lab), Humphrey Shi (U of Oregon & UIUC & Picsart AI Research), Bo Li (Tencent Youtu Lab), Ling Shao (UCAS-Terminus AI Lab, UCAS), Nicu Sebe (University of Trento), Radu Timofte (CVL, ETH Zurich; University of Wurzburg), and Luc Van Gool (CVL, ETH Zurich; KU Leuven)</i>	

Automatic High Resolution Wire Segmentation and Removal	2183
<i>Mang Tik Chiu (UIUC; Adobe), Xuaner Zhang (Adobe), Zijun Wei (Adobe), Yuqian Zhou (Adobe), Eli Shechtman (Adobe), Connelly Barnes (Adobe), Zhe Lin (Adobe), Florian Kainz (Adobe), Sohrab Amirghodsi (Adobe), and Humphrey Shi (UIUC; University of Oregon)</i>	
Tree Instance Segmentation With Temporal Contour Graph	2193
<i>Adnan Firoze (Purdue University), Cameron Wingren (Purdue University), Raymond A. Yeh (Purdue University), Bedrich Benes (Purdue University), and Daniel Aliaga (Purdue University)</i>	
Dual-Path Adaptation From Image to Video Transformers	2203
<i>Jungin Park (Yonsei University), Jiyoung Lee (NAVER AI Lab), and Kwanghoon Sohn (Yonsei University; Korea Institute of Science and Technology (KIST))</i>	
Rethinking Video ViTs: Sparse Video Tubes for Joint Image and Video Learning	2214
<i>Aj Piergiovanni (Google Research), Weicheng Kuo (Google Research), and Anelia Angelova (Google Research)</i>	
Modeling Video As Stochastic Processes for Fine-Grained Video Representation Learning	2225
<i>Heng Zhang (Renmin University of China; Beijing Key Laboratory of Big Data Management and Analysis Methods), Daqing Liu (JD.com), Qi Zheng (The University of Sydney), and Bing Su (Renmin University of China; Beijing Key Laboratory of Big Data Management and Analysis Methods)</i>	
Masked Motion Encoding for Self-Supervised Video Representation Learning	2235
<i>Xinyu Sun (South China University of Technology; Information Technology R&D Innovation Center of Peking University), Peihao Chen (South China University of Technology), Liangwei Chen (South China University of Technology), Changhao Li (South China University of Technology), Thomas H. Li (Peking University Shenzhen Graduate School), Mingkui Tan (South China University of Technology; Key Laboratory of Big Data and Intelligent Robot, Ministry of Education; Pazhou Laboratory), and Chuang Gan (UMass Amherst; MIT-IBM Watson AI Lab)</i>	
Boosting Video Object Segmentation via Space-Time Correspondence Learning	2246
<i>Yurong Zhang (Shanghai Jiao Tong University), Liulei Li (Zhejiang University), Wenguan Wang (Zhejiang University), Rong Xie (Shanghai Jiao Tong University), Li Song (Shanghai Jiao Tong University), and Wenjun Zhang (Shanghai Jiao Tong University)</i>	
Two-Shot Video Object Segmentation	2257
<i>Kun Yan (Peking University), Xiao Li (Microsoft Research Asia), Fangyun Wei (Microsoft Research Asia), Jinglu Wang (Microsoft Research Asia), Chenbin Zhang (Peking University), Ping Wang (Peking University), and Yan Lu (Microsoft Research Asia)</i>	

Look Before You Match: Instance Understanding Matters in Video Object Segmentation	2268
<i>Junke Wang (Fudan University), Dongdong Chen (Microsoft Cloud + AI), Zuxuan Wu (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Chong Luo (Microsoft Research Asia), Chuanxin Tang (Microsoft Research Asia), Xiyang Dai (Microsoft Cloud + AI), Yucheng Zhao (Microsoft Research Asia), Yujia Xie (Microsoft Cloud + AI), Lu Yuan (Microsoft Cloud + AI), and Yu-Gang Jiang (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing)</i>	
Spatial-Then-Temporal Self-Supervised Learning for Video Correspondence	2279
<i>Rui Li (University of Science and Technology of China, China) and Dong Liu (University of Science and Technology of China, China)</i>	
Few-Shot Referring Relationships in Videos	2289
<i>Yogesh Kumar (Indian Institute of Technology Jodhpur) and Anand Mishra (Indian Institute of Technology Jodhpur)</i>	
Vision Transformers Are Parameter-Efficient Audio-Visual Learners	2299
<i>Yan-Bo Lin (Department of Computer Science, UNC Chapel Hill), Yi-Lin Sung (Department of Computer Science, UNC Chapel Hill), Jie Lei (Department of Computer Science, UNC Chapel Hill), Mohit Bansal (Department of Computer Science, UNC Chapel Hill), and Gedas Bertasius (Department of Computer Science, UNC Chapel Hill)</i>	
Egocentric Video Task Translation	2310
<i>Zihui Xue (The University of Texas at Austin; FAIR, Meta AI), Yale Song (FAIR, Meta AI), Kristen Grauman (The University of Texas at Austin; FAIR, Meta AI), and Lorenzo Torresani (FAIR, Meta AI)</i>	
QPGesture: Quantization-Based and Phase-Guided Motion Matching for Natural Speech-Driven Gesture Generation	2321
<i>Sicheng Yang (Tsinghua University), Zhiyong Wu (Tsinghua University; The Chinese University of Hong Kong), Minglei Li (Huawei Cloud Computing Technologies Co., Ltd), Zhensong Zhang (Huawei Noah's Ark Lab), Lei Hao (Huawei Noah's Ark Lab), Weihong Bao (Tsinghua University), and Haolin Zhuang (Tsinghua University)</i>	
Co-Speech Gesture Synthesis by Reinforcement Learning With Contrastive Pre-Trained Rewards...	2331
<i>Mingyang Sun (Dalian University of Technology, China), Mengchen Zhao (Noah's Ark Lab, China), Yaqing Hou (Dalian University of Technology, China), Minglei Li (Huawei Cloud Computing Technologies Co., Ltd, China), Huang Xu (Huawei Cloud Computing Technologies Co., Ltd, China), Songcen Xu (Noah's Ark Lab, China), and Jianye Hao (Noah's Ark Lab, China; Tianjin University, China)</i>	
TimeBalance: Temporally-Invariant and Temporally-Distinctive Video Representations for Semi-Supervised Action Recognition	2341
<i>Ishan Rajendrakumar Dave (University of Central Florida, USA), Mamshad Nayeem Rizve (University of Central Florida, USA), Chen Chen (University of Central Florida, USA), and Mubarak Shah (University of Central Florida, USA)</i>	
How Can Objects Help Action Recognition?	2353
<i>Xingyi Zhou (Meta AI), Anurag Arnab (Google Research), Chen Sun (Google Research), and Cordelia Schmid (Google Research)</i>	

Actionlet-Dependent Contrastive Learning for Unsupervised Skeleton-Based Action Recognition	2363
<i>Lilang Lin (Wangxuan Institute of Computer Technology, Peking University, China), Jiahang Zhang (Wangxuan Institute of Computer Technology, Peking University, China), and Jiaying Liu (Wangxuan Institute of Computer Technology, Peking University, China)</i>	
Decomposed Cross-Modal Distillation for RGB-Based Temporal Action Detection	2373
<i>Pilhyeon Lee (Yonsei University), Taeoh Kim (Naver Cloud, AI Tech.), Minho Shim (Naver Cloud, AI Tech.), Dongyoon Wee (Naver Cloud, AI Tech.), and Hyeran Byun (Yonsei University)</i>	
ASPnet: Action Segmentation With Shared-Private Representation of Multiple Data Sources .	2384
<i>Beatrice van Amsterdam (Wellcome/EPSRC Centre for Interventional and Surgical Sciences; Medtronic plc), Abdolrahim Kadkhodamohammadi (Medtronic plc), Imanol Luengo (Medtronic plc), and Danail Stoyanov (Wellcome/EPSRC Centre for InteMedtronic plcrventional and Surgical Sciences; Medtronic plc)</i>	
Proposal-Based Multiple Instance Learning for Weakly-Supervised Temporal Action Localization	2394
<i>Huan Ren (University of Science and Technology of China), Wenfei Yang (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China; Deep Space Exploration Lab), and Yongdong Zhang (University of Science and Technology of China)</i>	
LOGO: A Long-Form Video Dataset for Group Action Quality Assessment	2405
<i>Shiyi Zhang (Tsinghua University; Tsinghua University; Beijing National Research Center for Information Science and Technology), Wenxun Dai (Tsinghua University), Sujia Wang (Tsinghua University), Xiangwei Shen (Tsinghua University), Jiwen Lu (Tsinghua University; Beijing National Research Center for Information Science and Technology), Jie Zhou (Tsinghua University; Beijing National Research Center for Information Science and Technology), and Yansong Tang (Tsinghua University)</i>	
Use Your Head: Improving Long-Tail Video Recognition	2415
<i>Toby Perrett (University of Bristol, UK), Saptarshi Sinha (University of Bristol, UK), Tilo Burghardt (University of Bristol, UK), Majid Mirmehdi (University of Bristol, UK), and Dima Damen (University of Bristol, UK)</i>	
Conditional Generation of Audio From Video via Foley Analogies	2426
<i>Yuexi Du (University of Michigan; Yale University), Ziyang Chen (University of Michigan), Justin Salamon (Adobe Research), Bryan Russell (Adobe Research), and Andrew Owens (University of Michigan)</i>	

Weakly Supervised Video Representation Learning With Unaligned Text for Sequential Videos	2437
<i>Sixun Dong (ShanghaiTech University), Huazhang Hu (ShanghaiTech University), Dongze Lian (ShanghaiTech University; National University of Singapore), Weixin Luo (Meituan), Yicheng Qian (ShanghaiTech University), and Shenghua Gao (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging; Shanghai Engineering Research Center of Energy Efficient and Custom AICC)</i>	
You Can Ground Earlier Than See: An Effective and Efficient Pipeline for Temporal Sentence Grounding in Compressed Videos	2448
<i>Xiang Fang (Huazhong University of Science and Technology), Daizong Liu (Peking University), Pan Zhou (Huazhong University of Science and Technology), and Guoshun Nan (Beijing University of Posts and Telecommunications)</i>	
Connecting Vision and Language With Video Localized Narratives	2461
<i>Paul Voigtlaender (Google Research), Soravit Changpinyo (Google Research), Jordi Pont-Tuset (Google Research), Radu Soricut (Google Research), and Vittorio Ferrari (Google Research)</i>	
Video-Text As Game Players: Hierarchical Banzhaf Interaction for Cross-Modal Representation Learning	2472
<i>Peng Jin (Peking University, China), Jinfa Huang (Peking University, China), Pengfei Xiong (Shopee, China), Shangxuan Tian (Shopee, China), Chang Liu (Tsinghua University), Xiangyang Ji (Tsinghua University), Li Yuan (Peking University, China; Peng Cheng Laboratory), and Jie Chen (Peking University, China; Peng Cheng Laboratory, China)</i>	
Aligning Step-by-Step Instructional Diagrams to Video Demonstrations	2483
<i>Jiahao Zhang (The Australian National University), Anoop Cherian (Mitsubishi Electric Research Labs), Yanbin Liu (The Australian National University), Yizhak Ben-Shabat (The Australian National University; Technion Israel Institute of Technology), Cristian Rodriguez (The Australian Institute for Machine Learning), and Stephen Gould (The Australian National University)</i>	
Make-a-Story: Visual Memory Conditioned Consistent Story Generation	2493
<i>Tanzila Rahman (University of British Columbia; Vector Institute for AI), Hsin-Ying Lee (Snap Inc.), Jian Ren (Snap Inc.), Sergey Tulyakov (Snap Inc.), Shweta Mahajan (University of British Columbia; Vector Institute for AI), and Leonid Sigal (University of British Columbia; Vector Institute for AI; Canada CIFAR AI Chair)</i>	
Test of Time: Instilling Video-Language Models With a Sense of Time	2503
<i>Piyush Bagad (University of Amsterdam), Makarand Tapaswi (University of Amsterdam), and Cees G. M. Snoek (University of Amsterdam)</i>	
How You Feelin'? Learning Emotions and Mental States in Movie Scenes	2517
<i>Dhruv Srivastava (CVIT, IIIT Hyderabad, India), Aditya Kumar Singh (CVIT, IIIT Hyderabad, India), and Makarand Tapaswi (CVIT, IIIT Hyderabad, India)</i>	

Continuous Sign Language Recognition With Correlation Network	2529
<i>Lianyu Hu (Tianjin University, China), Liqing Gao (Tianjin University, China), Zekang Liu (Tianjin University, China), and Wei Feng (Tianjin University, China)</i>	
DIP: Dual Incongruity Perceiving Network for Sarcasm Detection	2540
<i>Changsong Wen (Nankai University, China), Guoli Jia (Nankai University, China), and Jufeng Yang (Nankai University, China)</i>	
Gloss Attention for Gloss-Free Sign Language Translation	2551
<i>Aoxiong Yin (Zhejiang University), Tianyun Zhong (Zhejiang University), Li Tang (Zhejiang University), Weike Jin (Zhejiang University), Tao Jin (Zhejiang University), and Zhou Zhao (Zhejiang University)</i>	
Object-Goal Visual Navigation via Effective Exploration of Relations Among Historical Navigation States	2563
<i>Heming Du (Australian National University; The University of Queensland), Lincheng Li (Netease Fuxi AI Lab), Zi Huang (The University of Queensland), and Xin Yu (The University of Queensland)</i>	
Behavioral Analysis of Vision-and-Language Navigation Agents	2574
<i>Zijiao Yang (Oregon State University), Arjun Majumdar (Georgia Institute of Technology), and Stefan Lee (Oregon State University)</i>	
KERM: Knowledge Enhanced Reasoning for Vision-and-Language Navigation	2583
<i>Xiangyang Li (Key Lab of Intelligent Information Processing Laboratory of the Chinese Academy of Sciences (CAS), Institute of Computing Technology, China; University of Chinese Academy of Sciences, China), Zihan Wang (Key Lab of Intelligent Information Processing Laboratory of the Chinese Academy of Sciences (CAS), Institute of Computing Technology, China; University of Chinese Academy of Sciences, China), Jiahao Yang (Key Lab of Intelligent Information Processing Laboratory of the Chinese Academy of Sciences (CAS), Institute of Computing Technology, China; University of Chinese Academy of Sciences, China), Yaowei Wang (Peng Cheng Laboratory, China), and Shuqiang Jiang (Key Lab of Intelligent Information Processing Laboratory of the Chinese Academy of Sciences (CAS), Institute of Computing Technology, China; University of Chinese Academy of Sciences, China; Peng Cheng Laboratory, China)</i>	
Where Is My Wallet? Modeling Object Proposal Sets for Egocentric Visual Query Localization.	2593
<i>Mengmeng Xu (Meta AI; KAUST, Saudi Arabia), Yanghao Li (Meta AI), Cheng-Yang Fu (Meta AI), Bernard Ghanem (KAUST, Saudi Arabia), Tao Xiang (Meta AI), and Juan-Manuel Pérez-Rúa (Meta AI)</i>	
Efficient Multimodal Fusion via Interactive Prompting	2604
<i>Yaowei Li (University of Technology Sydney), Ruijie Quan (Zhejiang University), Linchao Zhu (Zhejiang University), and Yi Yang (Zhejiang University)</i>	
NS3D: Neuro-Symbolic Grounding of 3D Objects and Relations	2614
<i>Joy Hsu (Stanford University), Jiayuan Mao (Stanford University), and Jiajun Wu (Stanford University)</i>	

Dynamic Inference With Grounding Based Vision and Language Models	2624
<i>Burak Uz Kent (Amazon Prime Video), Amanmeet Garg (Amazon Prime Video), Wentao Zhu (Amazon Prime Video), Keval Doshi (Amazon Prime Video), Jingru Yi (Amazon Prime Video), Xiaolong Wang (Amazon Prime Video), and Mohamed Omar (Amazon Prime Video)</i>	
Improving Commonsense in Vision-Language Models via Knowledge Graph Riddles	2634
<i>Shuquan Ye (City University of Hong Kong), Yujia Xie (Microsoft), Dongdong Chen (Microsoft), Yichong Xu (Microsoft), Lu Yuan (Microsoft), Chenguang Zhu (Microsoft), and Jing Liao (City University of Hong Kong)</i>	
S3C: Semi-Supervised VQA Natural Language Explanation via Self-Critical Learning	2646
<i>Wei Suo (Northwestern Polytechnical University, China), Mengyang Sun (Northwestern Polytechnical University, China), Weisong Liu (Northwestern Polytechnical University, China), Yiqi Gao (Northwestern Polytechnical University, China), Peng Wang (Northwestern Polytechnical University, China), Yanning Zhang (Northwestern Polytechnical University, China), and Qi Wu (University of Adelaide, Australia)</i>	
Teaching Structured Vision & Language Concepts to Vision & Language Models	2657
<i>Sivan Doveh (IBM Research; Weizmann Institute of Science), Assaf Arbelle (IBM Research), Sivan Harary (IBM Research), Eli Schwartz (IBM Research; Tel-Aviv University), Roei Herzig (IBM Research; Tel-Aviv University), Raja Giryes (Tel-Aviv University), Rogerio Feris (MIT-IBM Watson AI Lab), Rameswar Panda (MIT-IBM Watson AI Lab), Shimon Ullman (Weizmann Institute of Science), and Leonid Karlinsky (MIT-IBM Watson AI Lab)</i>	
FAME-ViL: Multi-Tasking Vision-Language Model for Heterogeneous Fashion Tasks	2669
<i>Xiao Han (CVSSP, University of Surrey; iFlyTek-Surrey Joint Research Centre on Artificial Intelligence), Xiatian Zhu (CVSSP, University of Surrey; Surrey Institute for People-Centred Artificial Intelligence), Licheng Yu (NA), Li Zhang (Surrey Institute for People-Centred Artificial Intelligence), Yi-Zhe Song (CVSSP, University of Surrey; iFlyTek-Surrey Joint Research Centre on Artificial Intelligence), and Tao Xiang (CVSSP, University of Surrey; iFlyTek-Surrey Joint Research Centre on Artificial Intelligence)</i>	
RefCLIP: A Universal Teacher for Weakly Supervised Referring Expression Comprehension ...	2681
<i>Lei Jin (Xiamen University, P.R. China.; Xiamen University, P.R. China.), Gen Luo (Xiamen University, P.R. China.), Yiyi Zhou (Xiamen University, P.R. China.; Xiamen University, P.R. China.), Xiaoshuai Sun (Xiamen University, P.R. China.; Xiamen University, P.R. China.), Guannan Jiang (Intelligent Manufacturing Department, Contemporary Amperex Technology Co. Limited (CATL)), Annan Shu (Intelligent Manufacturing Department, Contemporary Amperex Technology Co. Limited (CATL)), and Rongrong Ji (Xiamen University, P.R. China.; Xiamen University, P.R. China.)</i>	

Uni-Perceiver v2: A Generalist Model for Large-Scale Vision and Vision-Language Tasks	2691
<i>Hao Li (The Chinese University of Hong Kong), Jinguo Zhu (Xi'an Jiaotong University), Xiaohu Jiang (Tsinghua University), Xizhou Zhu (Tsinghua University; Shanghai Artificial Intelligence Laboratory), Hongsheng Li (The Chinese University of Hong Kong), Chun Yuan (Xi'an Jiaotong University), Xiaohua Wang (SenseTime Research), Yu Qiao (Shanghai Artificial Intelligence Laboratory), Xiaogang Wang (The Chinese University of Hong Kong), Wenhai Wang (Shanghai Artificial Intelligence Laboratory), and Jifeng Dai (Tsinghua University; Shanghai Artificial Intelligence Laboratory)</i>	
Learning From Unique Perspectives: User-Aware Saliency Modeling	2701
<i>Shi Chen (University of Minnesota), Nachiappan Valliappan (Google Research), Shaolei Shen (Google), Xinyu Ye (Google), Kai Kohlhoff (Google Research), and Junfeng He (Google Research)</i>	
CRAFT: Concept Recursive Activation Factorization for Explainability	2711
<i>Thomas Fel (Brown University, USA; Université de Toulouse, France; Innovation & Research Division, SNCF), Agustin Picard (Université de Toulouse, France; Scalian), Louis Béthune (Université de Toulouse, France), Thibaut Boissin (Université de Toulouse, France; Institut de Recherche Technologique Saint-Exupery, France), David Vigouroux (Université de Toulouse, France; Institut de Recherche Technologique Saint-Exupery, France), Julien Colin (Brown University, USA; Université de Toulouse, France), Rémi Cadène (Brown University, USA; Sorbonne Université, CNRS, France), and Thomas Serre (Brown University, USA; Université de Toulouse, France)</i>	
Doubly Right Object Recognition: A Why Prompt for Visual Rationales	2722
<i>Chengzhi Mao (Columbia University), Revant Teotia (Columbia University), Amrutha Sundar (Columbia University), Sachit Menon (Columbia University), Junfeng Yang (Columbia University), Xin Wang (Microsoft Research), and Carl Vondrick (Columbia University)</i>	
Sketch2Saliency: Learning To Detect Salient Objects From Human Drawings	2733
<i>Ayan Kumar Bhunia (University of Surrey, United Kingdom), Subhadeep Koley (University of Surrey, United Kingdom; iFlyTek-Surrey Joint Research Centre on Artificial Intelligence), Amandeep Kumar (University of Surrey, United Kingdom), Aneeshan Sain (University of Surrey, United Kingdom; iFlyTek-Surrey Joint Research Centre on Artificial Intelligence), Pinaki Nath Chowdhury (University of Surrey, United Kingdom; iFlyTek-Surrey Joint Research Centre on Artificial Intelligence), Tao Xiang (University of Surrey, United Kingdom; iFlyTek-Surrey Joint Research Centre on Artificial Intelligence), and Yi-Zhe Song (University of Surrey, United Kingdom; iFlyTek-Surrey Joint Research Centre on Artificial Intelligence)</i>	
PIP-Net: Patch-Based Intuitive Prototypes for Interpretable Image Classification	2744
<i>Meike Nauta (University of Twente, the Netherlands; University of Duisburg-Essen, Germany), Jörg Schlötterer (University of Duisburg-Essen, Germany), Maurice van Keulen (University of Twente, Netherlands), and Christin Seifert (University of Duisburg-Essen, Germany)</i>	

Photo Pre-Training, but for Sketch	2754
<i>Ke Li (Beijing University of Posts and Telecommunications, China; University of Surrey, United Kingdom), Kaiyue Pang (Beijing University of Posts and Telecommunications, China), and Yi-Zhe Song (Beijing University of Posts and Telecommunications, China)</i>	
CLIP for All Things Zero-Shot Sketch-Based Image Retrieval, Fine-Grained or Not	2765
<i>Aneeshan Sain (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Ayan Kumar Bhunia (University of Surrey, United Kingdom), Pinaki Nath Chowdhury (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Subhadeep Koley (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Tao Xiang (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), and Yi-Zhe Song (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence)</i>	
iCLIP: Bridging Image Classification and Contrastive Language-Image Pre-Training for Visual Recognition	2776
<i>Yixuan Wei (Tsinghua University; Microsoft Research Asia), Yue Cao (Microsoft Research Asia), Zheng Zhang (Microsoft Research Asia), Houwen Peng (Microsoft Research Asia), Zhuliang Yao (Tsinghua University; Microsoft Research Asia), Zhenda Xie (Tsinghua University; Microsoft Research Asia), Han Hu (Microsoft Research Asia), and Baining Guo (Microsoft Research Asia)</i>	
Cross-Modal Implicit Relation Reasoning and Aligning for Text-to-Image Person Retrieval	2787
<i>Ding Jiang (Wuhan University, China) and Mang Ye (Wuhan University, China; Hubei LuoJia Laboratory, China)</i>	
Multi-Modal Representation Learning With Text-Driven Soft Masks	2798
<i>Jaeyoo Park (Computer Vision Laboratory, ECE1 & IPAI) and Bohyung Han (Seoul National University)</i>	
Texts as Images in Prompt Tuning for Multi-Label Image Recognition	2808
<i>Zixian Guo (Harbin Institute of Technology; Tomorrow Advancing Life), Bowen Dong (Harbin Institute of Technology), Zhilong Ji (Tomorrow Advancing Life), Jinfeng Bai (Tomorrow Advancing Life), Yiwu Guo (Independent Researcher), and Wangmeng Zuo (Harbin Institute of Technology; Tomorrow Advancing Life)</i>	
Reproducible Scaling Laws for Contrastive Language-Image Learning	2818
<i>Mehdi Cherti (LAION; Juelich Supercomputing Center (JSC), Research Center Juelich (FZ)), Romain Beaumont (LAION), Ross Wightman (LAION; HuggingFace), Mitchell Wortsman (University of Washington), Gabriel Ilharco (University of Washington), Cade Gordon (UC Berkeley), Christoph Schuhmann (LAION), Ludwig Schmidt (LAION; University of Washington), and Jenia Jitsev (LAION; Juelich Supercomputing Center (JSC), Research Center Juelich (FZ))</i>	

Multilateral Semantic Relations Modeling for Image Text Retrieval	2830
<i>Zheng Wang (University of Electronic Science and Technology of China; Institute of Electronic and Information Engineering of UESTC in Guangdong), Zhenwei Gao (University of Electronic Science and Technology of China), Kangshuai Guo (University of Electronic Science and Technology of China), Yang Yang (University of Electronic Science and Technology of China), Xiaoming Wang (University of Electronic Science and Technology of China), and Heng Tao Shen (University of Electronic Science and Technology of China; Peng Cheng Lab)</i>	
SmallCap: Lightweight Image Captioning Prompted With Retrieval Augmentation	2840
<i>Rita Ramos (University of Lisbon), Bruno Martins (University of Lisbon), Desmond Elliott (University of Copenhagen; Pioneer Center for AI), and Yova Kementchedjheva (University of Copenhagen)</i>	
Probing Sentiment-Oriented Pre-Training Inspired by Human Sentiment Perception Mechanism ..	2850
<i>Tinglei Feng (Nankai University, China), Jiakuan Liu (Nankai University, China), and Jufeng Yang (Nankai University, China)</i>	
Prefix Conditioning Unifies Language and Label Supervision	2861
<i>Kuniaki Saito (Boston University; Google Cloud AI Research), Kihyuk Sohn (Google Research), Xiang Zhang (Google Cloud AI Research), Chun-Liang Li (Google Cloud AI Research), Chen-Yu Lee (Google Cloud AI Research), Kate Saenko (Boston University; MIT-IBM Watson AI Lab), and Tomas Pfister (Google Cloud AI Research)</i>	
Crossing the Gap: Domain Generalization for Image Captioning	2871
<i>Yuchen Ren (University of Science and Technology of China, China; Shanghai Artificial Intelligence Laboratory, Shanghai, China), Zhendong Mao (University of Science and Technology of China, China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, China), Shancheng Fang (University of Science and Technology of China, China), Yan Lu (Shanghai Artificial Intelligence Laboratory, Shanghai, China), Tong He (Shanghai Artificial Intelligence Laboratory, Shanghai, China), Hao Du (University of Science and Technology of China, Hefei, China), Yongdong Zhang (University of Science and Technology of China, Hefei, China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, China), and Wanli Ouyang (Shanghai Artificial Intelligence Laboratory, Shanghai, China)</i>	
A Bag-of-Prototypes Representation for Dataset-Level Applications	2881
<i>Weijie Tu (Australian National University), Weijian Deng (Australian National University), Tom Gedeon (Curtin University), and Liang Zheng (Australian National University)</i>	
CrowdCLIP: Unsupervised Crowd Counting via Vision-Language Model	2893
<i>Dingkang Liang (Huazhong University of Science and Technology), Jiahao Xie (Beijing University of Posts and Telecommunications), Zhikang Zou (Baidu Inc., China), Xiaoqing Ye (Baidu Inc., China), Wei Xu (Beijing University of Posts and Telecommunications), and Xiang Bai (Huazhong University of Science and Technology)</i>	

D2Former: Jointly Learning Hierarchical Detectors and Contextual Descriptors via Agent-Based Transformers	2904
<i>Jianfeng He (University of Science and Technology of China), Yuan Gao (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China; Deep Space Exploration Laboratory), Zhe Zhang (Deep Space Exploration Laboratory), and Feng Wu (University of Science and Technology of China)</i>	
Learning To Generate Language-Supervised and Open-Vocabulary Scene Graph Using Pre-Trained Visual-Semantic Space	2915
<i>Yong Zhang (The Chinese University of Hong Kong, Shenzhen), Yingwei Pan (HiDream.ai Inc.), Ting Yao (HiDream.ai Inc.), Rui Huang (The Chinese University of Hong Kong, Shenzhen), Tao Mei (HiDream.ai Inc.), and Chang-Wen Chen (The Hong Kong Polytechnic University)</i>	
Relational Context Learning for Human-Object Interaction Detection	2925
<i>Sanghyun Kim (Pohang University of Science and Technology (POSTECH), South Korea), Deunsol Jung (Pohang University of Science and Technology (POSTECH), South Korea), and Minsu Cho (Pohang University of Science and Technology (POSTECH), South Korea)</i>	
Learning Open-Vocabulary Semantic Segmentation Models From Natural Language Supervision ..	2935
<i>Jilan Xu (Fudan University; Shanghai AI Laboratory), Junlin Hou (Fudan University), Yuejie Zhang (Fudan University), Rui Feng (Fudan University), Yi Wang (Shanghai AI Laboratory), Yu Qiao (Shanghai AI Laboratory), and Weidi Xie (Shanghai AI Laboratory; Shanghai Jiao Tong University)</i>	
Side Adapter Network for Open-Vocabulary Semantic Segmentation	2945
<i>Mengde Xu (Huazhong University of Science and Technology; Microsoft Research Asia), Zheng Zhang (Huazhong University of Science and Technology; Microsoft Research Asia), Fangyun Wei (Microsoft Research Asia), Han Hu (Microsoft Research Asia), and Xiang Bai (Huazhong University of Science and Technology)</i>	
Open-Vocabulary Panoptic Segmentation With Text-to-Image Diffusion Models	2955
<i>Jiarui Xu (UC San Diego), Sifei Liu (NVIDIA), Arash Vahdat (NVIDIA), Wonmin Byeon (NVIDIA), Xiaolong Wang (UC San Diego), and Shalini De Mello (NVIDIA)</i>	
IFSeg: Image-Free Semantic Segmentation via Vision-Language Model	2967
<i>Sukmin Yun (Korea Advanced Institute of Science and Technology (KAIST); Mohamed bin Zayed University of Artificial Intelligence (MBZUAI)), Seong Hyeon Park (Korea Advanced Institute of Science and Technology (KAIST)), Paul Hongsuck Seo (Google Research), and Jinwoo Shin (Korea Advanced Institute of Science and Technology (KAIST))</i>	
PartManip: Learning Cross-Category Generalizable Part Manipulation Policy From Point Cloud Observations	2978
<i>Haoran Geng (Peking University), Ziming Li (Peking University), Yiran Geng (Peking University), Jiayi Chen (Peking University; Beijing Academy of Artificial Intelligence), Hao Dong (Peking University), and He Wang (Peking University)</i>	

OneFormer: One Transformer To Rule Universal Image Segmentation	2989
<i>Jitesh Jain (SHI Labs @ U of Oregon & UIUC; IIT Roorkee), Jiachen Li (SHI Labs @ U of Oregon & UIUC), Mang Tik Chiu (SHI Labs @ U of Oregon & UIUC), Ali Hassani (SHI Labs @ U of Oregon & UIUC), Nikita Orlov (Picsart AI Research (PAIR)), and Humphrey Shi (Picsart AI Research (PAIR))</i>	
Delving Into Shape-Aware Zero-Shot Semantic Segmentation	2999
<i>Xinyu Liu (Xidian University; Institute for AI Industry Research (AIR), Tsinghua University), Beiwen Tian (Institute for AI Industry Research (AIR), Tsinghua University; Department of Computer Science and Technology, Tsinghua University), Zhen Wang (Didi Chuxing), Rui Wang (Didi Chuxing), Kehua Sheng (Didi Chuxing), Bo Zhang (Didi Chuxing), Hao Zhao (Institute for AI Industry Research (AIR), Tsinghua University), and Guyue Zhou (Institute for AI Industry Research (AIR), Tsinghua University)</i>	
CoMFormer: Continual Learning in Semantic and Panoptic Segmentation	3010
<i>Fabio Cermelli (Politecnico di Torino; Italian Institute of Technology), Matthieu Cord (Sorbonne Université; valeo.ai), and Arthur Douillard (Sorbonne Université)</i>	
Learning To Segment Every Referring Object Point by Point	3021
<i>Mengxue Qu (Beijing Jiaotong University; Beijing Key Laboratory of Advanced Information Science and Network Technology), Yu Wu (Wuhan University), Yunchao Wei (Beijing Jiaotong University; Beijing Key Laboratory of Advanced Information Science and Network Technology), Wu Liu (JD Explore Academy), Xiaodan Liang (Sun Yat-sen University; MBZUAI), and Yao Zhao (Beijing Jiaotong University; Beijing Key Laboratory of Advanced Information Science and Network Technology)</i>	
Unsupervised Continual Semantic Adaptation Through Neural Rendering	3031
<i>Zhizheng Liu (ETH Zurich), Francesco Milano (ETH Zurich), Jonas Frey (ETH Zurich; Max Planck ETH Center for Learning Systems), Roland Siegwart (ETH Zurich), Hermann Blum (ETH Zurich), and Cesar Cadena (ETH Zurich)</i>	
Mask DINO: Towards a Unified Transformer-Based Framework for Object Detection and Segmentation	3041
<i>Feng Li (The Hong Kong University of Science and Technology; International Digital Economy Academy (IDEA)), Hao Zhang (The Hong Kong University of Science and Technology; International Digital Economy Academy (IDEA)), Huaizhe Xu (The Hong Kong University of Science and Technology; International Digital Economy Academy (IDEA)), Shilong Liu (Tsinghua University; The Hong Kong University of Science and Technology (Guangzhou)), Lei Zhang (International Digital Economy Academy (IDEA)), Lionel M. Ni (The Hong Kong University of Science and Technology; International Digital Economy Academy (IDEA)), and Heung-Yeung Shum (The Hong Kong University of Science and Technology; International Digital Economy Academy (IDEA))</i>	
Transformer Scale Gate for Semantic Segmentation	3051
<i>Hengcan Shi (Monash University, Australia), Munawar Hayat (Monash University, Australia), and Jianfei Cai (Monash University, Australia)</i>	

Style Projected Clustering for Domain Generalized Semantic Segmentation	3061
<i>Wei Huang (University of Science and Technology of China), Chang Chen (Huawei Noah's Ark Lab), Yong Li (Huawei Noah's Ark Lab), Jiacheng Li (University of Science and Technology of China), Cheng Li (Huawei Noah's Ark Lab), Fenglong Song (Huawei Noah's Ark Lab), Youliang Yan (Huawei Noah's Ark Lab), and Zhiwei Xiong (University of Science and Technology of China)</i>	
Rethinking Few-Shot Medical Segmentation: A Vector Quantization View	3072
<i>Shiqi Huang (Beijing Institute of Technology), Tingfa Xu (Beijing Institute of Technology), Ning Shen (Beijing Institute of Technology), Feng Mu (Beijing Institute of Technology), and Jianan Li (Beijing Institute of Technology)</i>	
Continual Semantic Segmentation With Automatic Memory Sample Selection	3082
<i>Lanyun Zhu (Singapore University of Technology and Design), Tianrun Chen (Zhejiang University), Jianxiong Yin (NVIDIA AI Tech Centre), Simon See (NVIDIA AI Tech Centre), and Jun Liu (Singapore University of Technology and Design)</i>	
Token Contrast for Weakly-Supervised Semantic Segmentation	3093
<i>Lixiang Ru (Wuhan University, China), Heliang Zheng (JD Explore Academy, China), Yibing Zhan (JD Explore Academy, China), and Bo Du (Wuhan University, China)</i>	
Multi-Granularity Archaeological Dating of Chinese Bronze Dings Based on a Knowledge-Guided Relation Graph	3103
<i>Rixin Zhou (Jilin University), Jiafu Wei (Jilin University), Qian Zhang (Jilin University), Ruihua Qi (Jilin University), Xi Yang (Jilin University; Engineering Research Center of Knowledge-Driven Human-Machine Intelligence, MoE, China), and Chuntao Li (Jilin University)</i>	
Hunting Sparsity: Density-Guided Contrastive Learning for Semi-Supervised Semantic Segmentation	3114
<i>Xiaoyang Wang (XJTLU; University of Liverpool; Metavisioncn), Bingfeng Zhang (China University of Petroleum (East China)), Limin Yu (XJTLU), and Jimin Xiao (XJTLU)</i>	
Cut and Learn for Unsupervised Object Detection and Instance Segmentation	3124
<i>Xudong Wang (FAIR, Meta AI; UC Berkeley / ICSI), Rohit Girdhar (FAIR, Meta AI), Stella X. Yu (UC Berkeley / ICSI; University of Michigan), and Ishan Misra (FAIR, Meta AI)</i>	
Extracting Class Activation Maps From Non-Discriminative Features As Well	3135
<i>Zhaozheng Chen (Singapore Management University) and Qianru Sun (Singapore Management University)</i>	
BoxTeacher: Exploring High-Quality Pseudo Labels for Weakly Supervised Instance Segmentation	3145
<i>Tianheng Cheng (Huazhong University of Science & Technology), Xinggang Wang (Huazhong University of Science & Technology), Shaoyu Chen (Huazhong University of Science & Technology), Qian Zhang (Horizon Robotics), and Wenyu Liu (Huazhong University of Science & Technology)</i>	

Hierarchical Fine-Grained Image Forgery Detection	3155
<i>Xiao Guo (Michigan State University), Xiaohong Liu (Shanghai Jiao Tong University), Zhiyuan Ren (Michigan State University), Steven Grosz (Michigan State University), Jacopo Masi (Sapienza University of Rome), and Xiaoming Liu (Michigan State University)</i>	
Towards Professional Level Crowd Annotation of Expert Domain Data	3166
<i>Pei Wang (UC San Diego) and Nuno Vasconcelos (UC San Diego)</i>	
Unsupervised Object Localization: Observing the Background To Discover Objects	3176
<i>Oriane Siméoni (valeo.ai, France), Chloé Sekkat (valeo.ai, France), Gilles Puy (valeo.ai, France), Antonín Vobecký (valeo.ai, France; CTU, Czech Republic), Éloi Éloi Zablocki (valeo.ai, France), and Patrick Pérez (valeo.ai, France)</i>	
Semi-Supervised Learning Made Simple With Self-Supervised Clustering	3187
<i>Enrico Fini (University of Trento), Pietro Astolfi (University of Trento; Inria), Karteek Alahari (Inria), Xavier Alameda-Pineda (Inria), Julien Mairal (Inria), Moin Nabi (SAP AI Research), and Elisa Ricci (University of Trento; Fondazione Bruno Kessler)</i>	
Unbalanced Optimal Transport: A Unified Framework for Object Detection	3198
<i>Henri De Plaen (ESAT-STADIUS, KU Leuven, Belgium), Pierre-François De Plaen (ESAT-PSI, Belgium), Johan A. K. Suykens (ESAT-STADIUS, KU Leuven, Belgium), Marc Proesmans (ESAT-PSI, Belgium), Tinne Tuytelaars (ESAT-PSI, Belgium), and Luc Van Gool (ESAT-PSI, Belgium; Computer Vision Lab, ETH Zurich, Switzerland)</i>	
DiGeo: Discriminative Geometry-Aware Learning for Generalized Few-Shot Object Detection	3208
<i>Jiawei Ma (Columbia University), Yulei Niu (Columbia University), Jincheng Xu (Columbia University), Shiyuan Huang (Columbia University), Guangxing Han (Columbia University), and Shih-Fu Chang (Columbia University)</i>	
CLIP the Gap: A Single Domain Generalization Approach for Object Detection	3219
<i>Vidit Vidit (EPFL), Martin Engilberge (EPFL), and Mathieu Salzmann (EPFL; ClearSpace SA)</i>	
Unknown Sniffer for Object Detection: Don't Turn a Blind Eye to Unknown Objects	3230
<i>Wenteng Liang (Beijing University of Posts and Telecommunications, China), Feng Xue (Beijing University of Posts and Telecommunications, China), Yihao Liu (Beijing University of Posts and Telecommunications, China), Guofeng Zhong (Chongqing University of Posts and Telecommunications, China), and Anlong Ming (Beijing University of Posts and Telecommunications, China)</i>	
Consistent-Teacher: Towards Reducing Inconsistent Pseudo-Targets in Semi-Supervised Object Detection	3240
<i>Xinjiang Wang (SenseTime Research), Xingyi Yang (National University of Singapore), Shilong Zhang (Shanghai AI Laboratory), Yijiang Li (SenseTime Research), Litong Feng (SenseTime Research), Shijie Fang (Peking University), Chengqi Lyu (Shanghai AI Laboratory), Kai Chen (Shanghai AI Laboratory), and Wayne Zhang (SenseTime Research)</i>	

Optimal Proposal Learning for Deployable End-to-End Pedestrian Detection	3250
<i>Xiaolin Song (Beijing University of Posts and Telecommunications, China), Binghui Chen (Beijing University of Posts and Telecommunications, China), Pengyu Li (Beijing University of Posts and Telecommunications, China), Jun-Yan He (Beijing University of Posts and Telecommunications, China), Biao Wang (Beijing University of Posts and Telecommunications, China), Yifeng Geng (Beijing University of Posts and Telecommunications, China), Xuansong Xie (Beijing University of Posts and Telecommunications, China), and Honggang Zhang (Beijing University of Posts and Telecommunications, China)</i>	
AsyFOD: An Asymmetric Adaptation Paradigm for Few-Shot Domain Adaptive Object Detection	3261
<i>Yipeng Gao (Sun Yat-sen University, China; Ministry of Education, China), Kun-Yu Lin (Sun Yat-sen University, China; Ministry of Education, China), Junkai Yan (Sun Yat-sen University, China; Ministry of Education, China), Yaowei Wang (Pengcheng Lab.), and Wei-Shi Zheng (Sun Yat-sen University, China; Pengcheng Lab.; Ministry of Education, China)</i>	
Where Is My Spot? Few-Shot Image Generation via Latent Subspace Optimization	3272
<i>Chenxi Zheng (South China University of Technology), Bangzhen Liu (South China University of Technology), Huaidong Zhang (South China University of Technology), Xuemiao Xu (South China University of Technology; State Key Laboratory of Subtropical Building Science; Ministry of Education Key Laboratory of Big Data and Intelligent Robot; Guangdong Provincial Key Lab of Computational Intelligence and Cyberspace Information), and Shengfeng He (Singapore Management University)</i>	
Uncertainty-Aware Optimal Transport for Semantically Coherent Out-of-Distribution Detection	3282
<i>Fan Lu (University of Science and Technology of China), Kai Zhu (University of Science and Technology of China), Wei Zhai (University of Science and Technology of China), Kecheng Zheng (Ant Group), and Yang Cao (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center)</i>	
MAESTER: Masked Autoencoder Guided Segmentation at Pixel Resolution for Accurate, Self-Supervised Subcellular Structure Recognition	3292
<i>Ronald Xie (University of Toronto; Vector Institute; University Health Network; The Donnelly Centre), Kuan Pang (University of Toronto; The Donnelly Centre), Gary D. Bader (University of Toronto; University Health Network; The Donnelly Centre), and Bo Wang (University of Toronto; Vector Institute; University Health Network)</i>	
Orthogonal Annotation Benefits Barely-Supervised Medical Image Segmentation	3302
<i>Heng Cai (Nanjing University), Shumeng Li (Nanjing University), Lei Qi (Southeast University), Qian Yu (Shandong Women's University), Yinghuan Shi (Nanjing University), and Yang Gao (Nanjing University)</i>	

RepMode: Learning to Re-Parameterize Diverse Experts for Subcellular Structure Prediction	3312
<i>Donghao Zhou (Guangdong Provincial Key Laboratory of Computer Vision and Virtual Reality Technology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences; The Chinese University of Hong Kong), Chunbin Gu (The Chinese University of Hong Kong), Junde Xu (Guangdong Provincial Key Laboratory of Computer Vision and Virtual Reality Technology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences; The Chinese University of Hong Kong), Furui Liu (Zhejiang Lab), Qiong Wang (Guangdong Provincial Key Laboratory of Computer Vision and Virtual Reality Technology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), Guangyong Chen (Zhejiang Lab), and Pheng-Ann Heng (Guangdong Provincial Key Laboratory of Computer Vision and Virtual Reality Technology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences)</i>	
Topology-Guided Multi-Class Cell Context Generation for Digital Pathology	3323
<i>Shahira Abousamra (Stony Brook University, Department of Computer Science, USA), Rajarsi Gupta (Stony Brook University, Department of Biomedical Informatics, USA), Tahsin Kurc (Stony Brook University, Department of Biomedical Informatics, USA), Dimitris Samaras (Stony Brook University, Department of Computer Science, USA), Joel Saltz (Stony Brook University, Department of Biomedical Informatics, USA), and Chao Chen (Stony Brook University, Department of Biomedical Informatics, USA)</i>	
Dynamic Graph Enhanced Contrastive Learning for Chest X-Ray Report Generation	3334
<i>Mingjie Li (University of Technology Sydney), Bingqian Lin (Sun Yat-Sen University), Zicong Chen (The University of Hong Kong), Haokun Lin (Sun Yat-Sen University), Xiaodan Liang (Sun Yat-Sen University; Mohamed bin Zayed University of Artificial Intelligence; Peng Cheng National Lab), and Xiaojun Chang (University of Technology Sydney)</i>	
Benchmarking Self-Supervised Learning on Diverse Pathology Datasets	3344
<i>Mingu Kang (Lunit Inc.), Heon Song (Lunit Inc.), Seonwook Park (Lunit Inc.), Donggeun Yoo (Lunit Inc.), and Sérgio Pereira (Lunit Inc.)</i>	
Multiple Instance Learning via Iterative Self-Paced Supervised Contrastive Learning	3355
<i>Kangning Liu (NYU Center for Data Science), Weicheng Zhu (NYU Center for Data Science), Yiqiu Shen (NYU Center for Data Science), Sheng Liu (NYU Center for Data Science), Narges Razavian (NYU Grossman School of Medicine), Krzysztof J. Geras (NYU Center for Data Science; NYU Grossman School of Medicine), and Carlos Fernandez-Granda (NYU Center for Data Science; Courant Institute of Mathematical Sciences)</i>	
Learning Expressive Prompting With Residuals for Vision Transformers	3366
<i>Rajshekhar Das (Carnegie Mellon University; AWS AI Labs), Yonatan Dukler (AWS AI Labs), Avinash Ravichandran (AWS AI Labs), and Ashwin Swaminathan (AWS AI Labs)</i>	

Detection of Out-of-Distribution Samples Using Binary Neuron Activation Patterns	3378
<i>Bartłomiej Olber (Warsaw University of Technology; Silesian University of Technology), Krystian Radlak (Warsaw University of Technology; Silesian University of Technology), Adam Popowicz (Silesian University of Technology), Michal Szczepankiewicz (NVIDIA), and Krystian Chachula (Warsaw University of Technology)</i>	
Decoupling MaxLogit for Out-of-Distribution Detection	3388
<i>Zihan Zhang (Key Lab of Image Processing and Intelligent Control, Ministry of Education School of Artificial Intelligence and Automation Huazhong University of Science and Technology, China) and Xiang Xiang (Key Lab of Image Processing and Intelligent Control, Ministry of Education School of Artificial Intelligence and Automation Huazhong University of Science and Technology, China)</i>	
Exploring Structured Semantic Prior for Multi Label Recognition With Incomplete Labels	3398
<i>Zixuan Ding (Xidian University; Hangzhou Zhuoxi Institute of Brain and Intelligence), Ao Wang (Tsinghua University; BNRist; Hangzhou Zhuoxi Institute of Brain and Intelligence), Hui Chen (Tsinghua University; BNRist), Qiang Zhang (Xidian University), Pengzhang Liu (JD.com), Yongjun Bao (JD.com), Weipeng Yan (JD.com), and Jungong Han (Department of Computer Science, the University of Sheffield, UK; Centre for Machine Intelligence, the University of Sheffield, UK)</i>	
Bridging the Gap Between Model Explanations in Partially Annotated Multi-Label Classification	3408
<i>Youngwook Kim (Seoul National University), Jae Myung Kim (PSL Research University), Jieun Jeong (University of Tübingen), Cordelia Schmid (University of Tübingen; MPI for Intelligent Systems), Zeynep Akata (Seoul National University; HodooAI Lab), and Jungwoo Lee (Seoul National University; HodooAI Lab)</i>	
DivClust: Controlling Diversity in Deep Clustering	3418
<i>Ioannis Maniadis Metaxas (Queen Mary University of London, UK), Georgios Tzimiropoulos (Queen Mary University of London, UK), and Ioannis Patras (Queen Mary University of London, UK)</i>	
Deep Semi-Supervised Metric Learning With Mixed Label Propagation	3429
<i>Furen Zhuang (University of Illinois at Urbana-Champaign, Illinois; Institute for Infocomm Research (I2R), A*STAR, Singapore) and Pierre Moulin (University of Illinois at Urbana-Champaign, Illinois)</i>	
Leveraging Inter-Rater Agreement for Classification in the Presence of Noisy Labels	3439
<i>Maria Sofia Bucarelli (Sapienza University of Rome), Lucas Cassano (Amazon), Federico Siciliano (Sapienza University of Rome), Amin Mantrach (Amazon), and Fabrizio Silvestri (Sapienza University of Rome; ISTI-CNR, Pisa, Italy)</i>	
Modeling Inter-Class and Intra-Class Constraints in Novel Class Discovery	3449
<i>Wenbin Li (Nanjing University, China), Zhichen Fan (Nanjing University, China), Jing Huo (Nanjing University, China), and Yang Gao (Nanjing University, China)</i>	
Bootstrap Your Own Prior: Towards Distribution-Agnostic Novel Class Discovery	3459
<i>Muli Yang (Xidian University, China), Liancheng Wang (Xidian University, China), Cheng Deng (Xidian University, China), and Hanwang Zhang (Nanyang Technological University, Singapore)</i>	

Towards Realistic Long-Tailed Semi-Supervised Learning: Consistency Is All You Need	3469
<i>Tong Wei (Southeast University, China) and Kai Gan (Southeast University, China)</i>	
PromptCAL: Contrastive Affinity Learning via Auxiliary Prompts for Generalized Novel Category Discovery	3479
<i>Sheng Zhang (Mohamed bin Zayed University of AI), Salman Khan (Mohamed bin Zayed University of AI; Australian National University), Zhiqiang Shen (Mohamed bin Zayed University of AI; Hong Kong University of Science and Technology), Muzammal Naseer (Mohamed bin Zayed University of AI), Guangyi Chen (Mohamed bin Zayed University of AI; Carnegie Mellon University), and Fahad Shahbaz Khan (Mohamed bin Zayed University of AI; Linköping University)</i>	
Probabilistic Knowledge Distillation of Face Ensembles	3489
<i>Jianqing Xu (n/a), Shen Li (n/a), Ailin Deng (IDS and SoC, National University of Singapore), Miao Xiong (IDS and SoC, National University of Singapore), Jiaying Wu (IDS and SoC, National University of Singapore), Jiaxiang Wu (Youtu Lab, Tencent), Shouhong Ding (Youtu Lab, Tencent), and Bryan Hooi (IDS and SoC, National University of Singapore)</i>	
Class-Conditional Sharpness-Aware Minimization for Deep Long-Tailed Recognition	3499
<i>Zhipeng Zhou (University of Science and Technology of China), Lanqing Li (The Chinese University of Hong Kong; Zhejiang Lab), Peilin Zhao (Tencent AI Lab), Pheng-Ann Heng (The Chinese University of Hong Kong), and Wei Gong (University of Science and Technology of China)</i>	
Promoting Semantic Connectivity: Dual Nearest Neighbors Contrastive Learning for Unsupervised Domain Generalization	3510
<i>Yuchen Liu (Shanghai Jiao Tong University, China), Yaoming Wang (Shanghai Jiao Tong University, China), Yabo Chen (Shanghai Jiao Tong University, China), Wenrui Dai (Shanghai Jiao Tong University, China), Chenglin Li (Shanghai Jiao Tong University, China), Junni Zou (Shanghai Jiao Tong University, China), and Hongkai Xiong (Shanghai Jiao Tong University, China)</i>	
Instance Relation Graph Guided Source-Free Domain Adaptive Object Detection	3520
<i>Vibashan VS (Johns Hopkins University, USA), Poojan Oza (Johns Hopkins University, USA), and Vishal M. Patel (Johns Hopkins University, USA)</i>	
MOT: Masked Optimal Transport for Partial Domain Adaptation	3531
<i>You-Wei Luo (Sun Yat-Sen University, China) and Chuan-Xian Ren (Sun Yat-Sen University, China)</i>	
TOPLight: Lightweight Neural Networks With Task-Oriented Pretraining for Visible-Infrared Recognition	3541
<i>Hao Yu (Nanjing University of Information Science and Technology, China), Xu Cheng (Nanjing University of Information Science and Technology, China), and Wei Peng (Stanford University)</i>	
OSAN: A One-Stage Alignment Network To Unify Multimodal Alignment and Unsupervised Domain Adaptation	3551
<i>Ye Liu (Tencent Youtu Lab), Lingfeng Qiao (Tencent Youtu Lab), Changchong Lu (Tencent Youtu Lab), Di Yin (Tencent Youtu Lab), Chen Lin (Tencent Youtu Lab), Haoyuan Peng (Tencent Youtu Lab), and Bo Ren (Tencent Youtu Lab)</i>	

Patch-Mix Transformer for Unsupervised Domain Adaptation: A Game Perspective	3561
<i>Jinjing Zhu (AI Thrust, HKUST(GZ)), Haotian Bai (AI Thrust, HKUST(GZ)), and Lin Wang (AI Thrust, HKUST(GZ); Dept. of CSE, HKUST)</i>	
ARO-Net: Learning Implicit Fields From Anchored Radial Observations	3572
<i>Yizhi Wang (Shenzhen University; Simon Fraser University), Zeyu Huang (Shenzhen University), Ariel Shamir (Reichman University), Hui Huang (Shenzhen University; Simon Fraser University), Hao Zhang (Simon Fraser University), and Ruizhen Hu (Shenzhen University)</i>	
A Probabilistic Framework for Lifelong Test-Time Adaptation	3582
<i>Dhanajit Brahma (Indian Institute of Technology Kanpur) and Piyush Rai (Indian Institute of Technology Kanpur)</i>	
Distribution Shift Inversion for Out-of-Distribution Prediction	3592
<i>Runpeng Yu (National University of Singapore), Songhua Liu (National University of Singapore), Xingyi Yang (National University of Singapore), and Xinchao Wang (National University of Singapore)</i>	
Learning Joint Latent Space EBM Prior Model for Multi-Layer Generator	3603
<i>Jiali Cui (Department of Computer Science, Stevens Institute of Technology), Ying Nian Wu (Department of Statistics, University of California, Los Angeles), and Tian Han (Department of Computer Science, Stevens Institute of Technology)</i>	
A Data-Based Perspective on Transfer Learning	3613
<i>Saachi Jain (MIT), Hadi Salman (MIT), Alaa Khaddaj (MIT), Eric Wong (University of Pennsylvania), Sung Min Park (MIT), and Aleksander Mądry (MIT)</i>	
A Meta-Learning Approach to Predicting Performance and Data Requirements	3623
<i>Achin Jain (AWS AI Labs), Gurusurthy Swaminathan (AWS AI Labs), Paolo Favaro (AWS AI Labs), Hao Yang (AWS AI Labs), Avinash Ravichandran (AWS AI Labs), Hrayr Harutyunyan (University of Southern California), Alessandro Achille (AWS AI Labs), Onkar Dabeer (AWS AI Labs), Bernt Schiele (AWS AI Labs), Ashwin Swaminathan (AWS AI Labs), and Stefano Soatto (AWS AI Labs)</i>	
Guided Recommendation for Model Fine-Tuning	3633
<i>Hao Li (AWS AI Labs), Charless Fowlkes (AWS AI Labs), Hao Yang (AWS AI Labs), Onkar Dabeer (AWS AI Labs), Zhuowen Tu (AWS AI Labs), and Stefano Soatto (AWS AI Labs)</i>	
EMT-NAS:Transferring Architectural Knowledge Between Tasks From Different Datasets	3643
<i>Peng Liao (Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, ECUST, China), Yaochu Jin (Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, ECUST, China; Bielefeld University, Germany), and Wenli Du (Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, ECUST, China)</i>	

AttriCLIP: A Non-Incremental Learner for Incremental Knowledge Learning	3654
<i>Runqi Wang (Beihang University; Huawei Noah's Ark Lab), Xiaoyue Duan (Beihang University), Guoliang Kang (Beihang University; Zhongguancun Laboratory, China), Jianzhuang Liu (Huawei Noah's Ark Lab), Shaohui Lin (East China Normal University), Songcen Xu (Huawei Noah's Ark Lab), Jinhui Lü (Beihang University; Zhongguancun Laboratory, China), and Baochang Zhang (Beihang University; Zhongguancun Laboratory, China)</i>	
Batch Model Consolidation: A Multi-Task Model Consolidation Framework	3664
<i>Iordanis Fostiroopoulos (University of Southern California, Los Angeles, United States), Jiaye Zhu (University of Southern California, Los Angeles, United States), and Laurent Itti (University of Southern California, Los Angeles, United States)</i>	
SmartAssign: Learning a Smart Knowledge Assignment Strategy for Deraining and Desnowing	3677
<i>Yinglong Wang (Huawei Noah's Ark Lab), Chao Ma (AI Institute, Shanghai Jiao Tong University), and Jianzhuang Liu (Huawei Noah's Ark Lab)</i>	
TinyMIM: An Empirical Study of Distilling MIM Pre-Trained Models	3687
<i>Sucheng Ren (Microsoft Research Asia), Fangyun Wei (Microsoft Research Asia), Zheng Zhang (Microsoft Research Asia), and Han Hu (Microsoft Research Asia)</i>	
Computationally Budgeted Continual Learning: What Does Matter?	3698
<i>Ameyaa Prabhu (University of Oxford), Hasan Abed Al Kader Hammoud (King Abdullah University of Science and Technology (KAUST)), Puneet K. Dokania (University of Oxford), Philip H.S. Torr (University of Oxford), Ser-Nam Lim (Meta AI), Bernard Ghanem (King Abdullah University of Science and Technology (KAUST)), and Adel Bibi (University of Oxford)</i>	
GradMA: A Gradient-Memory-Based Accelerated Federated Learning With Alleviated Catastrophic Forgetting	3708
<i>Kangyang Luo (East China Normal University Shanghai, China), Xiang Li (East China Normal University Shanghai, China), Yunshi Lan (East China Normal University Shanghai, China), and Ming Gao (East China Normal University Shanghai, China)</i>	
Rethinking Gradient Projection Continual Learning: Stability / Plasticity Feature Space Decoupling	3718
<i>Zhen Zhao (East China Normal University, China), Zhizhong Zhang (East China Normal University, China), Xin Tan (East China Normal University, China), Jun Liu (Tencent Youtu Lab), Yanyun Qu (Xiamen University, China), Yuan Xie (East China Normal University, China), and Lizhuang Ma (East China Normal University, China)</i>	

Neuro-Modulated Hebbian Learning for Fully Test-Time Adaptation	3728
<i>Yushun Tang (Southern University of Science and Technology, China; Advanced Computing and Storage Laboratory, Huawei Technologies Co., Ltd., China), Ce Zhang (Southern University of Science and Technology, China), Heng Xu (Southern University of Science and Technology, China), Shuoshuo Chen (Southern University of Science and Technology, China), Jie Cheng (Advanced Computing and Storage Laboratory, Huawei Technologies Co., Ltd., China), Luziwei Leng (Advanced Computing and Storage Laboratory, Huawei Technologies Co., Ltd., China), Qinghai Guo (Advanced Computing and Storage Laboratory, Huawei Technologies Co., Ltd., China), and Zhihai He (Southern University of Science and Technology, China; Pengcheng Laboratory, China)</i>	
Generalizing Dataset Distillation via Deep Generative Prior	3739
<i>George Cazenavette (Massachusetts Institute of Technology), Tongzhou Wang (Massachusetts Institute of Technology), Antonio Torralba (Massachusetts Institute of Technology), Alexei A. Efros (UC Berkeley), and Jun-Yan Zhu (Carnegie Mellon University)</i>	
Minimizing the Accumulated Trajectory Error To Improve Dataset Distillation	3749
<i>Jiawei Du (Technology and Research (A*STAR), Singapore; National University of Singapore), Yidi Jiang (National University of Singapore), Vincent Y. F. Tan (National University of Singapore), Joey Tianyi Zhou (Technology and Research (A*STAR), Singapore; National University of Singapore), and Haizhou Li (The Chinese University of Hong Kong, Shenzhen, China; National University of Singapore)</i>	
Slimmable Dataset Condensation	3759
<i>Songhua Liu (National University of Singapore), Jingwen Ye (National University of Singapore), Runpeng Yu (National University of Singapore), and Xinchao Wang (National University of Singapore)</i>	
Sharpness-Aware Gradient Matching for Domain Generalization	3769
<i>Pengfei Wang (The Hong Kong Polytechnic University; Center for Artificial Intelligence and Robotics, HKISI, CAS), Zhaoxiang Zhang (The Hong Kong Polytechnic University; Center for Artificial Intelligence and Robotics, HKISI, CAS; State Key Laboratory of Multimodal Artificial Intelligence Systems, CASIA), Zhen Lei (The Hong Kong Polytechnic University; Center for Artificial Intelligence and Robotics, HKISI, CAS; State Key Laboratory of Multimodal Artificial Intelligence Systems, CASIA), and Lei Zhang (The Hong Kong Polytechnic University)</i>	
Dynamic Neural Network for Multi-Task Learning Searching Across Diverse Network Topologies...	3779
<i>Wonhyeok Choi (Department of Electrical Engineering & Computer Science, DGIST, Korea) and Sunghoon Im (Department of Electrical Engineering & Computer Science, DGIST, Korea)</i>	
SplineCam: Exact Visualization and Characterization of Deep Network Geometry and Decision Boundaries	3789
<i>Ahmed Imtiaz Humayun (Rice University), Randall Balestriero (Meta AI, FAIR), Guha Balakrishnan (Rice University), and Richard G. Baraniuk (Rice University)</i>	

VNE: An Effective Method for Improving Deep Representation by Manipulating Eigenvalue Distribution	3799
<i>Jaeill Kim (Seoul National University), Suhyun Kang (Seoul National University), Duhun Hwang (Seoul National University), Jungwook Shin (Seoul National University), and Wonjong Rhee (Seoul National University; Seoul National University; Seoul National University)</i>	
Efficient On-Device Training via Gradient Filtering	3811
<i>Yuedong Yang (The University of Texas at Austin), Guihong Li (The University of Texas at Austin), and Radu Marculescu (The University of Texas at Austin)</i>	
Are Data-Driven Explanations Robust Against Out-of-Distribution Data?	3821
<i>Tang Li (University of Delaware), Fengchun Qiao (University of Delaware), Mengmeng Ma (University of Delaware), and Xi Peng (University of Delaware)</i>	
BiasAdv: Bias-Adversarial Augmentation for Model Debiasing	3832
<i>Jongin Lim (Samsung Advanced Institute of Technology (SAIT)), Youngdong Kim (Samsung Advanced Institute of Technology (SAIT)), Byungjai Kim (Samsung Advanced Institute of Technology (SAIT)), Chanho Ahn (Samsung Advanced Institute of Technology (SAIT)), Jinwoo Shin (Korea Advanced Institute of Science and Technology (KAIST)), Eunho Yang (Korea Advanced Institute of Science and Technology (KAIST)), and Seungju Han (Samsung Advanced Institute of Technology (SAIT))</i>	
Q-DETR: An Efficient Low-Bit Quantized Detection Transformer	3842
<i>Sheng Xu (Beihang University), Yanjing Li (Beihang University), Mingbao Lin (Tencent), Peng Gao (Shanghai AI Laboratory), Guodong Guo (UNIUBI Research, Universal Ubiquitous Co.), Jinhu Lü (Beihang University; Zhongguancun Laboratory), and Baochang Zhang (Beihang University; Zhongguancun Laboratory)</i>	
NIPQ: Noise Proxy-Based Integrated Pseudo-Quantization	3852
<i>Juncheol Shin (Graduate School of Artificial Intelligence, POSTECH), Junhyuk So (Department of Computer Science and Engineering, POSTECH), Sein Park (Graduate School of Artificial Intelligence, POSTECH), Seungyeop Kang (Seoul National University), Sungjoo Yoo (Seoul National University), and Eunhyeok Park (Department of Computer Science and Engineering, POSTECH; Graduate School of Artificial Intelligence, POSTECH)</i>	
CUDA: Convolution-Based Unlearnable Datasets	3862
<i>Vinu Sankar Sadasivan (University of Maryland), Mahdi Soltanolkotabi (University of Southern California), and Soheil Feizi (University of Maryland)</i>	
KD-DLGAN: Data Limited Image Generation via Knowledge Distillation	3872
<i>Kaiwen Cui (Nanyang Technological University), Yingchen Yu (Nanyang Technological University), Fangneng Zhan (Max Planck Institute for Informatics), Shengcai Liao (Inception Institute of Artificial Intelligence), Shijian Lu (Nanyang Technological University), and Eric P. Xing (Mohamed bin Zayed University of Artificial Intelligence)</i>	

Spider GAN: Leveraging Friendly Neighbors To Accelerate GAN Training	3883
<i>Siddarth Asokan (Robert Bosch Center for Cyber-Physical Systems, Indian Institute of Science, India) and Chandra Sekhar Seelamantula (Department of Electrical Engineering, Indian Institute of Science, India)</i>	
Efficient Verification of Neural Networks Against LVM-Based Specifications	3894
<i>Harleen Hanspal (Imperial College London & Safe Intelligence, United Kingdom) and Alessio Lomuscio (Safe Intelligence, United Kingdom)</i>	
Bi-Directional Feature Fusion Generative Adversarial Network for Ultra-High Resolution Pathological Image Virtual Re-Staining	3904
<i>Kexin Sun (Fudan University), Zhineng Chen (Fudan University; Shanghai Qi Zhi Institute), Gongwei Wang (Peking University People's Hospital), Jun Liu (Peking University People's Hospital), Xiongjun Ye (Chinese Academy of Medical Sciences and Peking Union Medical College), and Yu-Gang Jiang (Fudan University)</i>	
DeSTSeg: Segmentation Guided Denoising Student-Teacher for Anomaly Detection	3914
<i>Xuan Zhang (Tsinghua University), Shiyu Li (Apple), Xi Li (Apple), Ping Huang (Apple), Jiulong Shan (Apple), and Ting Chen (Tsinghua University)</i>	
OmniAL: A Unified CNN Framework for Unsupervised Anomaly Localization	3924
<i>Ying Zhao (Ricoh Software Research Center (Beijing) Co., Ltd.)</i>	
Federated Incremental Semantic Segmentation	3934
<i>Jiahua Dong (Shenyang Institute of Automation, Chinese Academy of Sciences, China; Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Duzhen Zhang (ETH Zürich, Switzerland), Yang Cong (Shenyang Institute of Automation, Chinese Academy of Sciences, China; Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, China), Wei Cong (Shenyang Institute of Automation, Chinese Academy of Sciences, China; Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, China), Henghui Ding (ETH Zürich, Switzerland), and Dengxin Dai (ETH Zürich, Switzerland)</i>	
Re-Thinking Federated Active Learning Based on Inter-Class Diversity	3944
<i>SangMook Kim (KAIST AI), Sangmin Bae (KAIST AI), Hwanjun Song (NAVER AI LAB), and Se-Young Yun (KAIST AI)</i>	
Federated Domain Generalization With Generalization Adjustment	3954
<i>Ruipeng Zhang (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University; Shanghai AI Laboratory), Qinwei Xu (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University; Shanghai AI Laboratory), Jiangchao Yao (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University; Shanghai AI Laboratory), Ya Zhang (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University; Shanghai AI Laboratory), Qi Tian (Huawei Cloud & AI), and Yanfeng Wang (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University; Shanghai AI Laboratory)</i>	

On the Effectiveness of Partial Variance Reduction in Federated Learning With Heterogeneous Data	3964
<i>Bo Li (Technical University of Denmark), Mikkel N. Schmidt (Technical University of Denmark), Tommy S. Alstrøm (Technical University of Denmark), and Sebastian U. Stich (CISPA)</i>	
The Resource Problem of Using Linear Layer Leakage Attack in Federated Learning	3974
<i>Joshua C. Zhao (Purdue University), Ahmed Roushdy Elkordy (University of Southern California), Atul Sharma (Purdue University), Yahya H. Ezzeldin (University of Southern California), Salman Avestimehr (University of Southern California), and Saurabh Bagchi (Purdue University)</i>	
Unlearnable Clusters: Towards Label-Agnostic Unlearnable Examples	3984
<i>Jiaming Zhang (Beijing Jiaotong University), Xingjun Ma (Fudan University), Qi Yi (Beijing Jiaotong University), Jitao Sang (Beijing Jiaotong University; Peng Cheng Lab), Yu-Gang Jiang (Fudan University), Yaowei Wang (Peng Cheng Lab), and Changsheng Xu (Chinese Academy of Sciences; Peng Cheng Lab)</i>	
Implicit Identity Leakage: The Stumbling Block to Improving Deepfake Detection Generalization	3994
<i>Shichao Dong (MEGVII Technology), Jin Wang (MEGVII Technology), Renhe Ji (MEGVII Technology), Jiajun Liang (MEGVII Technology), Haoqiang Fan (MEGVII Technology), and Zheng Ge (MEGVII Technology)</i>	
Backdoor Defense via Adaptively Splitting Poisoned Dataset	4005
<i>Kuofeng Gao (Tsinghua University), Yang Bai (Tencent Security Zhuque Lab), Jindong Gu (University of Oxford), Yong Yang (Tencent Security Platform Department), and Shu-Tao Xia (Peng Cheng Laboratory)</i>	
How to Backdoor Diffusion Models?	4015
<i>Sheng-Yen Chou (National Tsing Hua University, R.O.C; The Chinese University of Hong Kong, Hong Kong), Pin-Yu Chen (IBM Research, USA), and Tsung-Yi Ho (The Chinese University of Hong Kong, Hong Kong)</i>	
TrojViT: Trojan Insertion in Vision Transformers	4025
<i>Mengxin Zheng (Indiana University Bloomington), Qian Lou (University of Central Florida), and Lei Jiang (Indiana University Bloomington)</i>	
TrojDiff: Trojan Attacks on Diffusion Models With Diverse Targets	4035
<i>Weixin Chen (UIUC), Dawn Song (UC Berkeley), and Bo Li (UIUC)</i>	
Ensemble-Based Blackbox Attacks on Dense Prediction	4045
<i>Zikui Cai (University of California Riverside), Yaoteng Tan (University of California Riverside), and M. Salman Asif (University of California Riverside)</i>	
Efficient Loss Function by Minimizing the Detrimental Effect of Floating-Point Errors on Gradient-Based Attacks	4056
<i>Yunrui Yu (University of Macau, Macau SAR, China) and Cheng-Zhong Xu (University of Macau, Macau SAR, China)</i>	
The Best Defense Is a Good Offense: Adversarial Augmentation Against Adversarial Attacks .	4067
<i>Iuri Frosio (NVIDIA) and Jan Kautz (NVIDIA)</i>	
Adversarial Robustness via Random Projection Filters	4077
<i>Minjing Dong (University of Sydney) and Chang Xu (University of Sydney)</i>	

Jedi: Entropy-Based Localization and Removal of Adversarial Patches	4087
<i>Bilel Tarchoun (Université de Sousse, Ecole Nationale d'Ingénieurs de Sousse, LATIS, Tunisia), Anouar Ben Khalifa (Université de Sousse, Ecole Nationale d'Ingénieurs de Sousse, LATIS, Tunisia; Université de Jendouba, Institut National des Technologies et des Sciences du Kef, Tunisia), Mohamed Ali Mahjoub (Université de Sousse, Ecole Nationale d'Ingénieurs de Sousse, LATIS, Tunisia), Nael Abu-Ghazaleh (University of California Riverside, USA), and Ihsen Alouani (CSIT, Queen's University Belfast, UK' IEMN CNRS 8520, Université Polytechnique Hauts-de-France)</i>	
Exploring the Relationship Between Architectural Design and Adversarially Robust Generalization	4096
<i>Aishan Liu (Beihang University), Shiyu Tang (Beihang University), Siyuan Liang (Chinese Academy of Sciences), Ruihao Gong (Beihang University; SenseTime), Boxi Wu (Zhejiang University), Xianglong Liu (Beihang University; Zhongguancun Laboratory; Hefei Comprehensive National Science Center), and Dacheng Tao (JD Explore Academy)</i>	
Improving Robustness of Vision Transformers by Reducing Sensitivity To Patch Corruptions .	4108
<i>Yong Guo (Max Planck Institute for Informatics, Saarland Informatics Campus), David Stutz (Max Planck Institute for Informatics, Saarland Informatics Campus), and Bernt Schiele (Max Planck Institute for Informatics, Saarland Informatics Campus)</i>	
Towards Effective Adversarial Textured 3D Meshes on Physical Face Recognition	4119
<i>Xiao Yang (Dept, of Comp. Sei. and Tech., Institute for AI, BNRist Center, Tsinghua-Bosch Joint ML Center, THBI Lab, Tsinghua-China Mobile Communications Group Co., Ltd. Joint Institute, Tsinghua University), Chang Liu (Peking University), Longlong Xu (Dept, of Comp. Sei. and Tech., Institute for AI, BNRist Center, Tsinghua-Bosch Joint ML Center, THBI Lab, Tsinghua-China Mobile Communications Group Co., Ltd. Joint Institute, Tsinghua University), Yikai Wang (Dept, of Comp. Sei. and Tech., Institute for AI, BNRist Center, Tsinghua-Bosch Joint ML Center, THBI Lab, Tsinghua-China Mobile Communications Group Co., Ltd. Joint Institute, Tsinghua University), Yinpeng Dong (Dept, of Comp. Sei. and Tech., Institute for AI, BNRist Center, Tsinghua-Bosch Joint ML Center, THBI Lab, Tsinghua-China Mobile Communications Group Co., Ltd. Joint Institute, Tsinghua University; RealAI), Ning Chen (Dept, of Comp. Sei. and Tech., Institute for AI, BNRist Center, Tsinghua-Bosch Joint ML Center, THBI Lab, Tsinghua-China Mobile Communications Group Co., Ltd. Joint Institute, Tsinghua University), Hang Su (Dept, of Comp. Sei. and Tech., Institute for AI, BNRist Center, Tsinghua-Bosch Joint ML Center, THBI Lab, Tsinghua-China Mobile Communications Group Co., Ltd. Joint Institute, Tsinghua University; Zhongguancun Laboratory), and Jun Zhu (Dept, of Comp. Sei. and Tech., Institute for AI, BNRist Center, Tsinghua-Bosch Joint ML Center, THBI Lab, Tsinghua-China Mobile Communications Group Co., Ltd. Joint Institute, Tsinghua University; RealAI; Zhongguancun Laboratory)</i>	

AltFreezing for More General Video Face Forgery Detection	4129
<i>Zhendong Wang (EEIS Department, University of Science and Technology of China), Jianmin Bao (Microsoft Research Asia), Wengang Zhou (EEIS Department, University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Weilun Wang (EEIS Department, University of Science and Technology of China), and Houqiang Li (EEIS Department, University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center)</i>	

Poster-Tue-PM

Passive Micron-Scale Time-of-Flight With Sunlight Interferometry	4139
<i>Alankar Kotwal (Carnegie Mellon University), Anat Levin (Technion), and Ioannis Gkioulekas (Carnegie Mellon University)</i>	
F2-NeRF: Fast Neural Radiance Field Training With Free Camera Trajectories	4150
<i>Peng Wang (The University of Hong Kong; Nanyang Technological University), Yuan Liu (The University of Hong Kong), Zhaoxi Chen (Nanyang Technological University), Lingjie Liu (Max Planck Institute for Informatics), Ziwei Liu (Nanyang Technological University), Taku Komura (The University of Hong Kong), Christian Theobalt (Max Planck Institute for Informatics), and Wenping Wang (Texas A&M University)</i>	
NoPe-NeRF: Optimising Neural Radiance Field With No Pose Prior	4160
<i>Wenjing Bian (University of Oxford), Zirui Wang (University of Oxford), Kejie Li (University of Oxford), Jia-Wang Bian (University of Oxford), and Victor Adrian Prisacariu (University of Oxford)</i>	
BAD-NeRF: Bundle Adjusted Deblur Neural Radiance Fields	4170
<i>Peng Wang (Zhejiang University; Westlake University), Lingzhe Zhao (Westlake University), Ruijie Ma (Westlake University), and Peidong Liu (Westlake University)</i>	
DiffusioNeRF: Regularizing Neural Radiance Fields With Denoising Diffusion Models	4180
<i>Jamie Wynn (Niantic) and Daniyar Turmukhambetov (Niantic)</i>	
SPARF: Neural Radiance Fields From Sparse and Noisy Poses	4190
<i>Prune Truong (ETH Zurich; Google), Marie-Julie Rakotosaona (Google), Fabian Manhardt (Google), and Federico Tombari (Google; Technical University of Munich)</i>	
Interactive Segmentation of Radiance Fields	4201
<i>Rahul Goel (CVIT, Kohli Center on Intelligent Systems (KCIS), IIIT Hyderabad), Dhawal Sirikonda (CVIT, Kohli Center on Intelligent Systems (KCIS), IIIT Hyderabad), Saurabh Saini (CVIT, Kohli Center on Intelligent Systems (KCIS), IIIT Hyderabad), and P. J. Narayanan (CVIT, Kohli Center on Intelligent Systems (KCIS), IIIT Hyderabad)</i>	

Temporal Interpolation Is All You Need for Dynamic Neural Radiance Fields	4212
<i>Sunghoon Park (Samsung Advanced Institute of Technology (SAIT)), Minjung Son (Samsung Advanced Institute of Technology (SAIT)), Seokhwan Jang (Samsung Advanced Institute of Technology (SAIT)), Young Chun Ahn (Samsung Advanced Institute of Technology (SAIT)), Ji-Yeon Kim (Samsung Advanced Institute of Technology (SAIT)), and Nahyup Kang (Samsung Advanced Institute of Technology (SAIT))</i>	
Compressing Volumetric Radiance Fields to 1 MB	4222
<i>Lingzhi Li (Alibaba Group, Beijing, China), Zhen Shen (Alibaba Group, Beijing, China), Zhongshu Wang (Alibaba Group, Beijing, China), Li Shen (Alibaba Group, Beijing, China), and Liefeng Bo (Alibaba Group, Beijing, China)</i>	
Multiscale Tensor Decomposition and Rendering Equation Encoding for View Synthesis	4232
<i>Kang Han (James Cook University) and Wei Xiang (La Trobe University)</i>	
Ref-NPR: Reference-Based Non-Photorealistic Radiance Fields for Controllable Scene Stylization	4242
<i>Yuechen Zhang (The Chinese University of Hong Kong; SmartMore), Zexin He (The Chinese University of Hong Kong), Jinbo Xing (The Chinese University of Hong Kong), Xufeng Yao (The Chinese University of Hong Kong), and Jiaya Jia (The Chinese University of Hong Kong; SmartMore)</i>	
Representing Volumetric Videos As Dynamic MLP Maps	4252
<i>Sida Peng (Zhejiang University), Yunzhi Yan (Zhejiang University), Qing Shuai (Zhejiang University), Hujun Bao (Zhejiang University), and Xiaowei Zhou (Zhejiang University)</i>	
Fast Monocular Scene Reconstruction With Global-Sparse Local-Dense Grids	4263
<i>Wei Dong (NVIDIA Research), Christopher Choy (NVIDIA Research), Charles Loop (NVIDIA Research), Or Litany (NVIDIA Research), Yuke Zhu (NVIDIA Research), and Anima Anandkumar (NVIDIA Research)</i>	
DynIBaR: Neural Dynamic Image-Based Rendering	4273
<i>Zhengqi Li (Google Research), Qianqian Wang (Google Research; Cornell Tech), Forrester Cole (Google Research), Richard Tucker (Google Research), and Noah Snavely (Google Research)</i>	
Plateau-Reduced Differentiable Path Tracing	4285
<i>Michael Fischer (University College London) and Tobias Ritschel (University College London)</i>	
NeFII: Inverse Rendering for Reflectance Decomposition With Near-Field Indirect Illumination	4295
<i>Haoqian Wu (NetEase Fuxi AI Lab), Zhipeng Hu (NetEase Fuxi AI Lab; Zhejiang University), Lincheng Li (NetEase Fuxi AI Lab), Yongqiang Zhang (NetEase Fuxi AI Lab), Changjie Fan (NetEase Fuxi AI Lab), and Xin Yu (The University of Queensland)</i>	
WildLight: In-the-Wild Inverse Rendering With a Flashlight	4305
<i>Ziang Cheng (Australian National University), Junxuan Li (Australian National University), and Hongdong Li (Australian National University)</i>	

Relightable Neural Human Assets From Multi-View Gradient Illuminations	4315
<i>Taotao Zhou (ShanghaiTech University; LumiAni Technology), Kai He (ShanghaiTech University), Di Wu (Deemos Technology), Teng Xu (ShanghaiTech University; LumiAni Technology), Qixuan Zhang (ShanghaiTech University; Deemos Technology), Kuixiang Shao (ShanghaiTech University; LumiAni Technology), Wenzheng Chen (University of Toronto), Lan Xu (ShanghaiTech University), and Jingyi Yu (ShanghaiTech University)</i>	
DiffRF: Rendering-Guided 3D Radiance Field Diffusion	4328
<i>Norman Müller (Technical University of Munich; Meta Reality Labs Zurich), Yawar Siddiqui (Technical University of Munich; Meta Reality Labs Zurich), Lorenzo Porzi (Meta Reality Labs Zurich), Samuel Rota Bulò (Meta Reality Labs Zurich), Peter Kotschieder (Meta Reality Labs Zurich), and Matthias Nießner (Technical University of Munich)</i>	
Analyzing Physical Impacts Using Transient Surface Wave Imaging	4339
<i>Tianyuan Zhang (Carnegie Mellon University), Mark Sheinin (Carnegie Mellon University), Dorian Chan (Carnegie Mellon University), Mark Rau (Stanford University), Matthew O'Toole (Carnegie Mellon University), and Srinivasa G. Narasimhan (Carnegie Mellon University)</i>	
Neural Kaleidoscopic Space Sculpting	4349
<i>Byeongjoo Ahn (Carnegie Mellon University), Michael De Zeeuw (Carnegie Mellon University), Ioannis Gkioulekas (Carnegie Mellon University), and Aswin C. Sankaranarayanan (Carnegie Mellon University)</i>	
Towards Unbiased Volume Rendering of Neural Implicit Surfaces With Geometry Priors	4359
<i>Yongqiang Zhang (NetEase Fuxi AI Lab), Zhipeng Hu (NetEase Fuxi AI Lab), Haoqian Wu (NetEase Fuxi AI Lab), Minda Zhao (NetEase Fuxi AI Lab), Lincheng Li (NetEase Fuxi AI Lab), Zhengxia Zou (Beihang University), and Changjie Fan (NetEase Fuxi AI Lab)</i>	
Neural Kernel Surface Reconstruction	4369
<i>Jiahui Huang (NVIDIA), Zan Gojcic (NVIDIA), Matan Atzmon (NVIDIA), Or Litany (NVIDIA), Sanja Fidler (NVIDIA; University of Toronto; Vector Institute), and Francis Williams (NVIDIA)</i>	
MM-3DScene: 3D Scene Understanding by Customizing Masked Modeling With Informative-Preserved Reconstruction and Self-Distilled Consistency	4380
<i>Mingye Xu (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences; Shanghai Artificial Intelligence Laboratory), Mutian Xu (SSE, CUHKSZ; FNii, CUHKSZ), Tong He (Shanghai Artificial Intelligence Laboratory), Wanli Ouyang (Shanghai Artificial Intelligence Laboratory), Yali Wang (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China; Shanghai Artificial Intelligence Laboratory), Xiaoguang Han (SSE, CUHKSZ; FNii, CUHKSZ), and Yu Qiao (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China; Shanghai Artificial Intelligence Laboratory)</i>	
Shape, Pose, and Appearance From a Single Image via Bootstrapped Radiance Field Inversion	4391
<i>Dario Pavllo (ETH Zurich), David Joseph Tan (ETH Zurich; Google), Marie-Julie Rakotosaona (Google), and Federico Tombari (Google; TU Munich)</i>	

DisCoScene: Spatially Disentangled Generative Radiance Fields for Controllable 3D-Aware Scene Synthesis	4402
<i>Yinghao Xu (CUHK; Snap Inc.), Menglei Chai (Snap Inc.), Zifan Shi (HKUST), Sida Peng (KAUST), Ivan Skorokhodov (Snap Inc.; KAUST), Aliaksandr Siarohin (Snap Inc.), Ceyuan Yang (CUHK), Yujun Shen (CUHK), Hsin-Ying Lee (Snap Inc.), Bolei Zhou (UCLA), and Sergey Tulyakov (Snap Inc.)</i>	
Heat Diffusion Based Multi-Scale and Geometric Structure-Aware Transformer for Mesh Segmentation	4413
<i>Chi-Chong Wong (University of Macau)</i>	
Learning Detailed Radiance Manifolds for High-Fidelity and 3D-Consistent Portrait Synthesis From Monocular Image	4423
<i>Yu Deng (Xiaobing.AI), Baoyuan Wang (Xiaobing.AI), and Heung-Yeung Shum (Xiaobing.AI)</i>	
3D-Aware Conditional Image Synthesis	4434
<i>Kangle Deng (Carnegie Mellon University), Gengshan Yang (Carnegie Mellon University), Deva Ramanan (Carnegie Mellon University), and Jun-Yan Zhu (Carnegie Mellon University)</i>	
VIVE3D: Viewpoint-Independent Video Editing Using 3D-Aware GANs	4446
<i>Anna Frühstück (Meta Reality Labs Research), Nikolaos Sarafianos (Meta Reality Labs Research), Yuanlu Xu (Meta Reality Labs Research), Peter Wonka (KAUST), and Tony Tung (Meta Reality Labs Research)</i>	
SDFusion: Multimodal 3D Shape Completion, Reconstruction, and Generation	4456
<i>Yen-Chi Cheng (University of Illinois Urbana-Champaign), Hsin-Ying Lee (Snap Research), Sergey Tulyakov (Snap Research), Alexander G. Schwing (University of Illinois Urbana-Champaign), and Liang-Yan Gui (University of Illinois Urbana-Champaign)</i>	
Generating Part-Aware Editable 3D Shapes Without 3D Supervision	4466
<i>Konstantinos Tertikas (National and Kapodistrian University of Athens; Athena RC, Greece), Despoina Paschalidou (Stanford University), Boxiao Pan (Stanford University), Jeong Joon Park (Stanford University), Mikaela Angelina Uy (Stanford University), Ioannis Emiris (Athena RC, Greece; National and Kapodistrian University of Athens), Yannis Avrithis (Institute of Advanced Research in Artificial Intelligence (IARA)), and Leonidas Guibas (Stanford University)</i>	
NeuralLift-360: Lifting an In-the-Wild 2D Photo to a 3D Object With 360° Views	4479
<i>Dejia Xu (University of Texas at Austin), Yifan Jiang (University of Texas at Austin), Peihao Wang (University of Texas at Austin), Zhiwen Fan (University of Texas at Austin), Yi Wang (University of Texas at Austin), and Zhangyang Wang (University of Texas at Austin)</i>	
Implicit Identity Driven Deepfake Face Swapping Detection	4490
<i>Baojin Huang (Wuhan University), Zhongyuan Wang (Wuhan University), Jifan Yang (Wuhan University), Jiaxin Ai (Wuhan University), Qin Zou (Wuhan University), Qian Wang (Wuhan University), and Dengpan Ye (Wuhan University)</i>	

Canonical Fields: Self-Supervised Learning of Pose-Canonicalized Neural Fields	4500
<i>Rohith Agaram (RRC, IIIT-Hyderabad), Shaurya Dewan (RRC, IIIT-Hyderabad), Rahul Sajnani (Brown University), Adrien Poulernard (Stanford University), Madhava Krishna (RRC, IIIT-Hyderabad), and Srinath Sridhar (Brown University)</i>	
Improving Fairness in Facial Albedo Estimation via Visual-Textual Cues	4511
<i>Xingyu Ren (Shanghai Jiao Tong University), Jiankang Deng (Shanghai Jiao Tong University), Chao Ma (Shanghai Jiao Tong University), Yichao Yan (Shanghai Jiao Tong University), and Xiaokang Yang (Shanghai Jiao Tong University)</i>	
High-Fidelity 3D Face Generation From Natural Language Descriptions	4521
<i>Menghua Wu (Nanjing University, China), Hao Zhu (Nanjing University, China), Linjia Huang (Nanjing University, China), Yiyu Zhuang (Nanjing University, China), Yuanxun Lu (Nanjing University, China), and Xun Cao (Nanjing University, China)</i>	
DSFNet: Dual Space Fusion Network for Occlusion-Robust 3D Dense Face Alignment	4531
<i>Heyuan Li (National University of Singapore), Bo Wang (CtrsVision), Yu Cheng (National University of Singapore), Mohan Kankanhalli (National University of Singapore), and Robby T. Tan (National University of Singapore)</i>	
High-Fidelity Facial Avatar Reconstruction From Monocular Video With Generative Priors	4541
<i>Yunpeng Bai (Tsinghua Shenzhen International Graduate School), Yanbo Fan (Tencent AI Lab), Xuan Wang (Ant Group), Yong Zhang (Tencent AI Lab), Jingxiang Sun (Tsinghua University), Chun Yuan (Tsinghua Shenzhen International Graduate School; Peng Cheng Laboratory), and Ying Shan (Tencent AI Lab)</i>	
3DAvatarGAN: Bridging Domains for Personalized Editable Avatars	4552
<i>Rameen Abdal (KAUST), Hsin-Ying Lee (Snap Inc.), Peihao Zhu (KAUST), Menglei Chai (Snap Inc.), Aliaksandr Siarohin (Snap Inc.), Peter Wonka (KAUST), and Sergey Tulyakov (Snap Inc.)</i>	
RODIN: A Generative Model for Sculpting 3D Digital Avatars Using Diffusion	4563
<i>Tengfei Wang (HKUST), Bo Zhang (Microsoft Research), Ting Zhang (Microsoft Research), Shuyang Gu (Microsoft Research), Jianmin Bao (Microsoft Research), Tadas Baltrusaitis (Microsoft Research), Jingjing Shen (Microsoft Research), Dong Chen (Microsoft Research), Fang Wen (Microsoft Research), Qifeng Chen (HKUST), and Baining Guo (Microsoft Research)</i>	
Instant Volumetric Head Avatars	4574
<i>Wojciech Zielonka (Max Planck Institute for Intelligent Systems, Germany), Timo Bolkart (Max Planck Institute for Intelligent Systems, Germany), and Justus Thies (Max Planck Institute for Intelligent Systems, Germany)</i>	
Synthesizing Photorealistic Virtual Humans Through Cross-Modal Disentanglement	4585
<i>Siddarth Ravichandran (NEON, Samsung Research America), Ondřej Texler (NEON, Samsung Research America), Dimitar Dinev (NEON, Samsung Research America), and Hyun Jae Kang (NEON, Samsung Research America)</i>	

3D Cinemagraphy From a Single Image	4595
<i>Xingyi Li (Huazhong University of Science and Technology; Nanyang Technological University), Zhiguo Cao (Huazhong University of Science and Technology), Huiqiang Sun (Huazhong University of Science and Technology), Jianming Zhang (Adobe Research), Ke Xian (Nanyang Technological University), and Guosheng Lin (Nanyang Technological University)</i>	
TryOnDiffusion: A Tale of Two UNets	4606
<i>Luyang Zhu (University of Washington; Google Research), Dawei Yang (Google Research), Tyler Zhu (Google Research), Fitsum Reda (University of Washington; Google Research), William Chan (Google Research), Chitwan Saharia (Google Research), Mohammad Norouzi (Google Research), and Ira Kemelmacher-Shlizerman (University of Washington; Google Research)</i>	
Diverse 3D Hand Gesture Prediction From Body Dynamics by Bilateral Hand Disentanglement	4616
<i>Xingqun Qi (University of Technology Sydney; Netease Fuxi AI Lab), Chen Liu (The University of Queensland), Muyi Sun (CRIPAC, NLPR, Institute of Automation, Chinese Academy of Sciences), Lincheng Li (Netease Fuxi AI Lab), Changjie Fan (Netease Fuxi AI Lab), and Xin Yu (University of Technology Sydney; The University of Queensland)</i>	
Normal-Guided Garment UV Prediction for Human Re-Texturing	4627
<i>Yasamin Jafarian (University of Minnesota), Tuanfeng Y. Wang (Adobe Research), Duygu Ceylan (Adobe Research), Jimei Yang (Adobe Research), Nathan Carr (Adobe Research), Yi Zhou (Adobe Research), and Hyun Soo Park (University of Minnesota)</i>	
REC-MV: REconstructing 3D Dynamic Cloth From Monocular Videos	4637
<i>Lingteng Qiu (SSE, CUHKSZ), Guanying Chen (SSE, CUHKSZ; FNii, CUHKSZ), Jiapeng Zhou (SSE, CUHKSZ), Mutian Xu (SSE, CUHKSZ), Junle Wang (Tencent), and Xiaoguang Han (SSE, CUHKSZ; FNii, CUHKSZ)</i>	
SeSDF: Self-Evolved Signed Distance Field for Implicit 3D Clothed Human Reconstruction	4647
<i>Yukang Cao (The University of Hong Kong), Kai Han (The University of Hong Kong), and Kwan-Yee K. Wong (The University of Hong Kong)</i>	
Unsupervised Volumetric Animation	4658
<i>Aliaksandr Siarohin (Snap Inc.), Willi Menapace (University of Trento), Ivan Skorokhodov (KAUST), Kyle Olszewski (Snap Inc.), Jian Ren (Snap Inc.), Hsin-Ying Lee (Snap Inc.), Menglei Chai (Snap Inc.), and Sergey Tulyakov (Snap Inc.)</i>	
Handy: Towards a High Fidelity 3D Hand Shape and Appearance Model	4670
<i>Rolandos Alexandros Potamias (Imperial College London), Stylianos Ploumpis (Imperial College London), Stylianos Moschoglou (Imperial College London), Vasileios Triantafyllou (Cosmos Designs Ltd), and Stefanos Zafeiriou (Imperial College London)</i>	
Fantastic Breaks: A Dataset of Paired 3D Scans of Real-World Broken Objects and Their Complete Counterparts	4681
<i>Nikolas Lamb (Clarkson University, USA), Cameron Palmer (Clarkson University, USA), Benjamin Molloy (Clarkson University, USA), Sean Banerjee (Clarkson University, USA), and Natasha Kholgade Banerjee (Clarkson University, USA)</i>	

Distilling Neural Fields for Real-Time Articulated Shape Reconstruction	4692
<i>Jeff Tan (Carnegie Mellon University), Gengshan Yang (Carnegie Mellon University), and Deva Ramanan (Carnegie Mellon University)</i>	
GANmouflage: 3D Object Nondetection With Texture Fields	4702
<i>Rui Guo (University of Michigan; XMotors.ai), Jasmine Collins (UC Berkeley), Oscar de Lima (University of Michigan), and Andrew Owens (University of Michigan)</i>	
3D Human Pose Estimation via Intuitive Physics	4713
<i>Shashank Tripathi (Max Planck Institute for Intelligent Systems, Germany), Lea Müller (Max Planck Institute for Intelligent Systems, Germany), Chun-Hao P. Huang (Max Planck Institute for Intelligent Systems, Germany), Omid Taheri (Max Planck Institute for Intelligent Systems, Germany), Michael J. Black (Max Planck Institute for Intelligent Systems, Germany), and Dimitrios Tzionas (University of Amsterdam, the Netherlands)</i>	
Object Pop-Up: Can We Infer 3D Objects and Their Poses From Human Interactions Alone? ..	4726
<i>Ilya A. Petrov (University of Tübingen, Germany), Riccardo Marin (University of Tübingen, Germany), Julian Chibane (University of Tübingen, Germany), and Gerard Pons-Moll (University of Tübingen, Germany; Max Planck Institute for Informatics, Saarland Informatics Campus, Germany; Tübingen AI Center, Germany)</i>	
UniDexGrasp: Universal Robotic Dexterous Grasping via Learning Diverse Proposal Generation and Goal-Conditioned Policy	4737
<i>Yinzhen Xu (Center on Frontiers of Computing Studies, Peking University; Peking University; Beijing Institute for General AI), Weikang Wan (Center on Frontiers of Computing Studies, Peking University; Peking University), Jialiang Zhang (Center on Frontiers of Computing Studies, Peking University; Peking University), Haoran Liu (Center on Frontiers of Computing Studies, Peking University; Peking University), Zikang Shan (Center on Frontiers of Computing Studies, Peking University), Hao Shen (Center on Frontiers of Computing Studies, Peking University), Ruicheng Wang (Center on Frontiers of Computing Studies, Peking University), Haoran Geng (Center on Frontiers of Computing Studies, Peking University; Peking University), Yijia Weng (Stanford University), Jiayi Chen (Center on Frontiers of Computing Studies, Peking University), Tengyu Liu (Beijing Institute for General AI), Li Yi (Tsinghua University), and He Wang (Center on Frontiers of Computing Studies, Peking University; Peking University)</i>	
Constrained Evolutionary Diffusion Filter for Monocular Endoscope Tracking	4747
<i>Xiongbiao Luo (Xiamen University, China; National Institute for Data Science in Health and Medicine, Xiamen University, China)</i>	
Visibility Aware Human-Object Interaction Tracking From Single RGB Camera	4757
<i>Xianghui Xie (University of Tübingen, Germany), Bharat Lal Bhatnagar (University of Tübingen, Germany), and Gerard Pons-Moll (University of Tübingen, Germany)</i>	

Transformer-Based Unified Recognition of Two Hands Manipulating Objects	4769
<i>Hoseong Cho (UNIST, South Korea), Chanwoo Kim (UNIST, South Korea), Jihyeon Kim (UNIST, South Korea), Seongyeong Lee (UNIST, South Korea), Elkhan Ismayilzada (UNIST, South Korea), and Seungryul Baek (UNIST, South Korea)</i>	
HuManiFlow: Ancestor-Conditioned Normalising Flows on SO(3) Manifolds for Human Pose and Shape Distribution Estimation	4779
<i>Akash Sengupta (University of Cambridge), Ignas Budvytis (University of Cambridge), and Roberto Cipolla (University of Cambridge)</i>	
3D Human Pose Estimation With Spatio-Temporal Criss-Cross Attention	4790
<i>Zhenhua Tang (University of Science and Technology of China, China), Zhaofan Qiu (University of Science and Technology of China, China), Yanbin Hao (University of Science and Technology of China, China), Richang Hong (University of Science and Technology of China, China), and Ting Yao (University of Science and Technology of China, China)</i>	
GFPose: Learning 3D Human Pose Prior With Gradient Fields	4800
<i>Hai Ci (Nat'l Key Lab. of GAI & Beijing Institute for GAI (BIGAI)), Mingdong Wu (CFCS, Sch'l of CS; Nat'l Key Lab. of GAI & Beijing Institute for GAI (BIGAI)), Wentao Zhu (CFCS, Sch'l of CS), Xiaoxuan Ma (CFCS, Sch'l of CS), Hao Dong (CFCS, Sch'l of CS), Fangwei Zhong (Sch'l of IST; Nat'l Key Lab. of GAI & Beijing Institute for GAI (BIGAI)), and Yizhou Wang (CFCS, Sch'l of CS; Inst. for AI, Peking University; Nat'l Eng. Research Center of Visual Technology)</i>	
JRDB-Pose: A Large-Scale Dataset for Multi-Person Pose Estimation and Tracking	4811
<i>Edward Vendrow (Stanford University), Duy Tho Le (Monash University), Jianfei Cai (Monash University), and Hamid Reza Tofighi (Monash University)</i>	
Analyzing and Diagnosing Pose Estimation With Attributions	4821
<i>Qiyuan He (National University of Singapore), Linlin Yang (National University of Singapore), Kerui Gu (National University of Singapore), Qiuxia Lin (National University of Singapore), and Angela Yao (National University of Singapore)</i>	
Shape-Constraint Recurrent Flow for 6D Object Pose Estimation	4831
<i>Yang Hai (Xidian University), Rui Song (Xidian University), Jiaojiao Li (Xidian University), and Yinlin Hu (MagicLeap)</i>	
TexPose: Neural Texture Learning for Self-Supervised 6D Object Pose Estimation	4841
<i>Hanzhi Chen (Technical University of Munich), Fabian Manhardt (Google Inc.), Nassir Navab (Technical University of Munich), and Benjamin Busam (Technical University of Munich; 3Dwe.ai)</i>	
Hi-LASSIE: High-Fidelity Articulated Shape and Skeleton Discovery From Sparse Image Ensemble	4853
<i>Chun-Han Yao (UC Merced), Wei-Chih Hung (Waymo), Yuanzhen Li (Google Research), Michael Rubinstein (Google Research), Ming-Hsuan Yang (UC Merced; Google Research; Yonsei University), and Varun Jampani (Google Research)</i>	

Revisiting Rolling Shutter Bundle Adjustment: Toward Accurate and Fast Solution	4863
<i>Bangyan Liao (Hunan University), Delin Qu (Hunan University; Shanghai AI Laboratory), Yifei Xue (Jiangxi Provincial Natural Resources Cause Development Center), Huiqing Zhang (Hunan University), and Yizhen Lao (Hunan University)</i>	
Revisiting the P3P Problem	4872
<i>Yaqing Ding (Lund University), Jian Yang (Nanjing University of Science and Technology), Viktor Larsson (Lund University), Carl Olsson (Lund University), and Kalle Åström (Lund University)</i>	
Common Pets in 3D: Dynamic New-View Synthesis of Real-Life Deformable Categories	4881
<i>Samarth Sinha (University of Toronto), Roman Shapovalov (Meta AI), Jeremy Reizenstein (Meta AI), Ignacio Rocco (Meta AI), Natalia Neverova (Meta AI), Andrea Vedaldi (Meta AI), and David Novotny (Meta AI)</i>	
MobileBrick: Building LEGO for 3D Reconstruction on Mobile Devices	4892
<i>Kejie Li (University of Oxford), Jia-Wang Bian (University of Oxford), Robert Castle (Apple), Philip H.S. Torr (University of Oxford), and Victor Adrian Prisacariu (University of Oxford)</i>	
EFEM: Equivariant Neural Field Expectation Maximization for 3D Object Segmentation Without Scene Supervision	4902
<i>Jiahui Lei (University of Pennsylvania), Congyue Deng (Stanford University), Karl Schmeckpeper (University of Pennsylvania), Leonidas Guibas (Stanford University), and Kostas Daniilidis (University of Pennsylvania)</i>	
GINA-3D: Learning To Generate Implicit Neural Assets in the Wild	4913
<i>Bokui Shen (Stanford University), Xinchun Yan (Waymo LLC), Charles R. Qi (Waymo LLC), Mahyar Najibi (Waymo LLC), Boyang Deng (Stanford University; Waymo LLC), Leonidas Guibas (Google), Yin Zhou (Waymo LLC), and Dragomir Anguelov (Waymo LLC)</i>	
Habitat-Matterport 3D Semantics Dataset	4927
<i>Karmesh Yadav (Meta AI), Ram Ramrakhya (Georgia Tech), Santhosh Kumar Ramakrishnan (UT Austin), Theo Gervet (Carnegie Mellon University), John Turner (Meta AI), Aaron Gokaslan (Cornell University), Noah Maestre (Meta AI), Angel Xuan Chang (Simon Fraser University), Dhruv Batra (Meta AI; Georgia Tech), Manolis Savva (Simon Fraser University), Alexander William Clegg (Meta AI), and Devendra Singh Chaplot (Meta AI)</i>	
BUOL: A Bottom-Up Framework With Occupancy-Aware Lifting for Panoptic 3D Scene Reconstruction From a Single Image	4937
<i>Tao Chu (South China University of Technology; Shanghai AI Laboratory), Pan Zhang (Shanghai AI Laboratory), Qiong Liu (South China University of Technology), and Jiaqi Wang (Shanghai AI Laboratory)</i>	
Panoptic Compositional Feature Field for Editable Scene Rendering With Network-Inferred Labels via Metric Learning	4947
<i>Xinhua Cheng (Peking University, China), Yanmin Wu (Peking University, China), Mengxi Jia (Peking University, China), Qian Wang (Peking University, China), and Jian Zhang (Peking University, China)</i>	

A Light Touch Approach to Teaching Transformers Multi-View Geometry	4958
<i>Yash Bhalgat (Visual Geometry Group, University of Oxford), João F. Henriques (Visual Geometry Group, University of Oxford), and Andrew Zisserman (Visual Geometry Group, University of Oxford)</i>	
Learning To Render Novel Views From Wide-Baseline Stereo Pairs	4970
<i>Yilun Du (MIT CSAIL), Cameron Smith (MIT CSAIL), Ayush Tewari (MIT CSAIL), and Vincent Sitzmann (MIT CSAIL)</i>	
Spring: A High-Resolution High-Detail Dataset and Benchmark for Scene Flow, Optical Flow and Stereo	4981
<i>Lukas Mehl (University of Stuttgart), Jenny Schmalfluss (University of Stuttgart), Azin Jahedi (University of Stuttgart), Yaroslava Nalivayko (University of Stuttgart), and Andrés Bruhn (University of Stuttgart)</i>	
EventNeRF: Neural Radiance Fields From a Single Colour Event Camera	4992
<i>Viktor Rudnev (Max Planck Institute for Informatics, SIC; Saarland University, SIC), Mohamed Elgharib (Max Planck Institute for Informatics, SIC), Christian Theobalt (Max Planck Institute for Informatics, SIC), and Vladislav Golyanik (Max Planck Institute for Informatics, SIC)</i>	
LightedDepth: Video Depth Estimation in Light of Limited Inference View Angles	5003
<i>Shengjie Zhu (Department of Computer Science and Engineering, Michigan State University) and Xiaoming Liu (Department of Computer Science and Engineering, Michigan State University)</i>	
Generating Aligned Pseudo-Supervision From Non-Aligned Data for Image Restoration in Under-Display Camera	5013
<i>Ruicheng Feng (Nanyang Technological University), Chongyi Li (Nanyang Technological University), Huaijin Chen (SenseBrain Technology), Shuai Li (SenseBrain Technology), Jinwei Gu (The Chinese University of Hong Kong; Shanghai AI Laboratory), and Chen Change Loy (Nanyang Technological University)</i>	
Spatio-Focal Bidirectional Disparity Estimation From a Dual-Pixel Image	5023
<i>Donggun Kim (KAIST), Hyeonjoong Jang (KAIST), Inchul Kim (KAIST), and Min H. Kim (KAIST)</i>	
Trap Attention: Monocular Depth Estimation With Manual Traps	5033
<i>Chao Ning (Northwestern Polytechnical University, China) and Hongping Gan (Northwestern Polytechnical University, China)</i>	
Accelerated Coordinate Encoding: Learning to Relocalize in Minutes Using RGB and Poses ...	5044
<i>Eric Brachmann (University of Oxford), Tommaso Cavallari (University of Oxford), and Victor Adrian Prisacariu (University of Oxford)</i>	
Energy-Efficient Adaptive 3D Sensing	5054
<i>Brevin Tilmon (University of Florida), Zhanghao Sun (Stanford University), Sanjeev J. Koppal (University of Florida), Yicheng Wu (SnapInc.), Georgios Evangelidis (SnapInc.), Ramzi Zahreddine (SnapInc.), Gurunandan Krishnan (SnapInc.), Sizhuo Ma (SnapInc.), and Jian Wang (SnapInc.)</i>	
Incremental 3D Semantic Scene Graph Prediction From RGB Sequences	5064
<i>Shun-Cheng Wu (Technische Universität München), Keisuke Tateno (Google), Nassir Navab (Technische Universität München), and Federico Tombari (Technische Universität München; Google)</i>	

Consistent Direct Time-of-Flight Video Depth Super-Resolution	5075
<i>Zhanghao Sun (Stanford University), Wei Ye (Meta Reality Labs), Jinhui Xiong (Meta Reality Labs), Gyeongmin Choe (Meta Reality Labs), Jialiang Wang (Meta Research), Shuochen Su (Meta Reality Labs), and Rakesh Ranjan (Meta Reality Labs)</i>	
Learning To Zoom and Unzoom	5086
<i>Chittesh Thavamani (Carnegie Mellon University), Mengtian Li (Carnegie Mellon University), Francesco Ferroni (Argo AI), and Deva Ramanan (Carnegie Mellon University)</i>	
FrustumFormer: Adaptive Instance-Aware Resampling for Multi-View 3D Detection	5096
<i>Yuqi Wang (CRIPAC, Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS)), Yuntao Chen (Centre for Artificial Intelligence and Robotics, HKISLICAS), and Zhaoxiang Zhang (CRIPAC, Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS); Centre for Artificial Intelligence and Robotics, HKISLICAS)</i>	
3D Video Object Detection With Learnable Object-Centric Global Optimization	5106
<i>Jiawei He (CRIPAC, Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS)), Yuntao Chen (Centre for Artificial Intelligence and Robotics, HKISLICAS), Naiyan Wang (TuSimple), and Zhaoxiang Zhang (CRIPAC, Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS); Centre for Artificial Intelligence and Robotics, HKISLICAS)</i>	
UniDistill: A Universal Cross-Modality Knowledge Distillation Framework for 3D Object Detection in Bird's-Eye View	5116
<i>Shengchao Zhou (MEGVII Technology), Weizhou Liu (MEGVII Technology), Chen Hu (MEGVII Technology), Shuchang Zhou (MEGVII Technology), and Chao Ma (Shanghai Jiao Tong University)</i>	
ARKitTrack: A New Diverse Dataset for Tracking Using Mobile RGB-D Data	5126
<i>Haojie Zhao (Dalian University of Technology, China), Junsong Chen (Dalian University of Technology, China), Lijun Wang (Dalian University of Technology, China), and Huchuan Lu (Dalian University of Technology, China; Peng Cheng Laboratory, China)</i>	
Deep Dive Into Gradients: Better Optimization for 3D Object Detection With Gradient-Corrected IoU Supervision	5136
<i>Qi Ming (Beijing Institute of Technology, China; Intelligent Science & Technology Academy of CASIC), Lingjuan Miao (Beijing Institute of Technology, China; Intelligent Science & Technology Academy of CASIC), Zhe Ma (Beijing Institute of Technology, China; Intelligent Science & Technology Academy of CASIC), Lin Zhao (Beijing Institute of Technology, China; Intelligent Science & Technology Academy of CASIC), Zhiqiang Zhou (Beijing Institute of Technology, China; Intelligent Science & Technology Academy of CASIC), Xuhui Huang (Beijing Institute of Technology, China; Intelligent Science & Technology Academy of CASIC), Yuanpei Chen (Beijing Institute of Technology, China; Intelligent Science & Technology Academy of CASIC), and Yufei Guo (Beijing Institute of Technology, China; Intelligent Science & Technology Academy of CASIC)</i>	

SlowLiDAR: Increasing the Latency of LiDAR-Based Detection Using Adversarial Examples	5146
<i>Han Liu (Washington University in St. Louis), Yuhao Wu (Washington University in St. Louis), Zhiyuan Yu (Washington University in St. Louis), Yevgeniy Vorobeychik (Washington University in St. Louis), and Ning Zhang (Washington University in St. Louis)</i>	
Normalizing Flow Based Feature Synthesis for Outlier-Aware Object Detection	5156
<i>Nishant Kumar (TU Dresden), Siniša Šegvić (University of Zagreb - FER), Abouzar Eslami (Carl Zeiss Meditec AG), and Stefan Gumhold (TU Dresden)</i>	
OcTr: Octree-Based Transformer for 3D Object Detection	5166
<i>Chao Zhou (Beihang University, China), Yanan Zhang (Beihang University, China), Jiaxin Chen (Beihang University, China), and Di Huang (Beihang University, China)</i>	
HypLiLoc: Towards Effective LiDAR Pose Regression With Hyperbolic Fusion	5176
<i>Sijie Wang (Nanyang Technological University), Qiyu Kang (Nanyang Technological University), Rui She (Nanyang Technological University), Wei Wang (University of Oxford), Kai Zhao (Nanyang Technological University), Yang Song (C3.AI), and Wee Peng Tay (Nanyang Technological University)</i>	
LiDAR2Map: In Defense of LiDAR-Based Semantic Map Construction Using Online Camera Distillation	5186
<i>Song Wang (Zhejiang University), Wentong Li (Zhejiang University), Wenyu Liu (Zhejiang University), Xiaolu Liu (Zhejiang University), and Jianke Zhu (Zhejiang University)</i>	
MSF: Motion-Guided Sequential Fusion for Efficient 3D Object Detection From Point Cloud Sequences	5196
<i>Chenheng He (The Hong Kong Polytechnic University), Ruihuang Li (The Hong Kong Polytechnic University), Yabin Zhang (The Hong Kong Polytechnic University), Shuai Li (The Hong Kong Polytechnic University), and Lei Zhang (The Hong Kong Polytechnic University)</i>	
SFD2: Semantic-Guided Feature Detection and Description	5206
<i>Fei Xue (University of Cambridge; Durham University), Ignas Budvytis (University of Cambridge), and Roberto Cipolla (University of Cambridge)</i>	
Temporal Consistent 3D LiDAR Representation Learning for Semantic Perception in Autonomous Driving	5217
<i>Lucas Nunes (University of Bonn), Louis Wiesmann (University of Bonn), Rodrigo Marcuzzi (University of Bonn), Xieyuanli Chen (University of Bonn), Jens Behley (University of Bonn), and Cyrill Stachniss (University of Bonn; University of Oxford; Lamarr Institute)</i>	
Unsupervised 3D Point Cloud Representation Learning by Triangle Constrained Contrast for Autonomous Driving	5229
<i>Bo Pang (Shanghai Jiao Tong University), Hongchi Xia (Shanghai Jiao Tong University), and Cewu Lu (Shanghai Jiao Tong University)</i>	

RangeViT: Towards Vision Transformers for 3D Semantic Segmentation in Autonomous Driving ... 5240	
<i>Angelika Ando (Valeo.ai, France; Université PSL, France), Spyros Gidaris (Valeo.ai, France), Andrei Bursuc (Valeo.ai, France), Gilles Puy (Valeo.ai, France), Alexandre Boulch (Valeo.ai, France), and Renaud Marlet (Valeo.ai, France; Univ. Gustave Eiffel, CNRS, France)</i>	
Spatiotemporal Self-Supervised Learning for Point Clouds in the Wild 5251	
<i>Yanhao Wu (Xi'an Jiaotong University, China), Tong Zhang (EPFL, Switzerland), Wei Ke (Xi'an Jiaotong University, China), Sabine Süssstrunk (EPFL, Switzerland), and Mathieu Salzmann (EPFL, Switzerland)</i>	
Change-Aware Sampling and Contrastive Learning for Satellite Images 5261	
<i>Utkarsh Mall (Cornell University), Bharath Hariharan (Cornell University), and Kavita Bala (Cornell University)</i>	
Self-Supervised 3D Scene Flow Estimation Guided by Superpoints 5271	
<i>Yaqi Shen (Nanjing University of Science and Technology, Nanjing, China), Le Hui (Nanjing University of Science and Technology, Nanjing, China), Jin Xie (Nanjing University of Science and Technology, Nanjing, China), and Jian Yang (Nanjing University of Science and Technology, Nanjing, China)</i>	
SCOOP: Self-Supervised Correspondence and Optimization-Based Scene Flow 5281	
<i>Itai Lang (Tel Aviv University; Google Research), Dror Aiger (Google Research), Forrester Cole (Google Research), Shai Avidan (Tel Aviv University), and Michael Rubinstein (Google Research)</i>	
PiMAE: Point Cloud and Image Interactive Masked Autoencoders for 3D Object Detection 5291	
<i>Anthony Chen (Peking University), Kevin Zhang (Peking University), Renrui Zhang (The Chinese University of Hong Kong), Zihan Wang (Peking University), Yuheng Lu (Peking University; Wukong Lab, iKingtec), Yandong Guo (Beijing University of Posts and Telecommunications), and Shanghang Zhang (Peking University)</i>	
CP3: Channel Pruning Plug-In for Point-Based Networks 5302	
<i>Yaomin Huang (East China Normal University; Midea Group), Ning Liu (Midea Group), Zhengping Che (Midea Group), Zhiyuan Xu (Midea Group), Chaomin Shen (East China Normal University), Yaxin Peng (Shanghai University), Guixu Zhang (East China Normal University), Xinmei Liu (East China Normal University), Feifei Feng (Midea Group), and Jian Tang (Midea Group)</i>	
Binarizing Sparse Convolutional Networks for Efficient Point Cloud Analysis 5313	
<i>Xiuwei Xu (Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China), Ziwei Wang (Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China), Jie Zhou (Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China), and Jiwen Lu (Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China)</i>	

Hyperspherical Embedding for Point Cloud Completion	5323
<i>Junming Zhang (University of Michigan), Haomeng Zhang (University of Michigan), Ram Vasudevan (University of Michigan), and Matthew Johnson-Roberson (University of Michigan)</i>	
Attention-Based Point Cloud Edge Sampling	5333
<i>Chengzhi Wu (Karlsruhe Institute of Technology, Germany), Junwei Zheng (Karlsruhe Institute of Technology, Germany), Julius Pfommer (Fraunhofer IOSB, Germany; Fraunhofer Center for Machine Learning, Germany), and Jürgen Beyerer (Karlsruhe Institute of Technology, Germany; Fraunhofer IOSB, Germany; Fraunhofer Center for Machine Learning, Germany)</i>	
Starting From Non-Parametric Networks for 3D Point Cloud Analysis	5344
<i>Renrui Zhang (CUHK MMLab; Shanghai Artificial Intelligence Laboratory), Liuhui Wang (Peking University; Heisenberg Robotics), Yali Wang (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences; Shanghai Artificial Intelligence Laboratory), Peng Gao (Shanghai Artificial Intelligence Laboratory), Hongsheng Li (CUHK MMLab), and Jianbo Shi (CUHK MMLab)</i>	
Grad-PU: Arbitrary-Scale Point Cloud Upsampling via Gradient Descent With Learned Distance Functions	5354
<i>Yun He (Fudan University), Danhang Tang (Google), Yinda Zhang (Google), Xiangyang Xue (Fudan University), and Yanwei Fu (Fudan University)</i>	
SE-ORNet: Self-Ensembling Orientation-Aware Network for Unsupervised Point Cloud Shape Correspondence	5364
<i>Jiacheng Deng (University of Science and Technology of China), Chuxin Wang (University of Science and Technology of China), Jiahao Lu (University of Science and Technology of China), Jianfeng He (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China; Deep Space Exploration Lab), Jiyang Yu (China Academy of Space Technology), and Zhe Zhang (Deep Space Exploration Lab)</i>	
Robust 3D Shape Classification via Non-Local Graph Attention Network	5374
<i>Shengwei Qin (Zhejiang Sci-Tech University), Zhong Li (Huzhou University; Zhejiang Sci-Tech University), and Ligang Liu (University of Science and Technology of China)</i>	
Rotation-Invariant Transformer for Point Cloud Matching	5384
<i>Hao Yu (TU Munich), Zheng Qin (NUDT), Ji Hou (Meta Reality Labs), Mahdi Saleh (TU Munich; 3Dwe.ai), Dongsheng Li (NUDT), Benjamin Busam (TU Munich; 3Dwe.ai), and Slobodan Ilic (TU Munich; Siemens AG, Munich)</i>	
Deep Graph-Based Spatial Consistency for Robust Non-Rigid Point Cloud Registration	5394
<i>Zheng Qin (National University of Defense Technology), Hao Yu (Technical University of Munich), Changjian Wang (National University of Defense Technology), Yuxing Peng (National University of Defense Technology), and Kai Xu (National University of Defense Technology)</i>	

Efficient RGB-T Tracking via Cross-Modality Distillation	5404
<i>Tianlu Zhang (Xidian University, China), Hongyuan Guo (Xidian University, China), Qiang Jiao (Xidian University, China), Qiang Zhang (Xidian University, China), and Jungong Han (the University of Sheffield, UK; the University of Sheffield, UK)</i>	
Finding Geometric Models by Clustering in the Consensus Space	5414
<i>Daniel Barath (Computer Vision and Geometry Group, Switzerland), Denys Rozumnyi (VRG, Faculty of Electrical Engineering; Computer Vision and Geometry Group, Switzerland), Ivan Eichhardt (TMEIC Corporation Americas, USA; Eotvos Lorand University, Hungary), Levente Hajder (TMEIC Corporation Americas, USA), and Jiri Matas (VRG, Faculty of Electrical Engineering, CTU in Prague, Czech Republic)</i>	
Adaptive Assignment for Geometry Aware Local Feature Matching	5425
<i>Dihe Huang (Tsinghua Univeristy), Ying Chen (Tencent YouTu Lab), Yong Liu (Tencent YouTu Lab), Jianlin Liu (Tencent YouTu Lab), Shang Xu (Tencent YouTu Lab), Wenlong Wu (Tencent YouTu Lab), Yikang Ding (Tsinghua Univeristy), Fan Tang (Institute of Computing Technology, Chinese Academy of Sciences), and Chengjie Wang (Tencent YouTu Lab; Shanghai Jiao Tong University)</i>	
Masked Representation Learning for Domain Generalized Stereo Matching	5435
<i>Zhibo Rao (Nanchang Hangkong University, China; Northwestern Polytechnical University, China), Bangshu Xiong (Nanchang Hangkong University, China), Mingyi He (Northwestern Polytechnical University, China), Yuchao Dai (Northwestern Polytechnical University, China), Renjie He (Northwestern Polytechnical University, China), Zhelun Shen (Baidu Research, China), and Xing Li (Northwestern Polytechnical University, China)</i>	
Learning Optical Expansion From Scale Matching	5445
<i>Han Ling (Nanjing University of Science and Technology), Yinghui Sun (Nanjing University of Science and Technology), Quansen Sun (Nanjing University of Science and Technology), and Zhenwen Ren (Nanjing University of Science and Technology; Southwest University of Science and Technology; SongShan Laboratory)</i>	
AnyFlow: Arbitrary Scale Optical Flow With Implicit Neural Representation	5455
<i>Hyunyoung Jung (Seoul National University), Zhuo Hui (Meta Reality Labs), Lei Luo (Meta Reality Labs), Haitao Yang (Meta Reality Labs), Feng Liu (Meta Reality Labs), Sungjoo Yoo (Seoul National University), Rakesh Ranjan (Meta Reality Labs), and Denis Demandolx (Seoul National University; Meta Reality Labs)</i>	
HouseDiffusion: Vector Floorplan Generation via a Diffusion Model With Discrete and Continuous Denoising	5466
<i>Mohammad Amin Shabani (Simon Fraser University), Sepidehsadat Hosseini (Simon Fraser University), and Yasutaka Furukawa (Simon Fraser University)</i>	

Localized Semantic Feature Mixers for Efficient Pedestrian Detection in Autonomous Driving.....	5476
<i>Abdul Hannan Khan (Department of Computer Science, RPTU Kaiserslautern-Landau, Germany; German Research Center for Artificial Intelligence (DFKI GmbH), Germany), Mohammed Shariq Nawaz (Department of Computer Science, RPTU Kaiserslautern-Landau, Germany), and Andreas Dengel (Department of Computer Science, RPTU Kaiserslautern-Landau, Germany; German Research Center for Artificial Intelligence (DFKI GmbH), Germany)</i>	
V2X-Seq: A Large-Scale Sequential Dataset for Vehicle-Infrastructure Cooperative Perception and Forecasting	5486
<i>Haibao Yu (Tsinghua University; The University of Hong Kong), Wenxian Yang (Tsinghua University), Hongzhi Ruan (Tsinghua University; University of Chinese Academy of Science), Zhenwei Yang (Tsinghua University; University of Science and Technology Beijing), Yingjuan Tang (Tsinghua University; Beijing Institute of Technology), Xu Gao (Baidu Inc.), Xin Hao (Baidu Inc.), Yifeng Shi (Baidu Inc.), Yifeng Pan (Baidu Inc.), Ning Sun (Beijing Connected and Autonomous Vehicles Technology Co., Ltd), Juan Song (Beijing Connected and Autonomous Vehicles Technology Co., Ltd), Jirui Yuan (Tsinghua University), Ping Luo (The University of Hong Kong), and Zaiqing Nie (Tsinghua University)</i>	
ViP3D: End-to-End Visual Trajectory Prediction via 3D Agent Queries	5496
<i>Junru Gu (Tsinghua University), Chenxu Hu (Tsinghua University), Tianyuan Zhang (Shanghai Qi Zhi Institute; CMU), Xuanyao Chen (Shanghai Qi Zhi Institute; Fudan University), Yilun Wang (Li Auto), Yue Wang (MIT), and Hang Zhao (Tsinghua University; Shanghai Qi Zhi Institute)</i>	
IPCC-TP: Utilizing Incremental Pearson Correlation Coefficient for Joint Multi-Agent Trajectory Prediction	5507
<i>Dekai Zhu (Technical University of Munich), Guangyao Zhai (Technical University of Munich), Yan Di (Technical University of Munich), Fabian Manhardt (Google), Hendrik Berkemeyer (Robert Bosch GmbH; University of Osnabrueck), Tuan Tran (Robert Bosch GmbH), Nassir Navab (Technical University of Munich), Federico Tombari (Technical University of Munich; Google), and Benjamin Busam (Technical University of Munich; 3Dwe.ai)</i>	
Leapfrog Diffusion Model for Stochastic Trajectory Prediction	5517
<i>Weibo Mao (Shanghai Jiao Tong University), Chenxin Xu (Shanghai Jiao Tong University), Qi Zhu (Shanghai Jiao Tong University), Siheng Chen (Shanghai Jiao Tong University; Shanghai AI Laboratory), and Yanfeng Wang (Shanghai Jiao Tong University; Shanghai AI Laboratory)</i>	
DeFeeNet: Consecutive 3D Human Motion Prediction With Deviation Feedback	5527
<i>Xiaoning Sun (Nanjing University of Science and Technology, China), Huaijiang Sun (Nanjing University of Science and Technology, China), Bin Li (Tianjin AiForward Science and Technology Co., Ltd., China), Dong Wei (Nanjing University of Science and Technology, China), Weiqing Li (Nanjing University of Science and Technology, China), and Jianfeng Lu (Nanjing University of Science and Technology, China)</i>	

Self-Correctable and Adaptable Inference for Generalizable Human Pose Estimation	5537
<i>Zhehan Kan (Southern University of Science and Technology, China), Shuoshuo Chen (Southern University of Science and Technology, China), Ce Zhang (Southern University of Science and Technology, China), Yushun Tang (Southern University of Science and Technology, China), and Zhihai He (Southern University of Science and Technology, China; Pengcheng Laboratory, China)</i>	
ReDirTrans: Latent-to-Latent Translation for Gaze and Head Redirection	5547
<i>Shiwei Jin (ECE Dept. UC San Diego), Zhen Wang (Qualcomm Technologies, Inc.), Lei Wang (Qualcomm Technologies, Inc.), Ning Bi (Qualcomm Technologies, Inc.), and Truong Nguyen (ECE Dept. UC San Diego)</i>	
Feature Shrinkage Pyramid for Camouflaged Object Detection With Transformers	5557
<i>Zhou Huang (Sichuan Changhong Electric Co., Ltd.; UESTC), Hang Dai (University of Glasgow), Tian-Zhu Xiang (G42), Shuo Wang (ETH Zurich), Huai-Xin Chen (UESTC), Jie Qin (CCST, NUAA), and Huan Xiong (MBZUAI)</i>	
OVTrack: Open-Vocabulary Multiple Object Tracking	5567
<i>Siyuan Li (ETH Zurich), Tobias Fischer (ETH Zurich), Lei Ke (ETH Zurich), Henghui Ding (ETH Zurich), Martin Danelljan (ETH Zurich), and Fisher Yu (ETH Zurich)</i>	
GaitGCI: Generative Counterfactual Intervention for Gait Recognition	5578
<i>Huanzhang Dou (Zhejiang University), Pengyi Zhang (Zhejiang University), Wei Su (Zhejiang University), Yunlong Yu (Zhejiang University), Yining Lin (SupreMind), and Xi Li (Zhejiang University; Shanghai AI Laboratory; Singapore Innovation and AI Joint Research Lab)</i>	
Multi-Label Compound Expression Recognition: C-EXPR Database & Network	5589
<i>Dimitrios Kollias (Queen Mary University of London, UK)</i>	
Blemish-Aware and Progressive Face Retouching With Limited Paired Data	5599
<i>Lianxin Xie (South China University of Technology), Wen Xue (South China University of Technology), Zhen Xu (South China University of Technology), Si Wu (South China University of Technology; Peng Cheng Laboratory; City University of Hong Kong), Zhiwen Yu (South China University of Technology), and Hau San Wong (PAZHOU LAB)</i>	
High-Fidelity and Freely Controllable Talking Head Video Generation	5609
<i>Yue Gao (Microsoft Research), Yuan Zhou (Microsoft Research), Jinglu Wang (Microsoft Research), Xiao Li (Microsoft Research), Xiang Ming (Microsoft Research), and Yan Lu (Microsoft Research)</i>	
3Mformer: Multi-Order Multi-Mode Transformer for Skeletal Action Recognition	5620
<i>Lei Wang (Australian National University) and Piotr Koniusz (Australian National University)</i>	
UDE: A Unified Driving Engine for Human Motion Generation	5632
<i>Zixiang Zhou (n/a) and Baoyuan Wang (n/a)</i>	
Data-Driven Feature Tracking for Event Cameras	5642
<i>Nico Messikommer (University of Zurich, Switzerland), Carter Fang (University of Zurich, Switzerland), Mathias Gehrig (University of Zurich, Switzerland), and Davide Scaramuzza (University of Zurich, Switzerland)</i>	

MoStGAN-V: Video Generation With Temporal Motion Styles	5652
<i>Xiaoqian Shen (King Abdullah University of Science and Technology (KAUST)), Xiang Li (King Abdullah University of Science and Technology (KAUST)), and Mohamed Elhoseiny (King Abdullah University of Science and Technology (KAUST))</i>	
Two-Stage Co-Segmentation Network Based on Discriminative Representation for Recovering Human Mesh From Videos	5662
<i>Boyang Zhang (Ningxia University, China), Kehua Ma (Ningxia University, China), Suping Wu (Ningxia University, China), and Zhixiang Yuan (Ningxia University, China)</i>	
Joint Appearance and Motion Learning for Efficient Rolling Shutter Correction	5671
<i>Bin Fan (Northwestern Polytechnical University, China), Yuxin Mao (Northwestern Polytechnical University, China), Yuchao Dai (Northwestern Polytechnical University, China), Zhexiong Wan (Northwestern Polytechnical University, China), and Qi Liu (Northwestern Polytechnical University, China)</i>	
Extracting Motion and Appearance via Inter-Frame Attention for Efficient Video Frame Interpolation	5682
<i>Guozhen Zhang (Nanjing University, China), Yuhan Zhu (Nanjing University, China), Haonan Wang (Nanjing University, China), Youxin Chen (Samsung Electronics (China) R&D Centre, China), Gangshan Wu (Nanjing University, China), and Limin Wang (Nanjing University, China; Shanghai AI Lab, China)</i>	
Deep Stereo Video Inpainting	5693
<i>Zhiliang Wu (Nanjing University of Science and Technology, China), Changchang Sun (Illinois Institute of Technology, USA), Hanyu Xuan (Anhui University, China), and Yan Yan (Illinois Institute of Technology, USA)</i>	
Burstormer: Burst Image Restoration and Enhancement Transformer	5703
<i>Akshay Dudhane (Mohamed bin Zayed University of AI), Syed Waqas Zamir (Inception Institute of AI), Salman Khan (Australian National University), Fahad Shahbaz Khan (Mohamed bin Zayed University of AI; Linkoping University), and Ming-Hsuan Yang (University of California, Merced; Yonsei University; Google Research)</i>	
Blur Interpolation Transformer for Real-World Motion From Blur	5713
<i>Zhihang Zhong (The University of Tokyo, Japan; National Institute of Informatics, Japan), Mingdeng Cao (The University of Tokyo, Japan), Xiang Ji (The University of Tokyo, Japan), Yinqiang Zheng (The University of Tokyo, Japan), and Imari Sato (The University of Tokyo, Japan; National Institute of Informatics, Japan)</i>	
HDR Imaging With Spatially Varying Signal-to-Noise Ratios	5724
<i>Yiheng Chi (Purdue University), Xingguang Zhang (Purdue University), and Stanley H. Chan (Purdue University)</i>	
Light Source Separation and Intrinsic Image Decomposition Under AC Illumination	5735
<i>Yusaku Yoshida (Department of Artificial Intelligence, Kyushu Institute of Technology, Japan), Ryo Kawahara (Department of Artificial Intelligence, Kyushu Institute of Technology, Japan), and Takahiro Okabe (Department of Artificial Intelligence, Kyushu Institute of Technology, Japan)</i>	

Physics-Guided ISO-Dependent Sensor Noise Modeling for Extreme Low-Light Photography	5744
<i>Yue Cao (School of Computer Science and Technology, Harbin Institute of Technology, China), Ming Liu (School of Computer Science and Technology, Harbin Institute of Technology, China), Shuai Liu (School of Computer Science and Technology, Harbin Institute of Technology, China), Xiaotao Wang (School of Computer Science and Technology, Harbin Institute of Technology, China), Lei Lei (School of Computer Science and Technology, Harbin Institute of Technology, China), and Wangmeng Zuo (School of Computer Science and Technology, Harbin Institute of Technology, China; Peng Cheng Laboratory, China)</i>	
Neumann Network With Recursive Kernels for Single Image Defocus Deblurring	5754
<i>Yuhui Quan (South China University of Technology, China; Pazhou Lab, China), Zicong Wu (South China University of Technology, China; Pazhou Lab, China), and Hui Ji (National University of Singapore, Singapore)</i>	
UMat: Uncertainty-Aware Single Image High Resolution Material Capture	5764
<i>Carlos Rodriguez-Pardo (SEDDI, Spain; Universidad Rey Juan Carlos, Spain), Henar Domínguez-Elvira (SEDDI, Spain; Universidad Rey Juan Carlos, Spain), David Pascual-Hernández (SEDDI, Spain), and Elena Garces (SEDDI, Spain; Universidad Rey Juan Carlos, Spain)</i>	
SMAE: Few-Shot Learning for HDR Deghosting With Saturation-Aware Masked Autoencoders	5775
<i>Qingsen Yan (Northwestern Polytechnical University), Song Zhang (Xidian University), Weiye Chen (Xidian University), Hao Tang (CVL), Yu Zhu (Northwestern Polytechnical University), Jinqiu Sun (Northwestern Polytechnical University), Luc Van Gool (CVL), and Yanning Zhang (Northwestern Polytechnical University)</i>	
Curricular Contrastive Regularization for Physics-Aware Single Image Dehazing	5785
<i>Yu Zheng (Ocean University of China), Jiahui Zhan (Ocean University of China), Shengfeng He (Singapore Management University), Junyu Dong (Ocean University of China), and Yong Du (Ocean University of China)</i>	
Patch-Craft Self-Supervised Training for Correlated Image Denoising	5795
<i>Gregory Vaksman (CS Department - The Technion, Israel) and Michael Elad (CS Department - The Technion, Israel)</i>	
Spectral Enhanced Rectangle Transformer for Hyperspectral Image Denoising	5805
<i>Miaoyu Li (Beijing Institute of Technology), Ji Liu (Baidu Inc.), Ying Fu (Beijing Institute of Technology), Yulun Zhang (ETH Zürich), and Dejing Dou (BCG X)</i>	
All-in-One Image Restoration for Unknown Degradations Using Adaptive Discriminative Filters for Specific Degradations	5815
<i>Dongwon Park (INMC; UNIST, Republic of Korea), Byung Hyun Lee (Dept. of ECE), and Se Young Chun (INMC; Dept. of ECE; IPAI, Seoul National University, Republic of Korea)</i>	

Ingredient-Oriented Multi-Degradation Learning for Image Restoration	5825
<i>Jinghao Zhang (University of Science and Technology of China), Jie Huang (University of Science and Technology of China), Mingde Yao (University of Science and Technology of China), Zizheng Yang (University of Science and Technology of China), Hu Yu (University of Science and Technology of China), Man Zhou (University of Science and Technology of China), and Feng Zhao (University of Science and Technology of China)</i>	
CR-FIQA: Face Image Quality Assessment by Learning Sample Relative Classifiability	5836
<i>Fadi Boutros (Fraunhofer Institute for Computer Graphics Research IGD, Germany), Meiling Fang (Fraunhofer Institute for Computer Graphics Research IGD, Germany; Department of Computer Science, TU Darmstadt, Germany), Marcel Klemm (Fraunhofer Institute for Computer Graphics Research IGD, Germany), Biying Fu (Fraunhofer Institute for Computer Graphics Research IGD, Germany), and Naser Damer (Fraunhofer Institute for Computer Graphics Research IGD, Germany; Department of Computer Science, TU Darmstadt, Germany)</i>	
Re-IQA: Unsupervised Learning for Image Quality Assessment in the Wild	5846
<i>Avinab Saha (Laboratory of Image and Video Engineering, The University of Texas, Austin), Sandeep Mishra (Laboratory of Image and Video Engineering, The University of Texas, Austin), and Alan C. Bovik (Laboratory of Image and Video Engineering, The University of Texas, Austin)</i>	
Toward Accurate Post-Training Quantization for Image Super Resolution	5856
<i>Zhijun Tu (Huawei Noah's Ark Lab), Jie Hu (Huawei Noah's Ark Lab), Hanting Chen (Huawei Noah's Ark Lab), and Yunhe Wang (Huawei Noah's Ark Lab)</i>	
Learning Steerable Function for Efficient Image Resampling	5866
<i>Jiacheng Li (University of Science and Technology of China), Chang Chen (Huawei Noah's Ark Lab), Wei Huang (University of Science and Technology of China), Zhiqiang Lang (Huawei Noah's Ark Lab), Fenglong Song (Huawei Noah's Ark Lab), Youliang Yan (Huawei Noah's Ark Lab), and Zhiwei Xiong (University of Science and Technology of China)</i>	
ABCD: Arbitrary Bitwise Coefficient for De-Quantization	5876
<i>Woo Kyoung Han (Daegu Gyeongbuk Institute of Science and Technology (DGIST), Korea), Byeonghun Lee (Daegu Gyeongbuk Institute of Science and Technology (DGIST), Korea), Sang Hyun Park (Daegu Gyeongbuk Institute of Science and Technology (DGIST), Korea), and Kyong Hwan Jin (Daegu Gyeongbuk Institute of Science and Technology (DGIST), Korea)</i>	
Efficient Frequency Domain-Based Transformers for High-Quality Image Deblurring	5886
<i>Lingshun Kong (Nanjing University of Science and Technology), Jiangxin Dong (Nanjing University of Science and Technology), Jianjun Ge (China Electronics Technology Group Corporation), Mingqiang Li (China Electronics Technology Group Corporation), and Jinshan Pan (Nanjing University of Science and Technology)</i>	

Learning a Sparse Transformer Network for Effective Image Deraining	5896
<i>Xiang Chen (Nanjing University of Science and Technology), Hao Li (Nanjing University of Science and Technology), Mingqiang Li (Information Science Academy, China Electronics Technology Group Corporation), and Jinshan Pan (Nanjing University of Science and Technology)</i>	
CDDFuse: Correlation-Driven Dual-Branch Feature Decomposition for Multi-Modality Image Fusion	5906
<i>Zixiang Zhao (Xi'an Jiaotong University; ETH Zurich), Haowen Bai (Xi'an Jiaotong University), Jiangshe Zhang (Xi'an Jiaotong University), Yulun Zhang (ETH Zurich), Shuang Xu (Research and Development Institute of Northwestern Polytechnical University in Shenzhen; Northwestern Polytechnical University), Zudi Lin (Harvard University), Radu Timofte (ETH Zurich; University of Wurzburg), and Luc Van Gool (ETH Zurich)</i>	
PCT-Net: Full Resolution Image Harmonization Using Pixel-Wise Color Transformations	5917
<i>Julian Jorge Andrade Guerreiro (The University of Tokyo), Mitsuru Nakazawa (Rakuten Institute of Technology, Rakuten Group, Inc.), and Björn Stenger (Rakuten Institute of Technology, Rakuten Group, Inc.)</i>	
Semi-Supervised Parametric Real-World Image Harmonization	5927
<i>Ke Wang (Adobe Inc.; EECS, University of California, Berkeley), Michaël Gharbi (Adobe Inc.), He Zhang (Adobe Inc.), Zhihao Xia (Adobe Inc.), and Eli Shechtman (Adobe Inc.)</i>	
Towards Robust Tampered Text Detection in Document Image: New Dataset and New Solution ...	5937
<i>Chenfan Qu (South China University of Technology), Chongyu Liu (South China University of Technology), Yuliang Liu (Huazhong University of Science and Technology), Xinhong Chen (South China University of Technology), Dezhi Peng (South China University of Technology), Fengjun Guo (IntSig Information Co., Ltd), and Lianwen Jin (South China University of Technology)</i>	
QuantArt: Quantizing Image Style Transfer Towards High Visual Fidelity	5947
<i>Siyu Huang (Harvard University), Jie An (University of Rochester), Donglai Wei (Boston College), Jiebo Luo (University of Rochester), and Hanspeter Pfister (Harvard University)</i>	
Deep Curvilinear Editing: Commutative and Nonlinear Image Manipulation for Pretrained Deep Generative Model	5957
<i>Takehiro Aoshima (Osaka University) and Takashi Matsubara (Osaka University)</i>	
Person Image Synthesis via Denoising Diffusion Model	5968
<i>Ankan Kumar Bhunia (Mohamed bin Zayed University of AI, UAE), Salman Khan (Mohamed bin Zayed University of AI, UAE; Australian National University, Australia), Hisham Cholakkal (Mohamed bin Zayed University of AI, UAE), Rao Muhammad Anwer (Mohamed bin Zayed University of AI, UAE; Aalto University, Finland), Jorma Laaksonen (Aalto University, Finland), Mubarak Shah (University of Central Florida, USA), and Fahad Shahbaz Khan (Mohamed bin Zayed University of AI, UAE; Linköping University, Sweden)</i>	

Disentangling Writer and Character Styles for Handwriting Generation	5977
<i>Gang Dai (South China University of Technology), Yifan Zhang (National University of Singapore), Qingfeng Wang (South China University of Technology), Qing Du (South China University of Technology), Zhuliang Yu (South China University of Technology), Zhuoman Liu (The Hong Kong Polytechnic University), and Shuangping Huang (South China University of Technology; Pazhou Laboratory)</i>	
NoisyTwins: Class-Consistent and Diverse Image Generation Through StyleGANs	5987
<i>Harsh Rangwani (Vision and AI Lab, IISc Bangalore), Lavish Bansal (Vision and AI Lab, IISc Bangalore; IIT BHU Varanasi), Kartik Sharma (Vision and AI Lab, IISc Bangalore; BITS Pilani), Tejan Karmali (Google Research), Varun Jampani (Google Research), and R. Venkatesh Babu (Vision and AI Lab, IISc Bangalore)</i>	
High-Fidelity Guided Image Synthesis With Latent Diffusion Models	5997
<i>Jaskirat Singh (The Australian National University), Stephen Gould (The Australian National University; Australian Centre for Robotic Vision), and Liang Zheng (The Australian National University; Australian Centre for Robotic Vision)</i>	
Imagic: Text-Based Real Image Editing With Diffusion Models	6007
<i>Bahjat Kawar (Google Research; Technion), Shiran Zada (Google Research), Oran Lang (Google Research), Omer Tov (Google Research), Huiwen Chang (Google Research), Tali Dekel (Google Research; Weizmann Institute of Science), Inbar Mosseri (Google Research), and Michal Irani (Google Research; Weizmann Institute of Science)</i>	
PosterLayout: A New Benchmark and Approach for Content-Aware Visual-Textual Presentation Layout	6018
<i>Hsiao Yuan Hsu (Peking University; Peking University), Xiangteng He (Peking University; Peking University), Yuxin Peng (Peking University; Peking University), Hao Kong (Meituan), and Qing Zhang (Meituan)</i>	
SINE: SINGle Image Editing With Text-to-Image Diffusion Models	6027
<i>Zhixing Zhang (Rutgers University), Ligong Han (Rutgers University), Arnab Ghosh (Snap Inc.), Dimitris N. Metaxas (Rutgers University), and Jian Ren (Snap Inc.)</i>	
NULL-Text Inversion for Editing Real Images Using Guided Diffusion Models	6038
<i>Ron Mokady (Google Research; Tel Aviv University), Amir Hertz (Google Research; Tel Aviv University), Kfir Aberman (Google Research), Yael Pritch (Google Research), and Daniel Cohen-Or (Google Research)</i>	
Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models	6048
<i>Gowthami Somepalli (University of Maryland, College Park), Vasu Singla (University of Maryland, College Park), Micah Goldblum (New York University), Jonas Geiping (University of Maryland, College Park), and Tom Goldstein (University of Maryland, College Park)</i>	
Parallel Diffusion Models of Operator and Image for Blind Inverse Problems	6059
<i>Hyungjin Chung (Dept. of Bio & Brain Engineering), Jeongsol Kim (Dept. of Bio & Brain Engineering), Sehui Kim (Kim Jae Chul Graduate School of AI, KAIST), and Jong Chul Ye (Kim Jae Chul Graduate School of AI, KAIST)</i>	

Unite and Conquer: Plug & Play Multi-Modal Synthesis Using Diffusion Models	6070
<i>Nithin Gopalakrishnan Nair (Johns Hopkins University, USA), Wele Gedara Chaminda Bandara (Johns Hopkins University, USA), and Vishal M. Patel (Johns Hopkins University, USA)</i>	
Collaborative Diffusion for Multi-Modal Face Generation and Editing	6080
<i>Ziqi Huang (Nanyang Technological University), Kelvin C.K. Chan (Nanyang Technological University), Yuming Jiang (Nanyang Technological University), and Ziwei Liu (Nanyang Technological University)</i>	
Diffusion Video Autoencoders: Toward Temporally Consistent Face Video Editing via Disentangled Video Encoding	6091
<i>Gyeongman Kim (Korea Advanced Institute of Science and Technology (KAIST), South Korea), Hajin Shim (Korea Advanced Institute of Science and Technology (KAIST), South Korea), Hyunsu Kim (NAVER AI Lab), Yunjey Choi (NAVER AI Lab), Junho Kim (NAVER AI Lab), and Eunho Yang (Korea Advanced Institute of Science and Technology (KAIST), South Korea; AITRICS, South Korea)</i>	
NVTC: Nonlinear Vector Transform Coding	6101
<i>Runsen Feng (University of Science and Technology of China), Zongyu Guo (University of Science and Technology of China), Weiping Li (University of Science and Technology of China), and Zhibo Chen (University of Science and Technology of China)</i>	
Motion Information Propagation for Neural Video Compression	6111
<i>Linfeng Qi (University of Science and Technology of China), Jiahao Li (Microsoft Research Asia), Bin Li (Microsoft Research Asia), Houqiang Li (University of Science and Technology of China), and Yan Lu (Microsoft Research Asia)</i>	
A Dynamic Multi-Scale Voxel Flow Network for Video Prediction	6121
<i>Xiaotao Hu (Nankai University; Megvii Technology), Zhewei Huang (Megvii Technology), Ailin Huang (Megvii Technology; Wuhan University), Jun Xu (Nankai University), and Shuchang Zhou (Megvii Technology)</i>	
Towards Scalable Neural Representation for Diverse Videos	6132
<i>Bo He (University of Maryland, College Park), Xitong Yang (Meta AI), Hanyu Wang (University of Maryland, College Park), Zuxuan Wu (Fudan University), Hao Chen (University of Maryland, College Park), Shuaiyi Huang (University of Maryland, College Park), Yixuan Ren (University of Maryland, College Park), Ser-Nam Lim (Meta AI), and Abhinav Shrivastava (University of Maryland, College Park)</i>	
DINER: Disorder-Invariant Implicit Neural Representation	6143
<i>Shaowen Xie (Nanjing University, China), Hao Zhu (Nanjing University, China), Zhen Liu (Nanjing University, China), Qi Zhang (AI Lab, Tencent Company, China), You Zhou (Nanjing University, China), Xun Cao (Nanjing University, China), and Zhan Ma (Nanjing University, China)</i>	
SCConv: Spatial and Channel Reconstruction Convolution for Feature Redundancy	6153
<i>Jiafeng Li (East China Normal University, China), Ying Wen (East China Normal University, China), and Lianghua He (Tongji University, China)</i>	

DeepMAD: Mathematical Architecture Design for Deep Convolutional Neural Network	6163
<i>Xuan Shen (Northeastern University), Yaohua Wang (Alibaba Group), Ming Lin (Amazon), Yilun Huang (Alibaba Group), Hao Tang (ETH Zurich), Xiuyu Sun (Alibaba Group), and Yanzhi Wang (Northeastern University)</i>	
Optimization-Inspired Cross-Attention Transformer for Compressive Sensing	6174
<i>Jiechong Song (Peking University Shenzhen Graduate School, China; Peng Cheng Laboratory, China), Chong Mou (Peking University Shenzhen Graduate School, China), Shiqi Wang (City University of Hong Kong, China), Siwei Ma (Peking University, China; Peng Cheng Laboratory, China), and Jian Zhang (Peking University Shenzhen Graduate School, China; Peng Cheng Laboratory, China)</i>	
Neighborhood Attention Transformer	6185
<i>Ali Hassani (SHI Labs @ U of Oregon & UIUC), Steven Walton (SHI Labs @ U of Oregon & UIUC), Jiachen Li (SHI Labs @ U of Oregon & UIUC), Shen Li (Meta/Facebook AI), and Humphrey Shi (SHI Labs @ U of Oregon & UIUC; Picsart AI Research (PAIR))</i>	
Making Vision Transformers Efficient From a Token Sparsification View	6195
<i>Shuning Chang (National University of Singapore), Pichao Wang (Alibaba Group), Ming Lin (Alibaba Group), Fan Wang (Alibaba Group), David Junhao Zhang (National University of Singapore), Rong Jin (Alibaba Group), and Mike Zheng Shou (National University of Singapore)</i>	
Towards Efficient Use of Multi-Scale Features in Transformer-Based Object Detectors	6206
<i>Gongjie Zhang (Nanyang Technological University; Black Sesame Technologies), Zhipeng Luo (Nanyang Technological University; SenseTime Research), Zichen Tian (Nanyang Technological University), Jingyi Zhang (Nanyang Technological University), Xiaoqin Zhang (Wenzhou University), and Shijian Lu (Nanyang Technological University)</i>	
Neuralizer: General Neuroimage Analysis Without Re-Training	6217
<i>Steffen Czolbe (University of Copenhagen and MGH) and Adrian V. Dalca (MIT, and MGH, Harvard Medical School)</i>	
Learning Partial Correlation Based Deep Visual Representation for Image Classification	6231
<i>Saimunur Rahman (Data61vCSIRO; University of Wollongong), Piotr Koniusz (Data61vCSIRO; Australian National University), Lei Wang (University of Wollongong), Luping Zhou (University of Sydney), Peyman Moghadam (Data61vCSIRO; Queensland University of Technology), and Changming Sun (Data61vCSIRO)</i>	
Understanding Masked Image Modeling via Learning Occlusion Invariant Feature	6241
<i>Xiangwen Kong (MEGVII Technology) and Xiangyu Zhang (MEGVII Technology; Beijing Academy of Artificial Intelligence)</i>	
MixMAE: Mixed and Masked Autoencoder for Efficient Pretraining of Hierarchical Vision Transformers	6252
<i>Jihao Liu (CUHK MMLab; SenseTime Research), Xin Huang (SenseTime Research), Jinliang Zheng (SenseTime Research), Yu Liu (SenseTime Research), and Hongsheng Li (CUHK MMLab; CII under InnoHK)</i>	

Adaptive Graph Convolutional Subspace Clustering	6262
<i>Lai Wei (Shanghai Maritime University, China), Zhengwei Chen (Shanghai Maritime University, China), Jun Yin (Shanghai Maritime University, China), Changming Zhu (Shanghai Maritime University, China), Rigui Zhou (Shanghai Maritime University, China), and Jin Liu (Shanghai Maritime University, China)</i>	
Deep Learning of Partial Graph Matching via Differentiable Top-K	6272
<i>Runzhong Wang (Shanghai Jiao Tong University), Ziao Guo (Shanghai Jiao Tong University), Shaofei Jiang (Shanghai Jiao Tong University), Xiaokang Yang (Shanghai Jiao Tong University), and Junchi Yan (Shanghai Jiao Tong University)</i>	
DynamicDet: A Unified Dynamic Architecture for Object Detection	6282
<i>Zhihao Lin (Peking University), Yongtao Wang (Peking University), Jinhe Zhang (Peking University), and Xiaojie Chu (Peking University)</i>	
IS-GGT: Iterative Scene Graph Generation With Generative Transformers	6292
<i>Sanjoy Kundu (Oklahoma State University) and Sathyanarayanan N. Aakur (Oklahoma State University)</i>	
Fast Contextual Scene Graph Generation With Unbiased Context Augmentation	6302
<i>Tianlei Jin (Research Center for Intelligent Robotics, Zhejiang Lab), Fangtai Guo (Research Center for Intelligent Robotics, Zhejiang Lab), Qiwei Meng (Research Center for Intelligent Robotics, Zhejiang Lab), Shiqiang Zhu (Research Center for Intelligent Robotics, Zhejiang Lab), Xiangming Xi (Research Center for Intelligent Robotics, Zhejiang Lab), Wen Wang (Research Center for Intelligent Robotics, Zhejiang Lab), Zonghao Mu (Research Center for Intelligent Robotics, Zhejiang Lab), and Wei Song (Research Center for Intelligent Robotics, Zhejiang Lab)</i>	
Masked Video Distillation: Rethinking Masked Feature Modeling for Self-Supervised Video Representation Learning	6312
<i>Rui Wang (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Dongdong Chen (Microsoft Cloud + AI), Zuxuan Wu (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Yinpeng Chen (Microsoft Cloud + AI), Xiyang Dai (Microsoft Cloud + AI), Mengchen Liu (Microsoft Cloud + AI), Lu Yuan (Microsoft Cloud + AI), and Yu-Gang Jiang (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing)</i>	
MED-VT: Multiscale Encoder-Decoder Video Transformer With Application To Object Segmentation	6323
<i>Rezaul Karim (York University), He Zhao (York University), Richard P. Wildes (York University), and Mennatullah Siam (York University)</i>	
MOVES: Manipulated Objects in Video Enable Segmentation	6334
<i>Richard E. L. Higgins (University of Michigan) and David F. Fouhey (University of Michigan)</i>	
InstMove: Instance Motion for Object-Centric Video Segmentation	6344
<i>Qihao Liu (Johns Hopkins University), Junfeng Wu (Huazhong University of Science and Technology), Yi Jiang (ByteDance), Xiang Bai (Huazhong University of Science and Technology), Alan L. Yuille (Johns Hopkins University), and Song Bai (ByteDance)</i>	

ZBS: Zero-Shot Background Subtraction via Instance-Level Background Modeling and Foreground Selection	6355
<i>Yongqi An (National Laboratory of Pattern Recognition, Institute of Automation, CAS, China; University of Chinese Academy of Sciences, China), Xu Zhao (National Laboratory of Pattern Recognition, Institute of Automation, CAS, China), Tao Yu (National Laboratory of Pattern Recognition, Institute of Automation, CAS, China; University of Chinese Academy of Sciences, China), Haiyun Guo (National Laboratory of Pattern Recognition, Institute of Automation, CAS, China; University of Chinese Academy of Sciences, China), Chaoyang Zhao (National Laboratory of Pattern Recognition, Institute of Automation, CAS, China), Ming Tang (National Laboratory of Pattern Recognition, Institute of Automation, CAS, China; University of Chinese Academy of Sciences, China), and Jinqiao Wang (National Laboratory of Pattern Recognition, Institute of Automation, CAS, China; University of Chinese Academy of Sciences, China)</i>	
Feature Aggregated Queries for Transformer-Based Video Object Detectors	6365
<i>Yiming Cui (University of Florida)</i>	
Context-Aware Relative Object Queries To Unify Video Instance and Panoptic Segmentation	6377
<i>Anwesa Choudhuri (University of Illinois at Urbana-Champaign), Girish Chowdhary (University of Illinois at Urbana-Champaign), and Alexander G. Schwing (University of Illinois at Urbana-Champaign)</i>	
Selective Structured State-Spaces for Long-Form Video Understanding	6387
<i>Jue Wang (Amazon Prime Video), Wentao Zhu (Amazon Prime Video), Pichao Wang (Amazon Prime Video), Xiang Yu (Amazon Prime Video), Linda Liu (Amazon Prime Video), Mohamed Omar (Amazon Prime Video), and Raffay Hamid (Amazon Prime Video)</i>	
Relational Space-Time Query in Long-Form Videos	6398
<i>Xitong Yang (Meta AI), Fu-Jen Chu (Meta AI), Matt Feiszli (Meta AI), Raghav Goyal (Meta AI; University of British Columbia), Lorenzo Torresani (Meta AI), and Du Tran (Meta AI)</i>	
Novel-View Acoustic Synthesis	6409
<i>Changan Chen (University of Texas at Austin; FAIR, Meta AI), Alexander Richard (Reality Labs Research at Meta), Roman Shapovalov (FAIR, Meta AI), Vamsi Krishna Ithapu (Reality Labs Research at Meta), Natalia Neverova (FAIR, Meta AI), Kristen Grauman (University of Texas at Austin; FAIR, Meta AI), and Andrea Vedaldi (FAIR, Meta AI)</i>	
Learning Audio-Visual Source Localization via False Negative Aware Contrastive Learning	6420
<i>Weixuan Sun (Australian National University; OPPO Research Institute), Jiayi Zhang (Beihang University), Jianyuan Wang (The University of Oxford), Zheyuan Liu (Australian National University), Yiran Zhong (Shanghai AI Lab), Tianpeng Feng (OPPO Research Institute), Yandong Guo (OPPO Research Institute), Yanhao Zhang (OPPO Research Institute), and Nick Barnes (Australian National University)</i>	

Sound to Visual Scene Generation by Audio-to-Visual Latent Alignment	6430
<i>Kim Sung-Bin (Dept. of Electrical Engineering), Arda Senocak (Dept. of Electrical Engineering, KAIST), Hyunwoo Ha (Dept. of Electrical Engineering), Andrew Owens (University of Michigan), and Tae-Hyun Oh (Dept. of Electrical Engineering; Grad. School of Artificial Intelligence, POSTECH; Yonsei University)</i>	
CASP-Net: Rethinking Video Saliency Prediction From an Audio-Visual Consistency Perceptual Perspective	6441
<i>Junwen Xiong (Northwestern Polytechnical University), Ganglai Wang (Northwestern Polytechnical University), Peng Zhang (Northwestern Polytechnical University; Ningbo Institute of Northwestern Polytechnical University), Wei Huang (Nanchang University), Yufei Zha (Northwestern Polytechnical University; Ningbo Institute of Northwestern Polytechnical University), and Guangtao Zhai (Shanghai Jiao Tong University)</i>	
Decompose More and Aggregate Better: Two Closer Looks at Frequency Representation Learning for Human Motion Prediction	6451
<i>Xuehao Gao (Xi'an Jiaotong University), Shaoyi Du (Xi'an Jiaotong University), Yang Wu (Tencent AI Lab), and Yang Yang (Xi'an Jiaotong University)</i>	
TempSAL – Uncovering Temporal Information for Deep Saliency Prediction	6461
<i>Bahar Aydemir (School of Computer and Communication Sciences, Switzerland), Ludo Hoffstetter (School of Computer and Communication Sciences, Switzerland), Tong Zhang (School of Computer and Communication Sciences, Switzerland), Mathieu Salzmann (School of Computer and Communication Sciences, Switzerland), and Sabine Süsstrunk (School of Computer and Communication Sciences, Switzerland)</i>	
Prompt-Guided Zero-Shot Anomaly Action Recognition Using Pretrained Deep Skeleton Features..	6471
<i>Fumiaki Sato (Konica Minolta, Inc.), Ryo Hachiuma (Konica Minolta, Inc.), and Taiki Sekii (Konica Minolta, Inc.)</i>	
MMG-Ego4D: Multimodal Generalization in Egocentric Action Recognition	6481
<i>Xinyu Gong (The University of Texas at Austin; Meta Reality Labs), Sreyas Mohan (Meta Reality Labs), Naina Dhingra (Meta Reality Labs), Jean-Charles Bazin (Meta Reality Labs), Yilei Li (Meta Reality Labs), Zhangyang Wang (The University of Texas at Austin), and Rakesh Ranjan (Meta Reality Labs)</i>	
Active Exploration of Multimodal Complementarity for Few-Shot Action Recognition	6492
<i>Yuyang Wanyan (State Key Laboratory of Multimodal Artificial Intelligence Systems (MAIS), Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS)), Xiaoshan Yang (State Key Laboratory of Multimodal Artificial Intelligence Systems (MAIS), Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS)); Peng Cheng Laboratory, China), Chaofan Chen (University of Science and Technology of China (USTC)), and Changsheng Xu (State Key Laboratory of Multimodal Artificial Intelligence Systems (MAIS), Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS); Peng Cheng Laboratory, China)</i>	

Reducing the Label Bias for Timestamp Supervised Temporal Action Segmentation	6503
<i>Kaiyuan Liu (Dalian University of Technology, China), Yunheng Li (Dalian University of Technology, China), Shenglan Liu (Dalian University of Technology, China), Chenwei Tan (Dalian University of Technology, China), and Zihang Shao (Dalian University of Technology, China)</i>	
Soft-Landing Strategy for Alleviating the Task Discrepancy Problem in Temporal Action Localization Tasks	6514
<i>Hyolim Kang (Yonsei University), Hanjung Kim (Yonsei University), Joungbin An (Yonsei University), Minsu Cho (POSTECH), and Seon Joo Kim (Yonsei University)</i>	
Iterative Proposal Refinement for Weakly-Supervised Video Grounding	6524
<i>Meng Cao (Peking University), Fangyun Wei (Microsoft Research Asia), Can Xu (Microsoft), Xiubo Geng (Microsoft), Long Chen (The Hong Kong University of Science and Technology), Can Zhang (Peking University), Yuexian Zou (Peking University), Tao Shen (Microsoft), and Daxin Jiang (Microsoft)</i>	
Movies2Scenes: Using Movie Metadata To Learn Scene Representation	6535
<i>Shixing Chen (Amazon Prime Video), Chun-Hao Liu (Amazon Prime Video), Xiang Hao (Amazon Prime Video), Xiaohan Nie (Amazon Prime Video), Maxim Arap (Amazon Prime Video), and Raffay Hamid (Amazon Prime Video)</i>	
Fine-Tuned CLIP Models Are Efficient Video Learners	6545
<i>Hanoona Rasheed (Mohamed bin Zayed University of AI), Muhammad Uzair Khattak (Mohamed bin Zayed University of AI), Muhammad Maaz (Mohamed bin Zayed University of AI), Salman Khan (Australian National University), and Fahad Shahbaz Khan (Mohamed bin Zayed University of AI; Linköping University)</i>	
Revisiting Temporal Modeling for CLIP-Based Image-to-Video Knowledge Transferring	6555
<i>Ruyang Liu (Peking University; Peng Cheng Laboratory), Jingjia Huang (ByteDance Inc), Ge Li (Peking University), Jiashi Feng (ByteDance Inc), Xinglong Wu (ByteDance Inc), and Thomas H. Li (Peking University)</i>	
VoP: Text-Video Co-Operative Prompt Tuning for Cross-Modal Retrieval	6565
<i>Siteng Huang (Westlake University; Zhejiang University), Biao Gong (Alibaba Group), Yulin Pan (Alibaba Group), Jianwen Jiang (Alibaba Group), Yiliang Lv (Alibaba Group), Yuyuan Li (Zhejiang University), and Donglin Wang (Westlake University)</i>	
ProTéGé: Untrimmed Pretraining for Video Temporal Grounding by Video Temporal Grounding ...	6575
<i>Lan Wang (Michigan State University), Gaurav Mittal (Michigan State University), Sandra Sajeed (Microsoft), Ye Yu (Microsoft), Matthew Hall (Microsoft), Vishnu Naresh Boddeti (Microsoft), and Mei Chen (Microsoft)</i>	
Learning Video Representations From Large Language Models	6586
<i>Yue Zhao (FAIR, Meta AI; University of Texas, Austin), Ishan Misra (FAIR, Meta AI), Philipp Krähenbühl (FAIR, Meta AI; University of Texas, Austin), and Rohit Girdhar (FAIR, Meta AI)</i>	

All in One: Exploring Unified Video-Language Pre-Training	6598
<i>Jinpeng Wang (National University of Singapore), Yixiao Ge (ARC Lab), Rui Yan (National University of Singapore), Yuying Ge (The University of Hong Kong), Kevin Qinghong Lin (National University of Singapore), Satoshi Tsutsui (National University of Singapore), Xudong Lin (Columbia University), Guanyu Cai (Tongji University), Jianping Wu (Tsinghua University), Ying Shan (ARC Lab), Xiaohu Qie (Tencent PCG), and Mike Zheng Shou (National University of Singapore)</i>	
High-Fidelity Generalized Emotional Talking Face Generation With Multi-Modal Emotion Space Learning	6609
<i>Chao Xu (Zhejiang University), Junwei Zhu (Youtu Lab, Tencent), Jiangning Zhang (Youtu Lab, Tencent), Yue Han (Zhejiang University), Wenqing Chu (Youtu Lab, Tencent), Ying Tai (Youtu Lab, Tencent), Chengjie Wang (Youtu Lab, Tencent; Shanghai Jiao Tong University), Zhifeng Xie (Shanghai University), and Yong Liu (Zhejiang University)</i>	
Bidirectional Cross-Modal Knowledge Exploration for Video Recognition With Pre-Trained Vision-Language Models	6620
<i>Wenhao Wu (The University of Sydney; Baidu Inc.), Xiaohan Wang (Zhejiang University), Haipeng Luo (University of Chinese Academy of Sciences), Jingdong Wang (Baidu Inc.), Yi Yang (Zhejiang University), and Wanli Ouyang (The University of Sydney; Shanghai AI Laboratory)</i>	
Decoupled Multimodal Distilling for Emotion Recognition	6631
<i>Yong Li (Nanjing University of Science and Technology, China), Yuanzhi Wang (Nanjing University of Science and Technology, China), and Zhen Cui (Nanjing University of Science and Technology, China)</i>	
Affection: Learning Affective Explanations for Real-World Visual Data	6641
<i>Panos Achlioptas (Snap Inc.; Stanford University), Maks Ovsjanikov (LIX, Ecole Polytechnique, IP Paris), Leonidas Guibas (Stanford University), and Sergey Tulyakov (Snap Inc.)</i>	
An Actor-Centric Causality Graph for Asynchronous Temporal Inference in Group Activity	6652
<i>Zhao Xie (Hefei University of Technology), Tian Gao (Hefei University of Technology), Kewei Wu (Hefei University of Technology), and Jiao Chang (Hefei University of Technology)</i>	
VLPD: Context-Aware Pedestrian Detection via Vision-Language Semantic Self-Supervision ...	6662
<i>Mengyin Liu (University of Science and Technology Beijing, China), Jie Jiang (Data Platform Department, China), Chao Zhu (University of Science and Technology Beijing, China), and Xu-Cheng Yin (University of Science and Technology Beijing, China)</i>	
3D-Aware Object Goal Navigation via Simultaneous Exploration and Identification	6672
<i>Jiazhao Zhang (CFCS, Peking University; Beijing Academy of Artificial Intelligence), Liu Dai (CEIE, Tongji University), Fanpeng Meng (Huazhong University of Science and Technology), Qingnan Fan (Tencent AI Lab), Xuelin Chen (Tencent AI Lab), Kai Xu (National University of Defense Technology), and He Wang (CFCS, Peking University)</i>	
Meta-Explore: Exploratory Hierarchical Vision-and-Language Navigation Using Scene Object Spectrum Grounding	6683
<i>Minyoung Hwang (Seoul National University), Jaeyeon Jeong (Seoul National University), Minsoo Kim (Seoul National University), Yoonseon Oh (Hanyang University), and Songhwai Oh (Seoul National University)</i>	

NaQ: Leveraging Narrations As Queries To Supervise Episodic Memory	6694
<i>Santhosh Kumar Ramakrishnan (UT Austin), Ziad Al-Halah (University of Utah), and Kristen Grauman (UT Austin; FAIR, Meta AI)</i>	
EC2: Emergent Communication for Embodied Control	6704
<i>Yao Mu (The University of Hong Kong), Shunyu Yao (Princeton University), Mingyu Ding (The University of Hong Kong), Ping Luo (The University of Hong Kong), and Chuang Gan (UMass Amherst; MIT-IBM Watson AI Lab)</i>	
Abstract Visual Reasoning: An Algebraic Approach for Solving Raven's Progressive Matrices ..	6715
<i>Jingyi Xu (Singapore University of Technology and Design), Tushar Vaidya (Nanyang Technological University), Yufei Wu (Nanyang Technological University), Saket Chandra (Singapore University of Technology and Design), Zhangsheng Lai (Singapore Polytechnic), and Kai Fong Ernest Chong (Singapore University of Technology and Design)</i>	
Logical Implications for Visual Question Answering Consistency	6725
<i>Sergio Tascon-Morales (University of Bern), Pablo Márquez-Neila (University of Bern), and Raphael Sznitman (University of Bern)</i>	
Divide and Conquer: Answering Questions With Object Factorization and Compositional Reasoning	6736
<i>Shi Chen (University of Minnesota) and Qi Zhao (University of Minnesota)</i>	
The Dialog Must Go On: Improving Visual Dialog via Generative Self-Training	6746
<i>Gi-Cheon Kang (IPAI, Seoul National University; AIIS), Sungdong Kim (NAVER AI Lab), Jin-Hwa Kim (AIIS; NAVER AI Lab), Donghyun Kwak (NAVER Cloud CLOVA), and Byoung-Tak Zhang (IPAI, Seoul National University; AIIS)</i>	
Visual-Language Prompt Tuning With Knowledge-Guided Context Optimization	6757
<i>Hantao Yao (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, CAS), Rui Zhang (State Key Lab of Processors, Institute of Computing Technology, CAS), and Changsheng Xu (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, CAS; University of Chinese Academy of Sciences(CAS))</i>	
Probabilistic Prompt Learning for Dense Prediction	6768
<i>Hyeongjun Kwon (Yonsei University), Taeyong Song (Hyundai Motor Company R&D Division), Somi Jeong (NAVER LABS), Jin Kim (Yonsei University), Jinhyun Jang (Yonsei University), and Kwanghoon Sohn (Yonsei University)</i>	
Is BERT Blind? Exploring the Effect of Vision-and-Language Pretraining on Visual Language Understanding	6778
<i>Morris Alper (Tel Aviv University), Michael Fiman (Tel Aviv University), and Hadar Averbuch-Elor (Tel Aviv University)</i>	
Seeing What You Miss: Vision-Language Pre-Training With Semantic Completion Learning	6789
<i>Yatai Ji (Tsinghua University), Rongcheng Tu (Tencent), Jie Jiang (Tencent), Weijie Kong (Tencent), Chengfei Cai (Tencent), Wenzhe Zhao (Tencent), Hongfa Wang (Tencent), Yujiu Yang (Tsinghua University), and Wei Liu (Tencent)</i>	

Affordance Grounding From Demonstration Video To Target Image	6799
<i>Joya Chen (National University of Singapore), Difei Gao (National University of Singapore), Kevin Qinghong Lin (National University of Singapore), and Mike Zheng Shou (National University of Singapore)</i>	
Leverage Interactive Affinity for Affordance Learning	6809
<i>Hongchen Luo (University of Science and Technology of China), Wei Zhai (University of Science and Technology of China), Jing Zhang (The University of Sydney), Yang Cao (University of Science and Technology of China; The University of Sydney), and Dacheng Tao (JD Explore Academy; The University of Sydney)</i>	
DeAR: Debiasing Vision-Language Models With Additive Residuals	6820
<i>Ashish Seth (IIT Madras, India), Mayur Hemani (Adobe Inc.), and Chirag Agarwal (Adobe Inc.)</i>	
Images Speak in Images: A Generalist Painter for In-Context Visual Learning	6830
<i>Xinlong Wang (Beijing Academy of Artificial Intelligence, China), Wen Wang (Zhejiang University), Yue Cao (Beijing Academy of Artificial Intelligence, China), Chunhua Shen (Zhejiang University), and Tiejun Huang (Beijing Academy of Artificial Intelligence, China; Peking University)</i>	
Hyperbolic Contrastive Learning for Visual Representations Beyond Objects	6840
<i>Songwei Ge (University of Maryland, College Park), Shlok Mishra (University of Maryland, College Park), Simon Kornblith (Google Research), Chun-Liang Li (Google Research), and David Jacobs (University of Maryland, College Park; Meta)</i>	
Picture That Sketch: Photorealistic Image Generation From Abstract Sketches	6850
<i>Subhadeep Koley (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Ayan Kumar Bhunia (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Aneeshan Sain (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Pinaki Nath Chowdhury (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Tao Xiang (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), and Yi-Zhe Song (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence)</i>	
GeneCIS: A Benchmark for General Conditional Image Similarity	6862
<i>Sagar Vaze (FAIR, Meta AI; University of Oxford), Nicolas Carion (FAIR, Meta AI), and Ishan Misra (FAIR, Meta AI)</i>	

Exploiting Unlabelled Photos for Stronger Fine-Grained SBIR	6873
<i>Aneeshan Sain (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Ayan Kumar Bhunia (University of Surrey, United Kingdom), Subhadeep Koley (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Pinaki Nath Chowdhury (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Soumitri Chattopadhyay (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Tao Xiang (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), and Yi-Zhe Song (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence)</i>	
Parts2Words: Learning Joint Embedding of Point Clouds and Texts by Bidirectional Matching Between Parts and Words	6884
<i>Chuan Tang (Jilin University, Chin), Xi Yang (Jilin University, Chin), Bojian Wu (Zhejiang University), Zhizhong Han (Wayne State University, USA), and Yi Chang (Jilin University, Chin; Engineering Research Center of Knowledge-Driven Human-Machine Intelligence, MoE, China)</i>	
DeltaEdit: Exploring Text-Free Training for Text-Driven Image Manipulation	N/A
<i>Yueming Lyu (University of Chinese Academy of Sciences; Chinese Academy of Sciences), Tianwei Lin (Baidu Inc.), Fu Li (Baidu Inc.), Dongliang He (Baidu Inc.), Jing Dong (Chinese Academy of Sciences), and Tieniu Tan (Chinese Academy of Sciences; Nanjing University)</i>	
Detecting and Grounding Multi-Modal Media Manipulation	6904
<i>Rui Shao (School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen); Nanyang Technological University), Tianxing Wu (Nanyang Technological University), and Ziwei Liu (Nanyang Technological University)</i>	
Positive-Augmented Contrastive Learning for Image and Video Captioning Evaluation	6914
<i>Sara Sarto (University of Modena and Reggio Emilia, Italy), Manuele Barraco (University of Modena and Reggio Emilia, Italy), Marcella Cornia (University of Modena and Reggio Emilia, Italy), Lorenzo Baraldi (University of Modena and Reggio Emilia, Italy), and Rita Cucchiara (University of Modena and Reggio Emilia, Italy; IIT-CNR, Italy)</i>	
Similarity Maps for Self-Training Weakly-Supervised Phrase Grounding	6925
<i>Tal Shaharabany (Tel Aviv University) and Lior Wolf (Tel Aviv University)</i>	
Cross-Domain Image Captioning With Discriminative Finetuning	6935
<i>Roberto Dessì (Meta AI/UPF), Michele Bevilacqua (Samaya AI), Eleonora Gualdoni (UPF), Nathanaël Carraz Rakotonirina (UPF), Francesca Franzon (UPF), and Marco Baroni (UPF/ICREA)</i>	
EXIF As Language: Learning Cross-Modal Associations Between Images and Camera Metadata	6945
<i>Chenhao Zheng (University of Michigan), Ayush Shrivastava (University of Michigan), and Andrew Owens (University of Michigan)</i>	

Uncurated Image-Text Datasets: Shedding Light on Demographic Bias	6957
<i>Noa Garcia (Osaka University), Yusuke Hirota (Osaka University), Yankun Wu (Osaka University), and Yuta Nakashima (Osaka University)</i>	
Filtering, Distillation, and Hard Negatives for Vision-Language Pre-Training	6967
<i>Filip Radenovic (Meta AI), Abhimanyu Dubey (Meta AI), Abhishek Kadian (Meta AI), Todor Mihaylov (Meta AI), Simon Vandenhende (Meta AI), Yash Patel (CTU in Prague), Yi Wen (Meta AI), Vignesh Ramanathan (Meta AI), and Dhruv Mahajan (Meta AI)</i>	
Turning a CLIP Model Into a Scene Text Detector	6978
<i>Wenwen Yu (Huazhong University of Science and Technology), Yuliang Liu (Huazhong University of Science and Technology), Wei Hua (Huazhong University of Science and Technology), Deqiang Jiang (Tencent YouTu Lab), Bo Ren (Tencent YouTu Lab), and Xiang Bai (Huazhong University of Science and Technology)</i>	
ScanDMM: A Deep Markov Model of Scanpath Prediction for 360° Images	6989
<i>Xiangjie Sui (Jiangxi University of Finance and Economics), Yuming Fang (Jiangxi University of Finance and Economics), Hanwei Zhu (City University of Hong Kong), Shiqi Wang (City University of Hong Kong), and Zhou Wang (University of Waterloo)</i>	
CrOC: Cross-View Online Clustering for Dense Visual Representation Learning	7000
<i>Thomas Stegmüller (EPFL), Tim Lebailly (KU Leuven), Behzad Bozorgtabar (EPFL; CHUV), Tinne Tuytelaars (KU Leuven), and Jean-Philippe Thiran (EPFL; CHUV)</i>	
PLA: Language-Driven Open-Vocabulary 3D Scene Understanding	7010
<i>Runyu Ding (The University of Hong Kong), Jihan Yang (The University of Hong Kong), Chuhui Xue (ByteDance), Wenqing Zhang (ByteDance), Song Bai (ByteDance), and Xiaojuan Qi (The University of Hong Kong)</i>	
CLIP2Scene: Towards Label-Efficient 3D Scene Understanding by CLIP	7020
<i>Runnan Chen (The University of Hong Kong; Shanghai AI Laboratory), Youquan Liu (Shanghai AI Laboratory; Hochschule Bremerhaven), Lingdong Kong (Shanghai AI Laboratory; National University of Singapore), Xinge Zhu (The Chinese University of Hong Kong), Yuexin Ma (ShanghaiTech University), Yikang Li (Shanghai AI Laboratory), Yuenan Hou (Shanghai AI Laboratory), Yu Qiao (Shanghai AI Laboratory), and Wenping Wang (Texas A&M University)</i>	
CORA: Adapting CLIP for Open-Vocabulary Detection With Region Prompting and Anchor Pre-Matching	7031
<i>Xiaoshi Wu (The Chinese University of Hong Kong), Feng Zhu (SenseTime Research), Rui Zhao (SenseTime Research; SenseTime Research), and Hongsheng Li (The Chinese University of Hong Kong; Centre for Perceptual and Interactive Intelligence (CPII))</i>	
Open-Vocabulary Attribute Detection	7041
<i>María A. Bravo (University of Freiburg, Germany), Sudhanshu Mittal (University of Freiburg, Germany), Simon Ging (University of Freiburg, Germany), and Thomas Brox (University of Freiburg, Germany)</i>	
Learning To Detect and Segment for Open Vocabulary Object Detection	7051
<i>Tao Wang (Sichuan University)</i>	

Open-Vocabulary Semantic Segmentation With Mask-Adapted CLIP	7061
<i>Feng Liang (The University of Texas, Austin), Bichen Wu (Meta Reality Labs), Xiaoliang Dai (Meta Reality Labs), Kunpeng Li (Meta Reality Labs), Yanan Zhao (Meta Reality Labs), Hang Zhang (Cruise), Peizhao Zhang (Meta Reality Labs), Peter Vajda (Meta Reality Labs), and Diana Marculescu (The University of Texas, Austin)</i>	
A Simple Framework for Text-Supervised Semantic Segmentation	7071
<i>Muyang Yi (Shanghai Jiao Tong University), Quan Cui (Waseda University; ByteDance Inc.), Hao Wu (ByteDance Inc.), Cheng Yang (ByteDance Inc.), Osamu Yoshie (Waseda University), and Hongtao Lu (Shanghai Jiao Tong University)</i>	
GAPartNet: Cross-Category Domain-Generalizable Object Perception and Manipulation via Generalizable and Actionable Parts	7081
<i>Haoran Geng (Peking University; Beijing Institute for General Artificial Intelligence), Helin Xu (Tsinghua University), Chengyang Zhao (Peking University), Chao Xu (University of California, Los Angeles), Li Yi (Tsinghua University), Siyuan Huang (Beijing Institute for General Artificial Intelligence), and He Wang (Peking University)</i>	
GeoLayoutLM: Geometric Pre-Training for Visual Information Extraction	7092
<i>Chuwei Luo (DAMO Academy, Alibaba Group), Changxu Cheng (DAMO Academy, Alibaba Group), Qi Zheng (DAMO Academy, Alibaba Group), and Cong Yao (DAMO Academy, Alibaba Group)</i>	
Self-Supervised Image-to-Point Distillation via Semantically Tolerant Contrastive Loss	7102
<i>Anas Mahmoud (University of Toronto Robotics Institute), Jordan S. K. Hu (University of Toronto Robotics Institute), Tianshu Kuai (University of Toronto Robotics Institute), Ali Harakeh (Mila, Universite de Montreal), Liam Paull (Mila, Universite de Montreal), and Steven L. Waslander (University of Toronto Robotics Institute)</i>	
Generative Semantic Segmentation	7111
<i>Jiaqi Chen (Fudan University), Jiachen Lu (Fudan University), Xiatian Zhu (University of Surrey), and Li Zhang (University of Surrey)</i>	
MISC210K: A Large-Scale Dataset for Multi-Instance Semantic Correspondence	7121
<i>Yixuan Sun (Academy of Engineering & Technology, Fudan University, China), Yiwen Huang (Fudan University, China), Haijing Guo (Fudan University, China), Yuzhou Zhao (Fudan University, China), Runmin Wu (The University of Hong Kong, China), Yizhou Yu (The University of Hong Kong, China), Weifeng Ge (Fudan University, China), and Wenqiang Zhang (Academy of Engineering & Technology, Fudan University, China; Fudan University, China)</i>	
MIANet: Aggregating Unbiased Instance and General Information for Few-Shot Semantic Segmentation	7131
<i>Yong Yang (South China University of Technology), Qiong Chen (South China University of Technology), Yuan Feng (South China University of Technology; Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application), and Tianlin Huang (South China University of Technology)</i>	

PACO: Parts and Attributes of Common Objects	7141
<i>Vignesh Ramanathan (Meta AI), Anmol Kalia (Meta AI), Vladan Petrovic (Meta AI), Yi Wen (Meta AI), Baixue Zheng (Meta AI), Baishan Guo (Meta AI), Rui Wang (Meta AI), Aaron Marquez (Meta AI), Rama Kovvuri (Meta AI), Abhishek Kadian (Meta AI), Amir Mousavi (Simon Fraser University), Yiwen Song (Meta AI), Abhimanyu Dubey (Meta AI), and Dhruv Mahajan (Meta AI)</i>	
PartDistillation: Learning Parts From Instance Segmentation	7152
<i>Jang Hyun Cho (UT Austin), Philipp Krähenbühl (UT Austin), and Vignesh Ramanathan (Meta AI)</i>	
ACSeg: Adaptive Conceptualization for Unsupervised Semantic Segmentation	7162
<i>Kehan Li (Peking University, China; Peking University Shenzhen Graduate School, Shenzhen, China), Zhennan Wang (Peng Cheng Laboratory, Shenzhen, China), Zesen Cheng (Peking University, China; Peking University Shenzhen Graduate School, Shenzhen, China), Runyi Yu (Peking University, China; Peking University Shenzhen Graduate School, Shenzhen, China), Yian Zhao (Dalian University of Technology), Guoli Song (Peng Cheng Laboratory, Shenzhen, China), Chang Liu (Tsinghua University, Beijing, China), Li Yuan (Peking University, China; Peng Cheng Laboratory, China; Peking University Shenzhen Graduate School, Shenzhen, China), and Jie Chen (Peking University, China; Peng Cheng Laboratory, China; Peking University Shenzhen Graduate School, Shenzhen, China)</i>	
Reliability in Semantic Segmentation: Are We on the Right Track?	7173
<i>Pau de Jorge (University of Oxford NAVER LABS Europe, NAVER LABS Europe*), Riccardo Volpi (NAVER LABS Europe), Philip H.S. Torr (University of Oxford), and Grégory Rogez (NAVER LABS Europe)</i>	
Rethinking the Correlation in Few-Shot Segmentation: A Buoy's View	7183
<i>Yuan Wang (University of Science and Technology of China), Rui Sun (University of Science and Technology of China), and Tianzhu Zhang (University of Science and Technology of China)</i>	
SIM: Semantic-Aware Instance Mask Generation for Box-Supervised Instance Segmentation	7193
<i>Ruihuang Li (The Hong Kong Polytechnic University), Chenhang He (The Hong Kong Polytechnic University), Yabin Zhang (The Hong Kong Polytechnic University), Shuai Li (The Hong Kong Polytechnic University), Liyi Chen (The Hong Kong Polytechnic University), and Lei Zhang (The Hong Kong Polytechnic University)</i>	
Endpoints Weight Fusion for Class Incremental Semantic Segmentation	7204
<i>Jia-Wen Xiao (Nankai University), Chang-Bin Zhang (Nankai University), Jiekang Feng (Tianjin University), Xialei Liu (Nankai University), Joost van de Weijer (Universitat Autònoma de Barcelona), and Ming-Ming Cheng (Nankai University)</i>	

Incrementer: Transformer for Class-Incremental Semantic Segmentation With Knowledge Distillation Focusing on Old Class	7214
<i>Chao Shang (University of Electronic Science and Technology of China), Hongliang Li (University of Electronic Science and Technology of China), Fanman Meng (University of Electronic Science and Technology of China), Qingbo Wu (University of Electronic Science and Technology of China), Heqian Qiu (University of Electronic Science and Technology of China), and Lanxiao Wang (University of Electronic Science and Technology of China)</i>	
Continuous Pseudo-Label Rectified Domain Adaptive Semantic Segmentation With Implicit Neural Representations	7225
<i>Rui Gong (CVL, ETH Zürich), Qin Wang (CVL, ETH Zürich), Martin Danelljan (CVL, ETH Zürich), Dengxin Dai (Max Planck Institute for Informatics, Saarland Informatics Campus), and Luc Van Gool (CVL, ETH Zürich)</i>	
Revisiting Weak-to-Strong Consistency in Semi-Supervised Semantic Segmentation	7236
<i>Lihe Yang (Nanjing University), Lei Qi (Southeast University), Litong Feng (SenseTime Research), Wayne Zhang (SenseTime Research), and Yinghuan Shi (Nanjing University)</i>	
Discriminative Co-Saliency and Background Mining Transformer for Co-Salient Object Detection	7247
<i>Long Li (Northwestern Polytechnical University), Junwei Han (Northwestern Polytechnical University), Ni Zhang (Northwestern Polytechnical University), Nian Liu (Mohamed bin Zayed University of Artificial Intelligence), Salman Khan (Mohamed bin Zayed University of Artificial Intelligence; Australian National University 4CVL, Linköping University), Hisham Cholakkal (Mohamed bin Zayed University of Artificial Intelligence), Rao Muhammad Anwer (Mohamed bin Zayed University of Artificial Intelligence), and Fahad Shahbaz Khan (Mohamed bin Zayed University of Artificial Intelligence; Australian National University 4CVL, Linköping University)</i>	
Texture-Guided Saliency Distilling for Unsupervised Salient Object Detection	7257
<i>Huajun Zhou (Sun Yat-sen University, China), Bo Qiao (Sun Yat-sen University, China), Lingxiao Yang (Sun Yat-sen University, China), Jianhuang Lai (Sun Yat-sen University, China; Guangdong Province Key Laboratory of Information Security Technology, China; Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, China), and Xiaohua Xie (Sun Yat-sen University, China; Guangdong Province Key Laboratory of Information Security Technology, China; Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, China)</i>	
An Erudite Fine-Grained Visual Classification Model	7268
<i>Dongliang Chang (Beijing University of Posts and Telecommunications, China), Yujun Tong (Beijing University of Posts and Telecommunications, China), Ruoyi Du (Beijing University of Posts and Telecommunications, China), Timothy Hospedales (University of Edinburgh, UK), Yi-Zhe Song (SketchX, CVSSP, University of Surrey, UK), and Zhanyu Ma (Beijing University of Posts and Telecommunications, China)</i>	

Dynamic Graph Learning With Content-Guided Spatial-Frequency Relation Reasoning for Deepfake Detection	7278
<i>Yuan Wang (Institute of Automation, Chinese Academy of Sciences; Alibaba Group; University of Chinese Academy of Sciences), Kun Yu (Alibaba Group), Chen Chen (Institute of Automation, Chinese Academy of Sciences), Xiyuan Hu (Nanjing University of Science and Technology), and Silong Peng (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Beijing Visystem Co.Ltd)</i>	
ScaleDet: A Scalable Multi-Dataset Object Detector	7288
<i>Yanbei Chen (AWS AI Labs), Manchen Wang (AWS AI Labs), Abhay Mittal (AWS AI Labs), Zhenlin Xu (AWS AI Labs), Paolo Favaro (AWS AI Labs), Joseph Tighe (AWS AI Labs), and Davide Modolo (AWS AI Labs)</i>	
Multi-Centroid Task Descriptor for Dynamic Class Incremental Inference	7298
<i>Tenghao Cai (East China Normal University, Shanghai, China), Zhizhong Zhang (East China Normal University, Shanghai, China), Xin Tan (East China Normal University, Shanghai, China), Yanyun Qu (Xiamen University, Fujian, China), Guannan Jiang (CATL, China), Chengjie Wang (Tencent Youtu Lab), and Yuan Xie (East China Normal University, Shanghai, China)</i>	
Matching Is Not Enough: A Two-Stage Framework for Category-Agnostic Pose Estimation	7308
<i>Min Shi (Huazhong University of Science and Technology; Shanghai AI Laboratory), Zihao Huang (Huazhong University of Science and Technology), Xianzheng Ma (Shanghai AI Laboratory), Xiaowei Hu (Shanghai AI Laboratory), and Zhiguo Cao (Huazhong University of Science and Technology; Shanghai AI Laboratory)</i>	
Dynamic Coarse-To-Fine Learning for Oriented Tiny Object Detection	7318
<i>Chang Xu (Wuhan University), Jian Ding (Wuhan University), Jinwang Wang (Wuhan University), Wen Yang (Wuhan University), Huai Yu (Wuhan University), Lei Yu (Wuhan University), and Gui-Song Xia (Wuhan University)</i>	
Dense Distinct Query for End-to-End Object Detection	7329
<i>Shilong Zhang (Shanghai AI Laboratory; The University of Hong Kong), Xinjiang Wang (SenseTime Research), Jiaqi Wang (Shanghai AI Laboratory), Jiangmiao Pang (Shanghai AI Laboratory), Chengqi Lyu (Shanghai AI Laboratory), Wenwei Zhang (Nanyang Technological University), Ping Luo (Shanghai AI Laboratory; The University of Hong Kong), and Kai Chen (Shanghai AI Laboratory)</i>	
Meta-Tuning Loss Functions and Data Augmentation for Few-Shot Object Detection	7339
<i>Berkan Demirel (Middle East Technical University; HAVELSAN Inc.), Orhun Buğra Baran (Middle East Technical University), and Ramazan Gokberk Cinbis (Middle East Technical University)</i>	
One-to-Few Label Assignment for End-to-End Dense Detection	7350
<i>Shuai Li (The Hong Kong Polytechnic University; OPPO Research Institute), Minghan Li (The Hong Kong Polytechnic University), Ruihuang Li (The Hong Kong Polytechnic University), Chenhong He (The Hong Kong Polytechnic University; OPPO Research Institute), and Lei Zhang (OPPO Research Institute)</i>	

Test Time Adaptation With Regularized Loss for Weakly Supervised Salient Object Detection	7360
<i>Olga Veksler (University of Waterloo, Canada)</i>	
MixTeacher: Mining Promising Labels With Mixed Scale Teacher for Semi-Supervised Object Detection	7370
<i>Liang Liu (Youtu Lab, Tencent), Boshen Zhang (Youtu Lab, Tencent), Jiangning Zhang (Youtu Lab, Tencent), Wuhao Zhang (Youtu Lab, Tencent), Zhenye Gan (Youtu Lab, Tencent), Guanzhong Tian (Zhejiang University), Wenbing Zhu (Rongcheer Co., Ltd), Yabiao Wang (Youtu Lab, Tencent), and Chengjie Wang (Youtu Lab, Tencent; Shanghai Jiao Tong University)</i>	
Exploring Incompatible Knowledge Transfer in Few-Shot Image Generation	7380
<i>Yunqing Zhao (Singapore University of Technology and Design (SUTD)), Chao Du (Sea AI Lab), Milad Abdollahzadeh (Singapore University of Technology and Design (SUTD)), Tianyu Pang (Sea AI Lab), Min Lin (Sea AI Lab), Shuicheng Yan (Sea AI Lab), and Ngai-Man Cheung (Singapore University of Technology and Design (SUTD))</i>	
Exploring Intra-Class Variation Factors With Learnable Cluster Prompts for Semi-Supervised Image Synthesis	7392
<i>Yunfei Zhang (South China University of Technology), Xiaoyang Huo (South China University of Technology), Tianyi Chen (South China University of Technology), Si Wu (South China University of Technology; Peng Cheng Laboratory; PAZHOU LAB), and Hau San Wong (Department of Computer Science, City University of Hong Kong)</i>	
A Soma Segmentation Benchmark in Full Adult Fly Brain	7402
<i>Xiaoyu Liu (University of Science and Technology of China), Bo Hu (University of Science and Technology of China), Mingxing Li (University of Science and Technology of China), Wei Huang (University of Science and Technology of China), Yueyi Zhang (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), and Zhiwei Xiong (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center)</i>	
SDC-UDA: Volumetric Unsupervised Domain Adaptation Framework for Slice-Direction Continuous Cross-Modality Medical Image Segmentation	7412
<i>Hyungseob Shin (Yonsei University), Hyeongyu Kim (Yonsei University), Sewon Kim (Naver AI Lab; Naver Cloud), Yohan Jun (Martinos Center for Biomedical Imaging; Harvard Medical School), Taejoon Eo (Yonsei University; Probe Medical, Inc.), and Dosik Hwang (Yonsei University; Center for Healthcare Robotics, Korea Institute of Science and Technology)</i>	
Label-Free Liver Tumor Segmentation	7422
<i>Qixin Hu (Huazhong University of Science and Technology), Yixiong Chen (The Chinese University of Hong Kong, Shenzhen), Junfei Xiao (Johns Hopkins University), Shuwen Sun (The First Affiliated Hospital of Nanjing Medical University), Jieneng Chen (Johns Hopkins University), Alan L. Yuille (Johns Hopkins University), and Zongwei Zhou (Johns Hopkins University)</i>	

Interactive and Explainable Region-Guided Radiology Report Generation	7433
<i>Tim Tanida (Technical University of Munich), Philip Müller (Technical University of Munich), Georgios Kaissis (Technical University of Munich; Helmholtz Zentrum Munich), and Daniel Rueckert (Technical University of Munich; Imperial College London)</i>	
A Loopback Network for Explainable Microvascular Invasion Classification	7443
<i>Shengxuming Zhang (Zhejiang University), Tianqi Shi (Alibaba Group), Yang Jiang (Alibaba Group), Xiuming Zhang (Zhejiang University), Jie Lei (Zhejiang University of Technology), Zunlei Feng (Zhejiang University; Shanghai Institute for Advanced Study of Zhejiang University; Zhejiang Provincial Key Laboratory of Service Robot), and Mingli Song (Zhejiang University; Shanghai Institute for Advanced Study of Zhejiang University; Zhejiang Provincial Key Laboratory of Service Robot)</i>	
Task-Specific Fine-Tuning via Variational Information Bottleneck for Weakly-Supervised Pathology Whole Slide Image Classification	7454
<i>Honglin Li (Zhejiang University; Westlake University), Chenglu Zhu (Research Center for Industries of the Future; Westlake University), Yunlong Zhang (Zhejiang University; Westlake University), Yuxuan Sun (Zhejiang University; Westlake University), Zhongyi Shui (Zhejiang University; Westlake University), Wenwei Kuang (Westlake University; The University of Hong Kong), Sunyi Zheng (Research Center for Industries of the Future; Westlake University), and Lin Yang (Research Center for Industries of the Future; Westlake University)</i>	
YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors.....	7464
<i>Chien-Yao Wang (Academia Sinica, Taiwan), Alexey Bochkovskiy (Academia Sinica, Taiwan), and Hong-Yuan Mark Liao (Academia Sinica, Taiwan)</i>	
Two-Way Multi-Label Loss	7476
<i>Takumi Kobayashi (National Institute of Advanced Industrial Science and Technology, Japan; University of Tsukuba, Japan)</i>	
Teaching Matters: Investigating the Role of Supervision in Vision Transformers	7486
<i>Matthew Walmer (University of Maryland, College Park), Saksham Suri (University of Maryland, College Park), Kamal Gupta (University of Maryland, College Park), and Abhinav Shrivastava (University of Maryland, College Park)</i>	
Label Information Bottleneck for Label Enhancement	7497
<i>Qinghai Zheng (Fuzhou University, China), Jihua Zhu (Xi'an Jiaotong University, China), and Haoyu Tang (Shandong University, China)</i>	
Glocal Energy-Based Learning for Few-Shot Open-Set Recognition	7507
<i>Haoyu Wang (Northwestern Polytechnical University), Guansong Pang (Singapore Management University), Peng Wang (University of Wollongong), Lei Zhang (Northwestern Polytechnical University), Wei Wei (Northwestern Polytechnical University), and Yanning Zhang (Northwestern Polytechnical University)</i>	
Noisy Correspondence Learning With Meta Similarity Correction	7517
<i>Haochen Han (Xi'an Jiaotong University, China), Kaiyao Miao (Xi'an Jiaotong University, China), Qinghua Zheng (Xi'an Jiaotong University, China), and Minnan Luo (Xi'an Jiaotong University, China)</i>	

Hubs and Hyperspheres: Reducing Hubness and Improving Transductive Few-Shot Learning With Hyperspherical Embeddings	7527
<i>Daniel J. Trosten (UiT The Arctic University of Norway), Riwiddhi Chakraborty (UiT The Arctic University of Norway), Sigurd Løkse (UiT The Arctic University of Norway), Kristoffer Knutsen Wickstrøm (UiT The Arctic University of Norway), Robert Jenssen (UiT The Arctic University of Norway), and Michael C. Kampffmeyer (UiT The Arctic University of Norway)</i>	
Coreset Sampling From Open-Set for Fine-Grained Self-Supervised Learning	7537
<i>Sungnyun Kim (KAIST AI), Sangmin Bae (KAIST AI), and Se-Young Yun (KAIST AI)</i>	
Boosting Semi-Supervised Learning by Exploiting All Unlabeled Data	7548
<i>Yuhao Chen (MEGVII Technology), Xin Tan (East China Normal University), Borui Zhao (MEGVII Technology), Zhaowei Chen (MEGVII Technology), Renjie Song (MEGVII Technology), Jiajun Liang (MEGVII Technology), and Xuequan Lu (Deakin University)</i>	
Trade-Off Between Robustness and Accuracy of Vision Transformers	7558
<i>Yanxi Li (The University of Sydney, Australia) and Chang Xu (The University of Sydney, Australia)</i>	
Exploring and Utilizing Pattern Imbalance	7569
<i>Shibin Mei (Shanghai Jiao Tong University, China), Chenglong Zhao (Shanghai Jiao Tong University, China), Shengchao Yuan (Shanghai Jiao Tong University, China), and Bingbing Ni (Shanghai Jiao Tong University, China)</i>	
Dynamic Conceptual Contrastive Learning for Generalized Category Discovery	7579
<i>Nan Pu (University of Trento, Italy), Zhun Zhong (University of Trento, Italy), and Nicu Sebe (University of Trento, Italy)</i>	
Towards Better Decision Forests: Forest Alternating Optimization	7589
<i>Miguel Á. Carreira-Perpiñán (Dept. CSE, University of California, Merced), Magzhan Gabidolla (Dept. CSE, University of California, Merced), and Arman Zharmagambetov (Dept. CSE, University of California, Merced)</i>	
Learning Debaised Representations via Conditional Attribute Interpolation	7599
<i>Yi-Kai Zhang (Nanjing University), Qi-Wei Wang (Nanjing University), De-Chuan Zhan (Nanjing University), and Han-Jia Ye (Nanjing University)</i>	
On the Pitfall of Mixup for Uncertainty Calibration	7609
<i>Deng-Bao Wang (Southeast University, China), Lanqing Li (Tencent AI Lab; Zhejiang Lab; The Chinese University of Hong Kong), Peilin Zhao (Tencent AI Lab), Pheng-Ann Heng (The Chinese University of Hong Kong), and Min-Ling Zhang (Southeast University, Nanjing, China)</i>	
Class Relationship Embedded Learning for Source-Free Unsupervised Domain Adaptation ...	7619
<i>Yixin Zhang (Institute of Artificial Intelligence, Hefei Comprehensive National Science Center; University of Science and Technology of China), Zilei Wang (University of Science and Technology of China), and Weinan He (University of Science and Technology of China)</i>	

FeatureBooster: Boosting Feature Descriptors With a Lightweight Neural Network	7630
<i>Xinjiang Wang (Shanghai Jiao Tong University; SJTU SEIEE * G60 Yun Zhi AI Innovation and Application Research Center), Zeyu Liu (Shanghai Jiao Tong University; SJTU SEIEE * G60 Yun Zhi AI Innovation and Application Research Center), Yu Hu (Shanghai Jiao Tong University; SJTU SEIEE * G60 Yun Zhi AI Innovation and Application Research Center), Wei Xi (Intelligent Perception Institute, Midea Corporate Research Center), Wenxian Yu (SJTU SEIEE * G60 Yun Zhi AI Innovation and Application Research Center), and Danping Zou (Shanghai Jiao Tong University; SJTU SEIEE * G60 Yun Zhi AI Innovation and Application Research Center)</i>	
Guiding Pseudo-Labels With Uncertainty Estimation for Source-Free Unsupervised Domain Adaptation	7640
<i>Mattia Litrico (Pattern Analysis and Computer Vision (PAVIS) - Istituto Italiano di Tecnologia), Alessio Del Bue (Pattern Analysis and Computer Vision (PAVIS) - Istituto Italiano di Tecnologia), and Pietro Morerio (Pattern Analysis and Computer Vision (PAVIS) - Istituto Italiano di Tecnologia)</i>	
Divide and Adapt: Active Domain Adaptation via Customized Learning	7651
<i>Duojun Huang (Sun Yat-sen University, China), Jichang Li (The University of Hong Kong), Weikai Chen (Tencent America), Junshi Huang (Meituan), Zhenhua Chai (Meituan), and Guanbin Li (Sun Yat-sen University, China)</i>	
Understanding and Constructing Latent Modality Structures in Multi-Modal Representation Learning	7661
<i>Qian Jiang (University of Illinois at Urbana-Champaign), Changyou Chen (University at Buffalo; Amazon), Han Zhao (University of Illinois at Urbana-Champaign; Amazon), Liqun Chen (Amazon), Qing Ping (Amazon), Son Dinh Tran (Amazon), Yi Xu (Amazon), Belinda Zeng (Amazon), and Trishul Chilimbi (Amazon)</i>	
Deep Factorized Metric Learning	7672
<i>Chengkun Wang (Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China), Wenzhao Zheng (Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China), Junlong Li (Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China), Jie Zhou (Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China), and Jiwen Lu (Tsinghua University, China; Beijing National Research Center for Information Science and Technology, China)</i>	

Meta-Causal Learning for Single Domain Generalization	7683
<i>Jin Chen (Beijing Key Laboratory of Intelligent Information Technology, School of Computer Science & Technology, Beijing Institute of Technology, China), Zhi Gao (Beijing Key Laboratory of Intelligent Information Technology, School of Computer Science & Technology, Beijing Institute of Technology, China) Beijing Key Laboratory of Intelligent Information Technology, School of Computer Science & Technology, Beijing Institute of Technology, China Beijing Key Laboratory of Intelligent Information Technology, School of Computer Science & Technology, Beijing Institute of Technology, China Beijing Key Laboratory of Intelligent Information Technology, School of Computer Science & Technology, Beijing Institute of Technology, China) Xinxiao Wu (Beijing Key Laboratory of Intelligent Information Technology, School of Computer Science & Technology, Beijing Institute of Technology, China; Shenzhen MSU-BIT University, China), and Jiebo Luo (University of Rochester, USA)</i>	
Meta Omnium: A Benchmark for General-Purpose Learning-To-Learn	7693
<i>Ondrej Bohdal (The University of Edinburgh), Yinbing Tian (Beijing University of Posts and Telecommunications), Yongshuo Zong (The University of Edinburgh), Ruchika Chavhan (The University of Edinburgh), Da Li (Samsung AI Center, Cambridge), Henry Gouk (The University of Edinburgh), Li Guo (Beijing University of Posts and Telecommunications), and Timothy Hospedales (The University of Edinburgh; Samsung AI Center, Cambridge)</i>	
Robust Mean Teacher for Continual and Gradual Test-Time Adaptation	7704
<i>Mario Döbler (University of Stuttgart), Robert A. Marsden (University of Stuttgart), and Bin Yang (University of Stuttgart)</i>	
NAR-Former: Neural Architecture Representation Learning Towards Holistic Attributes Prediction	7715
<i>Yun Yi (Xidian University), Haokui Zhang (Intellifusion; Harbin Institute of Technology, Shenzhen), Wenze Hu (Intellifusion), Nannan Wang (Xidian University), and Xiaoyu Wang (Intellifusion)</i>	
Visual Query Tuning: Towards Effective Usage of Intermediate Representations for Parameter and Memory Efficient Transfer Learning	7725
<i>Cheng-Hao Tu (The Ohio State University), Zheda Mai (The Ohio State University), and Wei-Lun Chao (The Ohio State University)</i>	
Architecture, Dataset and Model-Scale Agnostic Data-Free Meta-Learning	7736
<i>Zixuan Hu (Tsinghua Shenzhen International Graduate School, China), Li Shen (JD Explore Academy, China), Zhenyi Wang (State University of New York at Buffalo, USA), Tongliang Liu (The University of Sydney, Australia), Chun Yuan (Tsinghua Shenzhen International Graduate School, China), and Dacheng Tao (JD Explore Academy, China)</i>	
GKEAL: Gaussian Kernel Embedded Analytic Learning for Few-Shot Class Incremental Task ..	7746
<i>Huiping Zhuang (South China University of Technology, China), Zhenyu Weng (Nanyang Technological University, Singapore), Run He (South China University of Technology, China), Zhiping Lin (Nanyang Technological University, Singapore), and Ziqian Zeng (South China University of Technology, China)</i>	

Mitigating Task Interference in Multi-Task Learning via Explicit Task Routing With Non-Learnable Primitives	7756
<i>Chuntao Ding (Beijing Jiaotong University), Zhichao Lu (Sun Yat-sen University), Shangguang Wang (Beijing University of Posts and Telecommunications), Ran Cheng (Southern University of Science and Technology), and Vishnu Naresh Boddeti (Michigan State University)</i>	
Boundary Unlearning: Rapid Forgetting of Deep Networks via Shifting the Decision Boundary	7766
<i>Min Chen (Huazhong University of Science and Technology, China), Weizhuo Gao (Huazhong University of Science and Technology, China), Gaoyang Liu (Huazhong University of Science and Technology, China), Kai Peng (Huazhong University of Science and Technology, China), and Chen Wang (Huazhong University of Science and Technology, China)</i>	
Task Difficulty Aware Parameter Allocation & Regularization for Lifelong Learning	7776
<i>Wenjin Wang (Zhejiang University, China), Yunqing Hu (Zhejiang University, China), Qianglong Chen (Zhejiang University, China), and Yin Zhang (Zhejiang University, China)</i>	
Learning To Retain While Acquiring: Combating Distribution-Shift in Adversarial Data-Free Knowledge Distillation	7786
<i>Gaurav Patel (Purdue University, Indian Institute of Technology Hyderabad), Konda Reddy Mopuri (Purdue University, Indian Institute of Technology Hyderabad), and Qiang Qiu (Purdue University, Indian Institute of Technology Hyderabad)</i>	
A Unified Knowledge Distillation Framework for Deep Directed Graphical Models	7795
<i>Yizhuo Chen (William and Mary & UIUC), Kaizhao Liang (UIUC), Zhe Zeng (UCLA), Shuochao Yao (George Mason University), and Huajie Shao (William and Mary)</i>	
Coaching a Teachable Student	7805
<i>Jimuyang Zhang (Boston University), Zanming Huang (Boston University), and Eshed Ohn-Bar (Boston University)</i>	
Adaptive Plasticity Improvement for Continual Learning	7816
<i>Yan-Shuo Liang (Nanjing University, P. R. China) and Wu-Jun Li (Nanjing University, P. R. China)</i>	
Improving Generalization of Meta-Learning With Inverted Regularization at Inner-Level	7826
<i>Lianzhe Wang (Tsinghua University), Shiji Zhou (Tsinghua University), Shanghang Zhang (Peking University), Xu Chu (Tsinghua University), Heng Chang (Tsinghua University), and Wenwu Zhu (Tsinghua University)</i>	
Trainable Projected Gradient Method for Robust Fine-Tuning	7836
<i>Junjiao Tian (Georgia Institute of Technology), Xiaoliang Dai (Meta), Chih-Yao Ma (Meta), Zecheng He (Meta), Yen-Cheng Liu (Georgia Institute of Technology), and Zsolt Kira (Georgia Institute of Technology)</i>	
Imitation Learning As State Matching via Differentiable Physics	7846
<i>Siwei Chen (National University of Singapore; Sea AI Lab), Xiao Ma (Sea AI Lab), and Zhongwen Xu (Sea AI Lab)</i>	

Improved Distribution Matching for Dataset Condensation	7856
<i>Ganlong Zhao (Sun Yat-sen University; The University of Hong Kong), Guanbin Li (Sun Yat-sen University, Shenzhen), Yipeng Qin (Cardiff University), and Yizhou Yu (The University of Hong Kong)</i>	
A General Regret Bound of Preconditioned Gradient Method for DNN Training	7866
<i>Hongwei Yong (The Hong Kong Polytechnic University), Ying Sun (The Hong Kong Polytechnic University), and Lei Zhang (The Hong Kong Polytechnic University)</i>	
From Node Interaction To Hop Interaction: New Effective and Scalable Graph Learning Paradigm	7876
<i>Jie Chen (Shanghai Key Lab of Intelligent Information Processing, School of Computer Science), Zilong Li (Shanghai Key Lab of Intelligent Information Processing, School of Computer Science), Yin Zhu (Shanghai Key Lab of Intelligent Information Processing, School of Computer Science), Junping Zhang (Shanghai Key Lab of Intelligent Information Processing, School of Computer Science), and Jian Pu (Fudan University, China)</i>	
Constructing Deep Spiking Neural Networks From Artificial Neural Networks With Knowledge Distillation	7886
<i>Qi Xu (Dalin University of Technology), Yaxin Li (Dalin University of Technology), Jiangrong Shen (Zhejiang University), Jian K. Liu (University of Leeds), Huajin Tang (Zhejiang University), and Gang Pan (Zhejiang University)</i>	
Rate Gradient Approximation Attack Threats Deep Spiking Neural Networks	7896
<i>Tong Bu (Peking University), Jianhao Ding (Peking University), Zecheng Hao (Peking University), and Zhaofei Yu (Peking University)</i>	
MobileOne: An Improved One Millisecond Mobile Backbone	7907
<i>Pavan Kumar Anasosalu Vasu (Apple), James Gabriel (Apple), Jeff Zhu (Apple), Oncel Tuzel (Apple), and Anurag Ranjan (Apple)</i>	
Understanding Masked Autoencoders via Hierarchical Latent Variable Models	7918
<i>Lingjing Kong (Carnegie Mellon University), Martin Q. Ma (Carnegie Mellon University), Guangyi Chen (Carnegie Mellon University; Mohamed bin Zayed University of Artificial Intelligence), Eric P. Xing (Carnegie Mellon University; Mohamed bin Zayed University of Artificial Intelligence), Yuejie Chi (Carnegie Mellon University), Louis-Philippe Morency (Carnegie Mellon University), and Kun Zhang (Carnegie Mellon University; Mohamed bin Zayed University of Artificial Intelligence)</i>	
Training Debiased Subnetworks With Contrastive Weight Pruning	7929
<i>Geon Yeong Park (Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), Korea), Sangmin Lee (Mathematical Sciences, Korea Advanced Institute of Science and Technology (KAIST), Korea), Sang Wan Lee (Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), Korea), and Jong Chul Ye (Bio and Brain Engineering, Mathematical Sciences, Kim Jaechul Graduate School of AI, Korea Advanced Institute of Science and Technology (KAIST), Korea)</i>	

One-Shot Model for Mixed-Precision Quantization	7939
<i>Ivan Koryakovskiy (n/a), Alexandra Yakovleva (n/a), Valentin Buchnev (n/a), Temur Isaev (Huawei Technologies Co. Ltd.), and Gleb Odnokikh (n/a)</i>	
Solving Oscillation Problem in Post-Training Quantization Through a Theoretical Perspective	7950
<i>Yuexiao Ma (Xiamen University, P.R. China), Huixia Li (ByteDance Inc.), Xiawu Zheng (Peng Cheng Laboratory, China), Xuefeng Xiao (ByteDance Inc.), Rui Wang (ByteDance Inc.), Shilei Wen (ByteDance Inc.), Xin Pan (ByteDance Inc.), Fei Chao (Xiamen University, P.R. China), and Rongrong Ji (Xiamen University, P.R. China; Shenzhen Research Institute of Xiamen University)</i>	
Adaptive Data-Free Quantization	7960
<i>Biao Qian (Hefei University of Technology, China), Yang Wang (Hefei University of Technology, China), Richang Hong (Hefei University of Technology, China), and Meng Wang (Hefei University of Technology, China)</i>	
Learning To Generate Image Embeddings With User-Level Differential Privacy	7969
<i>Zheng Xu (Google Research), Maxwell Collins (Google Research), Yuxiao Wang (Google Research), Liviu Panait (Google Research), Sewoong Oh (Google Research), Sean Augenstein (Google Research), Ting Liu (Google Research), Florian Schroff (Google Research), and H. Brendan McMahan (Google Research)</i>	
Cross-GAN Auditing: Unsupervised Identification of Attribute Level Similarities and Differences Between Pretrained Generative Models	7981
<i>Matthew L. Olson (Oregon State University - EECS), Shusen Liu (Lawrence Livermore National Laboratory - CASC), Rushil Anirudh (Lawrence Livermore National Laboratory - CASC), Jayaraman J. Thiagarajan (Lawrence Livermore National Laboratory - CASC), Peer-Timo Bremer (Lawrence Livermore National Laboratory - CASC), and Weng-Keen Wong (Oregon State University - EECS)</i>	
HandsOff: Labeled Dataset Generation With No Additional Human Annotations	7991
<i>Austin Xu (Georgia Institute of Technology; Amazon AWS), Mariya I. Vasileva (Georgia Institute of Technology; Amazon AWS), Achal Dave (Georgia Institute of Technology; Amazon AWS), and Arjun Seshadri (Georgia Institute of Technology; Amazon AWS)</i>	
Attribute-Preserving Face Dataset Anonymization via Latent Code Optimization	8001
<i>Simone Barattin (University of Trento), Christos Tzelepis (Queen Mary University of London), Ioannis Patras (Queen Mary University of London), and Nicu Sebe (University of Trento)</i>	
Fake It Till You Make It: Learning Transferable Representations From Synthetic ImageNet Clones	8011
<i>Mert Bülent Saryıldız (NAVER LABS Europe; Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK), Karteek Alahari (Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK), Diane Larlus (NAVER LABS Europe), and Yannis Kalantidis (NAVER LABS Europe)</i>	

Unbiased Multiple Instance Learning for Weakly Supervised Video Anomaly Detection	8022
<i>Hui Lv (PCALab, Nanjing University of Science and Technology; Singapore Management University; Nanyang Technological University), Zhongqi Yue (Nanyang Technological University), Qianru Sun (Singapore Management University), Bin Luo (Alibaba Group), Zhen Cui (PCALab, Nanjing University of Science and Technology), and Hanwang Zhang (Nanyang Technological University)</i>	
Multimodal Industrial Anomaly Detection via Hybrid Fusion	8032
<i>Yue Wang (Shanghai Jiao Tong University, China), Jinlong Peng (Youtu Lab, Tencent), Jiangning Zhang (Youtu Lab, Tencent), Ran Yi (Shanghai Jiao Tong University, China), Yabiao Wang (Youtu Lab, Tencent), and Chengjie Wang (Shanghai Jiao Tong University, China; Youtu Lab, Tencent)</i>	
FedSeg: Class-Heterogeneous Federated Learning for Semantic Segmentation	8042
<i>Jiaxu Miao (Zhejiang University), Zongxin Yang (Zhejiang University), Leilei Fan (Zhejiang University), and Yi Yang (Zhejiang University)</i>	
Decentralized Learning With Multi-Headed Distillation	8053
<i>Andrey Zhmoginov (Google AI, USA), Mark Sandler (Google AI, USA), Nolan Miller (Google AI, USA), Gus Kristiansen (Google AI, USA), and Max Vladymyrov (Google AI, USA)</i>	
Learning Federated Visual Prompt in Null Space for MRI Reconstruction	8064
<i>Chun-Mei Feng (Technology and Research (A*STAR), Singapore), Bangjun Li (Shangdong University, China), Xinxing Xu (Technology and Research (A*STAR), Singapore), Yong Liu (Technology and Research (A*STAR), Singapore), Huazhu Fu (Technology and Research (A*STAR), Singapore), and Wangmeng Zuo (Harbin Institute of Technology, Harbin, China)</i>	
Federated Learning With Data-Agnostic Distribution Fusion	8074
<i>Jian-hui Duan (Nanjing University Nanjing, China), Wenzhong Li (Nanjing University Nanjing, China), Derun Zou (Nanjing University Nanjing, China), Ruichen Li (Nanjing University Nanjing, China), and Sanglu Lu (Nanjing University Nanjing, China)</i>	
CaPriDe Learning: Confidential and Private Decentralized Learning Based on Encryption-Friendly Distillation Loss	8084
<i>Nurbek Tastan (Mohamed bin Zayed University of Artificial Intelligence, UAE) and Karthik Nandakumar (Mohamed bin Zayed University of Artificial Intelligence, UAE)</i>	
RiDDLE: Reversible and Diversified De-Identification With Latent Encryptor	8093
<i>Dongze Li (University of Chinese Academy of Sciences; Center for Research on Intelligent Perception and Computing, CASIA), Wei Wang (Center for Research on Intelligent Perception and Computing, CASIA), Kang Zhao (Alibaba Group), Jing Dong (Center for Research on Intelligent Perception and Computing, CASIA), and Tieniu Tan (Nanjing University)</i>	

Multi-View Adversarial Discriminator: Mine the Non-Causal Factors for Object Detection in Unseen Domains	8103
<i>Mingjun Xu (Chongqing University, China; Hikvision Research Institute, China), Lingyun Qin (Chongqing University, China; Hikvision Research Institute, China), Weijie Chen (Chongqing University, China; Hikvision Research Institute, China), Shiliang Pu (Chongqing University, China; Hikvision Research Institute, China), and Lei Zhang (Chongqing University, China; Hikvision Research Institute, China)</i>	
Single Image Backdoor Inversion via Robust Smoothed Classifiers	8113
<i>Mingjie Sun (Carnegie Mellon University) and Zico Kolter (Carnegie Mellon University; Bosch Center for AI)</i>	
Effective Ambiguity Attack Against Passport-Based DNN Intellectual Property Protection Schemes Through Fully Connected Layer Substitution	8123
<i>Yiming Chen (University of Macau), Jinyu Tian (Macau University of Science and Technology), Xiangyu Chen (Shenzhen Institutes of Advanced Technology; Chinese Academy of Sciences), and Jiantao Zhou (University of Macau)</i>	
Color Backdoor: A Robust Poisoning Attack in Color Space	8133
<i>Wenbo Jiang (University of Electronic Science and Technology of China), Hongwei Li (University of Electronic Science and Technology of China), Guowen Xu (Nanyang Technological University), and Tianwei Zhang (Nanyang Technological University)</i>	
Adversarially Robust Neural Architecture Search for Graph Neural Networks	8143
<i>Beini Xie (Tsinghua University), Heng Chang (Tsinghua University), Ziwei Zhang (Tsinghua University), Xin Wang (Tsinghua University), Daixin Wang (Ant Group), Zhiqiang Zhang (Ant Group), Rex Ying (Yale University), and Wenwu Zhu (Tsinghua University)</i>	
Minimizing Maximum Model Discrepancy for Transferable Black-Box Targeted Attacks	8153
<i>Anqi Zhao (University of Electronic Science and Technology of China), Tong Chu (University of Electronic Science and Technology of China), Yahao Liu (University of Electronic Science and Technology of China), Wen Li (University of Electronic Science and Technology of China), Jingjing Li (University of Electronic Science and Technology of China), and Lixin Duan (University of Electronic Science and Technology of China)</i>	
StyLess: Boosting the Transferability of Adversarial Examples	8163
<i>Kaisheng Liang (The Hong Kong Polytechnic University) and Bin Xiao (The Hong Kong Polytechnic University)</i>	
Improving the Transferability of Adversarial Samples by Path-Augmented Method	8173
<i>Jianping Zhang (The Chinese University of Hong Kong), Jen-tse Huang (The Chinese University of Hong Kong), Wenxuan Wang (The Chinese University of Hong Kong), Yichen Li (The Chinese University of Hong Kong), Weibin Wu (Sun Yat-sen University), Xiaosen Wang (Huawei Singular Security Lab, China), Yuxin Su (Sun Yat-sen University), and Michael R. Lyu (The Chinese University of Hong Kong)</i>	

Feature Separation and Recalibration for Adversarial Robustness	8183
<i>Woo Jae Kim (Korea Advanced Institute of Science and Technology (KAIST)), Yoonki Cho (Korea Advanced Institute of Science and Technology (KAIST)), Junsik Jung (Korea Advanced Institute of Science and Technology (KAIST)), and Sung-Eui Yoon (Korea Advanced Institute of Science and Technology (KAIST))</i>	
CFA: Class-Wise Calibrated Fair Adversarial Training	8193
<i>Zeming Wei (Peking University), Yifei Wang (Peking University), Yiwen Guo (Independent Researcher), and Yisen Wang (Peking University)</i>	
Revisiting Residual Networks for Adversarial Robustness	8202
<i>Shihua Huang (Michigan State University), Zhichao Lu (Sun Yat-sen University), Kalyanmoy Deb (Michigan State University), and Vishnu Naresh Boddeti (Michigan State University)</i>	
Privacy-Preserving Adversarial Facial Features	8212
<i>Zhibo Wang (Zhejiang University, P. R. China; ZJU-Hangzhou Global Scientific and Technological Innovation Center), He Wang (Zhejiang University, P. R. China), Shuai fan Jin (Zhejiang University, P. R. China), Wenwen Zhang (Alibaba Group, P. R. China), Jiahui Hu (Zhejiang University, P. R. China), Yan Wang (Alibaba Group, P. R. China), Peng Sun (Hunan University, P. R. China), Wei Yuan (Hunan University, P. R. China), Kaixin Liu (Hunan University, P. R. China), and Kui Ren (Zhejiang University, P. R. China)</i>	
Edge-Aware Regional Message Passing Controller for Image Forgery Localization	8222
<i>Dong Li (University of Science and Technology of China, China), Jiaying Zhu (University of Science and Technology of China, China), Menglu Wang (University of Science and Technology of China, China), Jiawei Liu (University of Science and Technology of China, China), Xueyang Fu (University of Science and Technology of China, China), and Zheng-Jun Zha (University of Science and Technology of China, China)</i>	

Poster-Wed-AM

Swept-Angle Synthetic Wavelength Interferometry	8233
<i>Alankar Kotwal (Carnegie Mellon University), Anat Levin (Technion), and Ioannis Gkioulekas (Carnegie Mellon University)</i>	
RefSR-NeRF: Towards High Fidelity and Super Resolution View Synthesis	8244
<i>Xudong Huang (Huawei Noah's Ark Lab), Wei Li (Huawei Noah's Ark Lab), Jie Hu (Huawei Noah's Ark Lab), Hanting Chen (Huawei Noah's Ark Lab), and Yunhe Wang (Huawei Noah's Ark Lab)</i>	
FreeNeRF: Improving Few-Shot Neural Rendering With Free Frequency Regularization	8254
<i>Jiawei Yang (Stanford University; Nvidia Research), Marco Pavone (Stanford University; Nvidia Research), and Yue Wang (Stanford University; Nvidia Research)</i>	

Local-to-Global Registration for Bundle-Adjusting Neural Radiance Fields	8264
<i>Yue Chen (National Key Laboratory of Human-Machine Hybrid Augmented Intelligence; Xi'an Jiaotong University), Xingyu Chen (National Key Laboratory of Human-Machine Hybrid Augmented Intelligence; Xi'an Jiaotong University), Xuan Wang (Ant Group; Tencent AI Lab), Qi Zhang (Ant Group Tencent AI Lab), Yu Guo (National Key Laboratory of Human-Machine Hybrid Augmented Intelligence; Xi'an Jiaotong University; Tencent AI Lab), Ying Shan (Tencent AI Lab), and Fei Wang (National Key Laboratory of Human-Machine Hybrid Augmented Intelligence; Xi'an Jiaotong University)</i>	
Nerflets: Local Radiance Fields for Efficient Structure-Aware 3D Scene Representation From 2D Supervision	8274
<i>Xiaoshuai Zhang (Google Research; University of California, San Diego), Abhijit Kundu (Google Research), Thomas Funkhouser (Google Research), Leonidas Guibas (Google Research; Stanford University), Hao Su (University of California, San Diego), and Kyle Genova (Google Research)</i>	
NeRF-DS: Neural Radiance Fields for Dynamic Specular Objects	8285
<i>Zhiwen Yan (National University of Singapore), Chen Li (National University of Singapore), and Gim Hee Lee (National University of Singapore)</i>	
Grid-Guided Neural Radiance Fields for Large Urban Scenes	8296
<i>Linning Xu (The Chinese University of Hong Kong), Yuanbo Xiangli (The Chinese University of Hong Kong), Sida Peng (Zhejiang University), Xingang Pan (Max Planck Institute for Informatics), Nanxuan Zhao (Adobe Research), Christian Theobalt (Max Planck Institute for Informatics), Bo Dai (Shanghai AI Laboratory), and Dahua Lin (The Chinese University of Hong Kong; Shanghai AI Laboratory)</i>	
Learning Neural Duplex Radiance Fields for Real-Time View Synthesis	8307
<i>Ziyu Wan (City University of Hong Kong), Christian Richardt (Meta Reality Labs), Aljaž Božič (Meta Reality Labs), Chao Li (Meta Reality Labs), Vijay Rengarajan (Meta Reality Labs), Seonghyeon Nam (Meta Reality Labs), Xiaoyu Xiang (Meta Reality Labs), Tuotuo Li (Meta Reality Labs), Bo Zhu (Meta Reality Labs), Rakesh Ranjan (Meta Reality Labs), and Jing Liao (City University of Hong Kong)</i>	
EditableNeRF: Editing Topologically Varying Neural Radiance Fields by Key Points	8317
<i>Chengwei Zheng (Tsinghua University), Wenbin Lin (Tsinghua University), and Feng Xu (Tsinghua University)</i>	
Real-Time Neural Light Field on Mobile Devices	8328
<i>Junli Cao (Snap Inc.), Huan Wang (Northeastern University), Pavlo Chemerys (Snap Inc.), Vladislav Shakhrai (Snap Inc.), Ju Hu (Snap Inc.), Yun Fu (Northeastern University), Denys Makoviichuk (Snap Inc.), Sergey Tulyakov (Snap Inc.), and Jian Ren (Snap Inc.)</i>	

StyleRF: Zero-Shot 3D Style Transfer of Neural Radiance Fields	8338
<i>Kunhao Liu (Nanyang Technological University; Max Planck Institute for Informatics), Fangneng Zhan (University of Ottawa; MBZUAI), Yiwen Chen (Nanyang Technological University), Jiahui Zhang (Nanyang Technological University), Yingchen Yu (Nanyang Technological University), Abdulmotaleb El Saddik (University of Ottawa; MBZUAI), Shijian Lu (Nanyang Technological University), and Eric P. Xing (Carnegie Mellon University; MBZUAI)</i>	
Point2Pix: Photo-Realistic Point Cloud Rendering via Neural Radiance Fields	8349
<i>Tao Hu (The Chinese University of Hong Kong), Xiaogang Xu (The Chinese University of Hong Kong), Shu Liu (SmartMore), and Jiaya Jia (The Chinese University of Hong Kong; SmartMore)</i>	
Pointersect: Neural Rendering With Cloud-Ray Intersection	8359
<i>Jen-Hao Rick Chang (Apple), Wei-Yu Chen (Apple; Carnegie Mellon University), Anurag Ranjan (Apple), Kwang Moo Yi (Apple; University of British Columbia), and Oncel Tuzel (Apple)</i>	
Neural Fields Meet Explicit Geometric Representations for Inverse Rendering of Urban Scenes	8370
<i>Zian Wang (NVIDIA; University of Toronto; Vector Institute), Tianchang Shen (NVIDIA; University of Toronto; Vector Institute), Jun Gao (NVIDIA; University of Toronto; Vector Institute), Shengyu Huang (NVIDIA; ETH Zürich), Jacob Munkberg (NVIDIA), Jon Hasselgren (NVIDIA), Zan Gojcic (NVIDIA), Wenzheng Chen (NVIDIA; University of Toronto; Vector Institute), and Sanja Fidler (NVIDIA; University of Toronto; Vector Institute)</i>	
DANI-Net: Uncalibrated Photometric Stereo by Differentiable Shadow Handling, Anisotropic Reflectance Modeling, and Neural Inverse Rendering	8381
<i>Zongrui Li (Nanyang Technological University, Singapore), Qian Zheng (The State Key Lab of Brain-Machine Intelligence, Zhejiang University, China; Zhejiang University, China), Boxin Shi (National Key Laboratory for Multimedia Information Processing, Peking University, China; National Engineering Research Center of Visual Technology, Peking University, China), Gang Pan (The State Key Lab of Brain-Machine Intelligence, Zhejiang University, China; Zhejiang University, China), and Xudong Jiang (Nanyang Technological University, Singapore)</i>	
MAIR: Multi-View Attention Inverse Rendering With 3D Spatially-Varying Lighting Estimation..	8392
<i>JunYong Choi (Korea Institute of Science and Technology(KIST); Korea University), SeokYeong Lee (Korea Institute of Science and Technology(KIST); Korea University), Haesol Park (Korea Institute of Science and Technology(KIST)), Seung-Won Jung (Korea University), Ig-Jae Kim (Korea Institute of Science and Technology(KIST); University of Science and Technology; Yonsei University), and Junghyun Cho (Korea Institute of Science and Technology(KIST))</i>	
Weakly-Supervised Single-View Image Relighting	8402
<i>Renjiao Yi (National University of Defense Technology), Chenyang Zhu (National University of Defense Technology), and Kai Xu (National University of Defense Technology)</i>	

Controllable Light Diffusion for Portraits	8412
<i>David Futschik (Google Research; CTU in Prague, FEE), Kelvin Ritland (Google Research), James Vecore (Google Research), Sean Fanella (Google Research), Sergio Orts-Escolano (Google Research), Brian Curless (Google Research; University of Washington), Daniel Sýkora (Google Research; CTU in Prague, FEE), and Rohit Pandey (Google Research)</i>	
RGBD2: Generative Scene Synthesis via Incremental View Inpainting Using RGBD Diffusion Models	8422
<i>Jiabao Lei (South China University of Technology), Jiapeng Tang (Technical University of Munich), and Kui Jia (South China University of Technology; Peng Cheng Laboratory)</i>	
Neural Lens Modeling	8435
<i>Wenqi Xian (Meta Reality Labs Research; Cornell University), Aljaž Božič (Meta Reality Labs Research), Noah Snavely (Cornell University), and Christoph Lassner (Meta Reality Labs Research)</i>	
RealFusion: 360° Reconstruction of Any Object From a Single Image	8446
<i>Luke Melas-Kyriazi (University of Oxford), Iro Laina (University of Oxford), Christian Rupprecht (University of Oxford), and Andrea Vedaldi (University of Oxford)</i>	
Neuralangelo: High-Fidelity Neural Surface Reconstruction	8456
<i>Zhaoshuo Li (NVIDIA Research; Johns Hopkins University), Thomas Müller (NVIDIA Research), Alex Evans (NVIDIA Research), Russell H. Taylor (Johns Hopkins University), Mathias Unberath (Johns Hopkins University), Ming-Yu Liu (NVIDIA Research), and Chen-Hsuan Lin (NVIDIA Research)</i>	
PermutoSDF: Fast Multi-View Reconstruction With Implicit Surfaces Using Permutohedral Lattices	8466
<i>Radu Alexandru Rosu (University of Bonn, Germany) and Sven Behnke (University of Bonn, Germany)</i>	
NeuDA: Neural Deformable Anchor for High-Fidelity Implicit Surface Reconstruction	8476
<i>Bowen Cai (Tao Technology Department, Alibaba Group), Jinchi Huang (Tao Technology Department, Alibaba Group), Rongfei Jia (Tao Technology Department, Alibaba Group), Chengfei Lv (Tao Technology Department, Alibaba Group), and Huan Fu (Tao Technology Department, Alibaba Group)</i>	
NEF: Neural Edge Fields for 3D Parametric Curve Reconstruction From Multi-View Images	8486
<i>Yunfan Ye (National University of Defense Technology), Renjiao Yi (National University of Defense Technology), Zhirui Gao (National University of Defense Technology), Chenyang Zhu (National University of Defense Technology), Zhiping Cai (National University of Defense Technology), and Kai Xu (National University of Defense Technology)</i>	
NeuralField-LDM: Scene Generation With Hierarchical Latent Diffusion Models	8496
<i>Seung Wook Kim (NVIDIA; University of Toronto; Vector Institute), Bradley Brown (NVIDIA; University of Waterloo), Kangxue Yin (NVIDIA), Karsten Kreis (NVIDIA), Katja Schwarz (University of Tübingen, Tübingen AI Center), Daiqing Li (University of Tübingen, Tübingen AI Center), Robin Rombach (LMU Munich), Antonio Torralba (CSAIL, MIT), and Sanja Fidler (NVIDIA; University of Toronto; Vector Institute)</i>	

SinGRAF: Learning a 3D Generative Radiance Field for a Single Scene	8507
<i>Minjung Son (Samsung Advanced Institute of Technology (SAIT); Stanford University), Jeong Joon Park (Stanford University), Leonidas Guibas (Stanford University), and Gordon Wetzstein (Stanford University)</i>	
Painting 3D Nature in 2D: View Synthesis of Natural Scenes From a Single Semantic Mask	8518
<i>Shangzhan Zhang (Zhejiang University), Sida Peng (Zhejiang University), Tianrun Chen (Zhejiang University), Linzhan Mou (Zhejiang University), Haotong Lin (Zhejiang University), Kaicheng Yu (Alibaba Group), Yiyi Liao (Zhejiang University), and Xiaowei Zhou (Zhejiang University)</i>	
Quantitative Manipulation of Custom Attributes on 3D-Aware Image Synthesis	8529
<i>Hoseok Do (AI Lab, CTO Division, LG Electronics, South Korea; Seoul National University, South Korea), EunKyung Yoo (AI Lab, CTO Division, LG Electronics, South Korea), Taehyeong Kim (Seoul National University, South Korea), Chul Lee (AI Lab, CTO Division, LG Electronics, South Korea), and Jin Young Choi (Seoul National University, South Korea)</i>	
NeRFInventor: High Fidelity NeRF-GAN Inversion for Single-Shot Real Image Animation	8539
<i>Yu Yin (Northeastern University), Kamran Ghasedi (Microsoft), HsiangTao Wu (Microsoft), Jiaolong Yang (Microsoft), Xin Tong (Microsoft), and Yun Fu (Northeastern University)</i>	
PREIM3D: 3D Consistent Precise Image Attribute Editing From a Single Image	8549
<i>Jianhui Li (Tsinghua University; State Key Laboratory of Astronautic dynamics, Xi'an Satellite Control Center), Jianmin Li (Tsinghua University), Haoji Zhang (Tsinghua University), Shilong Liu (Tsinghua University), Zhengyi Wang (Tsinghua University), Zihao Xiao (RealAI), Kaiwen Zheng (Tsinghua University), and Jun Zhu (Tsinghua University)</i>	
Unsupervised 3D Shape Reconstruction by Part Retrieval and Assembly	8559
<i>Xianghao Xu (Brown University (USA)), Paul Guerrero (Adobe Research (UK, USA, India)), Matthew Fisher (Adobe Research (UK, USA, India)), Siddhartha Chaudhuri (Adobe Research (UK, USA, India)), and Daniel Ritchie (Brown University (USA))</i>	
DiffSwap: High-Fidelity and Controllable Face Swapping via 3D-Aware Masked Diffusion	8568
<i>Wenliang Zhao (Tsinghua University), Yongming Rao (Tsinghua University), Weikang Shi (Tsinghua University), Zuyan Liu (Tsinghua University), Jie Zhou (Tsinghua University), and Jiwen Lu (Tsinghua University)</i>	
Fine-Grained Face Swapping via Regional GAN Inversion	8578
<i>Zhian Liu (South China University of Technology), Maomao Li (Tencent AI Lab), Yong Zhang (Tencent AI Lab), Cairong Wang (Tsinghua Shenzhen International Graduate School), Qi Zhang (Tencent AI Lab), Jue Wang (Tencent AI Lab), and Yongwei Nie (South China University of Technology)</i>	
Logical Consistency and Greater Descriptive Power for Facial Hair Attribute Learning	8588
<i>Haiyu Wu (University of Notre Dame), Grace Bezdold (University of Notre Dame), Aman Bhatta (University of Notre Dame), and Kevin W. Bowyer (University of Notre Dame)</i>	

Learning a 3D Morphable Face Reflectance Model From Low-Cost Data	8598
<i>Yuxuan Han (Tsinghua University), Zhibo Wang (SenseTime Research), and Feng Xu (Tsinghua University)</i>	
StyleGAN Salon: Multi-View Latent Optimization for Pose-Invariant Hairstyle Transfer	8609
<i>Sasikarn Khwanmuang (VISTEC, Thailand), Pakkapon Phongthawee (VISTEC, Thailand), Patsorn Sangkloy (Phranakhon Rajabhat University, Thailand), and Supasorn Suwajanakorn (VISTEC, Thailand)</i>	
FaceLit: Neural 3D Relightable Faces	8619
<i>Anurag Ranjan (Apple), Kwang Moo Yi (Apple; The University of British Columbia), Jen-Hao Rick Chang (Apple), and Oncel Tuzel (Apple)</i>	
FitMe: Deep Photorealistic 3D Morphable Model Avatars	8629
<i>Alexandros Lattas (Imperial College London, UK), Stylianos Moschoglou (Imperial College London, UK), Stylianos Ploumpis (Imperial College London, UK), Baris Gecer (Imperial College London, UK), Jiankang Deng (Imperial College London, UK), and Stefanos Zafeiriou (Imperial College London, UK)</i>	
NeuWigs: A Neural Dynamic Model for Volumetric Hair Capture and Animation	8641
<i>Ziyan Wang (Carnegie Mellon University; Meta Reality Labs), Giljoo Nam (Meta Reality Labs), Tuur Stuyck (Meta Reality Labs), Stephen Lombardi (Meta Reality Labs), Chen Cao (Meta Reality Labs), Jason Saragih (Meta Reality Labs), Michael Zollhöfer (Meta Reality Labs), Jessica Hodgins (Carnegie Mellon University), and Christoph Lassner (Meta Reality Labs)</i>	
SadTalker: Learning Realistic 3D Motion Coefficients for Stylized Audio-Driven Single Image Talking Face Animation	8652
<i>Wenxuan Zhang (Xi'an Jiaotong University; National Key Laboratory of Human-Machine Hybrid Augmented Intelligence), Xiaodong Cun (Tencent AI Lab), Xuan Wang (Ant Group), Yong Zhang (Tencent AI Lab), Xi Shen (Tencent AI Lab), Yu Guo (National Key Laboratory of Human-Machine Hybrid Augmented Intelligence; IAIR, Xi'an Jiaotong University), Ying Shan (Tencent AI Lab), and Fei Wang (National Key Laboratory of Human-Machine Hybrid Augmented Intelligence; IAIR, Xi'an Jiaotong University)</i>	
High-Fidelity Clothed Avatar Reconstruction From a Single Image	8662
<i>Tingting Liao (University of Chinese Academy of Sciences, China; MAIS, Institute of Automation, Chinese Academy of Sciences, China), Xiaomei Zhang (University of Chinese Academy of Sciences, China; MAIS, Institute of Automation, Chinese Academy of Sciences, China), Yuliang Xiu (Max Planck Institute for Intelligent Systems, Germany), Hongwei Yi (Max Planck Institute for Intelligent Systems, Germany), Xudong Liu (OPPO Research), Guo-Jun Qi (OPPO Research; Westlake University), Yong Zhang (Tencent AI Lab), Xuan Wang (Tencent AI Lab), Xiangyu Zhu (University of Chinese Academy of Sciences, China; MAIS, Institute of Automation, Chinese Academy of Sciences, China), and Zhen Lei (University of Chinese Academy of Sciences, China; MAIS, Institute of Automation, Chinese Academy of Sciences, China; CAIR, HKISI, CAS)</i>	

Music-Driven Group Choreography	8673
<i>Nhat Le (AIOZ, Singapore), Thang Pham (AIOZ, Singapore), Tuong Do (AIOZ, Singapore), Erman Tjiputra (AIOZ, Singapore), Quang D. Tran (AIOZ, Singapore), and Anh Nguyen (University of Liverpool, UK)</i>	
Hand Avatar: Free-Pose Hand Animation and Rendering From Monocular Video	8683
<i>Xingyu Chen (Xiaobing.AI), Baoyuan Wang (Xiaobing.AI), and Heung-Yeung Shum (Xiaobing.AI)</i>	
Biomechanics-Guided Facial Action Unit Detection Through Force Modeling	8694
<i>Zijun Cui (Rensselaer Polytechnic Institute), Chenyi Kuang (Rensselaer Polytechnic Institute), Tian Gao (IBM Research), Kartik Talamadupula (IBM Research), and Qiang Ji (Rensselaer Polytechnic Institute)</i>	
Zero-Shot Pose Transfer for Unrigged Stylized 3D Characters	8704
<i>Jiashun Wang (Carnegie Mellon University), Xueting Li (NVIDIA), Sifei Liu (NVIDIA), Shalini De Mello (NVIDIA), Orazio Gallo (NVIDIA), Xiaolong Wang (UC San Diego), and Jan Kautz (NVIDIA)</i>	
Invertible Neural Skinning	8715
<i>Yash Kant (University of Toronto; Snap Research), Aliaksandr Siarohin (Snap Research), Riza Alp Guler (Snap Research), Menglei Chai (Snap Research), Jian Ren (Snap Research), Sergey Tulyakov (Snap Research), and Igor Gilitschenski (University of Toronto)</i>	
BEDLAM: A Synthetic Dataset of Bodies Exhibiting Detailed Lifelike Animated Motion	8726
<i>Michael J. Black (Max Planck Institute for Intelligent Systems, Germany), Priyanka Patel (Max Planck Institute for Intelligent Systems, Germany), Joachim Tesch (Max Planck Institute for Intelligent Systems, Germany), and Jinlong Yang (Google)</i>	
DiFu: Depth-Guided Implicit Function for Clothed Human Reconstruction	8738
<i>Dae-Young Song (Electronics and Telecommunications Research Institute, South Korea; Chungnam National University, South Korea), HeeKyung Lee (Electronics and Telecommunications Research Institute, South Korea), Jeongil Seo (Electronics and Telecommunications Research Institute, South Korea), and Donghyeon Cho (Chungnam National University, South Korea)</i>	
Complete 3D Human Reconstruction From a Single Incomplete Image	8748
<i>Junying Wang (University of Southern California), Jae Shin Yoon (Adobe Research), Tuanfeng Y. Wang (Adobe Research), Krishna Kumar Singh (Adobe Research), and Ulrich Neumann (University of Southern California)</i>	
Learning Neural Volumetric Representations of Dynamic Humans in Minutes	8759
<i>Chen Geng (Zhejiang University), Sida Peng (Zhejiang University), Zhen Xu (Zhejiang University), Hujun Bao (Zhejiang University), and Xiaowei Zhou (Zhejiang University)</i>	
Marching-Primitives: Shape Abstraction From Signed Distance Function	8771
<i>Weixiao Liu (National University of Singapore; Johns Hopkins University), Yuwei Wu (National University of Singapore), Sipu Ruan (National University of Singapore), and Gregory S. Chirikjian (National University of Singapore)</i>	

Learning Analytical Posterior Probability for Human Mesh Recovery	8781
<i>Qi Fang (NetEase Games AI Lab), Kang Chen (NetEase Games AI Lab), Yinghui Fan (NetEase Games AI Lab), Qing Shuai (Zhejiang University), Jiefeng Li (Shanghai Jiao Tong University), and Weidong Zhang (NetEase Games AI Lab)</i>	
MagicPony: Learning Articulated 3D Animals in the Wild	8792
<i>Shangzhe Wu (University of Oxford), Ruining Li (University of Oxford), Tomas Jakab (University of Oxford), Christian Rupprecht (University of Oxford), and Andrea Vedaldi (University of Oxford)</i>	
Visual-Tactile Sensing for In-Hand Object Reconstruction	8803
<i>Wenqiang Xu (Shanghai Jiao Tong University; Shanghai Qi Zhi institute), Zhenjun Yu (Shanghai Jiao Tong University; Shanghai Qi Zhi institute), Han Xue (Shanghai Jiao Tong University), Ruolin Ye (Cornell University), Siqiong Yao (Shanghai Jiao Tong University), and Cewu Lu (Shanghai Jiao Tong University; Shanghai Qi Zhi institute)</i>	
Command-Driven Articulated Object Understanding and Manipulation	8813
<i>Ruihang Chu (CUHK), Zhengzhe Liu (CUHK), Xiaoqing Ye (Baidu Inc.), Xiao Tan (Baidu Inc.), Xiaojuan Qi (HKU), Chi-Wing Fu (CUHK), and Jiaya Jia (CUHK; SmartMore)</i>	
Target-Referenced Reactive Grasping for Dynamic Objects	8824
<i>Jirong Liu (Shanghai Qi Zhi institute; Shanghai Jiao Tong University), Ruo Zhang (Shanghai Jiao Tong University), Hao-Shu Fang (Shanghai Jiao Tong University), Minghao Gou (Shanghai Jiao Tong University), Hongjie Fang (Shanghai Jiao Tong University), Chenxi Wang (Shanghai Jiao Tong University; Flexiv Robotics, LTD), Sheng Xu (Shanghai Jiao Tong University), Hengxu Yan (Shanghai Qi Zhi institute; Shanghai Qi Zhi institute), and Cewu Lu (Shanghai Jiao Tong University)</i>	
NeuralDome: A Neural Modeling Pipeline on Multi-View Human-Object Interactions	8834
<i>Juze Zhang (ShanghaiTech University; Shanghai Advanced Research Institute, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Shanghai Engineering Research Center of Intelligent Vision and Imaging), Haimin Luo (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging; LumiAni Technology), Hongdi Yang (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging), Xinru Xu (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging), Qianyang Wu (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging), Ye Shi (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging), Jingyi Yu (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging), Lan Xu (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging), and Jingya Wang (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging)</i>	

- A2J-Transformer: Anchor-to-Joint Transformer Network for 3D Interacting Hand Pose Estimation From a Single RGB Image 8846
*Changlong Jiang (Huazhong University of Science and Technology, Wuhan 430074, China), Yang Xiao (Huazhong University of Science and Technology, Wuhan 430074, China), Cunlin Wu (Huazhong University of Science and Technology, Wuhan 430074, China), Mingyang Zhang (Alibaba Group), Jinghong Zheng (Huazhong University of Science and Technology, Wuhan 430074, China), Zhiguo Cao (Huazhong University of Science and Technology, Wuhan 430074, China), and Joey Tianyi Zhou (Technology and Research (A*STAR), Singapore; Technology and Research (A*STAR), Singapore)*
- TRACE: 5D Temporal Regression of Avatars With Dynamic Cameras in 3D Environments 8856
Yu Sun (Harbin Institute of Technology), Qian Bao (Explore Academy of JD.com), Wu Liu (Explore Academy of JD.com), Tao Mei (HiDream.ai Inc.), and Michael J. Black (Max Planck Institute for Intelligent Systems)
- BITE: Beyond Priors for Improved Three-D Dog Pose Estimation 8867
Nadine Rüegg (ETH Zürich, Switzerland; Max Planck Institute for Intelligent Systems, Germany), Shashank Tripathi (Max Planck Institute for Intelligent Systems, Germany), Konrad Schindler (ETH Zürich, Switzerland), Michael J. Black (Max Planck Institute for Intelligent Systems, Germany), and Silvia Zuffi (IMATI-CNR, Italy)
- PoseFormerV2: Exploring Frequency Domain for Efficient and Robust 3D Human Pose Estimation. 8877
Qitao Zhao (Shandong University), Ce Zheng (Center for Research in Computer Vision, University of Central Florida), Mengyuan Liu (Key Laboratory of Machine Perception, Peking University), Pichao Wang (Amazon Prime Video), and Chen Chen (Center for Research in Computer Vision, University of Central Florida)
- Global-to-Local Modeling for Video-Based 3D Human Pose and Shape Estimation 8887
Xiaolong Shen (Zhejiang University; DAMO Academy, Alibaba Group), Zongxin Yang (Zhejiang University), Xiaohan Wang (Zhejiang University), Jianxin Ma (DAMO Academy, Alibaba Group), Chang Zhou (DAMO Academy, Alibaba Group), and Yi Yang (Zhejiang University)
- TokenHPE: Learning Orientation Tokens for Efficient Head Pose Estimation via Transformers 8897
Cheng Zhang (Central China Normal University, China), Hai Liu (Central China Normal University, China), Yongjian Deng (Beijing University of Technology, China; Engineering Research Center of Intelligence Perception and Autonomous Control, Ministry of Education, China), Bochen Xie (City University of Hong Kong, China), and Youfu Li (City University of Hong Kong, China)
- GFIE: A Dataset and Baseline for Gaze-Following From 2D to 3D in Indoor Environments 8907
Zhengxi Hu (IRAIS, Nankai University; tjKLIR, Nankai University; TBI center, Nankai University), Yuxue Yang (IRAIS, Nankai University), Xiaolin Zhai (IRAIS, Nankai University; tjKLIR, Nankai University; TBI center, Nankai University), Dingye Yang (IRAIS, Nankai University; tjKLIR, Nankai University; TBI center, Nankai University), Bohan Zhou (IRAIS, Nankai University), and Jingtai Liu (IRAIS, Nankai University; tjKLIR, Nankai University; TBI center, Nankai University)

Robot Structure Prior Guided Temporal Attention for Camera-to-Robot Pose Estimation From Image Sequence	8917
<i>Yang Tian (CFCS, Peking University), Jiyao Zhang (CFCS, Peking University), Zekai Yin (CFCS, Peking University), and Hao Dong (CFCS, Peking University)</i>	
Rigidity-Aware Detection for 6D Object Pose Estimation	8927
<i>Yang Hai (Xidian University), Rui Song (Xidian University), Jiaojiao Li (Xidian University), Mathieu Salzmann (EPFL; ClearSpace), and Yinlin Hu (MagicLeap)</i>	
Crowd3D: Towards Hundreds of People Reconstruction From a Single Image	8937
<i>Hao Wen (Tianjin University, China), Jing Huang (Tianjin University, China), Huili Cui (Tianjin University, China), Haozhe Lin (Tsinghua University, China), Yu-Kun Lai (Cardiff University, United Kingdom), Lu Fang (Tsinghua University, China), and Kun Li (Tianjin University, China)</i>	
Object Pose Estimation With Statistical Guarantees: Conformal Keypoint Detection and Geometric Uncertainty Propagation	8947
<i>Heng Yang (NVIDIA Research) and Marco Pavone (NVIDIA Research)</i>	
expoSE: Accurate Initialization-Free Projective Factorization Using Exponential Regularization	8959
<i>José Pedro Iglesias (Chalmers University of Technology, Sweden), Amanda Nilsson (Lund University, Sweden), and Carl Olsson (Chalmers University of Technology, Sweden; Lund University, Sweden)</i>	
Neural Voting Field for Camera-Space 3D Hand Pose Estimation	8969
<i>Lin Huang (University at Buffalo), Chung-Ching Lin (Microsoft), Kevin Lin (Microsoft), Lin Liang (Microsoft), Lijuan Wang (Microsoft), Junsong Yuan (University at Buffalo), and Zicheng Liu (Microsoft)</i>	
Two-View Geometry Scoring Without Correspondences	8979
<i>Axel Barroso-Laguna (Niantic), Eric Brachmann (Niantic), Victor Adrian Prisacariu (Niantic; University of Oxford), Gabriel J. Brostow (Niantic; University College London), and Daniyar Turmukhambetov (Niantic)</i>	
Four-View Geometry With Unknown Radial Distortion	8990
<i>Petr Hruby (ETH Zürich), Viktor Korotynskiy (CIIRC, CTU in Prague), Timothy Duff (University of Washington), Luke Oeding (Auburn University), Marc Pollefeys (ETH Zürich), Tomas Pajdla (CIIRC, CTU in Prague), and Viktor Larsson (ETH Zürich)</i>	
BKinD-3D: Self-Supervised 3D Keypoint Discovery From Multi-View Videos	9001
<i>Jennifer J. Sun (Caltech), Lili Karashchuk (U Washington), Amil David (Northwestern), Serim Ryou (SAIT), Sonia Fereidooni (U Washington), John C. Tuthill (U Washington), Aggelos Katsaggelos (Northwestern), Bingni W. Brunton (U Washington), Georgia Gkioxari (Caltech), Ann Kennedy (Northwestern), Yisong Yue (Caltech), and Pietro Perona (Caltech)</i>	

BAAM: Monocular 3D Pose and Shape Reconstruction With Bi-Contextual Attention Module and Attention-Guided Modeling	9011
<i>Hyo-Jun Lee (Chungnam National University), Hanul Kim (Seoul National University of Science and Technology), Su-Min Choi (42dot Inc.), Seong-Gyun Jeong (42dot Inc.), and Yeong Jun Koh (Chungnam National University)</i>	
Multi-Object Manipulation via Object-Centric Neural Scattering Functions	9021
<i>Stephen Tian (Stanford University), Yancheng Cai (Stanford University), Hong-Xing Yu (Stanford University), Sergey Zakharov (Toyota Research Institute), Katherine Liu (Toyota Research Institute), Adrien Gaidon (Toyota Research Institute), Yunzhu Li (Stanford University), and Jiajun Wu (Stanford University)</i>	
Neural Part Priors: Learning To Optimize Part-Based Object Completion in RGB-D Scans	9032
<i>Aleksei Bokhovkin (Technical University of Munich) and Angela Dai (Technical University of Munich)</i>	
Panoptic Lifting for 3D Scene Understanding With Neural Fields	9043
<i>Yawar Siddiqui (Technical University of Munich; Meta Reality Labs Zurich), Lorenzo Porzi (Meta Reality Labs Zurich), Samuel Rota Bulò (Meta Reality Labs Zurich), Norman Müller (Technical University of Munich; Meta Reality Labs Zurich), Matthias Nießner (Technical University of Munich), Angela Dai (Technical University of Munich), and Peter Kotschieder (Meta Reality Labs Zurich)</i>	
Virtual Occlusions Through Implicit Depth	9053
<i>Jamie Watson (Niantic; UCL), Mohamed Sayed (Niantic; UCL), Zawar Qureshi (Niantic), Gabriel J. Brostow (Niantic; UCL), Sara Vicente (Niantic), Oisín Mac Aodha (University of Edinburgh), and Michael Firman (Niantic)</i>	
Multiview Compressive Coding for 3D Reconstruction	9065
<i>Chao-Yuan Wu (FAIR, Meta AI), Justin Johnson (FAIR, Meta AI), Jitendra Malik (FAIR, Meta AI), Christoph Feichtenhofer (FAIR, Meta AI), and Georgia Gkioxari (FAIR, Meta AI)</i>	
Behind the Scenes: Density Fields for Single View Reconstruction	9076
<i>Felix Wimbauer (Technical University of Munich; MCML), Nan Yang (Technical University of Munich), Christian Rupprecht (University of Oxford), and Daniel Cremers (Technical University of Munich; MCML; University of Oxford)</i>	
VoxFormer: Sparse Voxel Transformer for Camera-Based 3D Semantic Scene Completion	9087
<i>Yiming Li (NYU), Zhiding Yu (NVIDIA), Christopher Choy (NVIDIA), Chaowei Xiao (NVIDIA; ASU), Jose M. Alvarez (NVIDIA), Sanja Fidler (NVIDIA; University of Toronto; Vector Institute), Chen Feng (NYU), and Anima Anandkumar (NVIDIA; Caltech)</i>	
Renderable Neural Radiance Map for Visual Navigation	9099
<i>Obin Kwon (ASRI, Seoul National University), Jeongho Park (ASRI, Seoul National University), and Songhwai Oh (ASRI, Seoul National University)</i>	
Learning To Detect Mirrors From Videos via Dual Correspondences	9109
<i>Jiaying Lin (City University of Hong Kong), Xin Tan (East China Normal University), and Rynson W.H. Lau (City University of Hong Kong)</i>	

Temporally Consistent Online Depth Estimation Using Point-Based Fusion	9119
<i>Numair Khan (NA), Eric Penner (NA), Douglas Lanman (NA), and Lei Xiao (NA)</i>	
Zero-Shot Dual-Lens Super-Resolution	9130
<i>Ruikang Xu (University of Science and Technology of China), Mingde Yao (University of Science and Technology of China), and Zhiwei Xiong (University of Science and Technology of China)</i>	
Fully Self-Supervised Depth Estimation From Defocus Clue	9140
<i>Haozhe Si (Shanghai AI Laboratory; University of Illinois Urbana-Champaign), Bin Zhao (Shanghai AI Laboratory; Northwestern Polytechnical University), Dong Wang (Shanghai AI Laboratory), Yunpeng Gao (Northwestern Polytechnical University), Mulin Chen (Shanghai AI Laboratory; Northwestern Polytechnical University), Zhigang Wang (Shanghai AI Laboratory), and Xuelong Li (Shanghai AI Laboratory; Northwestern Polytechnical University)</i>	
MVImgNet: A Large-Scale Dataset of Multi-View Images	9150
<i>Xianggang Yu (FNii, CUHKSZ; SSE, CUHKSZ), Mutian Xu (SSE, CUHKSZ; FNii, CUHKSZ), Yidan Zhang (SSE, CUHKSZ; FNii, CUHKSZ), Haolin Liu (FNii, CUHKSZ; SSE, CUHKSZ), Chongjie Ye (SSE, CUHKSZ; FNii, CUHKSZ), Yushuang Wu (FNii, CUHKSZ; SSE, CUHKSZ), Zizheng Yan (FNii, CUHKSZ; SSE, CUHKSZ), Chenming Zhu (FNii, CUHKSZ; SSE, CUHKSZ), Zhangyang Xiong (FNii, CUHKSZ; SSE, CUHKSZ), Tianyou Liang (FNii, CUHKSZ; SSE, CUHKSZ), Guanying Chen (FNii, CUHKSZ; SSE, CUHKSZ), Shuguang Cui (FNii, CUHKSZ; SSE, CUHKSZ), and Xiaoguang Han (FNii, CUHKSZ; SSE, CUHKSZ)</i>	
Revisiting the Stack-Based Inverse Tone Mapping	9162
<i>Ning Zhang (Peking University; Peng Cheng Laboratory), Yuyao Ye (Peking University; Peng Cheng Laboratory), Yang Zhao (Hefei University of Technology; Peng Cheng Laboratory), and Ronggang Wang (Peking University; Peng Cheng Laboratory)</i>	
Combining Implicit-Explicit View Correlation for Light Field Semantic Segmentation	9172
<i>Ruixuan Cong (Beihang University; Beihang Hangzhou Innovation Institute Yuhang; Macao Polytechnic University), Da Yang (Beihang University; Beihang Hangzhou Innovation Institute Yuhang; Macao Polytechnic University), Rongshan Chen (Beihang University; Beihang Hangzhou Innovation Institute Yuhang; Macao Polytechnic University), Sizhe Wang (Beihang University; Beihang Hangzhou Innovation Institute Yuhang; Macao Polytechnic University), Zhenglong Cui (Beihang University; Beihang Hangzhou Innovation Institute Yuhang; Macao Polytechnic University), and Hao Sheng (Beihang University; Beihang Hangzhou Innovation Institute Yuhang; Macao Polytechnic University)</i>	
3D Spatial Multimodal Knowledge Accumulation for Scene Graph Prediction in Point Cloud ..	9182
<i>Mingtao Feng (Xidian University), Haoran Hou (Xidian University), Liang Zhang (Xidian University), Zijie Wu (Hunan University), Yulan Guo (Sun Yat-Sen University), and Ajmal Mian (The University of Western Australia)</i>	

Role of Transients in Two-Bounce Non-Line-of-Sight Imaging	9192
<i>Siddharth Somasundaram (Massachusetts Institute of Technology), Akshat Dave (Rice University), Connor Henley (Massachusetts Institute of Technology), Ashok Veeraraghavan (Rice University), and Ramesh Raskar (Massachusetts Institute of Technology)</i>	
3D Concept Learning and Reasoning From Multi-View Images	9202
<i>Yining Hong (UCLA), Chunru Lin (Shanghai Jiaotong University), Yilun Du (MIT CSAIL), Zhenfang Chen (MIT-IBM Watson AI Lab), Joshua B. Tenenbaum (MIT CSAIL), and Chuang Gan (UMass Amherst; MIT-IBM Watson AI Lab)</i>	
Viewpoint Equivariance for Multi-View 3D Object Detection	9213
<i>Dian Chen (Toyota Research Institute (TRI), Los Altos), Jie Li (Toyota Research Institute (TRI), Los Altos), Vitor Guizilini (Toyota Research Institute (TRI), Los Altos), Rares Andrei Ambrus (Toyota Research Institute (TRI), Los Altos), and Adrien Gaidon (Toyota Research Institute (TRI), Los Altos)</i>	
Tri-Perspective View for Vision-Based 3D Semantic Occupancy Prediction	9223
<i>Yuanhui Huang (Beijing National Research Center for Information Science and Technology, China; Department of Automation, Tsinghua University, China), Wenzhao Zheng (Beijing National Research Center for Information Science and Technology, China; Department of Automation, Tsinghua University, China), Yunpeng Zhang (Beijing National Research Center for Information Science and Technology, China; Department of Automation, Tsinghua University, China), Jie Zhou (Beijing National Research Center for Information Science and Technology, China; Department of Automation, Tsinghua University, China), and Jiwen Lu (Beijing National Research Center for Information Science and Technology, China; Department of Automation, Tsinghua University, China)</i>	
BEV@DC: Bird's-Eye View Assisted Training for Depth Completion	9233
<i>Wending Zhou (FNii, CUHK-Shenzhen; SSE, CUHK-Shenzhen), Xu Yan (FNii, CUHK-Shenzhen; SSE, CUHK-Shenzhen), Yinghong Liao (FNii, CUHK-Shenzhen; SSE, CUHK-Shenzhen), Yuankai Lin (Huazhong University of Science and Technology), Jin Huang (Cardiff University), Gangming Zhao (The University of Hong Kong), Shuguang Cui (FNii, CUHK-Shenzhen; SSE, CUHK-Shenzhen), and Zhen Li (FNii, CUHK-Shenzhen; SSE, CUHK-Shenzhen)</i>	
Collaboration Helps Camera Overtake LiDAR in 3D Detection	9243
<i>Yue Hu (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University), Yifan Lu (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University), Runsheng Xu (University of California, Los Angeles), Weidi Xie (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University; Shanghai AI Laboratory), Siheng Chen (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University; Shanghai AI Laboratory), and Yanfeng Wang (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University; Shanghai AI Laboratory)</i>	

Uni3D: A Unified Baseline for Multi-Dataset 3D Object Detection	9253
<i>Bo Zhang (Shanghai AI Laboratory), Jiakang Yuan (School of Information Science and Technology, Fudern University), Botian Shi (Shanghai AI Laboratory), Tao Chen (School of Information Science and Technology, Fudern University), Yikang Li (Shanghai AI Laboratory), and Yu Qiao (Shanghai AI Laboratory)</i>	
Towards Building Self-Aware Object Detectors via Reliable Uncertainty Quantification and Calibration	9263
<i>Kemal Oksuz (Five AI Ltd., United Kingdom), Tom Joy (Five AI Ltd., United Kingdom), and Puneet K. Dokania (Five AI Ltd., United Kingdom)</i>	
Depth Estimation From Camera Image and mmWave Radar Point Cloud	9275
<i>Akash Deep Singh (University of California, Los Angeles), Yunhao Ba (University of California, Los Angeles), Ankur Sarker (University of California, Los Angeles), Howard Zhang (University of California, Los Angeles), Achuta Kadambi (University of California, Los Angeles), Stefano Soatto (University of California, Los Angeles), Mani Srivastava (University of California, Los Angeles), and Alex Wong (Yale University)</i>	
SGLoc: Scene Geometry Encoding for Outdoor LiDAR Localization	9286
<i>Wen Li (Xiamen University, China), Shangshu Yu (Xiamen University, China), Cheng Wang (Xiamen University, China), Guosheng Hu (Oosto, UK), Siqu Shen (Xiamen University, China), and Chenglu Wen (Xiamen University, China)</i>	
ConQueR: Query Contrast Voxel-DETR for 3D Object Detection	9296
<i>Benjin Zhu (The Chinese University of Hong Kong), Zhe Wang (The Chinese University of Hong Kong), Shaoshuai Shi (Max Planck Institute for Informatics), Hang Xu (Huawei Noah's Ark Lab), Lanqing Hong (Huawei Noah's Ark Lab), and Hongsheng Li (The Chinese University of Hong Kong; Centre for Perceptual and Interactive Intelligence)</i>	
DeepMapping2: Self-Supervised Large-Scale LiDAR Map Optimization	9306
<i>Chao Chen (New York University), Xinhao Liu (New York University), Yiming Li (New York University), Li Ding (University of Rochester), and Chen Feng (New York University)</i>	
Towards Unsupervised Object Detection From LiDAR Point Clouds	9317
<i>Lunjun Zhang (University of Toronto), Anqi Joyce Yang (University of Toronto), Yuwen Xiong (University of Toronto), Sergio Casas (University of Toronto), Bin Yang (University of Toronto), Mengye Ren (University of Toronto), and Raquel Urtasun (University of Toronto)</i>	
MoDAR: Using Motion Forecasting for 3D Object Detection in Point Cloud Sequences	9329
<i>Yingwei Li (Waymo LLC), Charles R. Qi (Waymo LLC), Yin Zhou (Waymo LLC), Chenxi Liu (Waymo LLC), and Dragomir Anguelov (Waymo LLC)</i>	
Hidden Gems: 4D Radar Scene Flow Learning Using Cross-Modal Supervision	9340
<i>Fangqiang Ding (University of Edinburgh), Andras Palffy (Delft University of Technology), Dariu M. Gavrilă (Delft University of Technology), and Chris Xiaoxuan Lu (University of Edinburgh)</i>	
Instant Domain Augmentation for LiDAR Semantic Segmentation	9350
<i>Kwonyoung Ryu (POSTECH), Soonmin Hwang (Carnegie Mellon University), and Jaesik Park (POSTECH)</i>	

Less Is More: Reducing Task and Model Complexity for 3D Point Cloud Semantic Segmentation	9361
<i>Li Li (Durham University, UK), Hubert P. H. Shum (Durham University, UK), and Toby P. Breckon (Durham University, UK)</i>	
MarS3D: A Plug-and-Play Motion-Aware Model for Semantic Segmentation on Multi-Scan 3D Point Clouds	9372
<i>Jiahui Liu (The University of Hong Kong), Chirui Chang (The University of Hong Kong), Jianhui Liu (The University of Hong Kong), Xiaoyang Wu (The University of Hong Kong), Lan Ma (TCL AI Lab), and Xiaojuan Qi (The University of Hong Kong)</i>	
3D Semantic Segmentation in the Wild: Learning Generalized Models for Adverse-Condition Point Clouds	9382
<i>Aoran Xiao (Nanyang Technological University), Jiaying Huang (Nanyang Technological University), Weihao Xuan (Waseda University), Ruijie Ren (Technical University of Denmark), Kangcheng Liu (Nanyang Technological University), Dayan Guan (Mohamed bin Zayed University of Artificial Intelligence), Abdulmotaleb El Saddik (Mohamed bin Zayed University of Artificial Intelligence; University of Ottawa), Shijian Lu (Nanyang Technological University), and Eric P. Xing (Mohamed bin Zayed University of Artificial Intelligence; Carnegie Mellon University)</i>	
Novel Class Discovery for 3D Point Cloud Semantic Segmentation	9393
<i>Luigi Riz (Fondazione Bruno Kessler), Cristiano Saltori (University of Trento), Elisa Ricci (Fondazione Bruno Kessler; University of Trento), and Fabio Poiesi (Fondazione Bruno Kessler)</i>	
GD-MAE: Generative Decoder for MAE Pre-Training on LiDAR Point Clouds	9403
<i>Honghui Yang (Zhejiang University; Shanghai AI Laboratory), Tong He (Shanghai AI Laboratory), Jiaheng Liu (Shanghai AI Laboratory; Beihang University), Hua Chen (COMAC Beijing Aircraft Technology Research Institute), Boxi Wu (Zhejiang University), Binbin Lin (Zhejiang University), Xiaofei He (Zhejiang University), and Wanli Ouyang (Shanghai AI Laboratory)</i>	
Masked Scene Contrast: A Scalable Framework for Unsupervised 3D Representation Learning	9415
<i>Xiaoyang Wu (The University of Hong Kong), Xin Wen (The University of Hong Kong), Xihui Liu (The University of Hong Kong), and Hengshuang Zhao (The University of Hong Kong)</i>	
Open-Set Semantic Segmentation for Point Clouds via Adversarial Prototype Framework	9425
<i>Jianan Li (University of Chinese Academy of Sciences, China; Chinese Academy of Sciences, Beijing, China) and Qiulei Dong (University of Chinese Academy of Sciences, China; Chinese Academy of Sciences, China; Chinese Academy of Sciences, Beijing, China)</i>	
ACL-SPC: Adaptive Closed-Loop System for Self-Supervised Point Cloud Completion	9435
<i>Sangmin Hong (IPAI, Korea), Mohsen Yavartanoo (Seoul National University, Korea), Reyhaneh Neshatavar (Seoul National University, Korea), and Kyoung Mu Lee (IPAI, Korea; Seoul National University, Korea)</i>	

Fast Point Cloud Generation With Straight Flows	9445
<i>Lemeng Wu (University of Texas at Austin), Dilin Wang (Meta), Chengyue Gong (University of Texas at Austin), Xingchao Liu (University of Texas at Austin), Yunyang Xiong (Meta), Rakesh Ranjan (Meta), Raghuraman Krishnamoorthi (Meta), Vikas Chandra (Meta), and Qiang Liu (University of Texas at Austin)</i>	
PointVector: A Vector Representation in Point Cloud Analysis	9455
<i>Xin Deng (University of Science and Technology of China), WenYu Zhang (University of Science and Technology of China), Qing Ding (University of Science and Technology of China), and XinMing Zhang (University of Science and Technology of China)</i>	
ProxyFormer: Proxy Alignment Assisted Point Cloud Completion With Missing Part Sensitive Transformer	9466
<i>Shanshan Li (Nanjing University of Aeronautics and Astronautics), Pan Gao (Nanjing University of Aeronautics and Astronautics), Xiaoyang Tan (Nanjing University of Aeronautics and Astronautics), and Mingqiang Wei (Nanjing University of Aeronautics and Astronautics)</i>	
FAC: 3D Representation Learning via Foreground Aware Feature Contrast	9476
<i>Kangcheng Liu (Nanyang Technological University), Aoran Xiao (Nanyang Technological University), Xiaoqin Zhang (Wenzhou University), Shijian Lu (Nanyang Technological University), and Ling Shao (UCAS-Terminus AI Lab, UCAS)</i>	
Rethinking the Approximation Error in 3D Surface Fitting for Point Cloud Normal Estimation.	9486
<i>Hang Du (Hikvision Research Institute, Hangzhou, China), Xuejun Yan (Hikvision Research Institute, Hangzhou, China), Jingjing Wang (Hikvision Research Institute, Hangzhou, China), Di Xie (Hikvision Research Institute, Hangzhou, China), and Shiliang Pu (Hikvision Research Institute, Hangzhou, China)</i>	
PointCert: Point Cloud Classification With Deterministic Certified Robustness Guarantees	9496
<i>Jinghuai Zhang (Duke University), Jinyuan Jia (UIUC), Hongbin Liu (Duke University), and Neil Zhenqiang Gong (Duke University)</i>	
Robust Multiview Point Cloud Registration With Reliable Pose Graph Initialization and History Reweighting	9506
<i>Haiping Wang (Wuhan University The University of Hong Kong Wuhan University), Yuan Liu (Wuhan University The University of Hong Kong Wuhan University), Zhen Dong (Wuhan University The University of Hong Kong Wuhan University), Yulan Guo (Sun Yat-sen University Tsinghua University), Yu-Shen Liu (Sun Yat-sen University Tsinghua University), Wenping Wang (Texas A&M University Wuhan University), and Bisheng Yang (Texas A&M University Wuhan University)</i>	
Visual Prompt Multi-Modal Tracking	9516
<i>Jiawen Zhu (Dalian University of Technology, China), Simiao Lai (Dalian University of Technology, China), Xin Chen (Dalian University of Technology, China), Dong Wang (Dalian University of Technology, China), and Huchuan Lu (Dalian University of Technology, China; Peng Cheng Laboratory, China)</i>	
Progressive Neighbor Consistency Mining for Correspondence Pruning	9527
<i>Xin Liu (Nankai University, China) and Jufeng Yang (Nankai University, China)</i>	

Geometric Visual Similarity Learning in 3D Medical Image Self-Supervised Pre-Training	9538
<i>Yuting He (Southeast University), Guanyu Yang (Southeast University), Rongjun Ge (Nanjing University of Aeronautics and Astronautics), Yang Chen (Southeast University), Jean-Louis Coatrieux (University of Rennes), Boyu Wang (Western University), and Shuo Li (Case Western Reserve University)</i>	
Unsupervised Visible-Infrared Person Re-Identification via Progressive Graph Matching and Alternate Learning	9548
<i>Zesen Wu (Wuhan University, China) and Mang Ye (Wuhan University, China; Hubei LuoJia Laboratory, China)</i>	
Domain Generalized Stereo Matching via Hierarchical Visual Transformation	9559
<i>Tianyu Chang (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Xun Yang (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China), and Meng Wang (Hefei University of Technology)</i>	
Unsupervised Cumulative Domain Adaptation for Foggy Scene Optical Flow	9569
<i>Hanyu Zhou (Huazhong University of Science and Technology), Yi Chang (Huazhong University of Science and Technology), Wending Yan (Huawei International Co. Ltd.), and Luxin Yan (Huazhong University of Science and Technology)</i>	
PVO: Panoptic Visual Odometry	9579
<i>Weicai Ye (Zhejiang University; ZJU-SenseTime Joint Lab of 3D Vision), Xinyue Lan (ZJU-SenseTime Joint Lab of 3D Vision), Shuo Chen (ZJU-SenseTime Joint Lab of 3D Vision), Yuhang Ming (Hangzhou Dianzi University; University of Bristol), Xingyuan Yu (Zhejiang University; ZJU-SenseTime Joint Lab of 3D Vision), Hujun Bao (Zhejiang University; ZJU-SenseTime Joint Lab of 3D Vision), Zhaopeng Cui (Zhejiang University), and Guofeng Zhang (Zhejiang University; ZJU-SenseTime Joint Lab of 3D Vision)</i>	
BAEFormer: Bi-Directional and Early Interaction Transformers for Bird's Eye View Semantic Segmentation	9590
<i>Cong Pan (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Yonghao He (Horizon Robotics), Junran Peng (Huawei Inc.), Qian Zhang (Horizon Robotics), Wei Sui (Horizon Robotics), and Zhaoxiang Zhang (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Center for Artificial Intelligence and Robotics, HKISLCAS)</i>	
Are We Ready for Vision-Centric Driving Streaming Perception? The ASAP Benchmark	9600
<i>Xiaofeng Wang (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Zheng Zhu (PhiGent Robotics), Yunpeng Zhang (PhiGent Robotics), Guan Huang (PhiGent Robotics), Yun Ye (PhiGent Robotics), Wenbo Xu (PhiGent Robotics), Ziwei Chen (Southeast University), and Xingang Wang (Institute of Automation, Chinese Academy of Sciences)</i>	

Visual Exemplar Driven Task-Prompting for Unified Perception in Autonomous Driving	9611
<i>Xiwen Liang (Shenzhen Campus of Sun Yat-sen University), Minzhe Niu (Huawei Noah's Ark Lab), Jianhua Han (Huawei Noah's Ark Lab), Hang Xu (Huawei Noah's Ark Lab), Chunjing Xu (Huawei Noah's Ark Lab), and Xiaodan Liang (Shenzhen Campus of Sun Yat-sen University)</i>	
MixSim: A Hierarchical Framework for Mixed Reality Traffic Simulation	9622
<i>Simon Suo (Waabi; University of Toronto), Kelvin Wong (Waabi; University of Toronto), Justin Xu (University of Waterloo), James Tu (Waabi; University of Toronto), Alexander Cui (Waabi; University of Toronto), Sergio Casas (Waabi; University of Toronto), and Raquel Urtasun (Waabi; University of Toronto)</i>	
Uncovering the Missing Pattern: Unified Framework Towards Trajectory Imputation and Prediction	9632
<i>Yi Xu (Northeastern University; Honda Research Institute, USA), Armin Bazarjani (Honda Research Institute, USA; University of Southern California), Hyung-gun Chi (Honda Research Institute, USA; Purdue University), Chiho Choi (Honda Research Institute, USA; Samsung Semiconductor US), and Yun Fu (Northeastern University)</i>	
MotionDiffuser: Controllable Multi-Agent Motion Prediction Using Diffusion	9644
<i>Chiyu Max Jiang (Waymo LLC), Andre Cornman (Waymo LLC), Cheolho Park (Waymo LLC), Benjamin Sapp (Waymo LLC), Yin Zhou (Waymo LLC), and Dragomir Anguelov (Waymo LLC)</i>	
Learning Human-to-Robot Handovers From Point Clouds	9654
<i>Sammy Christen (ETH Zurich; NVIDIA), Wei Yang (NVIDIA), Claudia Pérez-D'Arpino (NVIDIA), Otmar Hilliges (ETH Zurich), Dieter Fox (NVIDIA; University of Washington), and Yu-Wei Chao (NVIDIA)</i>	
Phone2Proc: Bringing Robust Robots Into Our Chaotic World	9665
<i>Matt Deitke (PRIOR @ Allen Institute for AI; University of Washington, Seattle), Rose Hendrix (PRIOR @ Allen Institute for AI), Ali Farhadi (University of Washington, Seattle), Kiana Ehsani (PRIOR @ Allen Institute for AI), and Aniruddha Kembhavi (University of Washington, Seattle)</i>	
GazeNeRF: 3D-Aware Gaze Redirection With Neural Radiance Fields	9676
<i>Alessandro Ruzzi (Department of Computer Science, ETH Zürich), Xiangwei Shi (Delft University of Technology), Xi Wang (Department of Computer Science, ETH Zürich), Gengyan Li (Department of Computer Science, ETH Zürich), Shalini De Mello (NVIDIA), Hyung Jin Chang (University of Birmingham), Xucong Zhang (Delft University of Technology), and Otmar Hilliges (Department of Computer Science, ETH Zürich)</i>	
Observation-Centric SORT: Rethinking SORT for Robust Multi-Object Tracking	9686
<i>Jinkun Cao (Carnegie Mellon University), Jiangmiao Pang (Shanghai AI Laboratory), Xinshuo Weng (Nvidia), Rawal Khirodkar (Carnegie Mellon University), and Kris Kitani (Carnegie Mellon University)</i>	
Autoregressive Visual Tracking	9697
<i>Xing Wei (Xi'an Jiaotong University), Yifan Bai (Xi'an Jiaotong University), Yongchao Zheng (Xi'an Jiaotong University), Dahu Shi (Zhejiang University; Hikvision Research Institute), and Yihong Gong (Xi'an Jiaotong University)</i>	

OpenGait: Revisiting Gait Recognition Towards Better Practicality	9707
<i>Chao Fan (Southern University of Science and Technology), Junhao Liang (Southern University of Science and Technology), Chuanfu Shen (Southern University of Science and Technology; The University of Hong Kong), Saihui Hou (Southern University of Science and Technology; WATRIX.AI), Yongzhen Huang (Beijing Normal University; WATRIX.AI), and Shiqi Yu (Southern University of Science and Technology)</i>	
Pose-Disentangled Contrastive Learning for Self-Supervised Facial Representation	9717
<i>Yuanyuan Liu (China University of Geosciences, China), Wenbin Wang (China University of Geosciences, China), Yibing Zhan (JD Explore Academy, China), Shaoze Feng (China University of Geosciences, China), Kejun Liu (China University of Geosciences, China), and Zhe Chen (The University of Sydney, Australia)</i>	
Identity-Preserving Talking Face Generation With Landmark and Appearance Priors	9729
<i>Weizhi Zhong (Sun Yat-sen University), Chaowei Fang (Xidian University), Yinqi Cai (Sun Yat-sen University), Pengxu Wei (Sun Yat-sen University), Gangming Zhao (The University of Hong Kong), Liang Lin (Sun Yat-sen University), and Guanbin Li (Sun Yat-sen University)</i>	
DF-Platter: Multi-Face Heterogeneous Deepfake Dataset	9739
<i>Kartik Narayan (IIT Jodhpur, India), Harsh Agarwal (IIT Jodhpur, India), Kartik Thakral (IIT Jodhpur, India), Surbhi Mittal (IIT Jodhpur, India), Mayank Vatsa (IIT Jodhpur, India), and Richa Singh (IIT Jodhpur, India)</i>	
Physics-Driven Diffusion Models for Impact Sound Synthesis From Videos	9749
<i>Kun Su (University of Washington), Kaizhi Qian (MIT-IBM Watson AI Lab), Eli Shlizerman (University of Washington), Antonio Torralba (MIT), and Chuang Gan (MIT-IBM Watson AI Lab; UMass Amherst)</i>	
Mofusion: A Framework for Denoising-Diffusion-Based Motion Synthesis	9760
<i>Rishabh Dabral (Max Planck Institute for Informatics, SIC), Muhammad Hamza Mughal (Max Planck Institute for Informatics, SIC; Saarland University), Vladislav Golyanik (Max Planck Institute for Informatics, SIC), and Christian Theobalt (Max Planck Institute for Informatics, SIC)</i>	
Adaptive Global Decay Process for Event Cameras	9771
<i>Urbano Miguel Nunes (Sorbonne University, 4 place jussieu, 75005 Paris), Ryad Benosman (Sorbonne University, 4 place jussieu, 75005 Paris), and Sio-Hoi Ieng (Sorbonne University, 4 place jussieu, 75005 Paris)</i>	
Frame-Event Alignment and Fusion Network for High Frame Rate Tracking	9781
<i>Jiqing Zhang (Dalian University of Technology), Yuanchen Wang (Dalian University of Technology), Wenxi Liu (Fuzhou University), Meng Li (HiSilicon(Shanghai) Technologies Co., Ltd), Jinpeng Bai (Dalian University of Technology), Baocai Yin (Dalian University of Technology), and Xin Yang (Dalian University of Technology)</i>	

Exploring Discontinuity for Video Frame Interpolation	9791
<i>Sangjin Lee (Yonsei University), Hyeongmin Lee (Yonsei University), Chajin Shin (Yonsei University), Hanbin Son (Yonsei University), and Sangyoun Lee (Yonsei University; Korea Institute of Science and Technology (KIST))</i>	
AMT: All-Pairs Multi-Field Transforms for Efficient Frame Interpolation	9801
<i>Zhen Li (Nankai University), Zuo-Liang Zhu (Nankai University), Ling-Hao Han (Nankai University), Qibin Hou (Nankai University), Chun-Le Guo (Nankai University), and Ming-Ming Cheng (Nankai University)</i>	
Frame Interpolation Transformer and Uncertainty Guidance	9811
<i>Markus Plack (University of Bonn, Germany), Karlis Martins Briedis (Department of Computer Science, ETH Zürich, Switzerland; DisneyResearch\Studios, Zürich, Switzerland), Abdelaziz Djelouah (DisneyResearch\Studios, Zürich, Switzerland), Matthias B. Hullin (University of Bonn, Germany), Markus Gross (Department of Computer Science, ETH Zürich, Switzerland; Department of Computer Science, ETH Zürich, Switzerland), and Christopher Schroers (DisneyResearch\Studios, Zürich, Switzerland)</i>	
A Simple Baseline for Video Restoration With Grouped Spatial-Temporal Shift	9822
<i>Dasong Li (CUHK MMLab), Xiaoyu Shi (CUHK MMLab), Yi Zhang (CUHK MMLab), Ka Chun Cheung (NVIDIA AI Technology Center), Simon See (NVIDIA AI Technology Center), Xiaogang Wang (CUHK MMLab; CII under InnoHK), Hongwei Qin (SenseTime Research), and Hongsheng Li (CUHK MMLab; CII under InnoHK)</i>	
Recurrent Homography Estimation Using Homography-Guided Image Warping and Focus Transformer	9833
<i>Si-Yuan Cao (Zhejiang University), Runmin Zhang (Zhejiang University), Lun Luo (Zhejiang University), Beinan Yu (Zhejiang University), Zehua Sheng (Zhejiang University), Junwei Li (Zhejiang University), and Hui-Liang Shen (Zhejiang University)</i>	
HyperCUT: Video Sequence From a Single Blurry Image Using Unsupervised Ordering	9843
<i>Bang-Dang Pham (VinAI Research, Vietnam), Phong Tran (VinAI Research, Vietnam; MBZUAI, UAE), Anh Tran (VinAI Research, Vietnam), Cuong Pham (VinAI Research, Vietnam; Posts & Telecommunications Inst. of Tech, Vietnam), Rang Nguyen (VinAI Research, Vietnam), and Minh Hoai (VinAI Research, Vietnam; Stony Brook University, USA)</i>	
Indescribable Multi-Modal Spatial Evaluator	9853
<i>Lingke Kong (Manteia Tech), X. Sharon Qi (University of California, Los Angeles), Qijin Shen (Fuzhou University), Jiacheng Wang (Xiamen University), Jingyi Zhang (Xiamen University), Yanle Hu (Mayo Clinic Arizona), and Qichao Zhou (Manteia Tech)</i>	
Structured Kernel Estimation for Photon-Limited Deconvolution	9863
<i>Yash Sanghvi (Purdue University), Zhiyuan Mao (Purdue University), and Stanley H. Chan (Purdue University)</i>	
Polarized Color Image Denoising	9873
<i>Zhuoxiao Li (The University of Tokyo), Haiyang Jiang (The University of Tokyo), Mingdeng Cao (The University of Tokyo), and Yinqiang Zheng (The University of Tokyo)</i>	

Uncertainty-Aware Unsupervised Image Deblurring With Deep Residual Prior	9883
<i>Xiaole Tang (University of Electronic Science and Technology of China, China), Xile Zhao (University of Electronic Science and Technology of China, China), Jun Liu (Northeast Normal University, China), Jianli Wang (University of Electronic Science and Technology of China, China), Yuchun Miao (University of Electronic Science and Technology of China, China), and Tiejong Zeng (The Chinese University of Hong Kong, Hong Kong)</i>	
Low-Light Image Enhancement via Structure Modeling and Guidance	9893
<i>Xiaogang Xu (Zhejiang Lab), Ruixing Wang (Honor Device Co., Ltd.), and Jiangbo Lu (SmartMore Corporation)</i>	
Learning Sample Relationship for Exposure Correction	9904
<i>Jie Huang (University of Science and Technology of China), Feng Zhao (University of Science and Technology of China), Man Zhou (University of Science and Technology of China), Jie Xiao (University of Science and Technology of China), Naishan Zheng (University of Science and Technology of China), Kaiwen Zheng (University of Science and Technology of China), and Zhiwei Xiong (University of Science and Technology of China)</i>	
Spatially Adaptive Self-Supervised Learning for Real-World Image Denoising	9914
<i>Junyi Li (School of Computer Science and Technology, Harbin Institute of Technology, China), Zhilu Zhang (School of Computer Science and Technology, Harbin Institute of Technology, China), Xiaoyu Liu (School of Computer Science and Technology, Harbin Institute of Technology, China), Chaoyu Feng (School of Computer Science and Technology, Harbin Institute of Technology, China), Xiaotao Wang (School of Computer Science and Technology, Harbin Institute of Technology, China), Lei Lei (School of Computer Science and Technology, Harbin Institute of Technology, China), and Wangmeng Zuo (Peng Cheng Laboratory, China)</i>	
Quantum-Inspired Spectral-Spatial Pyramid Network for Hyperspectral Image Classification	9925
<i>Jie Zhang (University of Macau, China), Yongshan Zhang (China University of Geosciences, China; University of Macau, China), and Yicong Zhou (University of Macau, China)</i>	
Generative Diffusion Prior for Unified Image Restoration and Enhancement	9935
<i>Ben Fei (Fudan University; Shanghai AI Laboratory), Zhaoyang Lyu (Shanghai AI Laboratory), Liang Pan (S-Lab, Nanyang Technological University), Junzhe Zhang (S-Lab, Nanyang Technological University), Weidong Yang (Fudan University), Tianyue Luo (Fudan University), Bo Zhang (Shanghai AI Laboratory), and Bo Dai (Shanghai AI Laboratory)</i>	
Ground-Truth Free Meta-Learning for Deep Compressive Sampling	9947
<i>Xinran Qin (School of Computer Science and Engineering, South China University of Technology, China; Pazhou Lab, China), Yuhui Quan (School of Computer Science and Engineering, South China University of Technology, China; Pazhou Lab, China), Tongyao Pang (National University of Singapore, Singapore), and Hui Ji (National University of Singapore, Singapore)</i>	

Recognizability Embedding Enhancement for Very Low-Resolution Face Recognition and Quality Estimation	9957
<i>Jacky Chen Long Chai (Yonsei University), Tiong-Sik Ng (Yonsei University), Cheng-Yaw Low (Institute for Basic Science), Jaewoo Park (Yonsei University), and Andrew Beng Jin Teoh (Yonsei University)</i>	
An Image Quality Assessment Dataset for Portraits	9968
<i>Nicolas Chahine (DXOMARK; Departement d'informatique de l'Ecole normale superieure), Stefania Calarasanu (DXOMARK), Davide Garcia-Civiero (DXOMARK), Théo Cayla (DXOMARK), Sira Ferradans (DXOMARK), and Jean Ponce (Departement d'informatique de l'Ecole normale superieure; Institute of Mathematical Sciences and Center for Data Science, New York University)</i>	
Bitstream-Corrupted JPEG Images Are Restorable: Two-Stage Compensation and Alignment Framework for Image Restoration	9979
<i>Wenyang Liu (Nanyang Technological University, Singapore), Yi Wang (Nanyang Technological University, Singapore), Kim-Hui Yap (Nanyang Technological University, Singapore), and Lap-Pui Chau (The Hong Kong Polytechnic University, Hong Kong)</i>	
Image Super-Resolution Using T-Tetromino Pixels	9989
<i>Simon Grosche (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Andy Regensky (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Jürgen Seiler (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), and André Kaup (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)</i>	
CUF: Continuous Upsampling Filters	9999
<i>Cristina N. Vasconcelos (Milad Hashemi Kevin Swersky Andrea Tagliasacchi), Cengiz Oztireli (Milad Hashemi Kevin Swersky Andrea Tagliasacchi), Mark Matthews (Milad Hashemi Kevin Swersky Andrea Tagliasacchi), Milad Hashemi (Milad Hashemi Kevin Swersky Andrea Tagliasacchi), Kevin Swersky (Milad Hashemi Kevin Swersky Andrea Tagliasacchi), and Andrea Tagliasacchi (Milad Hashemi Kevin Swersky Andrea Tagliasacchi)</i>	
OPE-SR: Orthogonal Position Encoding for Designing a Parameter-Free Upsampling Module in Arbitrary-Scale Image Super-Resolution	10009
<i>Gaochao Song (Tianjin University), Qian Sun (Nanjing University of Information Science and Technology), Luo Zhang (Nanyang Technological University), Ran Su (Tianjin University), Jianfeng Shi (Nanjing University of Information Science and Technology), and Ying He (Nanyang Technological University)</i>	
Implicit Diffusion Models for Continuous Super-Resolution	10021
<i>Sicheng Gao (Beihang University), Xuhui Liu (Beihang University), Bohan Zeng (Beihang University), Sheng Xu (Beihang University), Yanjing Li (Beihang University), Xiaoyan Luo (Beihang University), Jianzhuang Liu (Shenzhen Institute of Advanced Technology, China), Xiantong Zhen (United Imaging), and Baochang Zhang (Beihang University; Zhongguancun Laboratory, China)</i>	

Pixels, Regions, and Objects: Multiple Enhancement for Salient Object Detection	10031
<i>Yi Wang (DUT-RU International School of Information Science and Engineering Dalian University of Technology, China), Ruili Wang (School of Mathematical and Computational Sciences Massey University, New Zealand), Xin Fan (DUT-RU International School of Information Science and Engineering Dalian University of Technology, China), Tianzhu Wang (School of Mathematical and Computational Sciences Massey University, New Zealand), and Xiangjian He (School of Computer Science University of Nottingham Ningbo China, China)</i>	
VILA: Learning Image Aesthetics From User Comments With Vision-Language Pretraining ...	10041
<i>Junjie Ke (Google Research), Keren Ye (Google Research), Jiahui Yu (Google Research), Yonghui Wu (Google Research), Peyman Milanfar (Google Research), and Feng Yang (Google Research)</i>	
Image Cropping With Spatial-Aware Feature and Rank Consistency	10052
<i>Chao Wang (Shanghai Jiao Tong University), Li Niu (Shanghai Jiao Tong University), Bo Zhang (Shanghai Jiao Tong University), and Liqing Zhang (Shanghai Jiao Tong University)</i>	
B-Spline Texture Coefficients Estimator for Screen Content Image Super-Resolution	10062
<i>Byeonghyun Pak (Daegu Gyeongbuk Institute of Science and Technology (DGIST), Korea), Jaewon Lee (Daegu Gyeongbuk Institute of Science and Technology (DGIST), Korea), and Kyong Hwan Jin (Daegu Gyeongbuk Institute of Science and Technology (DGIST), Korea)</i>	
Delving StyleGAN Inversion for Image Editing: A Foundation Latent Space Viewpoint	10072
<i>Hongyu Liu (Hong Kong University of Science and Technology), Yibing Song (AI), and Qifeng Chen (Institute, Fudan University)</i>	
Learning Dynamic Style Kernels for Artistic Style Transfer	10083
<i>Wenju Xu (OPPO US Research Center, USA), Chengjiang Long (Meta Reality Lab, USA), and Yongwei Nie (South China University of Technology, China)</i>	
SVGformer: Representation Learning for Continuous Vector Graphics Using Transformers .	10093
<i>Defu Cao (University of Southern California), Zhaowen Wang (Adobe Research, Inc.), Jose Echevarria (Adobe Research, Inc.), and Yan Liu (University of Southern California)</i>	
Learning Generative Structure Prior for Blind Text Image Super-Resolution	10103
<i>Xiaoming Li (Nanyang Technological University), Wangmeng Zuo (Harbin Institute of Technology), and Chen Change Loy (Nanyang Technological University)</i>	
Unsupervised Domain Adaption With Pixel-Level Discriminator for Image-Aware Layout Generation	10114
<i>Chenchen Xu (Zhejiang University; Alibaba Group), Min Zhou (Alibaba Group), Tiezheng Ge (Alibaba Group), Yuning Jiang (Alibaba Group), and Weiwei Xu (Zhejiang University)</i>	
Scaling Up GANs for Text-to-Image Synthesis	10124
<i>Minguk Kang (POSTECH; Adobe Research), Jun-Yan Zhu (Carnegie Mellon University), Richard Zhang (Adobe Research), Jaesik Park (POSTECH), Eli Shechtman (Adobe Research), Sylvain Paris (Adobe Research), and Taesung Park (Adobe Research)</i>	

ERNIE-ViLG 2.0: Improving Text-to-Image Diffusion Model With Knowledge-Enhanced Mixture-of-Denoising-Experts	10135
<i>Zhida Feng (Baidu Inc.; Wuhan University of Science and Technology), Zhenyu Zhang (Baidu Inc.), Xintong Yu (Baidu Inc.), Yewei Fang (Baidu Inc.), Lanxin Li (Baidu Inc.), Xuyi Chen (Baidu Inc.), Yuxiang Lu (Baidu Inc.), Jiaxiang Liu (Baidu Inc.), Weichong Yin (Baidu Inc.), Shikun Feng (Baidu Inc.), Yu Sun (Baidu Inc.), Li Chen (Wuhan University of Science and Technology), Hao Tian (Baidu Inc.), Hua Wu (Baidu Inc.), and Haifeng Wang (Baidu Inc.)</i>	
Inversion-Based Style Transfer With Diffusion Models	10146
<i>Yuxin Zhang (MAIS, Institute of Automation, Chinese Academy of Sciences; School of AI, UCAS), Nisha Huang (MAIS, Institute of Automation, Chinese Academy of Sciences; School of AI, UCAS), Fan Tang (Institute of Computing Technology, Chinese Academy of Sciences), Haibin Huang (Kuaishou Technology), Chongyang Ma (Kuaishou Technology), Weiming Dong (MAIS, Institute of Automation, Chinese Academy of Sciences; School of AI, UCAS), and Changsheng Xu (MAIS, Institute of Automation, Chinese Academy of Sciences; School of AI, UCAS)</i>	
Shifted Diffusion for Text-to-Image Generation	10157
<i>Yufan Zhou (State University of New York at Buffalo), Bingchen Liu (ByteDance), Yizhe Zhu (ByteDance), Xiao Yang (ByteDance), Changyou Chen (State University of New York at Buffalo), and Jinhui Xu (State University of New York at Buffalo)</i>	
LayoutDM: Discrete Diffusion Model for Controllable Layout Generation	10167
<i>Naoto Inoue (CyberAgent, Japan), Kotaro Kikuchi (CyberAgent, Japan), Edgar Simo-Serra (Waseda University, Japan), Mayu Otani (CyberAgent, Japan), and Kota Yamaguchi (CyberAgent, Japan)</i>	
Unpaired Image-to-Image Translation With Shortest Path Regularization	10177
<i>Shaoan Xie (Carnegie Mellon University), Yanwu Xu (Boston University), Mingming Gong (University of Melbourne; Mohamed bin Zayed University of Artificial Intelligence), and Kun Zhang (Carnegie Mellon University; Mohamed bin Zayed University of Artificial Intelligence)</i>	
DiffCollage: Parallel Generation of Large Content With Diffusion Models	10188
<i>Qinsheng Zhang (Georgia Institute of Technology, NVIDIA Corporation), Jiaming Song (Georgia Institute of Technology, NVIDIA Corporation), Xun Huang (Georgia Institute of Technology, NVIDIA Corporation), Yongxin Chen (Georgia Institute of Technology, NVIDIA Corporation), and Ming-Yu Liu (Georgia Institute of Technology, NVIDIA Corporation)</i>	
Wavelet Diffusion Models Are Fast and Scalable Image Generators	10199
<i>Hao Phung (VinAI Research), Quan Dao (VinAI Research), and Anh Tran (VinAI Research)</i>	

VideoFusion: Decomposed Diffusion Models for High-Quality Video Generation	N/A
<i>Zhengxiong Luo (University of Chinese Academy of Sciences (UCAS); Alibaba Group; Center for Research on Intelligent Perception and Computing (CRIPAC); Institute of Automation, Chinese Academy of Sciences (CASIA)), Dayou Chen (Alibaba Group), Yingya Zhang (Alibaba Group), Yan Huang (Center for Research on Intelligent Perception and Computing (CRIPAC); Institute of Automation, Chinese Academy of Sciences (CASIA)), Liang Wang (Center for Research on Intelligent Perception and Computing (CRIPAC); Institute of Automation, Chinese Academy of Sciences (CASIA)), Yujun Shen (Ant Group), Deli Zhao (Alibaba Group), Jingren Zhou (Alibaba Group), and Tieniu Tan (Center for Research on Intelligent Perception and Computing (CRIPAC); Institute of Automation, Chinese Academy of Sciences (CASIA); Nanjing University)</i>	
MM-Diffusion: Learning Multi-Modal Diffusion Models for Joint Audio and Video Generation	10219
<i>Ludan Ruan (Renmin University of China), Yiyang Ma (Peking University), Huan Yang (Microsoft Research), Huiguo He (Microsoft Research), Bei Liu (Microsoft Research), Jianlong Fu (Microsoft Research), Nicholas Jing Yuan (Microsoft Research), Qin Jin (Renmin University of China), and Baining Guo (Microsoft Research)</i>	
Adaptive Human Matting for Dynamic Videos	10229
<i>Chung-Ching Lin (Microsoft), Jiang Wang (Microsoft), Kun Luo (Microsoft), Kevin Lin (Microsoft), Linjie Li (Microsoft), Lijuan Wang (Microsoft), and Zicheng Liu (Microsoft)</i>	
LVQAC: Lattice Vector Quantization Coupled With Spatially Adaptive Companding for Efficient Learned Image Compression	10239
<i>Xi Zhang (Shanghai Jiao Tong University) and Xiaolin Wu (McMaster University)</i>	
Hierarchical B-Frame Video Coding Using Two-Layer CANF Without Motion Coding	10249
<i>David Alexandre (Electrical Engineering and Computer Science International Graduate Program), Hsueh-Ming Hang (Institute of Electronics), and Wen-Hsiao Peng (Department of Computer Science)</i>	
Towards High-Quality and Efficient Video Super-Resolution via Spatial-Temporal Data Overfitting	10259
<i>Gen Li (Clemson University), Jie Ji (Clemson University), Minghai Qin (n/a), Wei Niu (William & Mary), Bin Ren (William & Mary), Fatemeh Afghah (Clemson University), Linke Guo (Clemson University), and Xiaolong Ma (Clemson University)</i>	
HNeRV: A Hybrid Neural Representation for Videos	10270
<i>Hao Chen (University of Maryland, College Park), Matthew Gwilliam (University of Maryland, College Park), Ser-Nam Lim (Meta AI), and Abhinav Shrivastava (University of Maryland, College Park)</i>	
Regularize Implicit Neural Representation by Itself	10280
<i>Zheming Li (National University of Defense Technology), Hongxia Wang (National University of Defense Technology), and Deyu Meng (Xi'an Jiaotong University; Macau University of Science and Technology)</i>	

SMPCnv: Self-Moving Point Representations for Continuous Convolution	10289
<i>Sanghyeon Kim (Sungkyunkwan University) and Eunbyung Park (Sungkyunkwan University)</i>	
Long Range Pooling for 3D Large-Scale Scene Understanding	10300
<i>Xiang-Li Li (Tsinghua University), Meng-Hao Guo (Tsinghua University), Tai-Jiang Mu (Tsinghua University), Ralph R. Martin (Cardiff University), and Shi-Min Hu (Tsinghua University)</i>	
Progressive Random Convolutions for Single Domain Generalization	10312
<i>Seokeon Choi (Qualcomm AI research), Debasmit Das (Qualcomm AI research), Sungha Choi (Qualcomm AI research), Seunghan Yang (Qualcomm AI research), Hyunsin Park (Qualcomm AI research), and Sungrack Yun (Qualcomm AI research)</i>	
BiFormer: Vision Transformer With Bi-Level Routing Attention	10323
<i>Lei Zhu (City University of Hong Kong), Xinjiang Wang (SenseTime Research), Zhanghan Ke (City University of Hong Kong), Wayne Zhang (SenseTime Research), and Rynson W.H. Lau (City University of Hong Kong)</i>	
Beyond Attentive Tokens: Incorporating Token Importance and Diversity for Efficient Vision Transformers	10334
<i>Sifan Long (Jilin University, China; Jilin University, China; Baidu VIS), Zhen Zhao (Baidu VIS; University of Sydney), Jimin Pi (Baidu VIS), Shengsheng Wang (Jilin University, China; Jilin University, China), and Jingdong Wang (Baidu VIS)</i>	
BioNet: A Biologically-Inspired Network for Face Recognition	10344
<i>Pengyu Li (Terminus Labs, Terminus Group, China)</i>	
Dual-Bridging With Adversarial Noise Generation for Domain Adaptive rPPG Estimation	10355
<i>Jingda Du (Hong Kong Baptist University, Hong Kong), Si-Qi Liu (Hong Kong Baptist University, Hong Kong), Bochao Zhang (Hong Kong Baptist University, Hong Kong), and Pong C. Yuen (Hong Kong Baptist University, Hong Kong)</i>	
On Data Scaling in Masked Image Modeling	10365
<i>Zhenda Xie (Tsmghua University; Microsoft Research Asia), Zheng Zhang (Microsoft Research Asia), Yue Cao (Microsoft Research Asia), Yutong Lin (Xi'an Jiaotong University; Microsoft Research Asia), Yixuan Wei (Tsmghua University; Microsoft Research Asia), Qi Dai (Microsoft Research Asia), and Han Hu (Microsoft Research Asia)</i>	

Hard Patches Mining for Masked Image Modeling	10375
<i>Haochen Wang (Center for Research on Intelligent Perception and Computing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Kaiyou Song (Megvii Technology), Junsong Fan (Center for Research on Intelligent Perception and Computing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences; Centre for Artificial Intelligence and Robotics, Hong Kong Institute of Science & Innovation, Chinese Academy of Science), Yuxi Wang (Center for Research on Intelligent Perception and Computing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences; Centre for Artificial Intelligence and Robotics, Hong Kong Institute of Science & Innovation, Chinese Academy of Science), Jin Xie (Megvii Technology), and Zhaoxiang Zhang (Center for Research on Intelligent Perception and Computing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Centre for Artificial Intelligence and Robotics, Hong Kong Institute of Science & Innovation, Chinese Academy of Science)</i>	
Evolved Part Masking for Self-Supervised Learning	10386
<i>Zhanzhou Feng (Peking University) and Shiliang Zhang (Peking University; Peng Cheng Laboratory)</i>	
BASIS: Batch Aligned Spectral Embedding Space	10396
<i>Or Streicher (Viterbi Faculty of Electrical and Computer Engineering Technion - Israel Institute of Technology, Haifa, Israel), Ido Cohen (Viterbi Faculty of Electrical and Computer Engineering Technion - Israel Institute of Technology, Haifa, Israel), and Guy Gilboa (Viterbi Faculty of Electrical and Computer Engineering Technion - Israel Institute of Technology, Haifa, Israel)</i>	
OmniMAE: Single Model Masked Pretraining on Images and Videos	10406
<i>Rohit Girdhar (FAIR, Meta AI), Alaaeldin El-Nouby (FAIR, Meta AI), Mannat Singh (FAIR, Meta AI), Kalyan Vasudev Alwala (FAIR, Meta AI), Armand Joulin (FAIR, Meta AI), and Ishan Misra (FAIR, Meta AI)</i>	
ViTs for SITS: Vision Transformers for Satellite Image Time Series	10418
<i>Michail Tarasiou (Imperial College London), Erik Chavez (Imperial College London), and Stefanos Zafeiriou (Imperial College London)</i>	
Probabilistic Debiasing of Scene Graphs	10429
<i>Bashirul Azam Biswas (Rensselaer Polytechnic Institute, Troy) and Qiang Ji (Rensselaer Polytechnic Institute, Troy)</i>	
Blind Video Deflickering by Neural Filtering With a Flawed Atlas	10439
<i>Chenyang Lei (CAIR, HKISI-CAS; Princeton University), Xuanchi Ren (University of Toronto; Vector Institute), Zhaoxiang Zhang (CAIR, HKISI-CAS; CASIA), and Qifeng Chen (HKUST)</i>	

SCOTCH and SODA: A Transformer Video Shadow Detection Framework	10449
<i>Lihao Liu (University of Cambridge, United Kingdom), Jean Prost (Univ. Bordeaux, CNRS, France), Lei Zhu (The Hong Kong University of Science and Technology (Guangzhou), China; The Hong Kong University of Science and Technology, HK SAR, China), Nicolas Papadakis (Univ. Bordeaux, CNRS, France), Pietro Liò (University of Cambridge, United Kingdom), Carola-Bibiane Schönlieb (University of Cambridge, United Kingdom), and Angelica I. Aviles-Rivero (University of Cambridge, United Kingdom)</i>	
MAGVIT: Masked Generative Video Transformer	10459
<i>Lijun Yu (Carnegie Mellon University; Google Research), Yong Cheng (Google Research), Kihyuk Sohn (Google Research), José Lezama (Google Research), Han Zhang (Google Research), Huiwen Chang (Google Research), Alexander G. Hauptmann (Carnegie Mellon University), Ming-Hsuan Yang (Google Research), Yuan Hao (Google Research), Irfan Essa (Google Research; Georgia Institute of Technology), and Lu Jiang (Google Research)</i>	
Improving Robustness of Semantic Segmentation to Motion-Blur Using Class-Centric Augmentation	10470
<i>Aakanksha Aakanksha (Indian Institute of Technology Madras) and A. N. Rajagopalan (Indian Institute of Technology Madras)</i>	
MobileVOS: Real-Time Video Object Segmentation Contrastive Learning Meets Knowledge Distillation	10480
<i>Roy Miles (Samsung Research UK), Mehmet Kerim Yucel (Samsung Research UK), Bruno Manganelli (Samsung Research UK), and Albert Saà-Garriga (Samsung Research UK)</i>	
Self-Supervised Video Forensics by Audio-Visual Anomaly Detection	10491
<i>Chao Feng (University of Michigan), Ziyang Chen (University of Michigan), and Andrew Owens (University of Michigan)</i>	
Frame Flexible Network	10504
<i>Yitian Zhang (Northeastern University), Yue Bai (Northeastern University), Chang Liu (Northeastern University), Huan Wang (Northeastern University), Sheng Li (University of Virginia), and Yun Fu (Northeastern University)</i>	
System-Status-Aware Adaptive Network for Online Streaming Video Understanding	10514
<i>Lin Geng Foo (Singapore University of Technology and Design), Jia Gong (Singapore University of Technology and Design), Zhipeng Fan (New York University), and Jun Liu (Singapore University of Technology and Design)</i>	
MDQE: Mining Discriminative Query Embeddings To Segment Occluded Instances on Challenging Videos	10524
<i>Minghan Li (Hong Kong Polytechnic University), Shuai Li (Hong Kong Polytechnic University), Wangmeng Xiang (Hong Kong Polytechnic University), and Lei Zhang (Hong Kong Polytechnic University)</i>	
Spatio-Temporal Pixel-Level Contrastive Learning-Based Source-Free Domain Adaptation for Video Semantic Segmentation	10534
<i>Shao-Yuan Lo (Johns Hopkins University), Poojan Oza (Amazon), Sumanth Chennupati (Amazon), Alejandro Galindo (Amazon), and Vishal M. Patel (Johns Hopkins University)</i>	

Taming Diffusion Models for Audio-Driven Co-Speech Gesture Generation	10544
<i>Lingting Zhu (The University of Hong Kong), Xian Liu (The Chinese University of Hong Kong), Xuanyu Liu (The University of Hong Kong), Rui Qian (The Chinese University of Hong Kong), Ziwei Liu (Nanyang Technological University), and Lequan Yu (Nanyang Technological University)</i>	
Chat2Map: Efficient Scene Mapping From Multi-Ego Conversations	10554
<i>Sagnik Majumder (UT Austin; Reality Labs Research, Meta; FAIR), Hao Jiang (Reality Labs Research, Meta), Pierre Moulon (Reality Labs Research, Meta), Ethan Henderson (Reality Labs Research, Meta), Paul Calamia (Reality Labs Research, Meta), Kristen Grauman (UT Austin; FAIR), and Vamsi Krishna Ithapu (Reality Labs Research, Meta)</i>	
Audio-Visual Grouping Network for Sound Localization From Mixtures	10565
<i>Shentong Mo (Carnegie Mellon University) and Yapeng Tian (University of Texas at Dallas)</i>	
Language-Guided Audio-Visual Source Separation via Trimodal Consistency	10575
<i>Reuben Tan (Boston University), Arijit Ray (Boston University), Andrea Burns (Boston University), Bryan A. Plummer (Boston University), Justin Salamon (Adobe Research), Oriol Nieto (Adobe Research), Bryan Russell (Boston University; MIT-IBM Watson AI Lab, IBM Research), and Kate Saenko (Boston University; MIT-IBM Watson AI Lab, IBM Research)</i>	
Fine-Grained Audible Video Description	10585
<i>Xuyang Shen (OpenNLPLab), Dong Li (Shanghai Artificial Intelligence Laboratory), Jinxing Zhou (Hefei University of Technology), Zhen Qin (OpenNLPLab), Bowen He (OpenNLPLab), Xiaodong Han (OpenNLPLab), Aixuan Li (Northwestern Polytechnical University), Yuchao Dai (Northwestern Polytechnical University), Lingpeng Kong (The University of Hong Kong), Meng Wang (Hefei University of Technology), Yu Qiao (Shanghai Artificial Intelligence Laboratory), and Yiran Zhong (Shanghai Artificial Intelligence Laboratory)</i>	
Neural Koopman Pooling: Control-Inspired Temporal Dynamics Encoding for Skeleton-Based Action Recognition	10597
<i>Xinghan Wang (Peking University), Xin Xu (Peking University), and Yadong Mu (Peng Cheng Laboratory)</i>	
Learning Discriminative Representations for Skeleton Based Action Recognition	10608
<i>Huanyu Zhou (Beihang University, Beijing, China), Qingjie Liu (Beihang University, Beijing, China; Zhongguancun Laboratory; Hangzhou Innovation Institute of Beihang University), and Yunhong Wang (Beihang University, Beijing, China)</i>	
Therbligs in Action: Video Understanding Through Motion Primitives	10618
<i>Eadam Dessalene (University of Maryland, USA), Michael Maynard (University of Maryland, USA), Cornelia Fermüller (University of Maryland, USA), and Yiannis Aloimonos (University of Maryland, USA)</i>	
Search-Map-Search: A Frame Selection Paradigm for Action Recognition	10627
<i>Mingjun Zhao (University of Alberta), Yakun Yu (University of Alberta), Xiaoli Wang (Shenzhen, China), Lei Yang (Shenzhen, China), and Di Niu (Edmonton, Canada)</i>	

Re2TAL: Rewiring Pretrained Video Backbones for Reversible Temporal Action Localization	10637
<i>Chen Zhao (King Abdullah University of Science and Technology (KAUST), Saudi Arabia), Shuming Liu (King Abdullah University of Science and Technology (KAUST), Saudi Arabia), Karttikeya Mangalam (UC Berkeley, US), and Bernard Ghanem (King Abdullah University of Science and Technology (KAUST), Saudi Arabia)</i>	
Boosting Weakly-Supervised Temporal Action Localization With Text Information	10648
<i>Guozhang Li (Xidian University), De Cheng (Xidian University), Xinpeng Ding (The Hong Kong University of Science and Technology), Nannan Wang (Xidian University), Xiaoyu Wang (The Chinese University of Hong Kong (Shenzhen)), and Xinbo Gao (Xidian University; Chongqing University of Posts and Telecommunications)</i>	
Perception and Semantic Aware Regularization for Sequential Confidence Calibration	10658
<i>Zhenghua Peng (South China University of Technology), Yu Luo (South China University of Technology), Tianshui Chen (Guangdong University of Technology), Keke Xu (South China University of Technology), and Shuangping Huang (South China University of Technology; Pashou Laboratory)</i>	
NewsNet: A Novel Dataset for Hierarchical Temporal Segmentation	10669
<i>Haoqian Wu (Tencent), Keyu Chen (Tencent), Haozhe Liu (King Abdullah University of Science and Technology (KAUST)), Mingchen Zhuge (King Abdullah University of Science and Technology (KAUST)), Bing Li (King Abdullah University of Science and Technology (KAUST)), Ruizhi Qiao (Tencent), Xiujun Shu (Tencent), Bei Gan (Tencent), Liangsheng Xu (Tencent), Bo Ren (Tencent), Mengmeng Xu (King Abdullah University of Science and Technology (KAUST)), Wentian Zhang (King Abdullah University of Science and Technology (KAUST)), Raghavendra Ramachandra (Norwegian University of Science and Technology (NTNU)), Chia-Wen Lin (National Tsing Hua University (NTHU)), and Bernard Ghanem (King Abdullah University of Science and Technology (KAUST))</i>	
Tell Me What Happened: Unifying Text-Guided Video Completion via Multimodal Masked Video Generation	10681
<i>Tsu-Jui Fu (UC Santa Barbara), Licheng Yu (Meta), Ning Zhang (Meta), Cheng-Yang Fu (Meta), Jong-Chyi Su (NEC Laboratories America), William Yang Wang (UC Santa Barbara), and Sean Bell (Meta)</i>	
Leveraging Temporal Context in Low Representational Power Regimes	10693
<i>Camilo L. Fosco (MIT CSAIL), SouYoung Jin (MIT CSAIL), Emilie Josephs (MIT CSAIL), and Aude Oliva (MIT CSAIL)</i>	
Cap4Video: What Can Auxiliary Captions Do for Text-Video Retrieval?	10704
<i>Wenhao Wu (The University of Sydney; Baidu Inc.), Haipeng Luo (University of Chinese Academy of Sciences), Bo Fang (University of Chinese Academy of Sciences), Jingdong Wang (Baidu Inc.), and Wanli Ouyang (The University of Sydney; Shanghai AI Laboratory)</i>	
Vid2Seq: Large-Scale Pretraining of a Visual Language Model for Dense Video Captioning ...	10714
<i>Antoine Yang (PSL Research University), Arsha Nagrai (Google Research), Paul Hongsuck Seo (Google Research), Antoine Miech (DeepMind), Jordi Pont-Tuset (Google Research), Ivan Laptev (PSL Research University), Josef Sivic (Czech Technical University, Prague), and Cordelia Schmid (Google Research)</i>	

Procedure-Aware Pretraining for Instructional Video Understanding	10727
<i>Honglu Zhou (Salesforce Research; Rutgers University), Roberto Martín-Martín (Salesforce Research; UT Austin), Mubbasir Kapadia (Rutgers University), Silvio Savarese (Salesforce Research), and Juan Carlos Niebles (Salesforce Research)</i>	
VindLU: A Recipe for Effective Video-and-Language Pretraining	10739
<i>Feng Cheng (NC Chapel Hill), Xizi Wang (Indiana University), Jie Lei (NC Chapel Hill), David Crandall (Indiana University), Mohit Bansal (NC Chapel Hill), and Gedas Bertasius (NC Chapel Hill)</i>	
Modular Memorability: Tiered Representations for Video Memorability Prediction	10751
<i>Théo Dumont (Mines Paris, PSL - Research University Paris, France), Juan Segundo Hevia (Memorable AI Boston, USA), and Camilo L. Fosco (Memorable AI Boston, USA)</i>	
Multivariate, Multi-Frequency and Multimodal: Rethinking Graph Neural Networks for Emotion Recognition in Conversation	10761
<i>Feiyu Chen (University of Electronic Science and Technology of China, China; Sichuan Artificial Intelligence Research Institute, China), Jie Shao (University of Electronic Science and Technology of China, China; Sichuan Artificial Intelligence Research Institute, China), Shuyuan Zhu (University of Electronic Science and Technology of China, China), and Heng Tao Shen (University of Electronic Science and Technology of China, China; Sichuan Artificial Intelligence Research Institute, China)</i>	
Distilling Cross-Temporal Contexts for Continuous Sign Language Recognition	10771
<i>Leming Guo (Tianjin University of Technology), Wanli Xue (Tianjin University of Technology), Qing Guo (Centre for Frontier AI Research (CFAR), A*STAR, Singapore), Bo Liu (Walmart Global Tech, USA), Kaihua Zhang (Nanjing University of Information Science and Technology), Tiantian Yuan (Tianjin University of Technology), and Shengyong Chen (Tianjin University of Technology)</i>	
You Need Multiple Exiting: Dynamic Early Exiting for Accelerating Unified Vision Language Model	10781
<i>Shengkun Tang (North Carolina State University, USA), Yaqing Wang (Google Research, USA), Zhenglun Kong (Northeastern University, USA), Tianchi Zhang (University of Michigan, USA), Yao Li (The University of North Carolina at Chapel Hill, USA), Caiwen Ding (University of Connecticut, USA), Yanzhi Wang (Northeastern University, USA), Yi Liang (Google Research, USA), and Dongkuan Xu (North Carolina State University, USA)</i>	
Layout-Based Causal Inference for Object Navigation	10792
<i>Sixian Zhang (Institute of Computing Technology, Beijing; University of Chinese Academy of Sciences, Beijing), Xinhang Song (Institute of Computing Technology, Beijing; University of Chinese Academy of Sciences, Beijing), Weijie Li (Institute of Computing Technology, Beijing; University of Chinese Academy of Sciences, Beijing), Yubing Bai (Institute of Computing Technology, Beijing; University of Chinese Academy of Sciences, Beijing), Xinyao Yu (Institute of Computing Technology, Beijing; University of Chinese Academy of Sciences, Beijing), and Shuqiang Jiang (Institute of Computing Technology, Beijing; University of Chinese Academy of Sciences, Beijing)</i>	

Improving Vision-and-Language Navigation by Generating Future-View Image Semantics ...	10803
<i>Jialu Li (UNC Chapel Hill) and Mohit Bansal (UNC Chapel Hill)</i>	
A New Path: Scaling Vision-and-Language Navigation With Synthetic Instructions and Imitation Learning	10813
<i>Aishwarya Kamath (New York University), Peter Anderson (Google Research), Su Wang (Google Research), Jing Yu Koh (Carnegie Mellon University), Alexander Ku (Google Research), Austin Waters (Google Research), Yinfei Yang (Apple), Jason Baldridge (Google Research), and Zarana Parekh (Google Research)</i>	
A-Cap: Anticipation Captioning With Commonsense Knowledge	10824
<i>Duc Minh Vo (The University of Tokyo, Japan), Quoc-An Luong (The Graduate University for Advanced Studies, Japan), Akihiro Sugimoto (National Institute of Informatics, Japan), and Hideki Nakayama (The University of Tokyo, Japan)</i>	
Are Deep Neural Networks SMARTer Than Second Graders?	10834
<i>Anoop Cherian (Mitsubishi Electric Research Labs (MERL)), Kuan-Chuan Peng (Mitsubishi Electric Research Labs (MERL)), Suhas Lohit (Mitsubishi Electric Research Labs (MERL)), Kevin A. Smith (Massachusetts Institute of Technology (MIT)), and Joshua B. Tenenbaum (Massachusetts Institute of Technology (MIT))</i>	
Fusing Pre-Trained Language Models With Multimodal Prompts Through Reinforcement Learning	10845
<i>Youngjae Yu (Yonsei University), Jiwan Chung (Yonsei University), Heeseung Yun (Seoul National University), Jack Hessel (Allen Institute for Artificial Intelligence), Jae Sung Park (Allen Institute for Artificial Intelligence; University of Washington), Ximing Lu (Allen Institute for Artificial Intelligence; University of Washington), Rowan Zellers (OpenAI), Prithviraj Ammanabrolu (Allen Institute for Artificial Intelligence), Ronan Le Bras (Allen Institute for Artificial Intelligence), Gunhee Kim (Seoul National University), and Yejin Choi (Allen Institute for Artificial Intelligence; University of Washington)</i>	
Language Adaptive Weight Generation for Multi-Task Visual Grounding	10857
<i>Wei Su (Zhejiang University), Peihan Miao (Zhejiang University), Huanzhang Dou (Zhejiang University), Gaoang Wang (Zhejiang University-University of Illinois Urbana-Champaign Institute, Zhejiang University), Liang Qiao (Zhejiang University; Hikvision Research Institute), Zheyang Li (Zhejiang University; Hikvision Research Institute), and Xi Li (Zhejiang University; Shanghai AI Laboratory; Shanghai Institute for Advanced Study of Zhejiang University)</i>	
From Images to Textual Prompts: Zero-Shot Visual Question Answering With Frozen Large Language Models	10867
<i>Jiaxian Guo (The University of Sydney), Junnan Li (Salesforce Research), Dongxu Li (Salesforce Research), Anthony Meng Huat Tiong (Salesforce Research; Nanyang Technological University), Boyang Li (Nanyang Technological University), Dacheng Tao (The University of Sydney), and Steven Hoi (Salesforce Research)</i>	

Diversity-Aware Meta Visual Prompting	10878
<i>Qidong Huang (University of Science and Technology of China), Xiaoyi Dong (University of Science and Technology of China), Dongdong Chen (Microsoft Cloud AI), Weiming Zhang (University of Science and Technology of China), Feifei Wang (University of Science and Technology of China), Gang Hua (Wormpex AI Research), and Nenghai Yu (University of Science and Technology of China)</i>	
Hierarchical Prompt Learning for Multi-Task Learning	10888
<i>Yajing Liu (JD Logistics), Yuning Lu (University of Science and Technology of China), Hao Liu (JD Logistics), Yaozu An (JD Logistics), Zhuoran Xu (JD Logistics), Zhuokun Yao (JD Logistics), Baofeng Zhang (JD Logistics), Zhiwei Xiong (University of Science and Technology of China), and Chenguang Gui (JD Logistics)</i>	
Task Residual for Tuning Vision-Language Models	10899
<i>Tao Yu (National University of Singapore; University of Science and Technology of China), Zhihe Lu (National University of Singapore), Xin Jin (National University of Singapore; University of Science and Technology of China), Zhibo Chen (University of Science and Technology of China), and Xinchao Wang (National University of Singapore)</i>	
CREPE: Can Vision-Language Foundation Models Reason Compositionally?	10910
<i>Zixian Ma (Stanford University), Jerry Hong (Stanford University), Mustafa Omer Gul (Cornell University), Mona Gandhi (University of Pennsylvania), Irena Gao (Stanford University), and Ranjay Krishna (University of Washington)</i>	
LOCATE: Localize and Transfer Object Parts for Weakly Supervised Affordance Grounding ..	10922
<i>Gen Li (University of Edinburgh), Varun Jampani (Google Research), Deqing Sun (Google Research), and Laura Sevilla-Lara (University of Edinburgh)</i>	
Overlooked Factors in Concept-Based Explanations: Dataset Choice, Concept Learnability, and Human Capability	10932
<i>Vikram V. Ramaswamy (Princeton University), Sunnie S. Y. Kim (Princeton University), Ruth Fong (Princeton University), and Olga Russakovsky (Princeton University)</i>	
Grounding Counterfactual Explanation of Image Classifiers to Textual Concept Space	10942
<i>Siwon Kim (Data Science and Artificial Intelligence Lab, Seoul National University, Korea), Jinhoh Oh (Amazon Alexa AI, USA), Sungjin Lee (Amazon Alexa AI, USA), Seunghak Yu (NAVER Search US), Jaeyoung Do (Amazon Alexa AI, USA), and Tara Taghavi (Amazon Alexa AI, USA)</i>	
GIVL: Improving Geographical Inclusivity of Vision-Language Models With Pre-Training Methods	10951
<i>Da Yin (University of California, Los Angeles), Feng Gao (Amazon Alexa AI), Govind Thattai (Amazon Alexa AI), Michael Johnston (Amazon Alexa AI), and Kai-Wei Chang (University of California, Los Angeles; Amazon Alexa AI)</i>	
Learning Bottleneck Concepts in Image Classification	10962
<i>Bowen Wang (Osaka University, Japan), Liangzhi Li (Osaka University, Japan), Yuta Nakashima (Osaka University, Japan), and Hajime Nagahara (Osaka University, Japan)</i>	

SceneTrilogy: On Human Scene-Sketch and Its Complementarity With Photo and Text	10972
<i>Pinaki Nath Chowdhury (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Ayan Kumar Bhunia (University of Surrey, United Kingdom), Aneeshan Sain (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Subhadeep Koley (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), Tao Xiang (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence), and Yi-Zhe Song (University of Surrey, United Kingdom; FlyTek-Surrey Joint Research Centre on Artificial Intelligence)</i>	
Context-Aware Alignment and Mutual Masking for 3D-Language Pre-Training	10984
<i>Zhao Jin (Sichuan University), Munawar Hayat (Monash University), Yuwei Yang (Sichuan University), Yulan Guo (Sun Yat-sen University), and Yinjie Lei (Sichuan University)</i>	
MaskCLIP: Masked Self-Distillation Advances Contrastive Language-Image Pretraining	10995
<i>Xiaoyi Dong (University of Science and Technology of China), Jianmin Bao (Microsoft Research Asia), Yinglin Zheng (Xiamen University), Ting Zhang (University of Science and Technology of China), Dongdong Chen (Microsoft Cloud + AI), Hao Yang (Microsoft Research Asia), Ming Zeng (Xiamen University), Weiming Zhang (University of Science and Technology of China), Lu Yuan (Microsoft Cloud + AI), Dong Chen (Microsoft Research Asia), Fang Wen (Microsoft Research Asia), and Nenghai Yu (University of Science and Technology of China)</i>	
CLIPPO: Image-and-Language Understanding From Pixels Only	11006
<i>Michael Tschannen (Google Research, Brain Team, Zurich), Basil Mustafa (Google Research, Brain Team, Zurich), and Neil Houlsby (Google Research, Brain Team, Zurich)</i>	
ViLEM: Visual-Language Error Modeling for Image-Text Retrieval	11018
<i>Yuxin Chen (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; ARC Lab; University of Chinese Academy of Sciences), Zongyang Ma (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; ARC Lab; University of Chinese Academy of Sciences), Ziqi Zhang (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Zhongang Qi (ARC Lab), Chunfeng Yuan (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences), Ying Shan (ARC Lab), Bing Li (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences), Weiming Hu (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; CAS Center for Excellence in Brain Science and Intelligence Technology), Xiaohu Qie (Tencent PCG), and Jianping Wu (Tsinghua University)</i>	
Non-Contrastive Learning Meets Language-Image Pre-Training	11028
<i>Jinghao Zhou (Microsoft), Li Dong (Microsoft), Zhe Gan (Microsoft), Lijuan Wang (Microsoft), and Furu Wei (Microsoft)</i>	

HAAV: Hierarchical Aggregation of Augmented Views for Image Captioning	11039
<i>Chia-Wen Kuo (Georgia Tech) and Zsolt Kira (Georgia Tech)</i>	
Learning Attribute and Class-Specific Representation Duet for Fine-Grained Fashion Analysis	11050
<i>Yang Jiao (Amazon), Yan Gao (Amazon), Jingjing Meng (Amazon), Jin Shang (Amazon), and Yi Sun (Amazon)</i>	
Learning Instance-Level Representation for Large-Scale Multi-Modal Pretraining in E-Commerce	11060
<i>Yang Jin (Peking University; ByteDance Inc.), Yongzhi Li (ByteDance Inc.), Zehuan Yuan (ByteDance Inc.), and Yadong Mu (Peking University)</i>	
Cross-Image-Attention for Conditional Embeddings in Deep Metric Learning	11070
<i>Dmytro Kotovenko (LMU Munich, IWR Heidelberg, MCML), Pingchuan Ma (LMU Munich, IWR Heidelberg, MCML), Timo Milbich (LMU Munich, IWR Heidelberg, MCML), and Björn Ommer (LMU Munich, IWR Heidelberg, MCML)</i>	
Asymmetric Feature Fusion for Image Retrieval	11082
<i>Hui Wu (University of Science and Technology of China), Min Wang (Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Wengang Zhou (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Zhenbo Lu (Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), and Houqiang Li (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center)</i>	
Improving Zero-Shot Generalization and Robustness of Multi-Modal Models	11093
<i>Yunhao Ge (Google Research; University of Southern California), Jie Ren (Google Research), Andrew Gallagher (Google Research), Yuxiao Wang (Google Research), Ming-Hsuan Yang (Google Research), Hartwig Adam (Google Research), Laurent Itti (University of Southern California), Balaji Lakshminarayanan (Google Research), and Jiaping Zhao (Google Research)</i>	
Hint-Aug: Drawing Hints From Foundation Vision Transformers Towards Boosted Few-Shot Parameter-Efficient Tuning	11102
<i>Zhongzhi Yu (Georgia Institute of Technology), Shang Wu (Rice University), Yonggan Fu (Georgia Institute of Technology), Shunyao Zhang (Rice University), and Yingyan Celine Lin (Georgia Institute of Technology)</i>	
Visual DNA: Representing and Comparing Images Using Distributions of Neuron Activations	11113
<i>Benjamin Ramtoula (Mobile Robotics Group, University of Oxford), Matthew Gadd (Mobile Robotics Group, University of Oxford), Paul Newman (Mobile Robotics Group, University of Oxford), and Daniele De Martini (Mobile Robotics Group, University of Oxford)</i>	
End-to-End 3D Dense Captioning With Vote2Cap-DETR	11124
<i>Sijin Chen (Fudan University), Hongyuan Zhu (A*STAR, Singapore), Xin Chen (Tencent PCG), Yinjie Lei (Sichuan University), Gang Yu (Tencent PCG), and Tao Chen (Fudan University)</i>	

Improving Table Structure Recognition With Visual-Alignment Sequential Coordinate Modeling.....	11134
<i>Yongshuai Huang (Huawei Technologies Ltd.), Ning Lu (Huawei Technologies Ltd.), Dapeng Chen (Huawei Technologies Ltd.), Yibo Li (Peking University), Zecheng Xie (Huawei Technologies Ltd.), Shenggao Zhu (Huawei Technologies Ltd.), Liangcai Gao (Peking University), and Wei Peng (Huawei Technologies Ltd.)</i>	
Region-Aware Pretraining for Open-Vocabulary Object Detection With Vision Transformers	11144
<i>Dahun Kim (Google Research, Brain Team), Anelia Angelova (Google Research, Brain Team), and Weicheng Kuo (Google Research, Brain Team)</i>	
Mobile User Interface Element Detection via Adaptively Prompt Tuning	11155
<i>Zhangxuan Gu (Tiansuan Lab, Ant Group), Zhuoer Xu (Tiansuan Lab, Ant Group), Haoxing Chen (Tiansuan Lab, Ant Group), Jun Lan (Tiansuan Lab, Ant Group), Changhua Meng (Tiansuan Lab, Ant Group), and Weiqiang Wang (Tiansuan Lab, Ant Group)</i>	
Learning To Generate Text-Grounded Mask for Open-World Semantic Segmentation From Only Image-Text Pairs	11165
<i>Junbum Cha (Kakao Brain), Jonghwan Mun (Kakao Brain), and Byungseok Roh (Kakao Brain)</i>	
ZegCLIP: Towards Adapting CLIP for Zero-Shot Semantic Segmentation	11175
<i>Ziqin Zhou (The University of Adelaide, Australia), Yinjie Lei (Sichuan University, China), Bowen Zhang (The University of Adelaide, Australia), Lingqiao Liu (The University of Adelaide, Australia), and Yifan Liu (The University of Adelaide, Australia)</i>	
Object-Aware Distillation Pyramid for Open-Vocabulary Object Detection	11186
<i>Luting Wang (Beihang University; Beihang University), Yi Liu (Beihang University; Beihang University), Penghui Du (Beihang University; Beihang University), Zihan Ding (Beihang University; Beihang University), Yue Liao (Beihang University; Beihang University), Qiaosong Qi (Alibaba Group), Biaolong Chen (Alibaba Group), and Si Liu (Beihang University; Beihang University)</i>	
Learning Conditional Attributes for Compositional Zero-Shot Learning	11197
<i>Qingsheng Wang (Northwestern Polytechnical University, China), Lingqiao Liu (University of Adelaide, Australia), Chenchen Jing (Zhejiang University, China), Hao Chen (Zhejiang University, China), Guoqiang Liang (Northwestern Polytechnical University, China), Peng Wang (Northwestern Polytechnical University, China), and Chunhua Shen (Zhejiang University, China)</i>	
CLIP-S4: Language-Guided Self-Supervised Semantic Segmentation	11207
<i>Wenbin He (Bosch Research North America & Bosch Center for Artificial Intelligence (BCAI)), Suphanut Jamonnak (Bosch Research North America & Bosch Center for Artificial Intelligence (BCAI)), Liang Gou (Bosch Research North America & Bosch Center for Artificial Intelligence (BCAI)), and Liu Ren (Bosch Research North America & Bosch Center for Artificial Intelligence (BCAI))</i>	

StructVPR: Distill Structural Knowledge With Weighting Samples for Visual Place Recognition	11217
<i>Yanqing Shen (Xi'an Jiaotong University), Sanping Zhou (Xi'an Jiaotong University), Jingwen Fu (Xi'an Jiaotong University), Ruotong Wang (Xi'an Jiaotong University), Shitao Chen (Xi'an Jiaotong University), and Nanning Zheng (Xi'an Jiaotong University)</i>	
UniDAformer: Unified Domain Adaptive Panoptic Segmentation Transformer via Hierarchical Mask Calibration	11227
<i>Jingyi Zhang (Nanyang Technological University), Jiaying Huang (Nanyang Technological University), Xiaoqin Zhang (Wenzhou University), and Shijian Lu (Nanyang Technological University)</i>	
Primitive Generation and Semantic-Related Alignment for Universal Zero-Shot Segmentation	11238
<i>Shuting He (Zhejiang University), Henghui Ding (Nanyang Technological University), and Wei Jiang (Zhejiang University; Nanyang Technological University)</i>	
Inferring and Leveraging Parts From Object Shape for Improving Semantic Image Synthesis	11248
<i>Yuxiang Wei (Harbin Institute of Technology; The Hong Kong Polytechnic University), Zhilong Ji (Tomorrow Advancing Life), Xiaohe Wu (Harbin Institute of Technology), Jinfeng Bai (Harbin Institute of Technology), Lei Zhang (The Hong Kong Polytechnic University), and Wangmeng Zuo (Harbin Institute of Technology)</i>	
Compositor: Bottom-Up Clustering and Compositing for Robust Part and Object Segmentation	11259
<i>Ju He (Johns Hopkins University), Jieneng Chen (Johns Hopkins University), Ming-Xian Lin (Chinese Academy of Sciences), Qihang Yu (Johns Hopkins University), and Alan L. Yuille (Johns Hopkins University)</i>	
A Strong Baseline for Generalized Few-Shot Semantic Segmentation	11269
<i>Sina Hajimiri (ETS Montreal), Malik Boudiaf (ETS Montreal), Ismail Ben Ayed (ETS Montreal), and Jose Dolz (ETS Montreal)</i>	
DynaMask: Dynamic Mask Selection for Instance Segmentation	11279
<i>Ruihuang Li (The Hong Kong Polytechnic University), Chenhang He (The Hong Kong Polytechnic University), Shuai Li (The Hong Kong Polytechnic University), Yabin Zhang (The Hong Kong Polytechnic University), and Lei Zhang (The Hong Kong Polytechnic University)</i>	
Focus on Details: Online Multi-Object Tracking With Diverse Fine-Grained Representation ..	11289
<i>Hao Ren (Huazhong University of Science and Technology), Shoudong Han (Huazhong University of Science and Technology), Huilin Ding (Huazhong University of Science and Technology), Ziwen Zhang (Huazhong University of Science and Technology), Hongwei Wang (Huazhong University of Science and Technology), and Faquan Wang (Huazhong University of Science and Technology)</i>	

Dynamic Focus-Aware Positional Queries for Semantic Segmentation	11299
<i>Haoyu He (ZIP Lab, Monash University), Jianfei Cai (ZIP Lab, Monash University), Zizheng Pan (ZIP Lab, Monash University), Jing Liu (The University of Sydney), Jing Zhang (The University of Sydney), Dacheng Tao (ZIP Lab, Monash University), and Bohan Zhuang (ZIP Lab, Monash University)</i>	
Beyond mAP: Towards Better Evaluation of Instance Segmentation	11309
<i>Rohit Jena (University of Pennsylvania), Lukas Zhornyak (University of Pennsylvania), Nehal Doiphode (University of Pennsylvania), Pratik Chaudhari (University of Pennsylvania), Vivek Buch (Stanford University), James Gee (University of Pennsylvania), and Jianbo Shi (University of Pennsylvania)</i>	
Learning Orthogonal Prototypes for Generalized Few-Shot Semantic Segmentation	11319
<i>Sun-Ao Liu (University of Science and Technology of China), Yiheng Zhang (University of Science and Technology of China), Zhaofan Qiu (University of Science and Technology of China), Hongtao Xie (University of Science and Technology of China), Yongdong Zhang (HiDream.ai Inc.), and Ting Yao (HiDream.ai Inc.)</i>	
Weakly Supervised Semantic Segmentation via Adversarial Learning of Classifier and Reconstructor	11329
<i>Hyeokjun Kweon (Visual Intelligence Lab., KAIST, Korea), Sung-Hoon Yoon (Visual Intelligence Lab., KAIST, Korea), and Kuk-Jin Yoon (Visual Intelligence Lab., KAIST, Korea)</i>	
SemiCVT: Semi-Supervised Convolutional Vision Transformer for Semantic Segmentation ..	11340
<i>Huimin Huang (Zhejiang University), Shiao Xie (Zhejiang University), Lanfen Lin (Zhejiang University), Ruofeng Tong (Zhejiang University; Zhejiang Lab), Yen-Wei Chen (Ritsumeikan University), Yuexiang Li (Tencent Jarvis Lab), Hong Wang (Tencent Jarvis Lab), Yawen Huang (Tencent Jarvis Lab), and Yefeng Zheng (Tencent Jarvis Lab)</i>	
Augmentation Matters: A Simple-Yet-Effective Approach to Semi-Supervised Semantic Segmentation	11350
<i>Zhen Zhao (University of Sydney), Lihe Yang (Nanjing University), Sifan Long (Jilin University), Jimin Pi (Baidu Inc.), Luping Zhou (University of Sydney), and Jingdong Wang (Baidu Inc.)</i>	
The Devil Is in the Points: Weakly Semi-Supervised Instance Segmentation via Point-Guided Mask Representation	11360
<i>Beomyoung Kim (NAVER Cloud, ImageVision; KAIST), Joonhyun Jeong (NAVER Cloud, ImageVision; KAIST), Dongyoon Han (NAVER AI Lab), and Sung Ju Hwang (KAIST)</i>	
Class-Incremental Exemplar Compression for Class-Incremental Learning	11371
<i>Zilin Luo (Singapore Management University), Yaoyao Liu (Max Planck Institute for Informatics, Saarland Informatics Campus), Bernt Schiele (Max Planck Institute for Informatics, Saarland Informatics Campus), and Qianru Sun (Singapore Management University)</i>	

Full or Weak Annotations? An Adaptive Strategy for Budget-Constrained Annotation Campaigns.... 11381	
<i>Javier Gamazo Tejero (University of Bern), Martin S. Zinkernagel (Inselspital Bern, Switzerland), Sebastian Wolf (Inselspital Bern, Switzerland), Raphael Sznitman (University of Bern), and Pablo Márquez-Neila (University of Bern)</i>	
Learning Common Rationale To Improve Self-Supervised Representation for Fine-Grained Visual Recognition Problems	11392
<i>Yangyang Shu (The University of Adelaide), Anton van den Hengel (The University of Adelaide), and Lingqiao Liu (The University of Adelaide)</i>	
Detection Hub: Unifying Object Detection Datasets via Query Adaptation on Language Embedding	11402
<i>Lingchen Meng (Shanghai Key Lab of Intell. Info. Processing, Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Xiyang Dai (Microsoft), Yinpeng Chen (Microsoft), Pengchuan Zhang (Microsoft), Dongdong Chen (Microsoft), Mengchen Liu (Microsoft), Jianfeng Wang (Microsoft), Zuxuan Wu (Shanghai Key Lab of Intell. Info. Processing, Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Lu Yuan (Microsoft), and Yu-Gang Jiang (Shanghai Key Lab of Intell. Info. Processing, Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing)</i>	
Self-Supervised AutoFlow	11412
<i>Hsin-Ping Huang (Google Research; University of California, Merced), Charles Herrmann (Google Research), Junhwa Hur (Google Research), Erika Lu (Google Research), Kyle Sargent (Google Research), Austin Stone (Google Research), Ming-Hsuan Yang (Google Research; University of California, Merced), and Deqing Sun (Google Research)</i>	
DETR With Additional Global Aggregation for Cross-Domain Weakly Supervised Object Detection	11422
<i>Zongheng Tang (Beihang University; Baidu Inc), Yifan Sun (Baidu Inc), Si Liu (Beihang University), and Yi Yang (Zhejiang University)</i>	
Detecting Everything in the Open World: Towards Universal Object Detection	11433
<i>Zhenyu Wang (Tsinghua University; Beijing National Research Center for Information Science and Technology (BNRist)), Yali Li (Tsinghua University; Beijing National Research Center for Information Science and Technology (BNRist)), Xi Chen (The University of Hong Kong), Ser-Nam Lim (Meta AI), Antonio Torralba (Massachusetts Institute of Technology), Hengshuang Zhao (The University of Hong Kong), and Shengjin Wang (Tsinghua University; Beijing National Research Center for Information Science and Technology (BNRist))</i>	
PROB: Probabilistic Objectness for Open World Object Detection	11444
<i>Orr Zohar (Stanford University), Kuan-Chieh Wang (Stanford University), and Serena Yeung (Stanford University)</i>	

Annealing-Based Label-Transfer Learning for Open World Object Detection	11454
<i>Yuqing Ma (SKLSDE Lab, Beihang University), Hainan Li (Peking University), Zhange Zhang (SKLSDE Lab, Beihang University), Jinyang Guo (SKLSDE Lab, Beihang University), Shanghang Zhang (Peking University), Ruihao Gong (SKLSDE Lab, Beihang University), and Xianglong Liu (SKLSDE Lab, Beihang University; Zhongguancun Laboratory; Institute of Data Space, Hefei Comprehensive National Science Center)</i>	
Learning Transformation-Predictive Representations for Detection and Description of Local Features	11464
<i>Zihao Wang (Peking University), Chunxu Wu (School of Automation, Beijing Institute of Technology), Yifei Yang (School of Automation, Beijing Institute of Technology), and Zhen Li (School of Automation, Beijing Institute of Technology)</i>	
Bridging Precision and Confidence: A Train-Time Loss for Calibrating Object Detection	11474
<i>Muhammad Akhtar Munir (Mohamed bin Zayed University of AI; Information Technology University), Muhammad Haris Khan (Mohamed bin Zayed University of AI), Salman Khan (Mohamed bin Zayed University of AI; Australian National University), and Fahad Shahbaz Khan (Mohamed bin Zayed University of AI; Linkoping University)</i>	
2PCNet: Two-Phase Consistency Training for Day-to-Night Unsupervised Domain Adaptive Object Detection	11484
<i>Mikhail Kennerley (National University of Singapore, Department of Electrical and Computer Engineering; Institute for Infocomm Research, A*STAR), Jian-Gang Wang (National University of Singapore, Department of Electrical and Computer Engineering), Bharadwaj Veeravalli (Institute for Infocomm Research, A*STAR), and Robby T. Tan (National University of Singapore, Department of Electrical and Computer Engineering)</i>	
Zero-Shot Generative Model Adaptation via Image-Specific Prompt Learning	11494
<i>Jiayi Guo (Tsinghua University, BNRist), Chaofei Wang (Tsinghua University, BNRist), You Wu (UCAS), Eric Zhang (SHI Labs @ Oregon & UIUC), Kai Wang (SHI Labs @ Oregon & UIUC), Xingqian Xu (SHI Labs @ Oregon & UIUC), Shiji Song (Tsinghua University, BNRist), Humphrey Shi (Tsinghua University, BNRist), and Gao Huang (Tsinghua University, BNRist)</i>	
AutoLabel: CLIP-Based Framework for Open-Set Video Domain Adaptation	11504
<i>Giacomo Zara (University of Trento, Italy), Subhankar Roy (Institut polytechnique de Paris, France), Paolo Rota (University of Trento, Italy), and Elisa Ricci (University of Trento, Italy; Fondazione Bruno Kessler, Italy)</i>	
Bidirectional Copy-Paste for Semi-Supervised Medical Image Segmentation	11514
<i>Yunhao Bai (East China Normal University), Duowen Chen (East China Normal University), Qingli Li (East China Normal University), Wei Shen (Shanghai Jiao Tong University), and Yan Wang (East China Normal University)</i>	
Directional Connectivity-Based Segmentation of Medical Images	11525
<i>Ziyun Yang (Duke University, United States) and Sina Farsiu (Duke University, United States)</i>	

Ambiguous Medical Image Segmentation Using Diffusion Models	11536
<i>Aimon Rahman (Johns Hopkins University), Jeya Maria Jose Valanarasu (Johns Hopkins University), Ilker Hacihaliloglu (University of British Columbia), and Vishal M. Patel (Johns Hopkins University)</i>	
Sparse Multi-Modal Graph Transformer With Shared-Context Processing for Representation Learning of Giga-Pixel Images	11547
<i>Ramin Nakhli (University of British Columbia), Puria Azadi Moghadam (University of British Columbia), Haoyang Mi (University of British Columbia), Hossein Farahani (University of British Columbia), Alexander Baras (University of British Columbia), Blake Gilks (University of British Columbia), and Ali Bashashati (University of British Columbia)</i>	
METransformer: Radiology Report Generation by Transformer With Multiple Learnable Expert Tokens	11558
<i>Zhanyu Wang (University of Sydney), Lingqiao Liu (University of Adelaide), Lei Wang (University of Wollongong), and Luping Zhou (University of Sydney)</i>	
Towards Trustable Skin Cancer Diagnosis via Rewriting Model's Decision	11568
<i>Siyuan Yan (Monash University; Monash Medical AI Group), Zhen Yu (Monash University; Monash Medical AI Group), Xuelin Zhang (Monash University; Monash Medical AI Group), Dwarikanath Mahapatra (Inception Institute of AI, UAE), Shekhar S. Chandra (The University of Queensland), Monika Janda (The University of Queensland), Peter Soyer (The University of Queensland), and Zongyuan Ge (Monash University; Monash Medical AI Group)</i>	
Rethinking Out-of-Distribution (OOD) Detection: Masked Image Modeling Is All You Need ...	11578
<i>Jingyao Li (The Chinese University of Hong Kong), Pengguang Chen (SmartMore), Zexin He (The Chinese University of Hong Kong), Shaozuo Yu (The Chinese University of Hong Kong), Shu Liu (SmartMore), and Jiaya Jia (The Chinese University of Hong Kong; SmartMore)</i>	
MetaViewer: Towards a Unified Multi-View Representation	11590
<i>Ren Wang (Shandong University), Haoliang Sun (Shandong University), Yuling Ma (Shandong Jianzhu University), Xiaoming Xi (Shandong Jianzhu University), and Yilong Yin (Shandong University)</i>	
Deep Incomplete Multi-View Clustering With Cross-View Partial Sample and Prototype Alignment	11600
<i>Jiaqi Jin (National University of Defense Technology, China), Siwei Wang (National University of Defense Technology, China), Zhibin Dong (National University of Defense Technology, China), Xinwang Liu (National University of Defense Technology, China), and En Zhu (National University of Defense Technology, China)</i>	
RONO: Robust Discriminative Learning With Noisy Labels for 2D-3D Cross-Modal Retrieval .	11610
<i>Yanglin Feng (Sichuan University), Hongyuan Zhu (Sichuan Zhiqian Technology Co., Ltd), Dezhong Peng (Sichuan University), Xi Peng (Sichuan University), and Peng Hu (Sichuan University)</i>	

Mind the Label Shift of Augmentation-Based Graph OOD Generalization	11620
<i>Junchi Yu (MAIS&CRIPAC, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Jian Liang (MAIS&CRIPAC, Institute of Automation, Chinese Academy of Sciences, China), and Ran He (MAIS&CRIPAC, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China; ShanghaiTech University, China)</i>	
Zero-Shot Model Diagnosis	11631
<i>Jinqi Luo (Carnegie Mellon University), Zhaoning Wang (Carnegie Mellon University), Chen Henry Wu (Carnegie Mellon University), Dong Huang (Carnegie Mellon University), and Fernando De la Torre (Carnegie Mellon University)</i>	
ProtoCon: Pseudo-Label Refinement via Online Clustering and Prototypical Consistency for Efficient Semi-Supervised Learning	11641
<i>Islam Nassar (Data Science and AI Department, Monash University, Australia), Munawar Hayat (Data Science and AI Department, Monash University, Australia), Ehsan Abbasnejad (Australian Institute for Machine Learning, The University of Adelaide, Australia), Hamid Rezatofighi (Data Science and AI Department, Monash University, Australia), and Gholamreza Haffari (Data Science and AI Department, Monash University, Australia)</i>	
Fine-Grained Classification With Noisy Labels	11651
<i>Qi Wei (Shandong University, China), Lei Feng (Nanyang Technological University, Singapore), Haoliang Sun (Shandong University, China), Ren Wang (Shandong University, China), Chenhui Guo (Shandong University, China), and Yilong Yin (Shandong University, China)</i>	
Twin Contrastive Learning With Noisy Labels	11661
<i>Zhizhong Huang (Fudan University, China), Junping Zhang (Fudan University, China), and Hongming Shan (Fudan University, China; Shanghai Center for Brain Science and Brain-inspired Technology, China)</i>	
RMLVQA: A Margin Loss Approach for Visual Question Answering With Language Biases	11671
<i>Abhipsa Basu (Vision and AI Lab, Indian Institute of Science, Bangalore), Sravanti Addepalli (Vision and AI Lab, Indian Institute of Science, Bangalore), and R. Venkatesh Babu (Vision and AI Lab, Indian Institute of Science, Bangalore)</i>	
Generative Bias for Robust Visual Question Answering	11681
<i>Jae Won Cho (KAIST), Dong-Jin Kim (Hanyang University), Hyeonggon Ryu (KAIST), and In So Kweon (KAIST)</i>	
On-the-Fly Category Discovery	11691
<i>Ruoyi Du (Beijing University of Posts and Telecommunications, China), Dongliang Chang (Beijing University of Posts and Telecommunications, China), Kongming Liang (Beijing University of Posts and Telecommunications, China), Timothy Hospedales (University of Edinburgh, UK), Yi-Zhe Song (University of Surrey, UK), and Zhanyu Ma (Beijing University of Posts and Telecommunications, China)</i>	

Co-Training 2L Submodels for Visual Recognition	11701
<i>Hugo Touvron (Meta AI/FAIR Paris; Sorbonne University), Matthieu Cord (Sorbonne University), Maxime Oquab (Meta AI/FAIR Paris), Piotr Bojanowski (Meta AI/FAIR Paris), Jakob Verbeek (Meta AI/FAIR Paris), and Hervé Jégou (Meta AI/FAIR Paris)</i>	
Neural Dependencies Emerging From Learning Massive Categories	11711
<i>Ruili Feng (University of Science and Technology of China, China; Alibaba Group, China), Kecheng Zheng (Ant Group), Kai Zhu (University of Science and Technology of China, China), Yujun Shen (Ant Group), Jian Zhao (University of Science and Technology of China, China), Yukun Huang (University of Science and Technology of China, China), Deli Zhao (Alibaba Group, China), Jingren Zhou (Alibaba Group, China), Michael Jordan (University of California, Berkeley), and Zheng-Jun Zha (University of Science and Technology of China, China)</i>	
MIC: Masked Image Consistency for Context-Enhanced Domain Adaptation	11721
<i>Lukas Hoyer (ETH Zurich), Dengxin Dai (Max Planck Institute for Informatics, Saarland Informatics Campus), Haoran Wang (Max Planck Institute for Informatics, Saarland Informatics Campus), and Luc Van Gool (ETH Zurich; KU Leuven)</i>	
Towards Better Stability and Adaptability: Improve Online Self-Training for Model Adaptation in Semantic Segmentation	11733
<i>Dong Zhao (Xidian University, China), Shuang Wang (Xidian University, China), Qi Zang (Xidian University, China), Dou Quan (Xidian University, China), Xiutiao Ye (Xidian University, China), and Licheng Jiao (Xidian University, China)</i>	
DARE-GRAM: Unsupervised Domain Adaptation Regression by Aligning Inverse Gram Matrices	11744
<i>Ismail Nejjar (EPFL, Switzerland), Qin Wang (ETH Zurich, Switzerland), and Olga Fink (EPFL, Switzerland)</i>	
Equiangular Basis Vectors	11755
<i>Yang Shen (Nanjing University of Science and Technology, China), Xuhao Sun (Nanjing University of Science and Technology, China), and Xiu-Shen Wei (Nanjing University of Science and Technology, China)</i>	
Enhanced Multimodal Representation Learning With Cross-Modal KD	11766
<i>Mengxi Chen (Shanghai Jiao Tong University), Linyu Xing (Shanghai Jiao Tong University), Yu Wang (Shanghai Jiao Tong University; Shanghai AI Laboratory), and Ya Zhang (Shanghai Jiao Tong University; Shanghai AI Laboratory)</i>	
Decompose, Adjust, Compose: Effective Normalization by Playing With Frequency for Domain Generalization	11776
<i>Sangrok Lee (Yonsei University), Jongseong Bae (Yonsei University), and Ha Young Kim (Yonsei University)</i>	
Back to the Source: Diffusion-Driven Adaptation To Test-Time Corruption	11786
<i>Jin Gao (Shanghai Jiao Tong University), Jialing Zhang (Shanghai Jiao Tong University), Xihui Liu (The University of Hong Kong), Trevor Darrell (University of California, Berkeley), Evan Shelhamer (DeepMind), and Dequan Wang (Shanghai Jiao Tong University; Shanghai Artificial Intelligence Laboratory)</i>	

Deep Frequency Filtering for Domain Generalization	11797
<i>Shiqi Lin (University of Science and Technology of China), Zhizheng Zhang (Microsoft), Zhipeng Huang (University of Science and Technology of China; Microsoft), Yan Lu (Microsoft), Cuiling Lan (Microsoft), Peng Chu (Microsoft), Quanzeng You (Microsoft), Jiang Wang (Microsoft), Zicheng Liu (Microsoft), Amey Parulkar (Microsoft), Viraj Navkal (Microsoft), and Zhibo Chen (Microsoft)</i>	
Generalizable Implicit Neural Representations via Instance Pattern Composers	11808
<i>Chiheon Kim (Kakao Brain), Doyup Lee (Kakao Brain), Saehoon Kim (Kakao Brain), Minsu Cho (POSTECH), and Wook-Shin Han (POSTECH)</i>	
Train-Once-for-All Personalization	11818
<i>Hong-You Chen (The Ohio State University), Yandong Li (Google Research), Yin Cui (Google Research), Mingda Zhang (Google Research), Wei-Lun Chao (The Ohio State University), and Li Zhang (Google Research)</i>	
Mod-Squad: Designing Mixtures of Experts As Modular Multi-Task Learners	11828
<i>Zitian Chen (University of Massachusetts Amherst), Yikang Shen (MIT-IBM Watson AI Lab), Mingyu Ding (The University of Hong Kong), Zhenfang Chen (MIT-IBM Watson AI Lab), Hengshuang Zhao (The University of Hong Kong), Erik G. Learned-Miller (University of Massachusetts Amherst), and Chuang Gan (University of Massachusetts Amherst; MIT-IBM Watson AI Lab)</i>	
Few-Shot Class-Incremental Learning via Class-Aware Bilateral Distillation	11838
<i>Linglan Zhao (Shanghai Jiao Tong University), Jing Lu (Hikvision Research Institute), Yunlu Xu (Hikvision Research Institute), Zhanzhan Cheng (Hikvision Research Institute), Dashan Guo (Shanghai Jiao Tong University), Yi Niu (Hikvision Research Institute), and Xiangzhong Fang (Shanghai Jiao Tong University)</i>	
Multi-Mode Online Knowledge Distillation for Self-Supervised Visual Representation Learning	11848
<i>Kaiyou Song (Megvii Technology), Jin Xie (Megvii Technology), Shan Zhang (Megvii Technology), and Zimeng Luo (Megvii Technology)</i>	
Dense Network Expansion for Class Incremental Learning	11858
<i>Zhiyuan Hu (UC San Diego), Yunsheng Li (Microsoft Cloud + AI), Jiancheng Lyu (Qualcomm AI Research), Dashan Gao (Qualcomm AI Research), and Nuno Vasconcelos (UC San Diego)</i>	
Class Attention Transfer Based Knowledge Distillation	11868
<i>Ziyao Guo (Xidian University), Haonan Yan (Xidian University; University of Guelph), Hui Li (Xidian University), and Xiaodong Lin (University of Guelph)</i>	
Dealing With Cross-Task Class Discrimination in Online Continual Learning	11878
<i>Yiduo Guo (Peking University), Bing Liu (University of Illinois Chicago), and Dongyan Zhao (Peking University; BIGAI, China; National Key Laboratory of General Artificial Intelligence)</i>	

Real-Time Evaluation in Online Continual Learning: A New Hope	11888
<i>Yasir Ghunaim (King Abdullah University of Science and Technology (KAUST)), Adel Bibi (University of Oxford), Kumail Alhamoud (King Abdullah University of Science and Technology (KAUST)), Motasem Alfarra (King Abdullah University of Science and Technology (KAUST)), Hasan Abed Al Kader Hammoud (King Abdullah University of Science and Technology (KAUST)), Ameya Prabhu (University of Oxford), Philip H.S. Torr (University of Oxford), and Bernard Ghanem (King Abdullah University of Science and Technology (KAUST))</i>	
DisWOT: Student Architecture Search for Distillation WithOut Training	11898
<i>Peijie Dong (National University of Defense Technology), Lujun Li (Chinese Academy of Sciences), and Zimian Wei (National University of Defense Technology)</i>	
CODA-Prompt: COntinual Decomposed Attention-Based Prompting for Rehearsal-Free Continual Learning	11909
<i>James Seale Smith (Georgia Institute of Technology; MIT-IBM Watson AI Lab), Leonid Karlinsky (MIT-IBM Watson AI Lab; IBM Research), Vyshnavi Gutta (Georgia Institute of Technology), Paola Cascante-Bonilla (MIT-IBM Watson AI Lab; Rice University), Donghyun Kim (MIT-IBM Watson AI Lab; IBM Research), Assaf Arbelle (IBM Research), Rameswar Panda (MIT-IBM Watson AI Lab; IBM Research), Rogerio Feris (MIT-IBM Watson AI Lab; IBM Research), and Zsolt Kira (Georgia Institute of Technology)</i>	
EcoTTA: Memory-Efficient Continual Test-Time Adaptation via Self-Distilled Regularization ..	11920
<i>Junha Song (Qualcomm AI Research; KAIST), Jungsoo Lee (Qualcomm AI Research), In So Kweon (KAIST), and Sungha Choi (Qualcomm AI Research)</i>	
Achieving a Better Stability-Plasticity Trade-Off via Auxiliary Networks in Continual Learning	11930
<i>Sanghwan Kim (Zürich, Switzerland), Lorenzo Noci (Zürich, Switzerland), Antonio Orvieto (Zürich, Switzerland), and Thomas Hofmann (Zürich, Switzerland)</i>	
PA&DA: Jointly Sampling Path and Data for Consistent NAS	11940
<i>Shun Lu (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Yu Hu (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Longxing Yang (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Zihao Sun (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Jilin Mei (Institute of Computing Technology, Chinese Academy of Sciences), Jianchao Tan (Kuaishou Technology), and Chengru Song (Kuaishou Technology)</i>	
Accelerating Dataset Distillation via Model Augmentation	11950
<i>Lei Zhang (Zhejiang University), Jie Zhang (Zhejiang University), Bowen Lei (Texas A&M University), Subhabrata Mukherjee (Microsoft Research), Xiang Pan (New York University), Bo Zhao (Beijing Academy of Artificial Intelligence), Caiwen Ding (University of Connecticut), Yao Li (University of North Carolina, Chapel Hill), and Dongkuan Xu (North Carolina State University)</i>	

Multi-Agent Automated Machine Learning	11960
<i>Zhaozhi Wang (Peking University; Peng Cheng Lab; University of Chinese Academy of Sciences), Kefan Su (Peking University), Jian Zhang (Huawei), Huizhu Jia (Peking University), Qixiang Ye (Peng Cheng Lab; University of Chinese Academy of Sciences), Xiaodong Xie (Peking University), and Zongqing Lu (Peking University)</i>	
Transformer-Based Learned Optimization	11970
<i>Erik Gärtner (Google Research; Lund University), Luke Metz (Google Research), Mykhaylo Andriluka (Google Research), C. Daniel Freeman (Google Research), and Cristian Sminchisescu (Google Research)</i>	
Solving Relaxations of MAP-MRF Problems: Combinatorial In-Face Frank-Wolfe Directions ..	11980
<i>Vladimir Kolmogorov (Institute of Science and Technology Austria (ISTA), Austria)</i>	
HOTNAS: Hierarchical Optimal Transport for Neural Architecture Search	11990
<i>Jiechao Yang (Renmin University of China, China; Beijing Key Laboratory of Big Data Management and Analysis Methods), Yong Liu (Renmin University of China, China; Beijing Key Laboratory of Big Data Management and Analysis Methods), and Hongteng Xu (Renmin University of China, China; Beijing Key Laboratory of Big Data Management and Analysis Methods)</i>	
Disentangled Representation Learning for Unsupervised Neural Quantization	12001
<i>Haechan Noh (Sungkyunkwan University), Sangeek Hyun (Sungkyunkwan University), Woojin Jeong (Sungkyunkwan University), Hanshin Lim (Electronics and Telecommunications Research Institute), and Jae-Pil Heo (Sungkyunkwan University)</i>	
FFCV: Accelerating Training by Removing Data Bottlenecks	12011
<i>Guillaume Leclerc (Massachusetts Institute of Technology), Andrew Ilyas (Massachusetts Institute of Technology), Logan Engstrom (Massachusetts Institute of Technology), Sung Min Park (Massachusetts Institute of Technology), Hadi Salman (Massachusetts Institute of Technology), and Aleksander Mądry (Massachusetts Institute of Technology)</i>	
Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks	12021
<i>Jierun Chen (HKUST), Shiu-hong Kao (HKUST), Hao He (HKUST), Weipeng Zhuo (HKUST), Song Wen (Rutgers University), Chul-Ho Lee (Texas State University), and S.-H. Gary Chan (HKUST)</i>	
FIANCEE: Faster Inference of Adversarial Networks via Conditional Early Exits	12032
<i>Polina Karpikova (Samsung AI Center - Moscow, Russia; Higher School of Economics - Moscow, Russia), Ekaterina Radionova (Samsung AI Center - Moscow, Russia), Anastasia Yaschenko (Samsung AI Center - Moscow, Russia; Higher School of Economics - Moscow, Russia), Andrei Spiridonov (Samsung AI Center - Moscow, Russia), Leonid Kostyushko (Lomonosov Moscow State University, Russia), Riccardo Fabbriatore (Samsung AI Center - Moscow, Russia), and Aleksei Ivakhnenko (Samsung AI Center - Moscow, Russia)</i>	
Gradient-Based Uncertainty Attribution for Explainable Bayesian Deep Learning	12044
<i>Hanjing Wang (Rensselaer Polytechnic Institute), Dhiraj Joshi (IBM Research), Shiqiang Wang (IBM Research), and Qiang Ji (Rensselaer Polytechnic Institute)</i>	

How To Prevent the Continuous Damage of Noises To Model Training?	12054
<i>Xiaotian Yu (Zhejiang University), Yang Jiang (Alibaba Group), Tianqi Shi (Alibaba Group), Zunlei Feng (Zhejiang University; Shanghai Institute for Advanced Study of Zhejiang University), Yuexuan Wang (University of Hong Kong), Mingli Song (Zhejiang University; Shanghai Institute for Advanced Study of Zhejiang University), and Li Sun (Ningbo Innovation Center Zhejiang University)</i>	
Genie: Show Me the Data for Quantization	12064
<i>Yongkweon Jeon (Samsung Research), Chungman Lee (Samsung Research), and Ho-young Kim (Samsung Research)</i>	
OpenMix: Exploring Outlier Samples for Misclassification Detection	12074
<i>Fei Zhu (MAIS, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Zhen Cheng (MAIS, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Xu-Yao Zhang (MAIS, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), and Cheng-Lin Liu (MAIS, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China)</i>	
Data-Free Sketch-Based Image Retrieval	12084
<i>Abhra Chaudhuri (University of Exeter, UK), Ayan Kumar Bhunia (University of Surrey, UK), Yi-Zhe Song (University of Surrey, UK), and Anjan Dutta (University of Surrey, UK)</i>	
GLeaD: Improving GANs With a Generator-Leading Task	12094
<i>Qingyan Bai (Tsinghua University), Ceyuan Yang (Shanghai AI Laboratory), Yinghao Xu (The Chinese University of Hong Kong), Xihui Liu (The University of Hong Kong), Yujiu Yang (Tsinghua University), and Yujun Shen (Ant Group)</i>	
Learning on Gradients: Generalized Artifacts Representation for GAN-Generated Images Detection	12105
<i>Chuangchuan Tan (Beijing Jiaotong University; Beijing Key Laboratory of Advanced Information Science and Network Technology), Yao Zhao (Beijing Jiaotong University; Beijing Key Laboratory of Advanced Information Science and Network Technology), Shikui Wei (Beijing Jiaotong University; Beijing Key Laboratory of Advanced Information Science and Network Technology), Guanghua Gu (Yanshan University; Hebei Key Laboratory of Information Transmission and Signal Processing), and Yunchao Wei (Beijing Jiaotong University; Beijing Key Laboratory of Advanced Information Science and Network Technology)</i>	
Adversarial Normalization: I Can Visualize Everything (ICE)	12115
<i>Hoyoung Choi (Hanyang University), Seungwan Jin (Hanyang University), and Kyungsik Han (Hanyang University)</i>	
Semi-Supervised Hand Appearance Recovery via Structure Disentanglement and Dual Adversarial Discrimination	12125
<i>Zimeng Zhao (Southeast University, China), Binghui Zuo (Southeast University, China), Zhiyu Long (Southeast University, China), and Yangang Wang (Southeast University, China)</i>	

Look Around for Anomalies: Weakly-Supervised Anomaly Detection via Context-Motion Relational Learning	12137
<i>MyeongAh Cho (Yonsei University), Minjung Kim (Yonsei University), Sangwon Hwang (Hyundai Motor Company), Chaewon Park (Yong In University), Kyungjae Lee (Yong In University), and Sangyoun Lee (Yonsei University)</i>	
Diversity-Measurable Anomaly Detection	12147
<i>Wenrui Liu (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Hong Chang (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Bingpeng Ma (University of Chinese Academy of Sciences, China), Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China)</i>	
Cloud-Device Collaborative Adaptation to Continual Changing Environments in the Real-World..... 12157	
<i>Yulu Gan (Peking University), Mingjie Pan (Peking University), Rongyu Zhang (The Chinese University of Hong Kong, Shenzhen), Zijian Ling (Imperial College London), Lingran Zhao (Peking University), Jiaming Liu (Peking University), and Shanghang Zhang (Peking University)</i>	
How To Prevent the Poor Performance Clients for Personalized Federated Learning?	12167
<i>Zhe Qu (Central South University), Xingyu Li (Mississippi State University), Xiao Han (University of South Florida), Rui Duan (Central South University), Chengchao Shen (Central South University), and Lixing Chen (Shanghai Jiaotong University)</i>	
DynaFed: Tackling Client Data Heterogeneity With Global Dynamics	12177
<i>Renjie Pi (The Hong Kong University of Science and Technology), Weizhong Zhang (Fudan University), Yueqi Xie (The Hong Kong University of Science and Technology), Jiahui Gao (The University of Hong Kong), Xiaoyu Wang (The Hong Kong University of Science and Technology), Sunghun Kim (The Hong Kong University of Science and Technology), and Qifeng Chen (The Hong Kong University of Science and Technology)</i>	
Elastic Aggregation for Federated Optimization	12187
<i>Dengsheng Chen (Meituan), Jie Hu (State Key Laboratory of Computer Science, ISCAS; University of Chinese Academy of Sciences), Vince Junkai Tan (Bytedance Inc.), Xiaoming Wei (Meituan), and Enhua Wu (State Key Laboratory of Computer Science, ISCAS; University of Chinese Academy of Sciences; University of Macau)</i>	
Breaching FedMD: Image Recovery via Paired-Logits Inversion Attack	12198
<i>Hideaki Takahashi (The University of Tokyo), Jingjing Liu (Tsinghua University), and Yang Liu (Tsinghua University)</i>	
Learning To Measure the Point Cloud Reconstruction Loss in a Representation Space	12208
<i>Tianxin Huang (Zhejiang University), Zhonggan Ding (Tencent YouTu Lab), Jiangning Zhang (Tencent YouTu Lab), Ying Tai (Tencent YouTu Lab), Zhenyu Zhang (Tencent YouTu Lab), Mingang Chen (Shanghai Development Center of Computer Software Technology), Chengjie Wang (Tencent YouTu Lab), and Yong Liu (Zhejiang University)</i>	

Backdoor Cleansing With Unlabeled Data	12218
<i>Lu Pang (Stony Brook University), Tao Sun (Stony Brook University), Haibin Ling (Stony Brook University), and Chao Chen (Stony Brook University)</i>	
Backdoor Defense via Deconfounded Representation Learning	12228
<i>Zaixi Zhang (University of Science and Technology of China; State Key Laboratory of Cognitive Intelligence, China), Qi Liu (University of Science and Technology of China; State Key Laboratory of Cognitive Intelligence, China), Zhicai Wang (University of Science and Technology of China), Zepu Lu (University of Science and Technology of China), and Qingyong Hu (State Key Laboratory of Cognitive Intelligence, China)</i>	
Defending Against Patch-Based Backdoor Attacks on Self-Supervised Learning	12239
<i>Ajinkya Tejankar (University of California, Davis), Maziar Sanjabi (Meta AI), Qifan Wang (Meta AI), Sinong Wang (Meta AI), Hamed Firooz (Meta AI), Hamed Pirsiavash (University of California, Davis), and Liang Tan (Meta AI)</i>	
Backdoor Attacks Against Deep Image Compression via Adaptive Frequency Trigger	12250
<i>Yi Yu (Nanyang Technological University; IGP-ROSE, NTU), Yufei Wang (Nanyang Technological University), Wenhan Yang (Peng Cheng Laboratory), Shijian Lu (Nanyang Technological University), Yap-Peng Tan (Nanyang Technological University), and Alex C. Kot (Nanyang Technological University)</i>	
CAP: Robust Point Cloud Classification via Semantic and Structural Modeling	12260
<i>Daizong Ding (Fudan University, China), Erling Jiang (Fudan University, China), Yuanmin Huang (Fudan University, China), Mi Zhang (Fudan University, China), Wenxuan Li (Fudan University, China), and Min Yang (Fudan University, China)</i>	
Evading DeepFake Detectors via Adversarial Statistical Consistency	12271
<i>Yang Hou (Kyushu University, Japan), Qing Guo (Centre for Frontier AI Research (CFAR), A*STAR, Singapore), Yihao Huang (Nanyang Technological University, Singapore), Xiaofei Xie (Singapore Management University, Singapore), Lei Ma (University of Alberta, Canada; The University of Tokyo, Japan), and Jianjun Zhao (Kyushu University, Japan)</i>	
Enhancing the Self-Universality for Transferable Targeted Attacks	12281
<i>Zhipeng Wei (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Jingjing Chen (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Zuxuan Wu (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), and Yu-Gang Jiang (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing)</i>	
Black-Box Sparse Adversarial Attack via Multi-Objective Optimisation	12291
<i>Phoenix Neale Williams (Department of Computer Science, University of Exeter) and Ke Li (Department of Computer Science, University of Exeter)</i>	

Demystifying Causal Features on Adversarial Examples and Causal Inoculation for Robust Network by Adversarial Instrumental Variable Regression	12302
<i>Junho Kim (Image and Video Systems Lab, School of Electrical Engineering, KAIST, South Korea), Byung-Kwan Lee (Image and Video Systems Lab, School of Electrical Engineering, KAIST, South Korea), and Yong Man Ro (Image and Video Systems Lab, School of Electrical Engineering, KAIST, South Korea)</i>	
Seasoning Model Soups for Robustness to Adversarial and Natural Distribution Shifts	12313
<i>Francesco Croce (University of Tübingen), Sylvestre-Alvise Rebuffi (DeepMind), Evan Shelhamer (DeepMind), and Sven Gowal (DeepMind)</i>	
Towards Benchmarking and Assessing Visual Naturalness of Physical World Adversarial Attacks	12324
<i>Simin Li (Beihang University), Shuning Zhang (Tsinghua University), Gujun Chen (Tsinghua University), Dong Wang (Tsinghua University), Pu Feng (Beihang University), Jiakai Wang (Zhongguancun Laboratory), Aishan Liu (Beihang University), Xin Yi (Tsinghua University; Zhongguancun Laboratory), and Xianglong Liu (Beihang University; Zhongguancun Laboratory; Institute of data space, Hefei Comprehensive National Science Center)</i>	
Physically Adversarial Infrared Patches With Learnable Shapes and Locations	12334
<i>Xingxing Wei (Institute of Artificial Intelligence, Beihang University, China), Jie Yu (Institute of Artificial Intelligence, Beihang University, China), and Yao Huang (Institute of Artificial Intelligence, Beihang University, China)</i>	
MaLP: Manipulation Localization Using a Proactive Scheme	12343
<i>Vishal Asnani (Michigan State University), Xi Yin (Meta AI), Tal Hassner (Meta AI), and Xiaoming Liu (Michigan State University)</i>	

Poster-Wed-PM

Polarimetric iToF: Measuring High-Fidelity Depth Through Scattering Media	12353
<i>Daniel S. Jeon (KAIST), Andréas Meuleman (KAIST), Seung-Hwan Baek (POSTECH), and Min H. Kim (KAIST)</i>	
NeRFLix: High-Quality Neural View Synthesis by Learning a Degradation-Driven Inter-Viewpoint MiXer	12363
<i>Kun Zhou (CUHK-Shenzhen; SmartMore Corporation), Wenbo Li (CUHK), Yi Wang (Shanghai AI Laboratory), Tao Hu (CUHK), Nianjuan Jiang (SmartMore Corporation), Xiaoguang Han (CUHK-Shenzhen), and Jiangbo Lu (SmartMore Corporation)</i>	
SUDS: Scalable Urban Dynamic Scenes	12375
<i>Haithem Turki (Carnegie Mellon University), Jason Y. Zhang (Carnegie Mellon University), Francesco Ferroni (Argo AI), and Deva Ramanan (Carnegie Mellon University)</i>	
DP-NeRF: Deblurred Neural Radiance Field With Physical Scene Priors	12386
<i>Dogyoon Lee (Yonsei University), Minhyeok Lee (Yonsei University), Chajin Shin (Yonsei University), and Sangyoun Lee (Korea Institute of Science and Technology (KIST))</i>	

DyLiN: Making Light Field Networks Dynamic	12397
<i>Heng Yu (Robotics Institute, Carnegie Mellon University), Joel Julin (Robotics Institute, Carnegie Mellon University), Zoltán Á. Milacski (Robotics Institute, Carnegie Mellon University), Koichiro Niinuma (Fujitsu Research of America), and László A. Jeni (Robotics Institute, Carnegie Mellon University)</i>	
Multi-Space Neural Radiance Fields	12407
<i>Ze-Xin Yin (Nankai University), Jiaxiong Qiu (Nankai University), Ming-Ming Cheng (Nankai University), and Bo Ren (Nankai University)</i>	
NeRFLight: Fast and Light Neural Radiance Fields Using a Shared Feature Grid	12417
<i>Fernando Rivas-Manzanegue (Arquimea Research Center; Universidad Politécnica de Madrid), Jorge Sierra-Acosta (Arquimea Research Center), Adrian Penate-Sanchez (Universidad de las Palmas de Gran Canaria, IUSIANI), Francesc Moreno-Noguer (Institut de Robòtica i Informàtica Industrial, CSIC-UPC), and Angela Ribeiro (Centre for Automation and Robotics, CSIC-UPM)</i>	
Cross-Guided Optimization of Radiance Fields With Multi-View Image Super-Resolution for High-Resolution Novel View Synthesis	12428
<i>Youngho Yoon (Visual Intelligence Lab., KAIST, Korea) and Kuk-Jin Yoon (Visual Intelligence Lab., KAIST, Korea)</i>	
NeuralEditor: Editing Neural Radiance Fields via Manipulating Point Clouds	12439
<i>Jun-Kun Chen (University of Illinois, Urbana-Champaign), Jipeng Lyu (Peking University), and Yu-Xiong Wang (University of Illinois, Urbana-Champaign)</i>	
DINER: Depth-Aware Image-Based NEural Radiance Fields	12449
<i>Malte Prinzler (Max Planck Institute for Intelligent Systems, Tübingen, Germany; Max Planck ETH Center for Learning Systems), Otmar Hilliges (ETH Zürich), and Justus Thies (Max Planck Institute for Intelligent Systems, Tübingen, Germany)</i>	
Modernizing Old Photos Using Multiple References via Photorealistic Style Transfer	12460
<i>Agus Gunawan (KAIST), Soo Ye Kim (KAIST; Adobe Research), Hyeonjun Sim (KAIST), Jae-Ho Lee (KAIST), and Munchurl Kim (ETRI)</i>	
Efficient Map Sparsification Based on 2D and 3D Discretized Grids	12470
<i>Xiaoyu Zhang (Chinese University of Hong Kong, Hong Kong Centre for Logistics Robotics) and Yun-Hui Liu (Chinese University of Hong Kong, Hong Kong Centre for Logistics Robotics)</i>	
K-Planes: Explicit Radiance Fields in Space, Time, and Appearance	12479
<i>Sara Fridovich-Keil (UC Berkeley), Giacomo Meanti (Istituto Italiano di Tecnologia), Frederik Rahbæk Warburg (Technical University of Denmark), Benjamin Recht (UC Berkeley), and Angjoo Kanazawa (UC Berkeley)</i>	

I2-SDF: Intrinsic Indoor Scene Reconstruction and Editing via Raytracing in Neural SDFs	12489
<i>Jingsen Zhu (State Key Lab of CAD&CG, Zhejiang University), Yuchi Huo (State Key Lab of CAD&CG, Zhejiang University; Zhejiang Lab), Qi Ye (Zhejiang University; Key Lab of CS&AUS of Zhejiang Province), Fujun Luan (Adobe Research), Jifan Li (State Key Lab of CAD&CG, Zhejiang University), Dianbing Xi (State Key Lab of CAD&CG, Zhejiang University), Lisha Wang (State Key Lab of CAD&CG, Zhejiang University), Rui Tang (KooLab, Manycore), Wei Hua (Zhejiang Lab), Hujun Bao (State Key Lab of CAD&CG, Zhejiang University), and Rui Wang (State Key Lab of CAD&CG, Zhejiang University)</i>	
Multi-View Inverse Rendering for Large-Scale Real-World Indoor Scenes	12499
<i>Zhen Li (Realsee), Lingli Wang (Realsee), Mofang Cheng (Realsee), Cihui Pan (Realsee), and Jiaqi Yang (Northwestern Polytechnical University)</i>	
Inverse Rendering of Translucent Objects Using Physical and Neural Renderers	12510
<i>Chenhao Li (Osaka University), Trung Thanh Ngo (Osaka University), and Hajime Nagahara (Osaka University)</i>	
Accidental Light Probes	12521
<i>Hong-Xing Yu (Stanford University), Samir Agarwala (Stanford University), Charles Herrmann (Google Research), Richard Szeliski (Google Research), Noah Snavely (Google Research), Jiajun Wu (Stanford University), and Deqing Sun (Google Research)</i>	
Humans As Light Bulbs: 3D Human Reconstruction From Thermal Reflection	12531
<i>Ruoshi Liu (Columbia University) and Carl Vondrick (Columbia University)</i>	
HumanGen: Generating Human Radiance Fields With Explicit Priors	12543
<i>Suyi Jiang (ShanghaiTech University), Haoran Jiang (ShanghaiTech University), Ziyu Wang (ShanghaiTech University), Haimin Luo (ShanghaiTech University; LumiAni Technology), Wenzheng Chen (University of Toronto), and Lan Xu (ShanghaiTech University)</i>	
Seeing Through the Glass: Neural 3D Reconstruction of Object Inside a Transparent Container	12555
<i>Jinguang Tong (The Australian National University), Sundaram Muthu (Data61, CSIRO), Fahira Afzal Maken (The Australian National University), Chuong Nguyen (Data61, CSIRO), and Hongdong Li (The Australian National University)</i>	
3D Shape Reconstruction of Semi-Transparent Worms	12565
<i>Thomas P. Ilett (University of Leeds, United Kingdom), Omer Yuval (University of Leeds, United Kingdom), Thomas Ranner (University of Leeds, United Kingdom), Netta Cohen (University of Leeds, United Kingdom), and David C. Hogg (University of Leeds, United Kingdom)</i>	
Dionysus: Recovering Scene Structures by Dividing Into Semantic Pieces	12576
<i>Likang Wang (The Hong Kong University of Science and Technology) and Lei Chen (The Hong Kong University of Science and Technology (Guangzhou))</i>	
SparseFusion: Distilling View-Conditioned Diffusion for 3D Reconstruction	12588
<i>Zhizhuo Zhou (Carnegie Mellon University) and Shubham Tulsiani (Carnegie Mellon University)</i>	

PET-NeuS: Positional Encoding Tri-Planes for Neural Surfaces	12598
<i>Yiqun Wang (Chongqing University and KAUST), Ivan Skorokhodov (Chongqing University and KAUST), and Peter Wonka (Chongqing University and KAUST)</i>	
RenderDiffusion: Image Diffusion for 3D Reconstruction, Inpainting and Generation	12608
<i>Titas Anciukevičius (University of Edinburgh; Adobe Research), Zexiang Xu (Adobe Research), Matthew Fisher (Adobe Research), Paul Henderson (University of Glasgow), Hakan Bilen (University of Edinburgh), Niloy J. Mitra (Adobe Research; UCL), and Paul Guerrero (Adobe Research)</i>	
Score Jacobian Chaining: Lifting Pretrained 2D Diffusion Models for 3D Generation	12619
<i>Haochen Wang (TTI-Chicago), Xiaodan Du (TTI-Chicago), Jiahao Li (TTI-Chicago), Raymond A. Yeh (Purdue University), and Greg Shakhnarovich (TTI-Chicago)</i>	
Infinite Photorealistic Worlds Using Procedural Generation	12630
<i>Alexander Raistrick (Princeton University), Lahav Lipson (Princeton University), Zeyu Ma (Princeton University), Lingjie Mei (Princeton University), Mingzhe Wang (Princeton University), Yiming Zuo (Princeton University), Karhan Kayan (Princeton University), Hongyu Wen (Princeton University), Beining Han (Princeton University), Yihan Wang (Princeton University), Alejandro Newell (Princeton University), Hei Law (Princeton University), Ankit Goyal (Princeton University), Kaiyu Yang (Princeton University), and Jia Deng (Princeton University)</i>	
Diffusion-SDF: Text-To-Shape via Voxelized Diffusion	12642
<i>Muheng Li (Tsinghua University), Yueqi Duan (Tsinghua University), Jie Zhou (Tsinghua University), and Jiwen Lu (Tsinghua University)</i>	
3D-Aware Multi-Class Image-to-Image Translation With NeRFs	12652
<i>Senmao Li (Nankai University), Joost van de Weijer (Universitat Autònoma de Barcelona), Yaxing Wang (Nankai University), Fahad Shahbaz Khan (Mohamed bin Zayed University of AI; Linköping University), Meiqin Liu (Beijing Jiaotong University), and Jian Yang (Nankai University)</i>	
Latent-NeRF for Shape-Guided Generation of 3D Shapes and Textures	12663
<i>Gal Metzer (Tel Aviv University), Elad Richardson (Tel Aviv University), Or Patashnik (Tel Aviv University), Raja Giryes (Tel Aviv University), and Daniel Cohen-Or (Tel Aviv University)</i>	
Local 3D Editing via 3D Distillation of CLIP Knowledge	12674
<i>Junha Hyung (KAIST AI; Kakao Enterprise Corp.), Sungwon Hwang (KAIST AI), Daejin Kim (Scatter Lab), Hyunji Lee (KAIST AI), and Jaegul Choo (KAIST AI)</i>	
ShapeTalk: A Language Dataset and Framework for 3D Shape Edits and Deformations	12685
<i>Panos Achlioptas (Stanford University; Snap Inc.), Ian Huang (Stanford University), Minhyuk Sung (KAIST), Sergey Tulyakov (Snap Inc.), and Leonidas Guibas (Stanford University)</i>	
CoralStyleCLIP: Co-Optimized Region and Layer Selection for Image Editing	12695
<i>Ambareesh Revanur (Adobe), Debraj Basu (Adobe), Shradha Agrawal (Adobe), Dhwanit Agarwal (Adobe), and Deepak Pai (Adobe)</i>	

3D-Aware Face Swapping	12705
<i>Yixuan Li (Shanghai Jiao Tong University, China), Chao Ma (Shanghai Jiao Tong University, China), Yichao Yan (Shanghai Jiao Tong University, China), Wenhan Zhu (Shanghai Jiao Tong University, China), and Xiaokang Yang (Shanghai Jiao Tong University, China)</i>	
DCFace: Synthetic Face Generation With Dual Condition Diffusion Model	12715
<i>Minchul Kim (Michigan State University), Feng Liu (Michigan State University), Anil Jain (Michigan State University), and Xiaoming Liu (Michigan State University)</i>	
HairStep: Transfer Synthetic to Real Using Strand and Depth Maps for Single-View 3D Hair Modeling	12726
<i>Yujian Zheng (FNii, CUHKSZ; SSE, CUHKSZ), Zirong Jin (SSE, CUHKSZ), Moran Li (Kuaishou Technology), Haibin Huang (Kuaishou Technology), Chongyang Ma (Kuaishou Technology), Shuguang Cui (SSE, CUHKSZ), and Xiaoguang Han (FNii, CUHKSZ; SSE, CUHKSZ)</i>	
DiffusionRig: Learning Personalized Priors for Facial Appearance Editing	12736
<i>Zheng Ding (UC San Diego), Xuaner Zhang (Adobe), Zhihao Xia (Adobe), Lars Jebe (Adobe), Zhuowen Tu (UC San Diego), and Xiuming Zhang (Adobe)</i>	
3D-Aware Facial Landmark Detection via Multi-View Consistent Training on Synthetic Data .	12747
<i>Libing Zeng (Texas A&M University), Lele Chen (OPPO US Research Center, InnoPeak Technology, Inc), Wentao Bao (Michigan State University), Zhong Li (OPPO US Research Center, InnoPeak Technology, Inc), Yi Xu (OPPO US Research Center, InnoPeak Technology, Inc), Junsong Yuan (University at Buffalo), and Nima Khademi Kalantari (Texas A&M University)</i>	
Parametric Implicit Face Representation for Audio-Driven Facial Reenactment	12759
<i>Ricong Huang (Sun Yat-sen University), Peiwen Lai (Sun Yat-sen University), Yipeng Qin (Cardiff University), and Guanbin Li (Sun Yat-sen University)</i>	
MEGANE: Morphable Eyeglass and Avatar Network	12769
<i>Junxuan Li (Australian National University; Meta Reality Labs Research), Shunsuke Saito (Meta Reality Labs Research), Tomas Simon (Meta Reality Labs Research), Stephen Lombardi (Australian National University), Hongdong Li (Australian National University), and Jason Saragih (Meta Reality Labs Research)</i>	
CodeTalker: Speech-Driven 3D Facial Animation With Discrete Motion Prior	12780
<i>Jinbo Xing (The Chinese University of Hong Kong), Menghan Xia (Tencent AI Lab), Yuechen Zhang (The Chinese University of Hong Kong), Xiaodong Cun (Tencent AI Lab), Jue Wang (Tencent AI Lab), and Tien-Tsin Wong (The Chinese University of Hong Kong)</i>	

Reconstructing Signing Avatars From Video Using Linguistic Priors	12791
<i>Maria-Paola Forte (Max Planck Institute for Intelligent Systems, Germany), Peter Kulits (Max Planck Institute for Intelligent Systems, Germany), Chun-Hao P. Huang (Max Planck Institute for Intelligent Systems, Germany), Vasileios Choutas (Max Planck Institute for Intelligent Systems, Germany), Dimitrios Tzionas (Max Planck Institute for Intelligent Systems, Germany), Katherine J. Kuchenbecker (Max Planck Institute for Intelligent Systems, Germany), and Michael J. Black (Max Planck Institute for Intelligent Systems, Germany)</i>	
HARP: Personalized Hand Reconstruction From a Monocular RGB Video	12802
<i>Korrawe Karunratanakul (ETH Zurich, Switzerland), Sergey Prokudin (ETH Zurich, Switzerland), Otmar Hilliges (ETH Zurich, Switzerland), and Siyu Tang (ETH Zurich, Switzerland)</i>	
OmniAvatar: Geometry-Guided Controllable 3D Head Synthesis	12814
<i>Hongyi Xu (ByteDance Inc), Guoxian Song (ByteDance Inc), Zihang Jiang (ByteDance Inc; National University of Singapore), Jianfeng Zhang (ByteDance Inc; National University of Singapore), Yichun Shi (ByteDance Inc), Jing Liu (ByteDance Inc), Wanchun Ma (ByteDance Inc), Jiashi Feng (ByteDance Inc), and Linjie Luo (ByteDance Inc)</i>	
RaBit: Parametric Modeling of 3D Biped Cartoon Characters With a Topological-Consistent Dataset	12825
<i>Zhongjin Luo (SSE, CUHKSZ), Shengcai Cai (SSE, CUHKSZ; Huawei Technologies Co., Ltd.), Jinguo Dong (SSE, CUHKSZ), Ruibo Ming (Tsinghua University; Tsinghua University), Liangdong Qiu (SSE, CUHKSZ; FNii, CUHKSZ), Xiaohang Zhan (Huawei Technologies Co., Ltd.), and Xiaoguang Han (SSE, CUHKSZ; FNii, CUHKSZ)</i>	
Transfer4D: A Framework for Frugal Motion Capture and Deformation Transfer	12836
<i>Shubh Maheshwari (TCS Research, India), Rahul Narain (Indian Institute of Technology Delhi, India), and Ramya Hebbalaguppe (TCS Research, India; Indian Institute of Technology Delhi, India)</i>	
CLOTH4D: A Dataset for Clothed Human Reconstruction	12847
<i>Xingxing Zou (Laboratory for Artificial Intelligence in Design; The Hong Kong Polytechnic University), Xintong Han (Huya Inc.), and Waikeung Wong (The Hong Kong Polytechnic University; Laboratory for Artificial Intelligence in Design)</i>	
Vid2Avatar: 3D Avatar Reconstruction From Videos in the Wild via Self-Supervised Scene Decomposition	12858
<i>Chen Guo (ETH Zurich), Tianjian Jiang (ETH Zurich), Xu Chen (ETH Zurich; Max Planck Institute for Intelligent Systems, Tübingen), Jie Song (ETH Zurich), and Otmar Hilliges (ETH Zurich)</i>	
High-Fidelity 3D Human Digitization From Single 2K Resolution Images	12869
<i>Sang-Hun Han (Gwangju Institute of Science and Technology (GIST)), Min-Gyu Park (Korea Electronics Technology Institute (KETI)), Ju Hong Yoon (Korea Electronics Technology Institute (KETI)), Ju-Mi Kang (Korea Electronics Technology Institute (KETI)), Young-Jae Park (Gwangju Institute of Science and Technology (GIST)), and Hae-Gon Jeon (Gwangju Institute of Science and Technology (GIST))</i>	

Sampling Is Matter: Point-Guided 3D Human Mesh Reconstruction	12880
<i>Jeonghwan Kim (Konkuk University), Mi-Gyeong Gwon (Konkuk University), Hyunwoo Park (Konkuk University), Hyukmin Kwon (Electronics and Telecommunications Research Institute), Gi-Mun Um (Electronics and Telecommunications Research Institute), and Wonjun Kim (Konkuk University)</i>	
gSDF: Geometry-Driven Signed Distance Functions for 3D Hand-Object Reconstruction	12890
<i>Zerui Chen (PSL Research Univ., 75005 Paris, France), Shizhe Chen (PSL Research Univ., 75005 Paris, France), Cordelia Schmid (PSL Research Univ., 75005 Paris, France), and Ivan Laptev (PSL Research Univ., 75005 Paris, France)</i>	
Human Body Shape Completion With Implicit Shape and Flow Learning	12901
<i>Boyao Zhou (Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France), Di Meng (Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France), Jean-Sébastien Franco (Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France), and Edmond Boyer (Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France)</i>	
ShapeClipper: Scalable 3D Shape Learning From Single-View Images via Geometric and CLIP-Based Consistency	12912
<i>Zixuan Huang (Georgia Institute of Technology), Varun Jampani (Google Research), Anh Thai (Georgia Institute of Technology), Yuanzhen Li (Google Research), Stefan Stojanov (Georgia Institute of Technology), and James M. Rehg (Georgia Institute of Technology)</i>	
PC2: Projection-Conditioned Point Cloud Diffusion for Single-Image 3D Reconstruction	12923
<i>Luke Melas-Kyriazi (Visual Geometry Group, Department of Engineering Science, University of Oxford), Christian Rupprecht (Visual Geometry Group, Department of Engineering Science, University of Oxford), and Andrea Vedaldi (Visual Geometry Group, Department of Engineering Science, University of Oxford)</i>	
NIKI: Neural Inverse Kinematics With Invertible Neural Networks for 3D Human Pose and Shape Estimation	12933
<i>Jiefeng Li (Shanghai Jiao Tong University), Siyuan Bian (Shanghai Jiao Tong University), Qi Liu (Shanghai Jiao Tong University), Jiasheng Tang (Alibaba Group), Fan Wang (Alibaba Group), and Cewu Lu (Shanghai Jiao Tong University)</i>	
ARCTIC: A Dataset for Dexterous Bimanual Hand-Object Manipulation	12943
<i>Zicong Fan (ETH Zurich, Switzerland; Max Planck Institute for Intelligent Systems, Germany), Omid Taheri (Max Planck Institute for Intelligent Systems, Germany), Dimitrios Tzionas (University of Amsterdam), Muhammed Kocabas (ETH Zurich, Switzerland; Max Planck Institute for Intelligent Systems, Germany), Manuel Kaufmann (ETH Zurich, Switzerland), Michael J. Black (Max Planck Institute for Intelligent Systems, Germany), and Otmar Hilliges (ETH Zurich, Switzerland)</i>	
ACR: Attention Collaboration-Based Regressor for Arbitrary Two-Hand Reconstruction	12955
<i>Zhengdi Yu (Tencent AI Lab; Durham University), Shaoli Huang (Tencent AI Lab), Chen Fang (Tencent AI Lab), Toby P. Breckon (Durham University), and Jue Wang (Tencent AI Lab)</i>	

MIME: Human-Aware 3D Scene Generation	12965
<i>Hongwei Yi (Max Planck Institute for Intelligent Systems, Germany), Chun-Hao P. Huang (Adobe Inc.), Shashank Tripathi (Max Planck Institute for Intelligent Systems, Germany), Lea Hering (Max Planck Institute for Intelligent Systems, Germany), Justus Thies (Max Planck Institute for Intelligent Systems, Germany), and Michael J. Black (Max Planck Institute for Intelligent Systems, Germany)</i>	
CIMI4D: A Large Multimodal Climbing Motion Dataset Under Human-Scene Interactions	12977
<i>Ming Yan (Xiamen University; Xiamen University; Xiamen University), Xin Wang (Xiamen University; Xiamen University), Yudi Dai (Xiamen University; Xiamen University), Siqi Shen (Xiamen University; Xiamen University), Chenglu Wen (Xiamen University; Xiamen University), Lan Xu (ShanghaiTech University), Yuexin Ma (ShanghaiTech University), and Cheng Wang (Xiamen University; Xiamen University)</i>	
Harmonious Feature Learning for Interactive Hand-Object Pose Estimation	12989
<i>Zhifeng Lin (South China University of Technology), Changxing Ding (South China University of Technology; Pazhou Lab, Guangzhou), Huan Yao (South China University of Technology), Zengsheng Kuang (South China University of Technology), and Shaoli Huang (Tencent AI-Lab, Shenzhen)</i>	
AssemblyHands: Towards Egocentric Activity Understanding via 3D Hand Pose Estimation .	12999
<i>Takehiko Ohkawa (Meta Reality Labs; The University of Tokyo), Kun He (Meta Reality Labs), Fadime Sener (Meta Reality Labs), Tomas Hodan (Meta Reality Labs), Luan Tran (Meta Reality Labs), and Cem Keskin (Meta Reality Labs)</i>	
A Characteristic Function-Based Method for Bottom-Up Human Pose Estimation	13009
<i>Haoxuan Qu (Singapore University of Technology and Design, Singapore), Yujun Cai (Nanyang Technological University, Singapore), Lin Geng Foo (Singapore University of Technology and Design, Singapore), Ajay Kumar (The Hong Kong Polytechnic University, Hong Kong), and Jun Liu (Singapore University of Technology and Design, Singapore)</i>	
Unified Pose Sequence Modeling	13019
<i>Lin Geng Foo (Singapore University of Technology & Design), Tianjiao Li (Singapore University of Technology & Design), Hossein Rahmani (Lancaster University), Qihong Ke (Monash University), and Jun Liu (Singapore University of Technology & Design)</i>	
Scene-Aware Egocentric 3D Human Pose Estimation	13031
<i>Jian Wang (MPI Informatics; Saarland Informatics Campus), Diogo Luvizon (MPI Informatics; Saarland Informatics Campus), Weipeng Xu (Meta Reality Labs), Lingjie Liu (MPI Informatics; Saarland Informatics Campus), Kripasindhu Sarkar (Google), and Christian Theobalt (MPI Informatics; Saarland Informatics Campus)</i>	
DiffPose: Toward More Reliable 3D Pose Estimation	13041
<i>Jia Gong (Singapore University of Technology and Design), Lin Geng Foo (Singapore University of Technology and Design), Zhipeng Fan (New York University), Qihong Ke (Monash University), Hossein Rahmani (Lancaster University), and Jun Liu (Singapore University of Technology and Design)</i>	

MammalNet: A Large-Scale Video Benchmark for Mammal Recognition and Behavior Understanding.....	13052
<i>Jun Chen (King Abdullah University of Science and Technology (KAUST)), Ming Hu (King Abdullah University of Science and Technology (KAUST)), Darren J. Coker (King Abdullah University of Science and Technology (KAUST)), Michael L. Berumen (King Abdullah University of Science and Technology (KAUST)), Blair Costelloe (Max Planck Institute of Animal Behavior; University of Konstanz), Sara Beery (Massachusetts Institute of Technology), Anna Rohrbach (University of California, Berkeley), and Mohamed Elhoseiny (King Abdullah University of Science and Technology (KAUST))</i>	
Learning 3D-Aware Image Synthesis With Unknown Pose Distribution	13062
<i>Zifan Shi (HKUST), Yujun Shen (Ant Group), Yinghao Xu (CUHK), Sida Peng (Zhejiang University), Yiyi Liao (Zhejiang University), Sheng Guo (Ant Group), Qifeng Chen (HKUST), and Dit-Yan Yeung (HKUST)</i>	
Pose Synchronization Under Multiple Pair-Wise Relative Poses	13072
<i>Yifan Sun (The University of Texas, Texas) and Qixing Huang (The University of Texas, Texas)</i>	
ObjectMatch: Robust Registration Using Canonical Object Correspondences	13082
<i>Can Gümeli (Technical University of Munich), Angela Dai (Technical University of Munich), and Matthias Nießner (Technical University of Munich)</i>	
Learning Articulated Shape With Keypoint Pseudo-Labels From Web Images	13092
<i>Anastasis Stathopoulos (Rutgers University), Georgios Pavlakos (UC Berkeley), Ligong Han (Rutgers University), and Dimitris N. Metaxas (Rutgers University)</i>	
Learning Correspondence Uncertainty via Differentiable Nonlinear Least Squares	13102
<i>Dominik Muhle (TU Munich; Munich Center for Machine Learning), Lukas Koestler (TU Munich; Munich Center for Machine Learning), Krishna Murthy Jatavallabhula (MIT), and Daniel Cremers (TU Munich; Munich Center for Machine Learning; University of Oxford)</i>	
Efficient Second-Order Plane Adjustment	13113
<i>Lipu Zhou (Meituan, China)</i>	
Learning a Depth Covariance Function	13122
<i>Eric Dexheimer (Dyson Robotics Lab, Imperial College London) and Andrew J. Davison (Dyson Robotics Lab, Imperial College London)</i>	
Privacy-Preserving Representations Are Not Enough: Recovering Scene Content From Camera Poses	13132
<i>Kunal Chelani (Chalmers University of Technology), Torsten Sattler (Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague), Fredrik Kahl (Chalmers University of Technology), and Zuzana Kukelova (Visual Recognition Group, Faculty of Electrical Engineering, Czech Technical University in Prague)</i>	

Objaverse: A Universe of Annotated 3D Objects	13142
<i>Matt Deitke (PRIOR @ Allen Institute for AI; University of Washington, Seattle), Dustin Schwenk (PRIOR @ Allen Institute for AI), Jordi Salvador (PRIOR @ Allen Institute for AI), Luca Weihs (PRIOR @ Allen Institute for AI), Oscar Michel (PRIOR @ Allen Institute for AI), Eli VanderBilt (PRIOR @ Allen Institute for AI), Ludwig Schmidt (University of Washington, Seattle), Kiana Ehsani (PRIOR @ Allen Institute for AI), Aniruddha Kembhavi (PRIOR @ Allen Institute for AI; University of Washington, Seattle), and Ali Farhadi (University of Washington, Seattle)</i>	
Omni3D: A Large Benchmark and Model for 3D Object Detection in the Wild	13154
<i>Garrick Brazil (Meta AI), Abhinav Kumar (Michigan State University), Julian Straub (Meta AI), Nikhila Ravi (Meta AI), Justin Johnson (Meta AI), and Georgia Gkioxari (Caltech)</i>	
HelixSurf: A Robust and Efficient Neural Implicit Surface Learning of Indoor Scenes With Iterative Intertwined Regularization	13165
<i>Zhihao Liang (South China University of Technology), Zhangjin Huang (South China University of Technology), Changxing Ding (South China University of Technology), and Kui Jia (South China University of Technology; Peng Cheng Laboratory)</i>	
Visual Localization Using Imperfect 3D Models From the Internet	13175
<i>Vojtech Panek (Czech Technical University in Prague), Zuzana Kukelova (Czech Technical University in Prague), and Torsten Sattler (Czech Technical University in Prague)</i>	
PRISE: Demystifying Deep Lucas-Kanade With Strongly Star-Convex Constraints for Multimodal Image Alignment	13187
<i>Yiqing Zhang (Worcester Polytechnic Institute, USA), Xinming Huang (Worcester Polytechnic Institute, USA), and Ziming Zhang (Worcester Polytechnic Institute, USA)</i>	
Scalable, Detailed and Mask-Free Universal Photometric Stereo	13198
<i>Satoshi Ikehata (National Institute of Informatics (NII))</i>	
Enhanced Stable View Synthesis	13208
<i>Nishant Jain (Indian Institute of Technology, Roorkee, India), Suryansh Kumar (ETH Zürich, Switzerland), and Luc Van Gool (ETH Zürich, Switzerland)</i>	
End-to-End Vectorized HD-Map Construction With Piecewise Bezier Curve	13218
<i>Limeng Qiao (MEGVII Technology), Wenjie Ding (MEGVII Technology), Xi Qiu (MEGVII Technology), and Chi Zhang (MEGVII Technology)</i>	
DynamicStereo: Consistent Dynamic Depth From Stereo Videos	13229
<i>Nikita Karaev (Meta AI; Visual Geometry Group, University of Oxford), Ignacio Rocco (Meta AI), Benjamin Graham (Meta AI), Natalia Neverova (Meta AI), Andrea Vedaldi (Meta AI), and Christian Rupprecht (Visual Geometry Group, University of Oxford)</i>	
Shakes on a Plane: Unsupervised Depth Estimation From Unstabilized Photography	13240
<i>Ilya Chugunov (Princeton University), Yuxuan Zhang (Princeton University), and Felix Heide (Princeton University)</i>	

Gated Stereo: Joint Depth Estimation From Gated and Wide-Baseline Active Stereo Cues	13252
<i>Stefanie Walz (Mercedes-Benz), Mario Bijelic (Princeton University), Andrea Ramazzina (Mercedes-Benz), Amanpreet Walia (Algolux), Fahim Mannan (Algolux), and Felix Heide (Mercedes-Benz)</i>	
K3DN: Disparity-Aware Kernel Estimation for Dual-Pixel Defocus Deblurring	13263
<i>Yan Yang (ANU), Liyuan Pan (BIT), Liu Liu (Cyberverse Dept., Huawei), and Miaomiao Liu (ANU)</i>	
HRDFuse: Monocular 360° Depth Estimation by Collaboratively Learning Holistic-With-Regional Depth Distributions	13273
<i>Hao Ai (AI Thrust, HKUST(GZ)), Zidong Cao (ARC Lab, Tencent PCG), Yan-Pei Cao (ARC Lab, Tencent PCG), Ying Shan (ARC Lab, Tencent PCG), and Lin Wang (AI Thrust, HKUST(GZ); Dept. of CSE, HKUST)</i>	
OSRT: Omnidirectional Image Super-Resolution With Distortion-Aware Transformer	13283
<i>Fanghua Yu (ShenZhen Key Lab of Computer Vision and Pattern Recognition, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), Xintao Wang (ARC, Tencent PCG), Mingdeng Cao (ARC, Tencent PCG; The University of Tokyo), Gen Li (Platform Technologies, Tencent Online Video), Ying Shan (ARC, Tencent PCG), and Chao Dong (ShenZhen Key Lab of Computer Vision and Pattern Recognition, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences; Shanghai Artificial Intelligence Laboratory)</i>	
Co-SLAM: Joint Coordinate and Sparse Parametric Encodings for Neural Real-Time SLAM ...	13293
<i>Hengyi Wang (University College London), Jingwen Wang (University College London), and Lourdes Agapito (University College London)</i>	
Few-Shot Non-Line-of-Sight Imaging With Signal-Surface Collaborative Regularization	13303
<i>Xintong Liu (Tsinghua University), Jianyu Wang (Tsinghua University), Leping Xiao (Tsinghua University), Xing Fu (Tsinghua University), Lingyun Qiu (Tsinghua University), Yanqi Lake Beijing Institute of Mathematical Sciences and Applications), and Zuoqiang Shi (Tsinghua University; Yanqi Lake Beijing Institute of Mathematical Sciences and Applications)</i>	
NLOST: Non-Line-of-Sight Imaging With Transformer	13313
<i>Yue Li (University of Science and Technology of China), Jiayong Peng (University of Science and Technology of China), Juntian Ye (University of Science and Technology of China), Yueyi Zhang (University of Science and Technology of China), Feihu Xu (University of Science and Technology of China), and Zhiwei Xiong (University of Science and Technology of China)</i>	
Listening Human Behavior: 3D Human Pose Estimation With Acoustic Signals	13323
<i>Yuto Shibata (Keio University), Yutaka Kawashima (Keio University), Mariko Isogawa (Keio University, JST Presto), Go Irie (Tokyo University of Science), Akisato Kimura (NTT Corporation), and Yoshimitsu Aoki (Aoki Keio University)</i>	

Towards Domain Generalization for Multi-View 3D Object Detection in Bird-Eye-View	13333
<i>Shuo Wang (University of Science and Technology of China), Xinhai Zhao (Huawei Noah's Ark Lab), Hai-Ming Xu (University of Adelaide), Zehui Chen (University of Science and Technology of China), Dameng Yu (Huawei Noah's Ark Lab), Jiahao Chang (University of Science and Technology of China), Zhen Yang (Huawei Noah's Ark Lab), and Feng Zhao (University of Science and Technology of China)</i>	
X3KD: Knowledge Distillation Across Modalities, Tasks and Stages for Multi-Camera 3D Object Detection	13343
<i>Marvin Klingner (Automated Driving, Qualcomm Technologies, Inc.), Shubhankar Borse (Qualcomm AI Research, an initiative of Qualcomm Technologies, Inc.), Varun Ravi Kumar (Automated Driving, Qualcomm Technologies, Inc.), Behnaz Rezaei (Automated Driving, Qualcomm Technologies, Inc.), Venkatraman Narayanan (Automated Driving, Qualcomm Technologies, Inc.), Senthil Yogamani (Automated Driving, QT Technologies Ireland Limited), and Fatih Porikli (Qualcomm AI Research, an initiative of Qualcomm Technologies, Inc.)</i>	
Phase-Shifting Coder: Predicting Accurate Orientation in Oriented Object Detection	13354
<i>Yi Yu (Southeast University, China) and Feipeng Da (Southeast University, China)</i>	
Learned Two-Plane Perspective Prior Based Image Resampling for Efficient Object Detection	13364
<i>Anurag Ghosh (Carnegie Mellon University), N. Dinesh Reddy (Carnegie Mellon University), Christoph Mertz (Carnegie Mellon University), and Srinivasa G. Narasimhan (Carnegie Mellon University)</i>	
Resource-Efficient RGBD Aerial Tracking	13374
<i>Jinyu Yang (Southern University of Science and Technology; University of Birmingham), Shang Gao (Southern University of Science and Technology), Zhe Li (Southern University of Science and Technology), Feng Zheng (Southern University of Science and Technology; Peng Cheng Laboratory), and Aleš Leonardis (University of Birmingham)</i>	
Toward RAW Object Detection: A New Benchmark and a New Model	13384
<i>Ruikang Xu (University of Science and Technology of China), Chang Chen (Huawei Noah's Ark Lab), Jingyang Peng (Huawei Noah's Ark Lab), Cheng Li (Huawei Noah's Ark Lab), Yibin Huang (Huawei Noah's Ark Lab), Fenglong Song (Huawei Noah's Ark Lab), Youliang Yan (Huawei Noah's Ark Lab), and Zhiwei Xiong (University of Science and Technology of China)</i>	
Bi-LRFusion: Bi-Directional LiDAR-Radar Fusion for 3D Dynamic Object Detection	13394
<i>Yingjie Wang (University of Science and Technology of China; Shanghai AI Laboratory), Jiajun Deng (University of Sydney), Yao Li (University of Science and Technology of China), Jinshui Hu (iFLYTEK), Cong Liu (iFLYTEK), Yu Zhang (University of Science and Technology of China), Jianmin Ji (University of Science and Technology of China), Wanli Ouyang (Shanghai AI Laboratory), and Yanyong Zhang (University of Science and Technology of China)</i>	
LiDAR-in-the-Loop Hyperparameter Optimization	13404
<i>Félix Goudreault (Algolux; Princeton University), Dominik Scheuble (Mercedes-Benz), Mario Bijelic (Mercedes-Benz), Nicolas Robidoux (Princeton University), and Felix Heide (Algolux)</i>	

Learning and Aggregating Lane Graphs for Urban Automated Driving	13415
<i>Martin Büchner (University of Freiburg), Jannik Zürn (University of Freiburg), Ion-George Todoran (Woven by Toyota), Abhinav Valada (University of Freiburg), and Wolfram Burgard (University of Technology Nuremberg)</i>	
Center Focusing Network for Real-Time LiDAR Panoptic Segmentation	13425
<i>Xiaoyan Li (Beijing Municipal Key Lab of Multimedia and Intelligent Software Technology; Beijing University of Technology, Beijing 100124, China), Gang Zhang (Mogo Auto Intelligence and Telematics Information Technology Co. Ltd.), Boyue Wang (Beijing Municipal Key Lab of Multimedia and Intelligent Software Technology; Beijing University of Technology, Beijing 100124, China), Yongli Hu (Beijing Municipal Key Lab of Multimedia and Intelligent Software Technology; Beijing University of Technology, Beijing 100124, China), and Baocai Yin (Beijing Municipal Key Lab of Multimedia and Intelligent Software Technology; Beijing University of Technology, Beijing 100124, China)</i>	
Adaptive Sparse Convolutional Networks With Global Context Enhancement for Faster Object Detection on Drone Images	13435
<i>Bowei Du (State Key Laboratory of Software Development Environment, Beihang University, China; School of Computer Science and Engineering, Beihang University, China), Yecheng Huang (State Key Laboratory of Software Development Environment, Beihang University, China; School of Computer Science and Engineering, Beihang University, China), Jiaxin Chen (School of Computer Science and Engineering, Beihang University, China), and Di Huang (State Key Laboratory of Software Development Environment, Beihang University, China; School of Computer Science and Engineering, Beihang University, China; Hangzhou Innovation Institute, Beihang University, China)</i>	
MV-JAR: Masked Voxel Jigsaw and Reconstruction for LiDAR-Based Self-Supervised Pre-Training	13445
<i>Runsen Xu (The Chinese University of Hong Kong; Shanghai AI Laboratory), Tai Wang (The Chinese University of Hong Kong; Shanghai AI Laboratory), Wenwei Zhang (Shanghai AI Laboratory; S-Lab, NTU), Runjian Chen (The University of Hong Kong), Jinkun Cao (Carnegie Mellon University), Jiangmiao Pang (Shanghai AI Laboratory), and Dahua Lin (The Chinese University of Hong Kong; Shanghai AI Laboratory)</i>	
ALSO: Automotive Lidar Self-Supervision by Occupancy Estimation	13455
<i>Alexandre Boulch (Waleo.ai, France), Corentin Sautier (Waleo.ai, France; Univ Gustave Eiffel, CNRS, Marne-la-Vallée, France), Björn Michele (Waleo.ai, France; CNRS, IRISA, Univ. Bretagne Sud, France), Gilles Puy (Waleo.ai, France), and Renaud Marlet (Waleo.ai, France; Univ Gustave Eiffel, CNRS, Marne-la-Vallée, France)</i>	
Unsupervised Intrinsic Image Decomposition With LiDAR Intensity	13466
<i>Shogo Sato (NTT Human Informatics Laboratories), Yasuhiro Yao (NTT Human Informatics Laboratories), Taiga Yoshida (NTT Human Informatics Laboratories), Takuhiro Kaneko (NTT Communication Science Laboratories), Shingo Ando (NTT Human Informatics Laboratories), and Jun Shimamura (NTT Human Informatics Laboratories)</i>	

PVT-SSD: Single-Stage 3D Object Detector With Point-Voxel Transformer	13476
<i>Honghui Yang (Zhejiang University; Shanghai AI Laboratory), Wenxiao Wang (Zhejiang University), Minghao Chen (Zhejiang University), Binbin Lin (Zhejiang University), Tong He (Shanghai AI Laboratory), Hua Chen (COMAC Beijing Aircraft Technology Research Institute), Xiaofei He (Zhejiang University), and Wanli Ouyang (Shanghai AI Laboratory)</i>	
LargeKernel3D: Scaling Up Kernels in 3D Sparse CNNs	13488
<i>Yukang Chen (The Chinese University of Hong Kong), Jianhui Liu (The University of Hong Kong), Xiangyu Zhang (MEGVII), Xiaojuan Qi (The University of Hong Kong), and Jiaya Jia (The Chinese University of Hong Kong; SmartMore)</i>	
WeatherStream: Light Transport Automation of Single Image Deweathering	13499
<i>Howard Zhang (University of California, Los Angeles), Yunhao Ba (University of California, Los Angeles), Ethan Yang (University of California, Los Angeles), Varan Mehra (University of California, Los Angeles), Blake Gella (University of California, Los Angeles), Akira Suzuki (University of California, Los Angeles), Arnold Pfahnl (University of California, Los Angeles), Chethan Chinder Chandrappa (University of California, Los Angeles), Alex Wong (Yale University), and Achuta Kadambi (University of California, Los Angeles)</i>	
Mask3D: Pre-Training 2D Vision Transformers by Learning Masked 3D Priors	13510
<i>Ji Hou (Meta Reality Labs), Xiaoliang Dai (Meta Reality Labs), Zijian He (Meta Reality Labs), Angela Dai (Technical University of Munich), and Matthias Nießner (Technical University of Munich)</i>	
DSVT: Dynamic Sparse Voxel Transformer With Rotated Sets	13520
<i>Haiyang Wang (Peking University), Chen Shi (Peking University), Shaoshuai Shi (Max Planck Institute for Informatics, Saarland Informatics Campus), Meng Lei (Peking University; Zhejiang Lab), Sen Wang (Huawei), Di He (Peking University), Bernt Schiele (Max Planck Institute for Informatics, Saarland Informatics Campus), and Liwei Wang (Peking University)</i>	
IterativePFN: True Iterative Point Cloud Filtering	13530
<i>Dasith de Silva Edirimuni (Deakin University), Xuequan Lu (Deakin University), Zhiwen Shao (Deakin University; China University of Mining and Technology), Gang Li (Deakin University), Antonio Robles-Kelly (Deakin University; Defense Science and Technology Group, Australia), and Ying He (Nanyang Technological University)</i>	
itKD: Interchange Transfer-Based Knowledge Distillation for 3D Object Detection	13540
<i>Hyeon Cho (Ajou University), Junyong Choi (Ajou University; Hyundai Motor Company), Geonwoo Baek (Ajou University), and Wonjun Hwang (Ajou University; Naver AI Lab)</i>	
ISBNet: A 3D Point Cloud Instance Segmentation Network With Instance-Aware Sampling and Box-Aware Dynamic Convolution	13550
<i>Tuan Duc Ngo (VinAI Research, Vietnam), Binh-Son Hua (VinAI Research, Vietnam), and Khoi Nguyen (VinAI Research, Vietnam)</i>	

Symmetric Shape-Preserving Autoencoder for Unsupervised Real Scene Point Cloud Completion .	13560
<i>Changfeng Ma (Nanjing University, China), Yinuo Chen (Nanjing University, China), Pengxiao Guo (Nanjing University, China), Jie Guo (Nanjing University, China), Chongjun Wang (Nanjing University, China), and Yanwen Guo (Nanjing University, China)</i>	
GeoMAE: Masked Geometric Target Prediction for Self-Supervised Point Cloud Pre-Training	13570
<i>Xiaoyu Tian (IIIS, Tsinghua University), Haoxi Ran (CMU), Yue Wang (NVIDIA), and Hang Zhao (IIIS, Tsinghua University)</i>	
AnchorFormer: Point Cloud Completion From Discriminative Nodes	13581
<i>Zhikai Chen (University of Science and Technology of China, China; University of Rochester, USA; HiDream.ai Inc.), Fuchen Long (University of Science and Technology of China, China; University of Rochester, USA; HiDream.ai Inc.), Zhaofan Qiu (University of Science and Technology of China, China; University of Rochester, USA; HiDream.ai Inc.), Ting Yao (University of Science and Technology of China, China; University of Rochester, USA; HiDream.ai Inc.), Wengang Zhou (University of Science and Technology of China, China; University of Rochester, USA; HiDream.ai Inc.), Jiebo Luo (University of Science and Technology of China, China; University of Rochester, USA; HiDream.ai Inc.), and Tao Mei (University of Science and Technology of China, China; University of Rochester, USA; HiDream.ai Inc.)</i>	
SHS-Net: Learning Signed Hyper Surfaces for Oriented Normal Estimation of Point Clouds .	13591
<i>Qing Li (Tsinghua University, China), Huifang Feng (Xiamen University, China), Kanle Shi (Kuaishou Technology, China), Yue Gao (Tsinghua University, China), Yi Fang (New York University Abu Dhabi, UAE), Yu-Shen Liu (Tsinghua University, China), and Zhizhong Han (Wayne State University, USA)</i>	
NerVE: Neural Volumetric Edges for Parametric Curve Extraction From Point Cloud	13601
<i>Xiangyu Zhu (CUHKSZ; CUHKSZ), Dong Du (CUHKSZ), Weikai Chen (Tencent America), Zhiyou Zhao (CUHKSZ), Yinyu Nie (Technical University of Munich), and Xiaoguang Han (CUHKSZ; CUHKSZ)</i>	
Unsupervised Deep Probabilistic Approach for Partial Point Cloud Registration	13611
<i>Guofeng Mei (GBDTC, UTS), Hao Tang (CVL, ETH Zurich), Xiaoshui Huang (Shanghai AI Lab), Weijie Wang (UNITN), Juan Liu (BIT), Jian Zhang (GBDTC, UTS), Luc Van Gool (CVL, ETH Zurich), and Qiang Wu (GBDTC, UTS)</i>	
Local Connectivity-Based Density Estimation for Face Clustering	13621
<i>Junho Shin (Chungnam National University), Hyo-Jun Lee (Chungnam National University), Hyunseop Kim (Chungnam National University), Jong-Hyeon Baek (Chungnam National University), Daehyun Kim (Chungnam National University), and Yeong Jun Koh (Chungnam National University)</i>	

Bridging Search Region Interaction With Template for RGB-T Tracking	13630
<i>Tianrui Hui (Institute of Information Engineering, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Meituan), Zizheng Xun (Beihang University; Beihang University), Fengguang Peng (Beihang University; Beihang University), Junshi Huang (Beihang University), Xiaoming Wei (Beihang University), Xiaolin Wei (Beihang University), Jiao Dai (Institute of Information Engineering, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Jizhong Han (Institute of Information Engineering, Chinese Academy of Sciences; University of Chinese Academy of Sciences), and Si Liu (Beihang University; Beihang University)</i>	
Quantum Multi-Model Fitting	13640
<i>Matteo Farina (University of Trento), Luca Magri (Politecnico di Milano), Willi Menapace (University of Trento), Elisa Ricci (Fondazione Bruno Kessler), Vladislav Golyanik (MPI for Informatics, SIC), and Federica Arrigoni (Politecnico di Milano)</i>	
Generalizable Local Feature Pre-Training for Deformable Shape Analysis	13650
<i>Souhaib Attaiki (LIX, École Polytechnique, IP Paris), Lei Li (LIX, École Polytechnique, IP Paris), and Maks Ovsjanikov (LIX, École Polytechnique, IP Paris)</i>	
Similarity Metric Learning for RGB-Infrared Group Re-Identification	13662
<i>Jianghao Xiong (Sun Yat-Sen University, China) and Jianhuang Lai (Sun Yat-Sen University, China; Guangdong Province Key Laboratory of Information Security Technology, China; Ministry of Education, China; Ministry of Public Security, China)</i>	
Unsupervised Deep Asymmetric Stereo Matching With Spatially-Adaptive Self-Similarity	13672
<i>Taeyong Song (Hyundai Motor Company R&D Division), Sunok Kim (Korea Aerospace University), and Kwanghoon Sohn (Yonsei University; Korea Institute of Science and Technology (KIST))</i>	
Sliced Optimal Partial Transport	13681
<i>Yikun Bai (Vanderbilt University), Bernhard Schmitzer (Göttingen University), Matthew Thorpe (University of Manchester; The Alan Turing Institute), and Soheil Kolouri (Vanderbilt University)</i>	
DistractFlow: Improving Optical Flow Estimation via Realistic Distractions and Pseudo-Labeling	13691
<i>Jisoo Jeong (Qualcomm AI Research), Hong Cai (Qualcomm AI Research), Rishiek Garrepalli (Qualcomm AI Research), and Fatih Porikli (Qualcomm AI Research)</i>	
Bayesian Posterior Approximation With Stochastic Ensembles	13701
<i>Oleksandr Balabanov (Stockholm University), Bernhard Mehlig (University of Gothenburg), and Hampus Linander (University of Gothenburg; Chalmers University of Technology)</i>	

V2V4Real: A Real-World Large-Scale Dataset for Vehicle-to-Vehicle Cooperative Perception .	13712
<i>Runsheng Xu (University of California, Los Angeles), Xin Xia (University of California, Los Angeles), Jinlong Li (Cleveland State University), Hanzhao Li (University of California, Los Angeles), Shuo Zhang (University of California, Los Angeles), Zhengzhong Tu (University of Texas at Austin), Zonglin Meng (University of California, Los Angeles), Hao Xiang (University of California, Los Angeles), Xiaoyu Dong (Northwestern University), Rui Song (Technical University of Munich; Fraunhofer Institute), Hongkai Yu (Cleveland State University), Bolei Zhou (University of California, Los Angeles), and Jiaqi Ma (University of California, Los Angeles)</i>	
ReasonNet: End-to-End Driving With Temporal and Global Reasoning	13723
<i>Hao Shao (SenseTime Research), Letian Wang (University of Toronto), Ruobing Chen (SenseTime Research), Steven L. Waslander (University of Toronto), Hongsheng Li (CUHK MMLab), and Yu Liu (SenseTime Research; Shanghai Artificial Intelligence Laboratory)</i>	
Open-World Multi-Task Control Through Goal-Aware Representation Learning and Adaptive Horizon Prediction	13734
<i>Shaofei Cai (Peking University), Zihao Wang (Peking University), Xiaojian Ma (University of California, Los Angeles), Anji Liu (University of California, Los Angeles), and Yitao Liang (Peking University; Beijing Institute for General Artificial Intelligence (BIGAI))</i>	
FJMP: Factorized Joint Multi-Agent Motion Prediction Over Learned Directed Acyclic Interaction Graphs	13745
<i>Luke Rowe (University of Waterloo), Martin Ethier (University of Waterloo), Eli-Henry Dykhne (University of Waterloo), and Krzysztof Czarnecki (University of Waterloo)</i>	
Trace and Pace: Controllable Pedestrian Animation via Guided Trajectory Diffusion	13756
<i>Davis Rempe (NVIDIA; Stanford University), Zhengyi Luo (NVIDIA; Carnegie Mellon University), Xue Bin Peng (NVIDIA; Simon Fraser University), Ye Yuan (NVIDIA), Kris Kitani (Carnegie Mellon University), Karsten Kreis (NVIDIA), Sanja Fidler (NVIDIA; University of Toronto; Vector Institute), and Or Litany (NVIDIA)</i>	
Galactic: Scaling End-to-End Reinforcement Learning for Rearrangement at 100k Steps-per-Second	13767
<i>Vincent-Pierre Berges (Meta AI (FAIR)), Andrew Szot (Georgia Tech), Devendra Singh Chaplot (Meta AI (FAIR)), Aaron Gokaslan (Cornell University), Roozbeh Mottaghi (Meta AI (FAIR)), Dhruv Batra (Meta AI (FAIR), Georgia Tech), and Eric Undersander (Meta AI (FAIR))</i>	
Affordances From Human Videos as a Versatile Representation for Robotics	13778
<i>Shikhar Bahl (CMU; Meta AI), Russell Mendonca (CMU), Lili Chen (CMU), Unnat Jain (CMU; Meta AI), and Deepak Pathak (CMU)</i>	
Indiscernible Object Counting in Underwater Scenes	13791
<i>Guolei Sun (ETH Zurich), Zhaochong An (ETH Zurich), Yun Liu (I2R, A*STAR), Ce Liu (ETH Zurich), Christos Sakaridis (ETH Zurich), Deng-Ping Fan (ETH Zurich), and Luc Van Gool (ETH Zurich; KU Leuven)</i>	

Tracking Through Containers and Occluders in the Wild	13802
<i>Basile Van Hoorick (Columbia University), Pavel Tokmakov (Toyota Research Institute), Simon Stent (Woven Planet), Jie Li (Toyota Research Institute), and Carl Vondrick (Columbia University)</i>	
Simple Cues Lead to a Strong Multi-Object Tracker	13813
<i>Jenny Seidenschwarz (Technical University of Munich), Guillem Brasó (Technical University of Munich; Munich Center for Machine Learning), Víctor Castro Serrano (Technical University of Munich), Ismail Elezi (Technical University of Munich), and Laura Leal-Taixé (Technical University of Munich)</i>	
An In-Depth Exploration of Person Re-Identification and Gait Recognition in Cloth-Changing Conditions	13824
<i>Weijia Li (Beijing Jiaotong University; Beijing Key Laboratory of Advanced Information Science and Network, China), Saihui Hou (Beijing Normal University; WATRIX.AI), Chunjie Zhang (Beijing Normal University; Beijing Key Laboratory of Advanced Information Science and Network, China), Chunshui Cao (WATRIX.AI), Xu Liu (Beijing Normal University; WATRIX.AI), Yongzhen Huang (WATRIX.AI), and Yao Zhao (Beijing Normal University; Beijing Key Laboratory of Advanced Information Science and Network, China)</i>	
SelfME: Self-Supervised Motion Learning for Micro-Expression Recognition	13834
<i>Xinqi Fan (City University of Hong Kong), Xueli Chen (City University of Hong Kong), Mingjie Jiang (City University of Hong Kong), Ali Raza Shahid (City University of Hong Kong; COMSATS University Islamabad), and Hong Yan (City University of Hong Kong)</i>	
LipFormer: High-Fidelity and Generalizable Talking Face Generation With a Pre-Learned Facial Codebook	13844
<i>Jiayu Wang (Alibaba Group), Kang Zhao (Alibaba Group), Shiwei Zhang (Alibaba Group), Yingya Zhang (Alibaba Group), Yujun Shen (Ant Group), Deli Zhao (Alibaba Group), and Jingren Zhou (Alibaba Group)</i>	
Real-Time Multi-Person Eyeblick Detection in the Wild for Untrimmed Video	13854
<i>Wenzheng Zeng (Huazhong University of Science and Technology, Wuhan 430074, China), Yang Xiao (Huazhong University of Science and Technology, Wuhan 430074, China), Sicheng Wei (Huazhong University of Science and Technology, Wuhan 430074, China), Jinfang Gan (Huazhong University of Science and Technology, Wuhan 430074, China), Xintao Zhang (Huazhong University of Science and Technology, Wuhan 430074, China), Zhiguo Cao (Huazhong University of Science and Technology, Wuhan 430074, China), Zhiwen Fang (Southern Medical University, China; Technology and Research (A*STAR), Singapore), and Joey Tianyi Zhou (Technology and Research (A*STAR), Singapore; Technology and Research (A*STAR), Singapore)</i>	
Skinned Motion Retargeting With Residual Perception of Motion Semantics & Geometry	13864
<i>Jiaxu Zhang (Wuhan University), Junwu Weng (Tencent AI Lab), Di Kang (Tencent AI Lab), Fang Zhao (Tencent AI Lab), Shaoli Huang (Tencent AI Lab), Xuefei Zhe (Tencent AI Lab), Linchao Bao (Tencent AI Lab), Ying Shan (Tencent AI Lab), Jue Wang (Tencent AI Lab), and Zhigang Tu (Wuhan University)</i>	

MoDi: Unconditional Motion Synthesis From Diverse Data	13873
<i>Sigal Raab (Tel-Aviv University), Inbal Leibovitch (Tel-Aviv University), Peizhuo Li (ETH ZurichGoogle Research), Kfir Aberman (ETH Zurich), Olga Sorkine-Hornung (ETH Zurich), and Daniel Cohen-Or (Tel-Aviv University)</i>	
Recurrent Vision Transformers for Object Detection With Event Cameras	13884
<i>Mathias Gehrig (University of Zurich, Switzerland) and Davide Scaramuzza (University of Zurich, Switzerland)</i>	
Continuous Intermediate Token Learning With Implicit Motion Manifold for Keyframe Based Motion Interpolation	13894
<i>Clinton A. Mo (The University of Sydney, Australia), Kun Hu (The University of Sydney, Australia), Chengjiang Long (Meta Reality Labs, USA), and Zhiyong Wang (The University of Sydney, Australia)</i>	
EvShutter: Transforming Events for Unconstrained Rolling Shutter Correction	13904
<i>Julius Erbach (Huawei Technologies, Zurich Research Center), Stepan Tulyakov (Huawei Technologies, Zurich Research Center), Patricia Vitoria (Huawei Technologies, Zurich Research Center), Alfredo Bochicchio (Huawei Technologies, Zurich Research Center), and Yuanyou Li (Huawei Technologies, Zurich Research Center)</i>	
Multi Domain Learning for Motion Magnification	13914
<i>Jasdeep Singh (CVPR Lab, Indian Institute of Technology Ropar, India), Subrahmanyam Murala (CVPR Lab, Indian Institute of Technology Ropar, India), and G. Sankara Raju Kosuru (CVPR Lab, Indian Institute of Technology Ropar, India)</i>	
Learning Event Guided High Dynamic Range Video Reconstruction	13924
<i>Yixin Yang (Peking University), Jin Han (The University of Tokyo; National Institute of Informatics), Jinxiu Liang (Peking University), Imari Sato (The University of Tokyo; National Institute of Informatics), and Boxin Shi (Peking University)</i>	
Joint Video Multi-Frame Interpolation and Deblurring Under Unknown Exposure Time	13935
<i>Wei Shang (Harbin Institute of Technology), Dongwei Ren (Harbin Institute of Technology), Yi Yang (Harbin Institute of Technology), Hongzhi Zhang (Harbin Institute of Technology), Kede Ma (City University of Hong Kong), and Wangmeng Zuo (Harbin Institute of Technology; Peng Cheng Laboratory, Shenzhen)</i>	
FeatER: An Efficient Network for Human Reconstruction via Feature Map-Based TransformerER	13945
<i>Ce Zheng (University of Central Florida), Matias Mendieta (University of Central Florida), Taojiannan Yang (University of Central Florida), Guo-Jun Qi (OPPO Seattle Research Center, USA; Westlake University), and Chen Chen (University of Central Florida)</i>	
MetaFusion: Infrared and Visible Image Fusion via Meta-Feature Embedding From Object Detection	13955
<i>Wenda Zhao (Dalian University of Technology, China), Shigeng Xie (Dalian University of Technology, China), Fan Zhao (Liaoning Normal University, China), You He (Tsinghua University, China), and Huchuan Lu (Dalian University of Technology, China)</i>	

Joint HDR Denoising and Fusion: A Real-World Mobile HDR Image Dataset	13966
<i>Shuaizheng Liu (The HongKong Polytechnic University; OPPO Research), Xindong Zhang (The HongKong Polytechnic University; OPPO Research), Lingchen Sun (The HongKong Polytechnic University; OPPO Research), Zhetong Liang (OPPO Research), Hui Zeng (OPPO Research), and Lei Zhang (The HongKong Polytechnic University; OPPO Research)</i>	
Visibility Constrained Wide-Band Illumination Spectrum Design for Seeing-in-the-Dark	13976
<i>Muyao Niu (The University of Tokyo), Zhuoxiao Li (The University of Tokyo), Zhihang Zhong (The University of Tokyo), and Yinqiang Zheng (The University of Tokyo)</i>	
Self-Supervised Blind Motion Deblurring With Deep Expectation Maximization	13986
<i>Ji Li (National University of Singapore, Singapore), Weixi Wang (National University of Singapore, Singapore), Yuesong Nan (National University of Singapore, Singapore), and Hui Ji (National University of Singapore, Singapore)</i>	
Structure Aggregation for Cross-Spectral Stereo Image Guided Denoising	13997
<i>Zehua Sheng (Zhejiang University), Zhu Yu (Zhejiang University), Xiongwei Liu (Zhejiang University), Si-Yuan Cao (Zhejiang University), Yuqi Liu (Zhejiang University), Hui-Liang Shen (Zhejiang University), and Huaqi Zhang (vivo Mobile Communication Company Ltd.)</i>	
Rawgmat: Noise-Accounted RAW Augmentation Enables Recognition in a Wide Variety of Environments	14007
<i>Masakazu Yoshimura (Sony Group Corporation), Junji Otsuka (Sony Group Corporation), Atsushi Irie (Sony Group Corporation), and Takeshi Ohashi (Sony Group Corporation)</i>	
Zero-Shot Noise2Noise: Efficient Image Denoising Without Any Data	14018
<i>Youssef Mansour (Technical University of Munich and Munich Center for Machine Learning, Germany) and Reinhard Heckel (Technical University of Munich and Munich Center for Machine Learning, Germany)</i>	
Real-Time Controllable Denoising for Image and Video	14028
<i>Zhaoyang Zhang (The Chinese University of Hong Kong), Yitong Jiang (The Chinese University of Hong Kong), Wenqi Shao (The Chinese University of Hong Kong; Shanghai AI Laboratory), Xiaogang Wang (The Chinese University of Hong Kong), Ping Luo (The Univesity of Hong Kong; Shanghai AI Laboratory), Kaimo Lin (SenseBrain), and Jinwei Gu (The Chinese University of Hong Kong; Shanghai AI Laboratory)</i>	
Probability-Based Global Cross-Modal Upsampling for Pansharpening	14039
<i>Zeyu Zhu (Xi'an Jiaotong University), Xiangyong Cao (Xi'an Jiaotong University), Man Zhou (Nanyang Technological University), Junhao Huang (Xi'an Jiaotong University), and Deyu Meng (Xi'an Jiaotong University; Macau University of Science and Technology)</i>	
ShadowDiffusion: When Degradation Prior Meets Diffusion Model for Shadow Removal	14049
<i>Lanqing Guo (Nanyang Technological University, Singapore), Chong Wang (Nanyang Technological University, Singapore), Wenhan Yang (Peng Cheng Laboratory, China), Siyu Huang (Harvard University, USA), Yufei Wang (Nanyang Technological University, Singapore), Hanspeter Pfister (Harvard University, USA), and Bihan Wen (Nanyang Technological University, Singapore)</i>	

Visual Recognition-Driven Image Restoration for Multiple Degradation With Intrinsic Semantics Recovery	14059
<i>Zizheng Yang (University of Science and Technology of China), Jie Huang (University of Science and Technology of China), Jiahao Chang (University of Science and Technology of China), Man Zhou (University of Science and Technology of China), Hu Yu (University of Science and Technology of China), Jinghao Zhang (University of Science and Technology of China), and Feng Zhao (University of Science and Technology of China)</i>	
Blind Image Quality Assessment via Vision-Language Correspondence: A Multitask Learning Perspective	14071
<i>Weixia Zhang (AI Institute, Shanghai Jiao Tong University), Guangtao Zhai (AI Institute, Shanghai Jiao Tong University), Ying Wei (Department of Computer Science, City University of Hong Kong), Xiaokang Yang (AI Institute, Shanghai Jiao Tong University), and Kede Ma (Department of Computer Science, City University of Hong Kong; Shenzhen Research Institute, City University of Hong Kong)</i>	
Human Guided Ground-Truth Generation for Realistic Image Super-Resolution	14082
<i>Du Chen (The Hong Kong Polytechnic University), Jie Liang (The Hong Kong Polytechnic University; OPPO Research Institute), Xindong Zhang (The Hong Kong Polytechnic University; Harbin Institute of Technology), Ming Liu (OPPO Research Institute), Hui Zeng (OPPO Research Institute), and Lei Zhang (The Hong Kong Polytechnic University; OPPO Research Institute)</i>	
Real-Time 6K Image Rescaling With Rate-Distortion Optimization	14092
<i>Chenyang Qi (Hong Kong University of Science and Technology), Xin Yang (Hong Kong University of Science and Technology), Ka Leong Cheng (Hong Kong University of Science and Technology), Ying-Cong Chen (Hong Kong University of Science and Technology), and Qifeng Chen (Hong Kong University of Science and Technology)</i>	
Equivalent Transformation and Dual Stream Network Construction for Mobile Image Super-Resolution	14102
<i>Jiahao Chao (East China Normal University), Zhou Zhou (East China Normal University), Hongfan Gao (East China Normal University), Jiali Gong (East China Normal University), Zhengfeng Yang (East China Normal University), Zhenbing Zeng (Chengdu Institute of Computer Applications of Chinese Academy of Sciences), and Lydia Dehbi (East China Normal University)</i>	
Ultrahigh Resolution Image/Video Matting With Spatio-Temporal Sparsity	14112
<i>Yanan Sun (HKUST), Chi-Keung Tang (HKUST), and Yu-Wing Tai (HKUST)</i>	
Comprehensive and Delicate: An Efficient Transformer for Image Restoration	14122
<i>Haiyu Zhao (Sichuan University), Yuanbiao Gou (Sichuan University), Boyun Li (Sichuan University), Dezhong Peng (Sichuan University), Jiancheng Lv (Sichuan University), and Xi Peng (Sichuan University)</i>	
PHA: Patch-Wise High-Frequency Augmentation for Transformer-Based Person Re-Identification...	14133
<i>Guiwei Zhang (Beihang University), Yongfei Zhang (Beihang University; Pengcheng Laboratory), Tianyu Zhang (Beihang University), Bo Li (Beihang University), and Shiliang Pu (Hikvision Research Institute)</i>	

PyramidFlow: High-Resolution Defect Contrastive Localization Using Pyramid Normalizing Flow	14143
<i>Jiarui Lei (Zhejiang University; ZJU-Hangzhou Global Scientific and Technological Innovation Center, China), Xiaobo Hu (Zhejiang University; ZJU-Hangzhou Global Scientific and Technological Innovation Center, China), Yue Wang (Zhejiang University; ZJU-Hangzhou Global Scientific and Technological Innovation Center, China), and Dong Liu (Zhejiang University; ZJU-Hangzhou Global Scientific and Technological Innovation Center, China; Jiaxing Key Laboratory of Photonic Sensing & Intelligent Imaging, China; Jiaxing Research Institute Zhejiang University)</i>	
Neural Fourier Filter Bank	14153
<i>Zhijie Wu (University of British Columbia), Yuhe Jin (University of British Columbia), and Kwang Moo Yi (University of British Columbia)</i>	
Restoration of Hand-Drawn Architectural Drawings Using Latent Space Mapping With Degradation Generator	14164
<i>Nakkwan Choi (Dept. of Electrical Engineering, Ulsan National Institute of Science and Technology, Korea), Seungjae Lee (Electronics and Telecommunications Research Institute, Korea), Yongsik Lee (Electronics and Telecommunications Research Institute, Korea), and Seungjoon Yang (Dept. of Electrical Engineering, Ulsan National Institute of Science and Technology, Korea)</i>	
Neural Preset for Color Style Transfer	14173
<i>Zhanghan Ke (City University of Hong Kong), Yuhao Liu (City University of Hong Kong), Lei Zhu (City University of Hong Kong), Nanxuan Zhao (Adobe Research), and Rynson W.H. Lau (City University of Hong Kong)</i>	
NÜWA-LIP: Language-Guided Image Inpainting With Defect-Free VQGAN	14183
<i>Minheng Ni (Harbin Institute of Technology), Xiaoming Li (Harbin Institute of Technology), and Wangmeng Zuo (Harbin Institute of Technology; Peng Cheng Laboratory)</i>	
DualVector: Unsupervised Vector Font Synthesis With Dual-Part Representation	14193
<i>Ying-Tian Liu (Tsinghua University), Zhifei Zhang (Adobe Research), Yuan-Chen Guo (Tsinghua University), Matthew Fisher (Adobe Research), Zhaowen Wang (Adobe Research), and Song-Hai Zhang (Tsinghua University)</i>	
DATID-3D: Diversity-Preserved Domain Adaptation Using Text-to-Image Diffusion for 3D Generative Model	14203
<i>Gwanghyun Kim (Dept. of Electrical and Computer Engineering) and Se Young Chun (Dept. of Electrical and Computer Engineering; INMC & IPAI Seoul National University, Republic of Korea)</i>	
GALIP: Generative Adversarial CLIPs for Text-to-Image Synthesis	14214
<i>Ming Tao (Nanjing University of Posts and Telecommunications; Peng Cheng Laboratory), Bing-Kun Bao (Nanjing University of Posts and Telecommunications; Peng Cheng Laboratory), Hao Tang (CVL, ETH Zurich), and Changsheng Xu (Peng Cheng Laboratory; MAIS, Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS))</i>	

Fix the Noise: Disentangling Source Feature for Controllable Domain Translation	14224
<i>Dongyeun Lee (KAIST; Klleon AI Research), Jae Young Lee (KAIST), Doyeon Kim (KAIST), Jaehyun Choi (KAIST), Jaejun Yoo (UNIST), and Junmo Kim (KAIST)</i>	
Conditional Text Image Generation With Diffusion Models	14235
<i>Yuanzhi Zhu (Alibaba DAMO Academy, China), Zhaohai Li (Alibaba DAMO Academy, China), Tianwei Wang (Alibaba DAMO Academy, China), Mengchao He (Alibaba DAMO Academy, China), and Cong Yao (Alibaba DAMO Academy, China)</i>	
ReCo: Region-Controlled Text-to-Image Generation	14246
<i>Zhengyuan Yang (Microsoft), Jianfeng Wang (Microsoft), Zhe Gan (Microsoft), Linjie Li (Microsoft), Kevin Lin (Microsoft), Chenfei Wu (Microsoft), Nan Duan (Microsoft), Zicheng Liu (Microsoft), Ce Liu (Microsoft), Michael Zeng (Microsoft), and Lijuan Wang (Microsoft)</i>	
Freestyle Layout-to-Image Synthesis	14256
<i>Han Xue (Shanghai Jiao Tong University; Singapore Management University), Zhiwu Huang (Singapore Management University; University of Southampton), Qianru Sun (Singapore Management University), Li Song (Shanghai Jiao Tong University; Shanghai Jiao Tong University), and Wenjun Zhang (Shanghai Jiao Tong University)</i>	
Specialist Diffusion: Plug-and-Play Sample-Efficient Fine-Tuning of Text-to-Image Diffusion Models To Learn Any Unseen Style	14267
<i>Haoming Lu (Picsart AI Research (PAIR)), Hazarapet Tunanyan (Picsart AI Research (PAIR)), Kai Wang (U of Oregon), Shant Navasardyan (Picsart AI Research), Zhangyang Wang (Picsart AI Research (PAIR); UT Austin), and Humphrey Shi (Picsart AI Research (PAIR))</i>	
Toward Verifiable and Reproducible Human Evaluation for Text-to-Image Generation	14277
<i>Mayu Otani (CyberAgent, Inc.), Riku Togashi (CyberAgent, Inc.), Yu Sawai (CyberAgent, Inc.), Ryosuke Ishigami (CyberAgent, Inc.), Yuta Nakashima (Osaka University), Esa Rahtu (Tampere University), Janne Heikkilä (University of Oulu), and Shin'ichi Satoh (CyberAgent, Inc.)</i>	
Towards Flexible Multi-Modal Document Models	14287
<i>Naoto Inoue (CyberAgent, Japan), Kotaro Kikuchi (CyberAgent, Japan), Edgar Simo-Serra (Waseda University, Japan), Mayu Otani (CyberAgent, Japan), and Kota Yamaguchi (CyberAgent, Japan)</i>	
On Distillation of Guided Diffusion Models	14297
<i>Chenlin Meng (Stanford University), Robin Rombach (Stability AI & LMU Munich), Ruiqi Gao (Google Research, Brain Team), Diederik Kingma (Google Research, Brain Team), Stefano Ermon (Stanford University), Jonathan Ho (Google Research, Brain Team), and Tim Salimans (Google Research, Brain Team)</i>	
Dimensionality-Varying Diffusion Process	14307
<i>Han Zhang (Shanghai Jiao Tong University; Alibaba Group), Ruili Feng (University of Science and Technology of China; Alibaba Group), Zhantao Yang (Shanghai Jiao Tong University; Alibaba Group), Lianghua Huang (Alibaba Group), Yu Liu (Alibaba Group), Yifei Zhang (Shanghai Jiao Tong University; Alibaba Group), Yujun Shen (Ant Group), Deli Zhao (Alibaba Group), Jingren Zhou (Alibaba Group), and Fan Cheng (Shanghai Jiao Tong University)</i>	

Shape-Aware Text-Driven Layered Video Editing	14317
<i>Yao-Chih Lee (University of Maryland, College Park), Ji-Ze Genevieve Jang (University of Maryland, College Park), Yi-Ting Chen (University of Maryland, College Park), Elizabeth Qiu (University of Maryland, College Park), and Jia-Bin Huang (University of Maryland, College Park)</i>	
Rethinking Image Super Resolution From Long-Tailed Distribution Learning Perspective	14327
<i>Yuanbiao Gou (Sichuan University, China), Peng Hu (Sichuan University, China), Jiancheng Lv (Sichuan University, China), Hongyuan Zhu (Institute for Infocomm Research (I2R), A*STAR, Singapore), and Xi Peng (Sichuan University, China)</i>	
End-to-End Video Matting With Trimap Propagation	14337
<i>Wei-Lun Huang (National Taiwan University) and Ming-Sui Lee (National Taiwan University)</i>	
Context-Based Trit-Plane Coding for Progressive Image Compression	14348
<i>Seungmin Jeon (Korea University), Kwang Pyo Choi (Samsung Electronics), Youngo Park (Samsung Electronics), and Chang-Su Kim (Korea University)</i>	
Complexity-Guided Slimmable Decoder for Efficient Deep Video Compression	14358
<i>Zhihao Hu (Beihang University, China) and Dong Xu (University of Hong Kong, China)</i>	
Efficient Hierarchical Entropy Model for Learned Point Cloud Compression	14368
<i>Rui Song (Peking University), Chunyang Fu (Peking University), Shan Liu (Tencent America), and Ge Li (Peking University)</i>	
NIRVANA: Neural Implicit Representations of Videos With Adaptive Networks and Autoregressive Patch-Wise Modeling	14378
<i>Shishira R. Maiya (University of Maryland; Snap Inc), Sharath Girish (University of Maryland), Max Ehrlich (University of Maryland), Hanyu Wang (University of Maryland), Kwot Sin Lee (Snap Inc), Patrick Poirson (Snap Inc), Pengxiang Wu (Snap Inc), Chen Wang (Snap Inc), and Abhinav Shrivastava (University of Maryland)</i>	
Learned Image Compression With Mixed Transformer-CNN Architectures	14388
<i>Jinming Liu (Waseda University, Japan; Shanghai Jiao Tong University, China), Heming Sun (Waseda University, Japan; JST PRESTO, 4-1-8 Honcho, Japan), and Jiro Katto (Waseda University, Japan; Shanghai Jiao Tong University, China)</i>	
Memory-Friendly Scalable Super-Resolution via Rewinding Lottery Ticket Hypothesis	14398
<i>Jin Lin (Xiamen University, China), Xiaotong Luo (Xiamen University, China), Ming Hong (Xiamen University, China), Yanyun Qu (Xiamen University, China), Yuan Xie (East China Normal University, China), and Zongze Wu (Shenzhen University, China)</i>	

InternImage: Exploring Large-Scale Vision Foundation Models With Deformable Convolutions	14408
<i>Wenhai Wang (Shanghai AI Laboratory), Jifeng Dai (Tsinghua University; Shanghai AI Laboratory), Zhe Chen (Nanjing University; Shanghai AI Laboratory), Zhenhang Huang (Shanghai AI Laboratory), Zhiqi Li (Nanjing University; Shanghai AI Laboratory), Xizhou Zhu (SenseTime Research), Xiaowei Hu (Shanghai AI Laboratory), Tong Lu (Nanjing University), Lewei Lu (SenseTime Research), Hongsheng Li (The Chinese University of Hong Kong), Xiaogang Wang (SenseTime Research; The Chinese University of Hong Kong), and Yu Qiao (Shanghai AI Laboratory)</i>	
EfficientViT: Memory Efficient Vision Transformer With Cascaded Group Attention	14420
<i>Xinyu Liu (The Chinese University of Hong Kong), Houwen Peng (Microsoft Research), Ningxin Zheng (Microsoft Research), Yuqing Yang (Microsoft Research), Han Hu (Microsoft Research), and Yixuan Yuan (The Chinese University of Hong Kong)</i>	
Castling-ViT: Compressing Self-Attention via Switching Towards Linear-Angular Attention at Vision Transformer Inference	14431
<i>Haoran You (Georgia Institute of Technology; Meta Research), Yunyang Xiong (Meta Research), Xiaoliang Dai (Meta Research), Bichen Wu (Meta Research), Peizhao Zhang (Meta Research), Haoqi Fan (Meta Research), Peter Vajda (Meta Research), and Yingyan Celine Lin (Georgia Institute of Technology)</i>	
RIFormer: Keep Your Vision Backbone Effective but Removing Token Mixer	14443
<i>Jiahao Wang (Shanghai AI Laboratory; The University of HongKong), Songyang Zhang (Shanghai AI Laboratory), Yong Liu (Tsinghua Shenzhen International Graduate School), Taiqiang Wu (Tsinghua Shenzhen International Graduate School), Yujiu Yang (Tsinghua Shenzhen International Graduate School), Xihui Liu (The University of HongKong), Kai Chen (Shanghai AI Laboratory), Ping Luo (The University of HongKong), and Dahua Lin (Shanghai AI Laboratory)</i>	
High-Resolution Image Reconstruction With Latent Diffusion Models From Human Brain Activity	14453
<i>Yu Takagi (Osaka University, Japan; CiNet, NICT, Japan) and Shinji Nishimoto (Osaka University, Japan; CiNet, NICT, Japan)</i>	
Non-Contrastive Unsupervised Learning of Physiological Signals From Video	14464
<i>Jeremy Speth (University of Notre Dame), Nathan Vance (University of Notre Dame), Patrick Flynn (University of Notre Dame), and Adam Czajka (University of Notre Dame)</i>	
Revealing the Dark Secrets of Masked Image Modeling	14475
<i>Zhenda Xie (Tsinghua University; Microsoft Research Asia), Zigang Geng (University of Science and Technology of China; Microsoft Research Asia), Jingcheng Hu (Tsinghua University; Microsoft Research Asia), Zheng Zhang (Microsoft Research Asia), Han Hu (Microsoft Research Asia), and Yue Cao (Microsoft Research Asia)</i>	
Improving Visual Representation Learning Through Perceptual Understanding	14486
<i>Samyakh Tukra (Tractable AI), Frederick Hoffman (Tractable AI), and Ken Chatfield (Tractable AI)</i>	

FlexiViT: One Model for All Patch Sizes	14496
<i>Lucas Beyler (Google Research), Pavel Izmailov (Google Research), Alexander Kolesnikov (Google Research), Mathilde Caron (Google Research), Simon Kornblith (Google Research), Xiaohua Zhai (Google Research), Matthias Minderer (Google Research), Michael Tschannen (Google Research), Ibrahim Alabdulmohsin (Google Research), and Filip Pavetic (Google Research)</i>	
AdaMAE: Adaptive Masking for Efficient Spatiotemporal Learning With Masked Autoencoders	14507
<i>Wele Gedara Chaminda Bandara (Johns Hopkins University), Naman Patel (Zippin), Ali Gholami (Zippin), Mehdi Nikkhah (Zippin), Motilal Agrawal (Zippin), and Vishal M. Patel (Johns Hopkins University)</i>	
SimpSON: Simplifying Photo Cleanup With Single-Click Distracting Object Segmentation Network	14518
<i>Chuong Huynh (University of Maryland, College Park), Yuqian Zhou (Adobe Research), Zhe Lin (Adobe Research), Connelly Barnes (Adobe Research), Eli Shechtman (Adobe Research), Sohrab Amirghodsi (Adobe Research), and Abhinav Shrivastava (University of Maryland, College Park)</i>	
Visual Dependency Transformers: Dependency Tree Emerges From Reversed Attention	14528
<i>Mingyu Ding (The University of Hong Kong; MIT), Yikang Shen (MIT-IBM Watson AI Lab), Lijie Fan (MIT), Zhenfang Chen (MIT-IBM Watson AI Lab), Zitian Chen (UMass Amherst), Ping Luo (The University of Hong Kong), Joshua B. Tenenbaum (MIT), and Chuang Gan (MIT-IBM Watson AI Lab; UMass Amherst)</i>	
Iterative Next Boundary Detection for Instance Segmentation of Tree Rings in Microscopy Images of Shrub Cross Sections	14540
<i>Alexander Gillert (Fraunhofer Institute for Computer Graphics Research (IGD), Rostock), Giulia Resente (Ernst Moritz Arndt University, Greifswald), Alba Anadon-Rosell (Centre for Research on Ecology and Forestry Applications (CREAF), Barcelona), Martin Wilmking (Institute of Botany and Landscape Ecology, Ernst Moritz Arndt University, Greifswald), and Uwe Freiherr von Lukas (Fraunhofer Institute for Computer Graphics Research (IGD), Rostock; Institute for Visual & Analytic Computing, University of Rostock)</i>	
VideoMAE V2: Scaling Video Masked Autoencoders With Dual Masking	14549
<i>Limin Wang (Nanjing University, China; Shanghai AI Lab, China), Bingkun Huang (Nanjing University, China; Shanghai AI Lab, China), Zhiyu Zhao (Nanjing University, China; Shanghai AI Lab, China), Zhan Tong (Nanjing University, China), Yinan He (Shanghai AI Lab, China), Yi Wang (Shanghai AI Lab, China), Yali Wang (Shanghai AI Lab, China; Shenzhen Institute of Advanced Technology, CAS, China), and Yu Qiao (Shanghai AI Lab, China; Shenzhen Institute of Advanced Technology, CAS, China)</i>	

DropMAE: Masked Autoencoders With Spatial-Attention Dropout for Tracking Tasks	14561
<i>Qiangqiang Wu (Department of Computer Science, City University of Hong Kong), Tianyu Yang (International Digital Economy Academy), Ziquan Liu (Department of Computer Science, City University of Hong Kong), Baoyuan Wu (The Chinese University of Hong Kong, Shenzhen), Ying Shan (Tencent AI Lab), and Antoni B. Chan (Department of Computer Science, City University of Hong Kong)</i>	
SeqTrack: Sequence to Sequence Learning for Visual Object Tracking	14572
<i>Xin Chen (Dalian University of Technology), Houwen Peng (Microsoft Research), Dong Wang (Dalian University of Technology), Huchuan Lu (Dalian University of Technology; Peng Cheng Laboratory), and Han Hu (Microsoft Research)</i>	
Bootstrapping Objectness From Videos by Relaxed Common Fate and Visual Grouping	14582
<i>Long Lian (UC Berkeley), Zhirong Wu (Microsoft Research Asia), and Stella X. Yu (UC Berkeley; University of Michigan)</i>	
Video Event Restoration Based on Keyframes for Video Anomaly Detection	14592
<i>Zhiwei Yang (Xidian University, China), Jing Liu (Xidian University, China), Zhaoyang Wu (Xidian University, China), Peng Wu (Northwestern Polytechnical University, China), and Xiaotao Liu (Xidian University, China)</i>	
Streaming Video Model	14602
<i>Yucheng Zhao (University of Science and Technology of China), Chong Luo (Microsoft Research Asia), Chuanxin Tang (Microsoft Research Asia), Dongdong Chen (Microsoft Cloud + AI), Noel Codella (Microsoft Cloud + AI), and Zheng-Jun Zha (University of Science and Technology of China)</i>	
LSTFE-Net: Long Short-Term Feature Enhancement Network for Video Small Object Detection	14613
<i>Jinsheng Xiao (Wuhan University, China), Yuanxu Wu (Wuhan University, China), Yunhua Chen (Guangdong University of Technology, China), Shurui Wang (Wuhan University, China), Zhongyuan Wang (Wuhan University, China), and Jiayi Ma (Wuhan University, China)</i>	
A Generalized Framework for Video Instance Segmentation	14623
<i>Miran Heo (Yonsei University), Sukjun Hwang (Yonsei University), Jeongseok Hyun (Yonsei University), Hanjung Kim (Yonsei University), Seoung Wug Oh (Adobe Research), Joon-Young Lee (Adobe Research), and Seon Joo Kim (Yonsei University)</i>	
Referring Multi-Object Tracking	14633
<i>Dongming Wu (Beijing Institute of Technology), Wencheng Han (University of Macau), Tiancai Wang (MEGVII Technology), Xingping Dong (Wuhan University), Xiangyu Zhang (University of Macau; Beijing Academy of Artificial Intelligence), and Jianbing Shen (University of Macau)</i>	
Source-Free Video Domain Adaptation With Spatial-Temporal-Historical Consistency Learning	14643
<i>Kai Li (NEC Labs, America), Deep Patel (NEC Labs, America), Erik Kruus (NEC Labs, America), and Martin Renqiang Min (NEC Labs, America)</i>	

Seeing What You Said: Talking Face Generation Guided by a Lip Reading Expert	14653
<i>Jiadong Wang (National University of Singapore), Xinyuan Qian (University of Science and Technology Beijing), Malu Zhang (University of Electronic Science and Technology of China), Robby T. Tan (National University of Singapore), and Haizhou Li (The Chinese University of Hong Kong, Shenzhen, China; National University of Singapore)</i>	
Egocentric Auditory Attention Localization in Conversations	14663
<i>Fiona Ryan (Georgia Institute of Technology; Meta Reality Labs Research), Hao Jiang (Meta Reality Labs Research), Abhinav Shukla (Meta Reality Labs Research), James M. Rehg (Georgia Institute of Technology; Meta Reality Labs Research), and Vamsi Krishna Ithapu (Meta Reality Labs Research)</i>	
iQuery: Instruments As Queries for Audio-Visual Sound Separation	14675
<i>Jiabao Chen (UC San Diego), Renrui Zhang (The Chinese University of Hong Kong), Dongze Lian (National University of Singapore), Jiaqi Yang (ShanghaiTech University), Ziyao Zeng (ShanghaiTech University), and Jianbo Shi (University of Pennsylvania)</i>	
Learning To Dub Movies via Hierarchical Prosody Models	14687
<i>Gaoxiang Cong (Shandong University), Liang Li (Institute of Computing Technology, Chinese Academy of Sciences; Lishui Institute of Hangzhou Dianzi University), Yuankai Qi (University of Adelaide), Zheng-Jun Zha (University of Science and Technology of China), Qi Wu (University of Adelaide), Wenyu Wang (Shandong University), Bin Jiang (Shandong University), Ming-Hsuan Yang (University of California; Yonsei University), and Qingming Huang (Institute of Computing Technology, Chinese Academy of Sciences)</i>	
A Large-Scale Robustness Analysis of Video Action Recognition Models	14698
<i>Madeline Chantry Schiappa (CRCV, University of Central Florida), Naman Biyani (IIT Kanpur), Prudvi Kamtam (CRCV, University of Central Florida), Shruti Vyas (CRCV, University of Central Florida), Hamid Palangi (Microsoft Research), Vibhav Vineet (Microsoft Research), and Yogesh S. Rawat (CRCV, University of Central Florida)</i>	
The Wisdom of Crowds: Temporal Progressive Attention for Early Action Prediction	14709
<i>Alexandros Stergiou (Vrije University of Brussels, Belgium; imec, Belgium) and Dima Damen (University of Bristol, UK)</i>	
STMixer: A One-Stage Sparse Action Detector	14720
<i>Tao Wu (Nanjing University), Mengqi Cao (Nanjing University), Ziteng Gao (Nanjing University), Gangshan Wu (Nanjing University), and Limin Wang (Nanjing University; Shanghai AI Lab)</i>	
Generating Human Motion From Textual Descriptions With Discrete Representations	14730
<i>Jianrong Zhang (Jilin University; Tencent AI Lab), Yangsong Zhang (Shanghai Jiao Tong University; Tencent AI Lab), Xiaodong Cun (Tencent AI Lab), Yong Zhang (Tencent AI Lab), Hongwei Zhao (Jilin University), Hongtao Lu (Shanghai Jiao Tong University), Xi Shen (Tencent AI Lab), and Ying Shan (Tencent AI Lab)</i>	

Cascade Evidential Learning for Open-World Weakly-Supervised Temporal Action Localization	14741
<i>Mengyuan Chen (Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS)), Junyu Gao (Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS)), and Changsheng Xu (Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS); Peng Cheng Laboratory, China)</i>	
Distilling Vision-Language Pre-Training To Collaborate With Weakly-Supervised Temporal Action Localization	14751
<i>Chen Ju (Shanghai Jiao Tong University), Kunhao Zheng (Shanghai Jiao Tong University), Jinxiang Liu (Shanghai Jiao Tong University), Peisen Zhao (Huawei Cloud), Ya Zhang (Shanghai Jiao Tong University; Shanghai AI Laboratory), Jianlong Chang (Huawei Cloud), Qi Tian (Huawei Cloud), and Yanfeng Wang (Shanghai Jiao Tong University; Shanghai AI Laboratory)</i>	
Simultaneously Short- and Long-Term Temporal Modeling for Semi-Supervised Video Semantic Segmentation	14763
<i>Jiangwei Lao (Ant Group), Weixiang Hong (Ant Group), Xin Guo (Ant Group), Yingying Zhang (Ant Group), Jian Wang (Ant Group), Jingdong Chen (Ant Group), and Wei Chu (Ant Group)</i>	
MIST: Multi-Modal Iterative Spatial-Temporal Transformer for Long-Form Video Question Answering	14773
<i>Difei Gao (National University of Singapore), Luowei Zhou (Microsoft), Lei Ji (Microsoft Research Asia), Linchao Zhu (Zhejiang University), Yi Yang (Zhejiang University), and Mike Zheng Shou (National University of Singapore)</i>	
Language-Guided Music Recommendation for Video via Prompt Analogies	14784
<i>Daniel McKee (University of Illinois at Urbana-Champaign), Justin Salamon (Adobe Research), Josef Sivic (Adobe Research; Czech Technical University), and Bryan Russell (University of Illinois at Urbana-Champaign; Adobe Research)</i>	
Text-Visual Prompting for Efficient 2D Temporal Video Grounding	14794
<i>Yimeng Zhang (Michigan State University; Applied AI, Intel), Xin Chen (Applied AI, Intel), Jinghan Jia (Michigan State University), Sijia Liu (Michigan State University), and Ke Ding (Applied AI, Intel)</i>	
CelebV-Text: A Large-Scale Facial Text-Video Dataset	14805
<i>Jianhui Yu (University of Sydney), Hao Zhu (SenseTime Research), Liming Jiang (Nanyang Technological University), Chen Change Loy (Nanyang Technological University), Weidong Cai (University of Sydney), and Wayne Wu (Shanghai AI Laboratory)</i>	
CNVid-3.5M: Build, Filter, and Pre-Train the Large-Scale Public Chinese Video-Text Dataset..	14815
<i>Tian Gan (Shandong University), Qing Wang (Ant Group), Xingning Dong (Shandong University), Xiangyuan Ren (Ant Group), Liqiang Nie (Harbin Institute of Technology (Shenzhen)), and Qingpei Guo (Ant Group)</i>	

Learning Procedure-Aware Video Representation From Instructional Videos and Their Narrations	14825
<i>Yiwu Zhong (University of Wisconsin-Madison), Licheng Yu (Meta AI), Yang Bai (Meta AI), Shangwen Li (Meta AI), Xueting Yan (Meta AI), and Yin Li (University of Wisconsin-Madison)</i>	
PDPP:Projected Diffusion for Procedure Planning in Instructional Videos	14836
<i>Hanlin Wang (Nanjing University, China), Yilu Wu (Nanjing University, China), Sheng Guo (MYbank, Ant Group, China), and Limin Wang (Nanjing University, China; Shanghai AI Lab, China)</i>	
Towards Fast Adaptation of Pretrained Contrastive Models for Multi-Channel Video-Language Retrieval	14846
<i>Xudong Lin (Columbia University), Simran Tiwari (Columbia University), Shiyuan Huang (Columbia University), Manling Li (UIUC), Mike Zheng Shou (National University of Singapore), Heng Ji (UIUC), and Shih-Fu Chang (Columbia University)</i>	
Clover: Towards a Unified Video-Language Alignment and Fusion Model	14856
<i>Jingjia Huang (Byte Dance Inc, 100043, China), Yanan Li (Xiamen University, 361005, China), Jiashi Feng (Byte Dance Inc, 100043, China), Xinglong Wu (Byte Dance Inc, 100043, China), Xiaoshuai Sun (Xiamen University, 361005, China), and Rongrong Ji (Xiamen University, 361005, China)</i>	
Align and Attend: Multimodal Summarization With Dual Contrastive Losses	14867
<i>Bo He (University of Maryland, College Park), Jun Wang (University of Maryland, College Park), Jieli Qiu (Carnegie Mellon University), Trung Bui (Adobe Research), Abhinav Shrivastava (University of Maryland, College Park), and Zhaowen Wang (Adobe Research)</i>	
Learning Situation Hyper-Graphs for Video Question Answering	14879
<i>Aisha Urooj (CRCV, University of Central Florida; Mayo Clinic, AZ), Hilde Kuehne (Mil-IBM Watson AI Lab; Goethe University Frankfurt Germany), Bo Wu (Mil-IBM Watson AI Lab), Kim Chheu (üMass Amherst), Walid Bouselham (Western Michigan University), Chuang Gan (Mil-IBM Watson AI Lab; üMass Amherst), Niels Lobo (CRCV, University of Central Florida), and Mubarak Shah (CRCV, University of Central Florida)</i>	
Natural Language-Assisted Sign Language Recognition	14890
<i>Ronglai Zuo (The Hong Kong University of Science and Technology), Fangyun Wei (Microsoft Research Asia), and Brian Mak (The Hong Kong University of Science and Technology)</i>	
SkyEye: Self-Supervised Bird's-Eye-View Semantic Mapping Using Monocular Frontal View Images	14901
<i>Nikhil Gosala (University of Freiburg), Kürsat Petek (University of Freiburg), Paulo L. J. Drews-Jr (University of Freiburg; Federal University of Rio Grande), Wolfram Burgard (Federal University of Rio Grande), and Abhinav Valada (University of Freiburg)</i>	
Adaptive Zone-Aware Hierarchical Planner for Vision-Language Navigation	14911
<i>Chen Gao (Beihang University), Xingyu Peng (Beihang University), Mi Yan (CFCS and School of EECS), He Wang (CFCS and School of EECS), Lirong Yang (Meituan), Haibing Ren (Meituan), Hongsheng Li (The Chinese University of Hong Kong), and Si Liu (Beihang University; Beihang University)</i>	

Iterative Vision-and-Language Navigation	14921
<i>Jacob Krantz (Oregon State University), Shurjo Banerjee (University of Michigan), Wang Zhu (University of Southern California), Jason Corso (University of Michigan), Peter Anderson (Google Research), Stefan Lee (Oregon State University), and Jesse Thomason (University of Southern California)</i>	
EXCALIBUR: Encouraging and Evaluating Embodied Exploration	14931
<i>Hao Zhu (Carnegie Mellon University), Raghav Kapoor (Carnegie Mellon University), So Yeon Min (Carnegie Mellon University), Winson Han (Allen Institute for Artificial Intelligence), Jiatai Li (Carnegie Mellon University), Kaiwen Geng (Carnegie Mellon University), Graham Neubig (Carnegie Mellon University), Yonatan Bisk (Carnegie Mellon University), Aniruddha Kembhavi (Allen Institute for Artificial Intelligence), and Luca Weihs (Allen Institute for Artificial Intelligence)</i>	
Multimodal Prompting With Missing Modalities for Visual Recognition	14943
<i>Yi-Lun Lee (National Yang Ming Chiao Tung University), Yi-Hsuan Tsai (Google), Wei-Chen Chiu (National Yang Ming Chiao Tung University), and Chen-Yu Lee (Google)</i>	
Visual Programming: Compositional Visual Reasoning Without Training	14953
<i>Tanmay Gupta (PRIOR @ Allen Institute for AI) and Aniruddha Kembhavi (PRIOR @ Allen Institute for AI)</i>	
Super-CLEVR: A Virtual Benchmark To Diagnose Domain Robustness in Visual Reasoning ...	14963
<i>Zhuowan Li (Johns Hopkins University), Xingrui Wang (University of Southern California), Elias Stengel-Eskin (Johns Hopkins University), Adam Kortylewski (Max Planck Institute for Informatics; University of Freiburg), Wufei Ma (Johns Hopkins University), Benjamin Van Durme (Johns Hopkins University), and Alan L. Yuille (Johns Hopkins University)</i>	
Prompting Large Language Models With Answer Heuristics for Knowledge-Based Visual Question Answering	14974
<i>Zhenwei Shao (Hangzhou Dianzi University, China), Zhou Yu (Hangzhou Dianzi University, China), Meng Wang (Hefei University of Technology, China), and Jun Yu (Hangzhou Dianzi University, China)</i>	
À-La-Carte Prompt Tuning (APT): Combining Distinct Data via Composable Prompting	14984
<i>Benjamin Bowman (AWS AI Labs; UCLA), Alessandro Achille (AWS AI Labs), Luca Zancato (AWS AI Labs), Matthew Trager (AWS AI Labs), Pramuditha Perera (AWS AI Labs), Giovanni Paolini (AWS AI Labs), and Stefano Soatto (AWS AI Labs)</i>	
ConStruct-VL: Data-Free Continual Structured VL Concepts Learning	14994
<i>James Seale Smith (MIT-IBM Watson AI Lab; Georgia Institute of Technology), Paola Cascante-Bonilla (MIT-IBM Watson AI Lab; Rice University), Assaf Arbelle (IBM Research), Donghyun Kim (MIT-IBM Watson AI Lab; IBM Research), Rameswar Panda (MIT-IBM Watson AI Lab; IBM Research), David Cox (MIT-IBM Watson AI Lab; IBM Research), Diyi Yang (Stanford University), Zsolt Kira (Georgia Institute of Technology), Rogerio Feris (MIT-IBM Watson AI Lab; IBM Research), and Leonid Karlinsky (MIT-IBM Watson AI Lab; IBM Research)</i>	

Q: How To Specialize Large Vision-Language Models to Data-Scarce VQA Tasks? A: Self-Train on Unlabeled Images!	15005
<i>Zaid Khan (Northeastern University), Vijay Kumar BG (NEC Labs America), Samuel Schulter (NEC Labs America), Xiang Yu (Amazon), Yun Fu (Northeastern University), and Manmohan Chandraker (NEC Labs America; UC San Diego)</i>	
Learning To Exploit Temporal Structure for Biomedical Vision–Language Processing	15016
<i>Shruthi Bannur (Microsoft Health Futures), Stephanie Hyland (Microsoft Health Futures), Qianchu Liu (Microsoft Health Futures), Fernando Pérez-García (Microsoft Health Futures), Maximilian Ilse (Microsoft Health Futures), Daniel C. Castro (Microsoft Health Futures), Benedikt Boecking (Microsoft Health Futures), Harshita Sharma (Microsoft Health Futures), Kenza Bouzid (Microsoft Health Futures), Anja Thieme (Microsoft Health Futures), Anton Schwaighofer (Microsoft Health Futures), Maria Wetscherek (Microsoft Health Futures), Matthew P. Lungren (Microsoft Health Futures), Aditya Nori (Microsoft Health Futures), Javier Alvarez-Valle (Microsoft Health Futures), and Ozan Oktay (Microsoft Health Futures)</i>	
FashionSAP: Symbols and Attributes Prompt for Fine-Grained Fashion Vision-Language Pre-Training	15028
<i>Yunpeng Han (Harbin Institute of Technology Shenzhen, China), Lisai Zhang (Harbin Institute of Technology Shenzhen, China), Qingcai Chen (Harbin Institute of Technology Shenzhen, China; PengCheng Laboratory, China), Zhijian Chen (Huawei Technologies Co., Ltd), Zhonghua Li (Huawei Technologies Co., Ltd), Jianxin Yang (Huawei Technologies Co., Ltd), and Zhao Cao (Huawei Technologies Co., Ltd)</i>	
Advancing Visual Grounding With Scene Knowledge: Benchmark and Method	15039
<i>Zhihong Chen (The Chinese University of Hong Kong, Shenzhen; Sun Yat-sen University; Shenzhen Research Institute of Big Data), Ruipei Zhang (Sun Yat-sen University), Yibing Song (Tencent AI Lab; Fudan University), Xiang Wan (Shenzhen Research Institute of Big Data), and Guanbin Li (Sun Yat-sen University)</i>	
Beyond Appearance: A Semantic Controllable Self-Supervised Learning Framework for Human-Centric Visual Tasks	15050
<i>Weihua Chen (Alibaba Group), Xianzhe Xu (Alibaba Group), Jian Jia (Alibaba Group), Hao Luo (Alibaba Group), Yaohua Wang (Alibaba Group), Fan Wang (Alibaba Group), Rong Jin (Alibaba Group), and Xiuyu Sun (Alibaba Group)</i>	
OCTET: Object-Aware Counterfactual Explanations	15062
<i>Mehdi Zemni (Valeo.ai, France), Mickaël Chen (Valeo.ai, France), Éloi Éloi Zablocki (Valeo.ai, France), Hédi Ben-Younes (Valeo.ai, France), Patrick Pérez (Valeo.ai, France), and Matthieu Cord (Valeo.ai, France; Sorbonne Université, France)</i>	
Local-Guided Global: Paired Similarity Representation for Visual Reinforcement Learning ...	15072
<i>Hyesong Choi (Ewha W. University), Hunsang Lee (Hyundai Motor Company), Wonil Song (Yonsei University), Sangryul Jeon (University of Michigan), Kwanghoon Sohn (Yonsei University), and Dongbo Min (Ewha W. University)</i>	

What Can Human Sketches Do for Object Detection?	15083
<i>Pinaki Nath Chowdhury (University of Surrey, United Kingdom), Ayan Kumar Bhunia (University of Surrey, United Kingdom), Aneeshan Sain (University of Surrey, United Kingdom), Subhadeep Koley (University of Surrey, United Kingdom), Tao Xiang (University of Surrey, United Kingdom), and Yi-Zhe Song (University of Surrey, United Kingdom)</i>	
Revisiting Multimodal Representation in Contrastive Learning: From Patch and Token Embeddings to Finite Discrete Tokens	15095
<i>Yuxiao Chen (Rutgers University), Jianbo Yuan (ByteDance Inc.), Yu Tian (ByteDance Inc.), Shijie Geng (Rutgers University; ByteDance Inc.), Xinyu Li (ByteDance Inc.), Ding Zhou (ByteDance Inc.), Dimitris N. Metaxas (Rutgers University), and Hongxia Yang (Zhejiang University)</i>	
Correlational Image Modeling for Self-Supervised Visual Pre-Training	15105
<i>Wei Li (Nanyang Technological University), Jiahao Xie (Nanyang Technological University), and Chen Change Loy (Nanyang Technological University)</i>	
Generalized Decoding for Pixel, Image, and Language	15116
<i>Xueyan Zou (Equal Technical Contribution; University of Wisconsin-Madison), Zi-Yi Dou (Equal Technical Contribution; UCLA), Jianwei Yang (Equal Technical Contribution; Microsoft Research at Redmond; Project Lead), Zhe Gan (Microsoft Cloud & AI), Linjie Li (Microsoft Cloud & AI), Chunyuan Li (Microsoft Research at Redmond), Xiyang Dai (Microsoft Cloud & AI), Harkirat Behl (Microsoft Research at Redmond), Jianfeng Wang (Microsoft Cloud & AI), Lu Yuan (Microsoft Cloud & AI), Nanyun Peng (UCLA), Lijuan Wang (Microsoft Cloud & AI), Yong Jae Lee (Equal Advisory Contribution; University of Wisconsin-Madison), and Jianfeng Gao (Equal Advisory Contribution; Microsoft Research at Redmond)</i>	
Towards Modality-Agnostic Person Re-Identification With Descriptive Query	15128
<i>Cuiqun Chen (Wuhan University, China), Mang Ye (Wuhan University, China; Hubei LuoJia Laboratory, China), and Ding Jiang (Wuhan University, China)</i>	
M6Doc: A Large-Scale Multi-Format, Multi-Type, Multi-Layout, Multi-Language, Multi-Annotation Category Dataset for Modern Document Layout Analysis	15138
<i>Hiuyi Cheng (South China University of Technology), Peirong Zhang (South China University of Technology), Sihang Wu (Huawei Cloud Computing Technologies Co., Ltd.), Jiabin Zhang (South China University of Technology), Qiyuan Zhu (Huawei Cloud Computing Technologies Co., Ltd.), Zecheng Xie (Huawei Cloud Computing Technologies Co., Ltd.), Jing Li (Huawei Cloud Computing Technologies Co., Ltd.), Kai Ding (IntSig Information Co., Ltd.), and Lianwen Jin (South China University of Technology)</i>	
Learning Customized Visual Models With Retrieval-Augmented Knowledge	15148
<i>Haotian Liu (University of Wisconsin-Madison), Kilho Son (Microsoft), Jianwei Yang (Microsoft), Ce Liu (Microsoft), Jianfeng Gao (Microsoft), Yong Jae Lee (University of Wisconsin-Madison), and Chunyuan Li (Microsoft)</i>	

Learning Semantic Relationship Among Instances for Image-Text Matching	15159
<i>Zheren Fu (University of Science and Technology of China, China), Zhendong Mao (University of Science and Technology of China, China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, China), Yan Song (University of Science and Technology of China, China), and Yongdong Zhang (University of Science and Technology of China, China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, China)</i>	
I2MVFormer: Large Language Model Generated Multi-View Document Supervision for Zero-Shot Image Classification	15169
<i>Muhammad Ferjad Naeem (ETH Zurich), Muhammad Gul Zain Ali Khan (TUKL; DFKI), Yongqin Xian (Google), Muhammad Zeshan Afzal (TUKL; DFKI), Didier Stricker (TUKL; DFKI), Luc Van Gool (ETH Zurich), and Federico Tombari (TUM; Google)</i>	
ImageBind: One Embedding Space To Bind Them All	15180
<i>Rohit Girdhar (FAIR, Meta AI), Alaaeldin El-Nouby (FAIR, Meta AI), Zhuang Liu (FAIR, Meta AI), Mannat Singh (FAIR, Meta AI), Kalyan Vasudev Alwala (FAIR, Meta AI), Armand Joulin (FAIR, Meta AI), and Ishan Misra (FAIR, Meta AI)</i>	
Model-Agnostic Gender Debaised Image Captioning	15191
<i>Yusuke Hirota (Osaka University), Yuta Nakashima (Osaka University), and Noa Garcia (Osaka University)</i>	
Boundary-Aware Backward-Compatible Representation via Adversarial Learning in Image Retrieval	15201
<i>Tan Pan (Ant Group), Furong Xu (Ant Group), Xudong Yang (Ant Group), Sifeng He (Ant Group), Chen Jiang (Ant Group), Qingpei Guo (Ant Group), Feng Qian (Ant Group), Xiaobo Zhang (Ant Group), Yuan Cheng (Ant Group), Lei Yang (Ant Group), and Wei Chu (Ant Group)</i>	
Prompt, Generate, Then Cache: Cascade of Foundation Models Makes Strong Few-Shot Learners . 15211	
<i>Renrui Zhang (CUHK MMLab; Shanghai Artificial Intelligence Laboratory), Xiangfei Hu (Shanghai Artificial Intelligence Laboratory), Bohao Li (Shanghai Artificial Intelligence Laboratory), Siyuan Huang (Shanghai Artificial Intelligence Laboratory), Hanqiu Deng (Shanghai Artificial Intelligence Laboratory), Yu Qiao (Shanghai Artificial Intelligence Laboratory), Peng Gao (Shenzhen Institute of Advanced Technology, Chinese Academy of Science; Shanghai Artificial Intelligence Laboratory), and Hongsheng Li (CUHK MMLab; CPII under InnoHK)</i>	
Towards Unified Scene Text Spotting Based on Sequence Generation	15223
<i>Taeho Kil (Naver Cloud), Seonghyeon Kim (Kakao Brain), Sukmin Seo (Naver Cloud), Yoonsik Kim (Naver Cloud), and Daehee Kim (Naver Cloud)</i>	
CapDet: Unifying Dense Captioning and Open-World Detection Pretraining	15233
<i>Yanxin Long (Shenzhen Campus of Sun Yat-sen University), Youpeng Wen (Shenzhen Campus of Sun Yat-sen University), Jianhua Han (Huawei Noah's Ark Lab), Hang Xu (Huawei Noah's Ark Lab), Pengzhen Ren (Shenzhen Campus of Sun Yat-sen University), Wei Zhang (Huawei Noah's Ark Lab), Shen Zhao (Shenzhen Campus of Sun Yat-sen University), and Xiaodan Liang (Shenzhen Campus of Sun Yat-sen University; MBZUAI)</i>	

CLIP2: Contrastive Language-Image-Point Pretraining From Real-World Point Cloud Data ...	15244
<i>Yihan Zeng (Huawei Noah's Ark Lab), Chenhan Jiang (Hong Kong University of Science and Technology), Jiageng Mao (The Chinese University of Hong Kong), Jianhua Han (Huawei Noah's Ark Lab), Chaoqiang Ye (Huawei Noah's Ark Lab), Qingqiu Huang (Huawei Noah's Ark Lab), Dit-Yan Yeung (Hong Kong University of Science and Technology), Zhen Yang (Huawei Noah's Ark Lab), Xiaodan Liang (Sun Yat-san University), and Hang Xu (Huawei Noah's Ark Lab)</i>	
Aligning Bag of Regions for Open-Vocabulary Object Detection	15254
<i>Size Wu (Nanyang Technological University), Wenwei Zhang (Nanyang Technological University), Sheng Jin (The University of Hong Kong; SenseTime Research and Tetras.AI), Wentao Liu (SenseTime Research and Tetras.AI; Shanghai AI Laboratory), and Chen Change Loy (Nanyang Technological University)</i>	
Visual Recognition by Request	15265
<i>Chufeng Tang (Tsinghua University), Lingxi Xie (Huawei Inc.), Xiaopeng Zhang (Huawei Inc.), Xiaolin Hu (Tsinghua University; Chinese Institute for Brain Research (CIBR); Tsinghua University), and Qi Tian (Huawei Inc.)</i>	
Category Query Learning for Human-Object Interaction Classification	15275
<i>Chi Xie (Tongji University), Fangao Zeng (MEGVII Technology), Yue Hu (Shanghai Jiao Tong University), Shuang Liang (Tongji University), and Yichen Wei (MEGVII Technology)</i>	
Self-Supervised Implicit Glyph Attention for Text Recognition	15285
<i>Tongkun Guan (Shanghai Jiao Tong University), Chaochen Gu (Shanghai Jiao Tong University), Jingzheng Tu (Shanghai Jiao Tong University), Xue Yang (Shanghai Jiao Tong University), Qi Feng (Shanghai Jiao Tong University), Yudi Zhao (Shanghai Jiao Tong University), and Wei Shen (Shanghai Jiao Tong University)</i>	
Enlarging Instance-Specific and Class-Specific Information for Open-Set Action Recognition.	15295
<i>Jun Cen (The Hong Kong University of Science and Technology; Alibaba Group), Shiwei Zhang (Alibaba Group), Xiang Wang (Huazhong University of Science and Technology), Yixuan Pei (Xi'an Jiaotong University), Zhiwu Qing (Huazhong University of Science and Technology), Yingya Zhang (Alibaba Group), and Qifeng Chen (The Hong Kong University of Science and Technology)</i>	
CLIP Is Also an Efficient Segmenter: A Text-Driven Approach for Weakly Supervised Semantic Segmentation	15305
<i>Yuqi Lin (Zhejiang University), Minghao Chen (Zhejiang University), Wenxiao Wang (Zhejiang University), Boxi Wu (Zhejiang University), Ke Li (Fullong Technology), Binbin Lin (Zhejiang University), Haifeng Liu (Zhejiang University), and Xiaofei He (Zhejiang University)</i>	
Learning Attention As Disentangler for Compositional Zero-Shot Learning	15315
<i>Shaozhe Hao (The University of Hong Kong), Kai Han (The University of Hong Kong), and Kwan-Yee K. Wong (The University of Hong Kong)</i>	

Universal Instance Perception As Object Discovery and Retrieval	15325
<i>Bin Yan (Dalian University of Technology, China), Yi Jiang (ByteDance), Jiannan Wu (The University of Hong Kong), Dong Wang (Dalian University of Technology, China), Ping Luo (The University of Hong Kong), Zehuan Yuan (ByteDance), and Huchuan Lu (Dalian University of Technology, China; Peng Cheng Laboratory)</i>	
Progressive Semantic-Visual Mutual Adaption for Generalized Zero-Shot Learning	15337
<i>Man Liu (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology, China), Feng Li (Hefei University of Technology, China), Chunjie Zhang (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology, China), Yunchao Wei (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology, China), Huihui Bai (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology, China), and Yao Zhao (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology, China)</i>	
DPF: Learning Dense Prediction Fields With Weak Supervision	15347
<i>Xiaoxue Chen (AIR, Tsinghua University), Yuhang Zheng (BUAA), Yupeng Zheng (CASIA), Qiang Zhou (AIR, Tsinghua University), Hao Zhao (AIR, Tsinghua University), Guyue Zhou (AIR, Tsinghua University), and Ya-Qin Zhang (AIR, Tsinghua University)</i>	
Modeling Entities As Semantic Points for Visual Information Extraction in the Wild	15358
<i>Zhibo Yang (Huazhong University of Science and Technology; Alibaba Group), Rujiao Long (Alibaba Group), Pengfei Wang (Alibaba Group), Sibo Song (Alibaba Group), Humen Zhong (Alibaba Group), Wenqing Cheng (Huazhong University of Science and Technology), Xiang Bai (Huazhong University of Science and Technology), and Cong Yao (Alibaba Group)</i>	
GeoNet: Benchmarking Unsupervised Adaptation Across Geographies	15368
<i>Tarun Kalluri (UC San Diego), Wangdong Xu (UC San Diego), and Manmohan Chandraker (UC San Diego)</i>	
SegLoc: Learning Segmentation-Based Representations for Privacy-Preserving Visual Localization	15380
<i>Maxime Pietrantoni (Czech Technical University in Prague; NAVER LABS Europe), Martin Humenberger (Czech Technical University in Prague), Torsten Sattler (Czech Technical University in Prague), and Gabriela Csurka (NAVER LABS Europe)</i>	
Towards Open-World Segmentation of Parts	15392
<i>Tai-Yu Pan (The Ohio State University), Qing Liu (Adobe Research), Wei-Lun Chao (The Ohio State University), and Brian Price (Adobe Research)</i>	

Pruning Parameterization With Bi-Level Optimization for Efficient Semantic Segmentation on the Edge	15402
<i>Changdi Yang (Northeastern University), Pu Zhao (Northeastern University), Yanyu Li (Northeastern University), Wei Niu (College of William & Mary), Jiexiong Guan (College of William & Mary), Hao Tang (CVL, ETH Zurich), Minghai Qin (Northeastern University), Bin Ren (College of William & Mary), Xue Lin (College of William & Mary), and Yanzhi Wang (College of William & Mary)</i>	
HGFormer: Hierarchical Grouping Transformer for Domain Generalized Semantic Segmentation ..	15413
<i>Jian Ding (Wuhan University, China; Wuhan University, China; Max Planck Institute for Informatics, Saarland Informatics Campus, Germany), Nan Xue (Wuhan University, China), Gui-Song Xia (Wuhan University, China; Wuhan University, China), Bernt Schiele (Max Planck Institute for Informatics, Saarland Informatics Campus, Germany), and Dengxin Dai (Max Planck Institute for Informatics, Saarland Informatics Campus, Germany)</i>	
Exemplar-FreeSOLO: Enhancing Unsupervised Instance Segmentation With Exemplars	15424
<i>Taoseef Ishtiaq (Carleton University, Canada), Qing En (Carleton University, Canada), and Yuhong Guo (Carleton University, Canada; CIFAR AI Chair, Canada)</i>	
Weakly-Supervised Domain Adaptive Semantic Segmentation With Prototypical Contrastive Learning	15434
<i>Anurag Das (MPI for Informatics, Saarland Informatics Campus), Yongqin Xian (ETH Zurich), Dengxin Dai (MPI for Informatics, Saarland Informatics Campus), and Bernt Schiele (MPI for Informatics, Saarland Informatics Campus)</i>	
Spatial-Temporal Concept Based Explanation of 3D ConvNets	15444
<i>Ying Ji (Nagoya University), Yu Wang (Hitotsubashi University), and Jien Kato (Ritsumeikan University)</i>	
Sparsely Annotated Semantic Segmentation With Adaptive Gaussian Mixtures	15454
<i>Linshan Wu (Hunan University), Zhun Zhong (University of Trento), Leyuan Fang (Hunan University), Xingxin He (Hunan University), Qiang Liu (Hunan University), Jiayi Ma (Wuhan University), and Hao Chen (Hong Kong University of Science and Technology)</i>	
Fuzzy Positive Learning for Semi-Supervised Semantic Segmentation	15465
<i>Pengchong Qiao (Peking University, China; Peng Cheng Laboratory, China; Peking University Shenzhen Graduate School, China), Zhidan Wei (Peking University, China; Peking University Shenzhen Graduate School, China), Yu Wang (Peking University, China; Peking University Shenzhen Graduate School, China), Zhennan Wang (Peng Cheng Laboratory, Shenzhen, China), Guoli Song (Peng Cheng Laboratory, Shenzhen, China), Fan Xu (Peng Cheng Laboratory, Shenzhen, China), Xiangyang Ji (Tsinghua University, Beijing, China), Chang Liu (Tsinghua University, Beijing, China), and Jie Chen (Peking University, China; Peng Cheng Laboratory, China; Peking University Shenzhen Graduate School, China)</i>	

STAR Loss: Reducing Semantic Ambiguity in Facial Landmark Detection	15475
<i>Zhenglin Zhou (Xiamen University), Huaxia Li (Tencent PCG), Hong Liu (National Institute of Information), Nanyang Wang (Tencent PCG), Gang Yu (Tencent PCG), and Rongrong Ji (Xiamen University)</i>	
Boosting Low-Data Instance Segmentation by Unsupervised Pre-Training With Saliency Prompt ...	15485
<i>Hao Li (Northwestern Polytechnical University; Zhejiang Lab), Dingwen Zhang (Northwestern Polytechnical University), Nian Liu (Mohamed bin Zayed University of Artificial Intelligence), Lechao Cheng (Zhejiang Lab), Yalun Dai (University of Chinese Academy of Sciences), Chao Zhang (NetEase), Xinggong Wang (Huazhong University of Science and Technology), and Junwei Han (Northwestern Polytechnical University)</i>	
Decoupled Semantic Prototypes Enable Learning From Diverse Annotation Types for Semi-Weakly Segmentation in Expert-Driven Domains	15495
<i>Simon Reiß (Karlsruhe Institute of Technology), Constantin Seibold (Karlsruhe Institute of Technology), Alexander Freytag (Carl Zeiss AG), Erik Rodner (University of Applied Sciences Berlin), and Rainer Stiefelwagen (Karlsruhe Institute of Technology)</i>	
The Treasure Beneath Multiple Annotations: An Uncertainty-Aware Edge Detector	15507
<i>Caixia Zhou (Beijing Jiaotong University, China), Yaping Huang (Beijing Jiaotong University, China), Mengyang Pu (North China Electric Power University, China), Qingji Guan (Beijing Jiaotong University, China), Li Huang (Beijing Jiaotong University, China), and Haibin Ling (Stony Brook University, USA)</i>	
Knowledge Combination To Learn Rotated Detection Without Rotated Annotation	15518
<i>Tianyu Zhu (Amazon IML.; Monash University), Bryce Ferenczi (Monash University), Pulak Purkait (Amazon IML. Monash University), Tom Drummond (Monash University), Hamid Reza Tofighi (Monash University), and Anton van den Hengel (Amazon IML.)</i>	
Mapping Degeneration Meets Label Evolution: Learning Infrared Small Target Detection With Single Point Supervision	15528
<i>Xinyi Ying (National University of Defense Technology), Li Liu (National University of Defense Technology), Yingqian Wang (National University of Defense Technology), Ruoqing Li (National University of Defense Technology), Nuo Chen (National University of Defense Technology), Zaiping Lin (National University of Defense Technology), Weidong Sheng (National University of Defense Technology), and Shilin Zhou (National University of Defense Technology)</i>	
SAP-DETR: Bridging the Gap Between Salient Points and Queries-Based Transformer Detector for Fast Model Convergency	15539
<i>Yang Liu (Institute of Computing Technology (ICT), Chinese Academy of Sciences; University of Chinese Academy of Sciences), Yao Zhang (Institute of Computing Technology (ICT), Chinese Academy of Sciences; University of Chinese Academy of Sciences), Yixin Wang (Stanford University), Yang Zhang (AI Lab, Lenovo Research), Jiang Tian (AI Lab, Lenovo Research), Zhongchao Shi (AI Lab, Lenovo Research), Jianping Fan (AI Lab, Lenovo Research), and Zhiqiang He (Institute of Computing Technology (ICT), Chinese Academy of Sciences; University of Chinese Academy of Sciences; Lenovo Ltd.)</i>	

Zero-Shot Object Counting	15548
<i>Jingyi Xu (Stony Brook University), Hieu Le (EPFL), Vu Nguyen (Stony Brook University), Viresh Ranjan (Amazon), and Dimitris Samaras (Stony Brook University)</i>	
SOOD: Towards Semi-Supervised Oriented Object Detection	15558
<i>Wei Hua (Huazhong University of Science and Technology), Dingkang Liang (Huazhong University of Science and Technology), Jingyu Li (Huazhong University of Science and Technology), Xiaolong Liu (Huazhong University of Science and Technology), Zhikang Zou (Baidu Inc., China), Xiaoqing Ye (Baidu Inc., China), and Xiang Bai (Huazhong University of Science and Technology)</i>	
Large-Scale Training Data Search for Object Re-Identification	15568
<i>Yue Yao (Australian National University), Tom Gedeon (Curtin University), and Liang Zheng (Australian National University)</i>	
Ambiguity-Resistant Semi-Supervised Learning for Dense Object Detection	15579
<i>Chang Liu (Shanghai University), Weiming Zhang (Baidu Inc), Xiangru Lin (Baidu Inc), Wei Zhang (Baidu Inc), Xiao Tan (Baidu Inc), Junyu Han (Baidu Inc), Xiaomao Li (Shanghai University; Shanghai Artificial Intelligence Laboratory), Errui Ding (Baidu Inc), and Jingdong Wang (Baidu Inc)</i>	
Towards Effective Visual Representations for Partial-Label Learning	15589
<i>Shiyu Xia (Southeast University), Jiaqi Lv (RIKEN Center for Advanced Intelligence Project), Ning Xu (Southeast University), Gang Niu (RIKEN Center for Advanced Intelligence Project; Southeast University), and Xin Geng (Southeast University)</i>	
Bi3D: Bi-Domain Active Learning for Cross-Domain 3D Object Detection	15599
<i>Jiakang Yuan (Fudan University), Bo Zhang (Shanghai AI Laboratory), Xiangchao Yan (Shanghai AI Laboratory), Tao Chen (Fudan University), Botian Shi (Shanghai AI Laboratory), Yikang Li (Shanghai AI Laboratory), and Yu Qiao (Shanghai AI Laboratory)</i>	
Boosting Detection in Crowd Analysis via Underutilized Output Features	15609
<i>Shaokai Wu (Jilin University) and Fengyu Yang (University of Michigan)</i>	
Self-Supervised Learning From Images With a Joint-Embedding Predictive Architecture	15619
<i>Mahmoud Assran (Meta AI (FAIR); McGill University; Mila, Quebec AI Institute), Quentin Duval (Meta AI (FAIR)), Ishan Misra (Meta AI (FAIR)), Piotr Bojanowski (Meta AI (FAIR)), Pascal Vincent (Meta AI (FAIR)), Michael Rabbat (Meta AI (FAIR); Mila, Quebec AI Institute), Yann LeCun (Meta AI (FAIR); New York University), and Nicolas Ballas (Meta AI (FAIR))</i>	
Weakly Supervised Segmentation With Point Annotations for Histopathology Images via Contrast-Based Variational Model	15630
<i>Hongrun Zhang (University of Liverpool, UK; University of Liverpool, Liverpool, UK), Liam Burrows (University of Liverpool, Liverpool, UK), Yanda Meng (University of Liverpool, Liverpool, UK), Declan Sculthorpe (University of Nottingham, Nottingham, UK), Abhik Mukherjee (University of Nottingham, Nottingham, UK), Sarah E. Coupland (University of Liverpool, Liverpool, UK), Ke Chen (University of Liverpool, Liverpool, UK), and Yalin Zheng (University of Liverpool, Liverpool, UK)</i>	

DoNet: Deep De-Overlapping Network for Cytology Instance Segmentation	15641
<i>Hao Jiang (The Hong Kong University of Science and Technology), Rushan Zhang (The Hong Kong University of Science and Technology), Yanning Zhou (Tencent AI Lab), Yumeng Wang (The Hong Kong University of Science and Technology), and Hao Chen (The Hong Kong University of Science and Technology)</i>	
MCF: Mutual Correction Framework for Semi-Supervised Medical Image Segmentation	15651
<i>Yongchao Wang (Chongqing University of Posts and Telecommunications, China), Bin Xiao (Chongqing University of Posts and Telecommunications, China), Xiuli Bi (Chongqing University of Posts and Telecommunications, China), Weisheng Li (Chongqing University of Posts and Telecommunications, China), and Xinbo Gao (Chongqing University of Posts and Telecommunications, China)</i>	
Histopathology Whole Slide Image Analysis With Heterogeneous Graph Representation Learning..	15661
<i>Tsai Hor Chan (The University of Hong Kong), Fernando Julio Cendra (The University of Hong Kong; TCL Corporate Research Hong Kong), Lan Ma (TCL Corporate Research Hong Kong), Guosheng Yin (The University of Hong Kong; Imperial College London), and Lequan Yu (The University of Hong Kong)</i>	
PEFAT: Boosting Semi-Supervised Medical Image Classification via Pseudo-Loss Estimation and Feature Adversarial Training	15671
<i>Qingjie Zeng (Northwestern Polytechnical University, China), Yutong Xie (The University of Adelaide, Australia), Zilin Lu (Northwestern Polytechnical University, China), and Yong Xia (Northwestern Polytechnical University, China)</i>	
Causally-Aware Intraoperative Imputation for Overall Survival Time Prediction	15681
<i>Xiang Li (Fudan University), Xuelin Qian (Fudan University), Litian Liang (Fudan University), Lingjie Kong (Fudan University), Qiaole Dong (Fudan University), Jiejun Chen (Fudan University), Dingxia Liu (Fudan University), Xiuzhong Yao (Fudan University), and Yanwei Fu (Fudan University)</i>	
Balanced Energy Regularization Loss for Out-of-Distribution Detection	15691
<i>Hyunjun Choi (ASRI, ECE., Seoul National University; RideFlux Inc.), Hawook Jeong (RideFlux Inc.), and Jin Young Choi (Seoul National University)</i>	
Block Selection Method for Using Feature Norm in Out-of-Distribution Detection	15701
<i>Yeonguk Yu (Gwangju Institute of Science and Technology), Sungho Shin (Gwangju Institute of Science and Technology), Seongju Lee (Gwangju Institute of Science and Technology), Changhyun Jun (Gwangju Institute of Science and Technology), and Kyoobin Lee (Gwangju Institute of Science and Technology)</i>	

Highly Confident Local Structure Based Consensus Graph Learning for Incomplete Multi-View Clustering	15712
<i>Jie Wen (Harbin Institute of Technology, China), Chengliang Liu (Harbin Institute of Technology, China), Gehui Xu (Harbin Institute of Technology, China), Zhihao Wu (Harbin Institute of Technology, China), Chao Huang (Shenzhen Campus of Sun Yat-sen University, China), Lunke Fei (Guangdong University of Technology, China), and Yong Xu (Harbin Institute of Technology, China; Pengcheng Laboratory, China)</i>	
Siamese DETR	15722
<i>Zeren Chen (SenseTime Research), Gengshi Huang (Beihang University; SenseTime Research), Wei Li (SenseTime Research), Jianing Teng (Nanyang Technological University), Kun Wang (SenseTime Research), Jing Shao (SenseTime Research), Chen Change Loy (Nanyang Technological University), and Lu Sheng (Beihang University)</i>	
Towards Bridging the Performance Gaps of Joint Energy-Based Models	15732
<i>Xiulong Yang (Georgia State University), Qing Su (Georgia State University), and Shihao Ji (Georgia State University)</i>	
Three Guidelines You Should Know for Universally Slimmable Self-Supervised Learning	15742
<i>Yun-Hao Cao (Nanjing University), Peiqin Sun (MEGVII Technology), and Shuchang Zhou (MEGVII Technology)</i>	
Boosting Transductive Few-Shot Fine-Tuning With Margin-Based Uncertainty Weighting and Probability Regularization	15752
<i>Ran Tao (Carnegie Mellon University), Hao Chen (Carnegie Mellon University), and Marios Savvides (Carnegie Mellon University)</i>	
CHMATCH: Contrastive Hierarchical Matching and Robust Adaptive Threshold Boosted Semi-Supervised Learning	15762
<i>Jianlong Wu (School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen)), Haozhe Yang (Shandong University), Tian Gan (Shandong University), Ning Ding (Alibaba Group), Feijun Jiang (Alibaba Group), and Liqiang Nie (School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen))</i>	
MarginMatch: Improving Semi-Supervised Learning with Pseudo-Margins	15773
<i>Tiberiu Sosea (University of Illinois Chicago) and Cornelia Caragea (University of Illinois Chicago)</i>	
Ranking Regularization for Critical Rare Classes: Minimizing False Positives at a High True Positive Rate	15783
<i>Kiarash Mohammadi (Borealis AI; Mila, Université de Montréal), He Zhao (Borealis AI), Mengyao Zhai (Borealis AI), and Frederick Tung (Borealis AI)</i>	
Learning Imbalanced Data With Vision Transformers	15793
<i>Zhengzhuo Xu (Tsinghua University, China), Ruikang Liu (Tsinghua University, China), Shuo Yang (Tsinghua University, China), Zenghao Chai (Tsinghua University, China), and Chun Yuan (Tsinghua University, China)</i>	
No One Left Behind: Improving the Worst Categories in Long-Tailed Learning	15804
<i>Yingxiao Du (Nanjing University, China) and Jianxin Wu (Nanjing University, China)</i>	

Global and Local Mixture Consistency Cumulative Learning for Long-Tailed Visual Recognitions	15814
<i>Fei Du (Yunnan University, China; Yunnan Key Laboratory of Software Engineering), Peng Yang (Yunnan University, China; Yunnan Key Laboratory of Software Engineering), Qi Jia (Yunnan University, China; Yunnan Key Laboratory of Software Engineering), Fengtao Nan (Yunnan University, China; Yunnan Key Laboratory of Software Engineering), Xiaoting Chen (Yunnan University, China; Yunnan Key Laboratory of Software Engineering), and Yun Yang (Yunnan University, China; Yunnan Key Laboratory of Software Engineering)</i>	
Curvature-Balanced Feature Manifold Learning for Long-Tailed Classification	15824
<i>Yanbiao Ma (Xidian University), Licheng Jiao (Xidian University), Fang Liu (Xidian University), Shuyuan Yang (Xidian University), Xu Liu (Xidian University), and Lingling Li (Xidian University)</i>	
DAA: A Delta Age AdaIN Operation for Age Estimation via Binary Code Transformer	15836
<i>Ping Chen (Jiayu Intelligent Technology Co., Ltd.), Xingpeng Zhang (Southwest Petroleum University, Chengdu, China), Ye Li (Jiayu Intelligent Technology Co., Ltd.), Ju Tao (Jiayu Intelligent Technology Co., Ltd.), Bin Xiao (Southwest Petroleum University, Chengdu, China), Bing Wang (Southwest Petroleum University, Chengdu, China), and Zongjie Jiang (Jiayu Intelligent Technology Co., Ltd.)</i>	
DLBD: A Self-Supervised Direct-Learned Binary Descriptor	15846
<i>Bin Xiao (Chongqing University of Posts and Telecommunications Chongqing, China), Yang Hu (Chongqing University of Posts and Telecommunications Chongqing, China), Bo Liu (Chongqing University of Posts and Telecommunications Chongqing, China), Xiuli Bi (Chongqing University of Posts and Telecommunications Chongqing, China), Weisheng Li (Chongqing University of Posts and Telecommunications Chongqing, China), and Xinbo Gao (Chongqing University of Posts and Telecommunications Chongqing, China)</i>	
Progressive Open Space Expansion for Open-Set Model Attribution	15856
<i>Tianyun Yang (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Danding Wang (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Fan Tang (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Xinying Zhao (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Juan Cao (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), and Sheng Tang (Institute of Computing Technology, Chinese Academy of Sciences, China; Research Institute of Intelligent Computing, Zhejiang Lab, China)</i>	
DiGA: Distil To Generalize and Then Adapt for Domain Adaptive Semantic Segmentation	15866
<i>Fengyi Shen (Technical University of Munich; Huawei Munich Research Center; Peking University), Akhil Gurram (Huawei Munich Research Center), Ziyuan Liu (Huawei Munich Research Center), He Wang (Peking University), and Alois Knoll (Technical University of Munich)</i>	

Multi-Modal Learning With Missing Modality via Shared-Specific Feature Modelling	15878
<i>Hu Wang (The University of Adelaide, Adelaide, Australia), Yuanhong Chen (The University of Adelaide, Adelaide, Australia), Congbo Ma (The University of Adelaide, Adelaide, Australia), Jodie Avery (The University of Adelaide, Adelaide, Australia), Louise Hull (The University of Adelaide, Adelaide, Australia), and Gustavo Carneiro (University of Surrey, UK)</i>	
Towards All-in-One Pre-Training via Maximizing Multi-Modal Mutual Information	15888
<i>Weijie Su (University of Science and Technology of China), Xizhou Zhu (SenseTime Research; Shanghai Artificial Intelligence Laboratory), Chenxin Tao (Tsinghua University), Lewei Lu (SenseTime Research), Bin Li (University of Science and Technology of China), Gao Huang (Tsinghua University), Yu Qiao (Shanghai Artificial Intelligence Laboratory), Xiaogang Wang (SenseTime Research; The Chinese University of Hong Kong), Jie Zhou (Tsinghua University), and Jifeng Dai (Tsinghua University; Shanghai Artificial Intelligence Laboratory)</i>	
Bi-Level Meta-Learning for Few-Shot Domain Generalization	15900
<i>Xiaorong Qin (Key Lab of Intelligent Information Processing Laboratory of the Chinese Academy of Sciences (CAS), Institute of Computing Technology, Beijing; University of Chinese Academy of Sciences, Beijing), Xinhang Song (Key Lab of Intelligent Information Processing Laboratory of the Chinese Academy of Sciences (CAS), Institute of Computing Technology, Beijing; University of Chinese Academy of Sciences, Beijing), and Shuqiang Jiang (Key Lab of Intelligent Information Processing Laboratory of the Chinese Academy of Sciences (CAS), Institute of Computing Technology, Beijing; University of Chinese Academy of Sciences, Beijing)</i>	
Train/Test-Time Adaptation With Retrieval	15911
<i>Luca Zancato (AWS AI Labs), Alessandro Achille (AWS AI Labs), Tian Yu Liu (AWS AI Labs), Matthew Trager (AWS AI Labs), Pramuditha Perera (AWS AI Labs), and Stefano Soatto (AWS AI Labs)</i>	
Robust Test-Time Adaptation in Dynamic Scenarios	15922
<i>Longhui Yuan (Beijing Institute of Technology), Binhui Xie (Beijing Institute of Technology), and Shuang Li (Beijing Institute of Technology)</i>	
Domain Expansion of Image Generators	15933
<i>Yotam Nitzan (Adobe Research; Tel-Aviv University), Michaël Gharbi (Adobe Research), Richard Zhang (Adobe Research), Taesung Park (Adobe Research), Jun-Yan Zhu (Carnegie Mellon University), Daniel Cohen-Or (Tel-Aviv University), and Eli Shechtman (Adobe Research)</i>	
Switchable Representation Learning Framework With Self-Compatibility	15943
<i>Shengsen Wu (Peng Cheng Laboratory, China; Peking University, China), Yan Bai (Peking University, China), Yihang Lou (Huawei Technologies), Xiongkun Linghu (Tsinghua University), Jianzhong He (Huawei Technologies), and Ling-Yu Duan (Peng Cheng Laboratory, China; Peking University, China)</i>	

A New Benchmark: On the Utility of Synthetic Data With Blender for Bare Supervised Learning and Downstream Domain Adaptation	15954
<i>Hui Tang (South China University of Technology; DexForce Co. Ltd.) and Kui Jia (South China University of Technology)</i>	
Adapting Shortcut With Normalizing Flow: An Efficient Tuning Framework for Visual Recognition	15965
<i>Yaoming Wang (Shanghai Jiao Tong University), Bowen Shi (Shanghai Jiao Tong University), Xiaopeng Zhang (Huawei Cloud), Jin Li (Shanghai Jiao Tong University), Yuchen Liu (Shanghai Jiao Tong University), Wenrui Dai (Shanghai Jiao Tong University), Chenglin Li (Shanghai Jiao Tong University), Hongkai Xiong (Shanghai Jiao Tong University), and Qi Tian (Huawei Cloud)</i>	
Manipulating Transfer Learning for Property Inference	15975
<i>Yulong Tian (Nanjing University, China), Fnu Suya (University of Virginia, USA), Anshuman Suri (Nanjing University, China), Fengyuan Xu (Nanjing University, China), and David Evans (University of Virginia, USA)</i>	
Heterogeneous Continual Learning	15985
<i>Divyam Madaan (NVIDIA; New York University), Hongxu Yin (NVIDIA), Wonmin Byeon (NVIDIA), Jan Kautz (NVIDIA), and Pavlo Molchanov (NVIDIA)</i>	
Generic-to-Specific Distillation of Masked Autoencoders	15996
<i>Wei Huang (University of Chinese Academy of Sciences), Zhiliang Peng (University of Chinese Academy of Sciences), Li Dong (Microsoft Research), Furu Wei (Microsoft Research), Jianbin Jiao (University of Chinese Academy of Sciences), and Qixiang Ye (University of Chinese Academy of Sciences)</i>	
Towards a Smaller Student: Capacity Dynamic Distillation for Efficient Image Retrieval	16006
<i>Yi Xie (South China University of Technology), Huaidong Zhang (South China University of Technology), Xuemiao Xu (South China University of Technology; State Key Laboratory of Subtropical Building Science; Ministry of Education Key Laboratory of Big Data and Intelligent Robot; Guangdong Provincial Key Lab of Computational Intelligence and Cyberspace Information), Jianqing Zhu (Huaqiao University), and Shengfeng He (Singapore Management University)</i>	
CafeBoost: Causal Feature Boost To Eliminate Task-Induced Bias for Class Incremental Learning	16016
<i>Benliu Qiu (University of Electronic Science and Technology of China, China), Hongliang Li (University of Electronic Science and Technology of China, China), Haitao Wen (University of Electronic Science and Technology of China, China), Heqian Qiu (University of Electronic Science and Technology of China, China), Lanxiao Wang (University of Electronic Science and Technology of China, China), Fanman Meng (University of Electronic Science and Technology of China, China), Qingbo Wu (University of Electronic Science and Technology of China, China), and Lili Pan (University of Electronic Science and Technology of China, China)</i>	

Bilateral Memory Consolidation for Continual Learning	16026
<i>Xing Nie (Chinese Academy of Sciences; University of Chinese Academy of Sciences), Shixiong Xu (Chinese Academy of Sciences; University of Chinese Academy of Sciences), Xiyan Liu (Baidu Inc., China.), Gaofeng Meng (Chinese Academy of Sciences; University of Chinese Academy of Sciences; HK Institute of Science & Innovation, CAS.), Chunlei Huo (Chinese Academy of Sciences; University of Chinese Academy of Sciences), and Shiming Xiang (Chinese Academy of Sciences; University of Chinese Academy of Sciences)</i>	
NICO++: Towards Better Benchmarking for Domain Generalization	16036
<i>Xingxuan Zhang (Tsinghua University), Yue He (Tsinghua University), Renzhe Xu (Tsinghua University), Han Yu (Tsinghua University), Zheyang Shen (Tsinghua University), and Peng Cui (Tsinghua University)</i>	
DART: Diversify-Aggregate-Repeat Training Improves Generalization of Neural Networks	16048
<i>Samyak Jain (Vision and AI Lab, Indian Institute of Science, Bangalore), Sravanti Addepalli (Vision and AI Lab, Indian Institute of Science, Bangalore), Pawan Kumar Sahu (Vision and AI Lab, Indian Institute of Science, Bangalore), Priyam Dey (Vision and AI Lab, Indian Institute of Science, Bangalore), and R. Venkatesh Babu (Vision and AI Lab, Indian Institute of Science, Bangalore)</i>	
Differentiable Architecture Search With Random Features	16060
<i>Xuanyang Zhang (MEGVII Technology), Yonggang Li (Peking University), Xiangyu Zhang (MEGVII Technology), Yongtao Wang (Peking University), and Jian Sun (MEGVII Technology)</i>	
Class Adaptive Network Calibration	16070
<i>Bingyuan Liu (ÉTS Montreal, Canada), Jérôme Rony (ÉTS Montreal, Canada), Adrian Galdran (Universitat Pompeu Fabra, Spain), Jose Dolz (ÉTS Montreal, Canada), and Ismail Ben Ayed (ÉTS Montreal, Canada)</i>	
Meta-Learning With a Geometry-Adaptive Preconditioner	16080
<i>Suhyun Kang (Seoul National University), Duhun Hwang (Seoul National University), Moonjung Eo (Seoul National University), Taesup Kim (Seoul National University), and Wonjong Rhee (IPAI and AIIS, Seoul National University)</i>	
DepGraph: Towards Any Structural Pruning	16091
<i>Gongfan Fang (National University of Singapore), Xinyin Ma (National University of Singapore), Mingli Song (Zhejiang University), Michael Bi Mi (Huawei Technologies Ltd.), and Xinchao Wang (National University of Singapore)</i>	
Stitchable Neural Networks	16102
<i>Zizheng Pan (Monash University), Jianfei Cai (Monash University), and Bohan Zhuang (Monash University)</i>	
Integral Neural Networks	16113
<i>Kirill Solodskikh (Huawei Noah's Ark Lab), Azim Kurbanov (Huawei Noah's Ark Lab), Ruslan Aydarkhanov (Huawei Noah's Ark Lab), Irina Zhelavskaya (Huawei Noah's Ark Lab), Yury Parfenov (Huawei Noah's Ark Lab), Dehua Song (Huawei Noah's Ark Lab), and Stamatios Lefkimmiatis (Huawei Noah's Ark Lab)</i>	

Regularization of Polynomial Networks for Image Recognition	16123
<i>Grigorios G. Chrysos (LIONS, EPFL, Lausanne, Switzerland), Bohan Wang (LIONS, EPFL, Lausanne, Switzerland), Jiankang Deng (Huawei UKRD), and Volkan Cevher (LIONS, EPFL, Lausanne, Switzerland)</i>	
ConvNeXt V2: Co-Designing and Scaling ConvNets With Masked Autoencoders	16133
<i>Sanghyun Woo (KAIST), Shoubhik Debnath (Meta AI, FAIR), Ronghang Hu (Meta AI, FAIR), Xinlei Chen (Meta AI, FAIR), Zhuang Liu (Meta AI, FAIR), In So Kweon (KAIST), and Saining Xie (New York University)</i>	
Shortcomings of Top-Down Randomization-Based Sanity Checks for Evaluations of Deep Neural Network Explanations	16143
<i>Alexander Binder (ICT Cluster, SIT Singapore; UiO Oslo; SUTD Singapore), Leander Weber (Fraunhofer HHI), Sebastian Lapuschkin (Fraunhofer HHI), Grégoire Montavon (FU Berlin; BIFOLD Berlin), Klaus-Robert Müller (BIFOLD Berlin; TU Berlin; Korea University Seoul; MPI Saarbrücken), and Wojciech Samek (Fraunhofer HHI; BIFOLD Berlin; TU Berlin)</i>	
Don't Lie to Me! Robust and Efficient Explainability With Verified Perturbation Analysis	16153
<i>Thomas Fel (Brown University, USA; Artificial and Natural Intelligence Toulouse Institute; IRT Saint-Exupery, France; Innovation & Research Division, SNCF), Melanie Ducoffe (Artificial and Natural Intelligence Toulouse Institute; IRT Saint-Exupery, France; Airbus AI Research), David Vigouroux (Artificial and Natural Intelligence Toulouse Institute; IRT Saint-Exupery, France), Rémi Cadène (Sorbonne Université, CNRS, France), Mikaël Capelle (Thales Alenia Space, France), Claire Nicodème (Innovation & Research Division, SNCF), and Thomas Serre (Brown University, USA; Artificial and Natural Intelligence Toulouse Institute)</i>	
OT-Filter: An Optimal Transport Filter for Learning With Noisy Labels	16164
<i>Chuanwen Feng (University of Science and Technology of China), Yilong Ren (University of Science and Technology of China), and Xike Xie (University of Science and Technology of China)</i>	
Robust Generalization Against Photon-Limited Corruptions via Worst-Case Sharpness Minimization	16175
<i>Zhuo Huang (Sydney AI Centre, The University of Sydney), Miaoxi Zhu (Wuhan University), Xiaobo Xia (Sydney AI Centre, The University of Sydney), Li Shen (JD Explore Academy), Jun Yu (University of Science and Technology of China), Chen Gong (Nanjing University of Science and Technology), Bo Han (Hong Kong Baptist University), Bo Du (Wuhan University), and Tongliang Liu (Sydney AI Centre, The University of Sydney)</i>	
Learning With Noisy Labels via Self-Supervised Adversarial Noisy Masking	16186
<i>Yuanpeng Tu (Tongji Univeristy, Shanghai), Boshen Zhang (YouTu Lab, Shanghai), Yuxi Li (YouTu Lab, Shanghai), Liang Liu (YouTu Lab, Shanghai), Jian Li (YouTu Lab, Shanghai), Jiangning Zhang (YouTu Lab, Shanghai), Yabiao Wang (YouTu Lab, Shanghai), Chengjie Wang (YouTu Lab, Shanghai; Shanghai Jiao Tong University), and Cai Rong Zhao (Tongji Univeristy, Shanghai)</i>	

Bit-Shrinking: Limiting Instantaneous Sharpness for Improving Post-Training Quantization .	16196
<i>Chen Lin (Hikvision Research Institute), Bo Peng (Hikvision Research Institute), Zheyang Li (Hikvision Research Institute), Wenming Tan (Hikvision Research Institute), Ye Ren (Hikvision Research Institute), Jun Xiao (Zhe Jiang University), and Shiliang Pu (Hikvision Research Institute)</i>	
Enhancing Multiple Reliability Measures via Nuisance-Extended Information Bottleneck	16206
<i>Jongheon Jeong (Korea Advanced Institute of Science and Technology (KAIST)), Sihyun Yu (Korea Advanced Institute of Science and Technology (KAIST)), Hankook Lee (LG AI Research), and Jinwoo Shin (Korea Advanced Institute of Science and Technology (KAIST))</i>	
AdaptiveMix: Improving GAN Training via Feature Space Shrinkage	16219
<i>Haozhe Liu (King Abdullah University of Science and Technology (KAUST)), Wentian Zhang (Jarvis Lab, Tencent), Bing Li (King Abdullah University of Science and Technology (KAUST)), Haoqian Wu (YouTu Lab, Tencent), Nanjun He (Jarvis Lab, Tencent), Yawen Huang (Jarvis Lab, Tencent), Yuexiang Li (Jarvis Lab, Tencent), Bernard Ghanem (King Abdullah University of Science and Technology (KAUST)), and Yefeng Zheng (Jarvis Lab, Tencent)</i>	
Re-GAN: Data-Efficient GANs Training via Architectural Reconfiguration	16230
<i>Divya Saxena (The Hong Kong Polytechnic University, Hong Kong), Jiannong Cao (The Hong Kong Polytechnic University, Hong Kong), Jiahao Xu (The Hong Kong Polytechnic University, Hong Kong), and Tarun Kulshrestha (The Hong Kong Polytechnic University, Hong Kong)</i>	
Soft Augmentation for Image Classification	16241
<i>Yang Liu (Argo AI), Shen Yan (Argo AI), Laura Leal-Taixé (Argo AI), James Hays (Argo AI), and Deva Ramanan (Argo AI)</i>	
Boosting Verified Training for Robust Image Classifications via Abstraction	16251
<i>Zhaodi Zhang (Shanghai Key Laboratory of Trustworthy Computing; East China Normal University; Chengdu Education Research Institute), Zhiyi Xue (East China Normal University), Yang Chen (East China Normal University), Si Liu (ETH Zurich), Yueling Zhang (East China Normal University), Jing Liu (Shanghai Key Laboratory of Trustworthy Computing), and Min Zhang (East China Normal University)</i>	
A New Dataset Based on Images Taken by Blind People for Testing the Robustness of Image Classification Models Trained for ImageNet Categories	16261
<i>Reza Akbarian Bafghi (University of Colorado, Boulder) and Danna Gurari (University of Colorado, Boulder)</i>	

Exploiting Completeness and Uncertainty of Pseudo Labels for Weakly Supervised Video Anomaly Detection	16271
<i>Chen Zhang (State Key Laboratory of Information Security, Institute of Information Engineering, CAS; University of Chinese Academy of Sciences), Guorong Li (University of Chinese Academy of Sciences), Yuankai Qi (Australian Institute for Machine Learning, The University of Adelaide), Shuhui Wang (Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, CAS; Peng Cheng Laboratory), Laiyun Qing (University of Chinese Academy of Sciences), Qingming Huang (State Key Laboratory of Information Security, Institute of Information Engineering, CAS; University of Chinese Academy of Sciences; Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, CAS), and Ming-Hsuan Yang (University of California, Merced; Yonsei University)</i>	
Prototypical Residual Networks for Anomaly Detection and Localization	16281
<i>Hui Zhang (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Zuxuan Wu (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Zheng Wang (Zhejiang University of Technology), Zhineng Chen (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), and Yu-Gang Jiang (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing)</i>	
Class Balanced Adaptive Pseudo Labeling for Federated Semi-Supervised Learning	16292
<i>Ming Li (East China Normal University), Qingli Li (East China Normal University), and Yan Wang (East China Normal University)</i>	
Fair Federated Medical Image Segmentation via Client Contribution Estimation	16302
<i>Meirui Jiang (The Chinese University of Hong Kong), Holger R. Roth (NVIDIA), Wenqi Li (NVIDIA), Dong Yang (NVIDIA), Can Zhao (NVIDIA), Vishwesh Nath (NVIDIA), Daguang Xu (NVIDIA), Qi Dou (The Chinese University of Hong Kong), and Ziyue Xu (NVIDIA)</i>	
Rethinking Federated Learning With Domain Shift: A Prototype View	16312
<i>Wenke Huang (Wuhan University, China), Mang Ye (Wuhan University, China; Hubei LuoJia Laboratory, China), Zekun Shi (Wuhan University, China), He Li (Wuhan University, China), and Bo Du (Wuhan University, China; Hubei LuoJia Laboratory, China)</i>	
FedDM: Iterative Distribution Matching for Communication-Efficient Federated Learning	16323
<i>Yuanhao Xiong (UCLA), Ruochen Wang (UCLA), Minhao Cheng (HKUST), Felix Yu (Google Research), and Cho-Jui Hsieh (UCLA)</i>	
Alias-Free Convnets: Fractional Shift Invariance via Polynomial Activations	16333
<i>Hagay Michaeli (Department of Electrical and Computer Engineering, Technion Haifa, Israel), Tomer Michaeli (Department of Electrical and Computer Engineering, Technion Haifa, Israel), and Daniel Soudry (Department of Electrical and Computer Engineering, Technion Haifa, Israel)</i>	

STDLens: Model Hijacking-Resilient Federated Learning for Object Detection	16343
<i>Ka-Ho Chow (Georgia Institute of Technology Atlanta, USA), Ling Liu (Georgia Institute of Technology Atlanta, USA), Wenqi Wei (Georgia Institute of Technology Atlanta, USA), Fatih Ilhan (Georgia Institute of Technology Atlanta, USA), and Yanzhao Wu (Georgia Institute of Technology Atlanta, USA)</i>	
Detecting Backdoors in Pre-Trained Encoders	16352
<i>Shiwei Feng (Purdue University), Guanhong Tao (Purdue University), Siyuan Cheng (Purdue University), Guangyu Shen (Purdue University), Xiangzhe Xu (Purdue University), Yingqi Liu (Purdue University), Kaiyuan Zhang (Purdue University), Shiqing Ma (Rutgers University), and Xiangyu Zhang (Purdue University)</i>	
Detecting Backdoors During the Inference Stage Based on Corruption Robustness Consistency	16363
<i>Xiaogeng Liu (Huazhong University of Science and Technology; National Engineering Research Center for Big Data Technology and System; Hubei Key Laboratory of Distributed System Security; Hubei Engineering Research Center on Big Data Security), Minghui Li (Huazhong University of Science and Technology), Haoyu Wang (Huazhong University of Science and Technology), Shengshan Hu (Huazhong University of Science and Technology; National Engineering Research Center for Big Data Technology and System; Hubei Key Laboratory of Distributed System Security; Hubei Engineering Research Center on Big Data Security), Dengpan Ye (Wuhan University), Hai Jin (Huazhong University of Science and Technology; National Engineering Research Center for Big Data Technology and System; Services Computing Technology and System Lab; Cluster and Grid Computing Lab), Libing Wu (Wuhan University), and Chaowei Xiao (Arizona State University)</i>	
Can't Steal? Cont-Steal! Contrastive Stealing Attacks Against Image Encoders	16373
<i>Zeyang Sha (CISPA Helmholtz Center for Information Security), Xinlei He (CISPA Helmholtz Center for Information Security), Ning Yu (Salesforce Research), Michael Backes (CISPA Helmholtz Center for Information Security), and Yang Zhang (CISPA Helmholtz Center for Information Security)</i>	
Re-Thinking Model Inversion Attacks Against Deep Neural Networks	16384
<i>Ngoc-Bao Nguyen (Singapore University of Technology and Design (SUTD)), Keshigeyan Chandrasegaran (Singapore University of Technology and Design (SUTD)), Milad Abdollahzadeh (Singapore University of Technology and Design (SUTD)), and Ngai-Man Cheung (Singapore University of Technology and Design (SUTD))</i>	
Turning Strengths Into Weaknesses: A Certified Robustness Inspired Attack Framework Against Graph Neural Networks	16394
<i>Binghui Wang (Department of Computer Science, Illinois Institute of Technology), Meng Pang (Nanchang University), and Yun Dong (Department of Visual and Performing Arts, Education, and Sciences, Waubensee Community College)</i>	

Dynamic Generative Targeted Attacks With Pattern Injection	16404
<i>Weiwei Feng (University of Science and Technology of China), Nanqing Xu (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China; Deep Space Exploration Lab), and Yongdong Zhang (University of Science and Technology of China)</i>	
Transferable Adversarial Attacks on Vision Transformers With Token Gradient Regularization.....	16415
<i>Jianping Zhang (The Chinese University of Hong Kong), Yizhan Huang (The Chinese University of Hong Kong), Weibin Wu (Sun Yat-sen University), and Michael R. Lyu (The Chinese University of Hong Kong)</i>	
Adversarial Counterfactual Visual Explanations	16425
<i>Guillaume Jeanneret (University of Caen Normandie, ENSICAEN, CNRS, France), Loïc Simon (University of Caen Normandie, ENSICAEN, CNRS, France), and Frédéric Jurie (University of Caen Normandie, ENSICAEN, CNRS, France)</i>	
TWINS: A Fine-Tuning Framework for Improved Transferability of Adversarial Robustness and Generalization	16436
<i>Ziquan Liu (City University of Hong Kong), Yi Xu (Dalian University of Technology), Xiangyang Ji (Tsinghua University), and Antoni B. Chan (City University of Hong Kong)</i>	
Randomized Adversarial Training via Taylor Expansion	16447
<i>Gaojie Jin (State Key Laboratory of Computer Science, Institute of Software, CAS, China; University of Liverpool, UK), Xinping Yi (University of Liverpool, UK), Dengyu Wu (University of Liverpool, UK), Ronghui Mu (Lancaster University, UK), and Xiaowei Huang (University of Liverpool, UK)</i>	
Improving Robust Generalization by Direct PAC-Bayesian Bound Minimization	16458
<i>Zifan Wang (Carnegie Mellon University), Nan Ding (Google Research), Tomer Levinboim (Google Research), Xi Chen (Google Research), and Radu Soricut (Google Research)</i>	
Evading Forensic Classifiers With Attribute-Conditioned Adversarial Faces	16469
<i>Fahad Shamshad (Mohamed bin Zayed University of AI, UAE), Koushik Srivatsan (Mohamed bin Zayed University of AI, UAE), and Karthik Nandakumar (Mohamed bin Zayed University of AI, UAE)</i>	
DartBlur: Privacy Preservation With Detection Artifact Suppression	16479
<i>Baowei Jiang (Tsinghua University), Bing Bai (Tsinghua University), Haozhe Lin (Tsinghua University), Yu Wang (Tsinghua University), Yuchen Guo (Tsinghua University), and Lu Fang (Tsinghua University)</i>	

Poster-Thu-AM

Fresnel Microfacet BRDF: Unification of Polari-Radiometric Surface-Body Reflection	16489
<i>Tomoki Ichikawa (Kyoto University), Yoshiki Fukao (Kyoto University), Shohei Nobuhara (Kyoto University), and Ko Nishino (Kyoto University)</i>	

JacobiNeRF: NeRF Shaping With Mutual Information Gradients	16498
<i>Xiaomeng Xu (Tsinghua University), Yanchao Yang (The University of Hong Kong; Stanford University), Kaichun Mo (Stanford University; NVIDIA Research), Boxiao Pan (Stanford University), Li Yi (Tsinghua University; Shanghai AI Laboratory; Shanghai Qizhi Institute), and Leonidas Guibas (Stanford University; Google Research)</i>	
ContraNeRF: Generalizable Neural Radiance Fields for Synthetic-to-Real Novel View Synthesis via Contrastive Learning	16508
<i>Hao Yang (Peking University; Zhejiang Lab), Lanqing Hong (Huawei Noah's Ark Lab), Aoxue Li (Huawei Noah's Ark Lab), Tianyang Hu (Huawei Noah's Ark Lab), Zhenguo Li (Huawei Noah's Ark Lab), Gim Hee Lee (National University of Singapore), and Liwei Wang (Peking University; Peking University)</i>	
SCADE: NeRFs from Space Carving With Ambiguity-Aware Depth Estimates	16518
<i>Mikaela Angelina Uy (Stanford University; Google), Ricardo Martin-Brualla (Google), Leonidas Guibas (Stanford University; Google), and Ke Li (Google; Simon Fraser University)</i>	
Removing Objects From Neural Radiance Fields	16528
<i>Silvan Weder (Niantic; ETH Zurich), Guillermo Garcia-Hernando (Niantic), Áron Monzspart (Niantic), Marc Pollefeys (ETH Zurich), Gabriel J. Brostow (Niantic; University College London), Michael Firman (Niantic), and Sara Vicente (Niantic)</i>	
Progressively Optimized Local Radiance Fields for Robust View Synthesis	16539
<i>Andreas Meuleman (KAIST), Yu-Lun Liu (National Taiwan University), Chen Gao (Meta), Jia-Bin Huang (Meta; University of Maryland, College Park), Changil Kim (Meta), Min H. Kim (KAIST), and Johannes Kopf (Meta)</i>	
NeRFVS: Neural Radiance Fields for Free View Synthesis via Geometry Scaffolds	16549
<i>Chen Yang (Shanghai Jiao Tong University), Peihao Li (Tsinghua University), Zanwei Zhou (Shanghai Jiao Tong University), Shanxin Yuan (Huawei Noah's Ark Lab), Bingbing Liu (Huawei Noah's Ark Lab), Xiaokang Yang (Shanghai Jiao Tong University), Weichao Qiu (Huawei Noah's Ark Lab), and Wei Shen (Shanghai Jiao Tong University)</i>	
ABLE-NeRF: Attention-Based Rendering With Learnable Embeddings for Neural Radiance Field	16559
<i>Zhe Jun Tang (Nanyang Technological University), Tat-Jen Cham (Nanyang Technological University), and Haiyu Zhao (SenseTime Research)</i>	
MobileNeRF: Exploiting the Polygon Rasterization Pipeline for Efficient Neural Field Rendering on Mobile Architectures	16569
<i>Zhiqin Chen (Google Research; Simon Fraser University), Thomas Funkhouser (Google Research), Peter Hedman (Google Research), and Andrea Tagliasacchi (Google Research; Simon Fraser University; University of Toronto)</i>	
pCON: Polarimetric Coordinate Networks for Neural Scene Representations	16579
<i>Henry Peters (University of California, Los Angeles), Yunhao Ba (Electrical and Computer Engineering Department, UCLA), and Achuta Kadambi (University of California, Los Angeles; Electrical and Computer Engineering Department, UCLA)</i>	

Balanced Spherical Grid for Egocentric View Synthesis	16590
<i>Changwoon Choi (Seoul National University, Korea), Sang Min Kim (Seoul National University, Korea), and Young Min Kim (Seoul National University)</i>	
Complementary Intrinsic From Neural Radiance Fields and CNNs for Outdoor Scene Relighting... 16600	
<i>Siqi Yang (Peking University), Xuanning Cui (Peking University), Yongjie Zhu (Beijing University of Posts and Telecommunications), Jiajun Tang (Peking University), Si Li (Beijing University of Posts and Telecommunications), Zhaofer Yu (Peking University), and Boxin Shi (Peking University)</i>	
HyperReel: High-Fidelity 6-DoF Video With Ray-Conditioned Sampling	16610
<i>Benjamin Attal (Carnegie Mellon University; Meta), Jia-Bin Huang (University of Maryland; Meta), Christian Richardt (Reality Labs Research), Michael Zollhöfer (Reality Labs Research), Johannes Kopf (Meta), Matthew O'Toole (Carnegie Mellon University), and Changil Kim (Meta)</i>	
UV Volumes for Real-Time Rendering of Editable Free-View Human Performance	16621
<i>Yue Chen (National Key Laboratory of Human-Machine Hybrid Augmented Intelligence; Xi'an Jiaotong University), Xuan Wang (Ant Group; Tencent AI Lab), Xingyu Chen (National Key Laboratory of Human-Machine Hybrid Augmented Intelligence; Xi'an Jiaotong University), Qi Zhang (Tencent AI Lab), Xiaoyu Li (Tencent AI Lab), Yu Guo (National Key Laboratory of Human-Machine Hybrid Augmented Intelligence; Xi'an Jiaotong University; Tencent AI Lab), Jue Wang (Tencent AI Lab), and Fei Wang (National Key Laboratory of Human-Machine Hybrid Augmented Intelligence; Xi'an Jiaotong University)</i>	
Tensor4D: Efficient Neural 4D Decomposition for High-Fidelity Dynamic Reconstruction and Rendering	16632
<i>Ruizhi Shao (Tsinghua University), Zerong Zheng (Tsinghua University; NNKosmos Technology), Hanzhang Tu (Tsinghua University), Boning Liu (Tsinghua University), Hongwen Zhang (Tsinghua University), and Yebin Liu (Tsinghua University)</i>	
PixHt-Lab: Pixel Height Based Light Effect Generation for Image Compositing	16643
<i>Yichen Sheng (Purdue University), Jianming Zhang (Adobe Inc.), Julien Philip (Adobe Inc.), Yannick Hold-Geoffroy (Adobe Inc.), Xin Sun (Adobe Inc.), He Zhang (Adobe Inc.), Lu Ling (Purdue University), and Bedrich Benes (Purdue University)</i>	
Computational Flash Photography Through Intrinsic	16654
<i>Sepideh Sarajian Maralan (Simon Fraser University), Chris Careaga (Simon Fraser University), and Yagiz Aksoy (Simon Fraser University)</i>	
RelightableHands: Efficient Neural Relighting of Articulated Hand Models	16663
<i>Shun Iwase (Carnegie Mellon University; Reality Labs Research), Shunsuke Saito (Reality Labs Research), Tomas Simon (Reality Labs Research), Stephen Lombardi (Reality Labs Research), Timur Bagautdinov (Reality Labs Research), Rohan Joshi (Reality Labs Research), Fabian Prada (Reality Labs Research), Takaaki Shiratori (Reality Labs Research), Yaser Sheikh (Reality Labs Research), and Jason Saragih (Reality Labs Research)</i>	

TMO: Textured Mesh Acquisition of Objects With a Mobile Device by Using Differentiable Rendering	16674
<i>Jaehoon Choi (NAVER LABS; University of Maryland), Dongki Jung (NAVER LABS), Taejae Lee (NAVER LABS), Sangwook Kim (NAVER LABS), Youngdong Jung (NAVER LABS), Dinesh Manocha (University of Maryland), and Donghwan Lee (NAVER LABS)</i>	
VolRecon: Volume Rendering of Signed Ray Distance Functions for Generalizable Multi-View Reconstruction	16685
<i>Yufan Ren (IVRF IC EPFL), Fangjinhua Wang (Department of Computer Science, ETH Zurich), Tong Zhang (IVRF IC EPFL), Marc Pollefeys (Department of Computer Science, ETH Zurich), and Sabine Süssstrunk (IVRF IC EPFL)</i>	
Multi-View Reconstruction Using Signed Ray Distance Functions (SRDF)	16696
<i>Pierre Zins (Inria centre at the University Grenoble Alpes; Meta Reality Labs, USA), Yuanlu Xu (Meta Reality Labs, USA), Edmond Boyer (Inria centre at the University Grenoble Alpes; Meta Reality Labs, Switzerland), Stefanie Wuhrer (Inria centre at the University Grenoble Alpes), and Tony Tung (Meta Reality Labs, USA)</i>	
Structural Multiplane Image: Bridging Neural View Synthesis and 3D Reconstruction	16707
<i>Mingfang Zhang (The University of Tokyo; Microsoft Research Asia), Jinglu Wang (Microsoft Research Asia), Xiao Li (Microsoft Research Asia), Yifei Huang (The University of Tokyo), Yoichi Sato (The University of Tokyo), and Yan Lu (Microsoft Research Asia)</i>	
Octree Guided Unoriented Surface Reconstruction	16717
<i>Chamin Hewa Koneputugodage (The Australian National University), Yizhak Ben-Shabat (The Australian National University; Technion Israel Institute of Technology), and Stephen Gould (The Australian National University)</i>	
Neural Vector Fields: Implicit Representation by Explicit Learning	16727
<i>Xianghui Yang (The University of Sydney), Guosheng Lin (Nanyang Technological University), Zhenghao Chen (The University of Sydney), and Luping Zhou (The University of Sydney)</i>	
DA Wand: Distortion-Aware Selection Using Neural Mesh Parameterization	16739
<i>Richard Liu (University of Chicago), Noam Aigerman (Adobe Research), Vladimir G. Kim (Adobe Research), and Rana Hanocka (University of Chicago)</i>	

Diffusion-Based Generation, Optimization, and Planning in 3D Scenes	16750
<i>Siyuan Huang (National Key Laboratory of General Artificial Intelligence, BIGAI), Zan Wang (National Key Laboratory of General Artificial Intelligence, BIGAI; School of Computer Science & Technology, Beijing Institute of Technology), Puhao Li (National Key Laboratory of General Artificial Intelligence, BIGAI; Dept. of Automation, Tsinghua University 4 Institute for AI, Peking University), Baoxiong Jia (National Key Laboratory of General Artificial Intelligence, BIGAI), Tengyu Liu (National Key Laboratory of General Artificial Intelligence, BIGAI), Yixin Zhu (Institute for AI, Peking University), Wei Liang (School of Computer Science & Technology, Beijing Institute of Technology; Yangtze Delta Region Academy of Beijing Institute of Technology, Jiaxing), and Song-Chun Zhu (National Key Laboratory of General Artificial Intelligence, BIGAI; Dept. of Automation, Tsinghua University 4 Institute for AI, Peking University; Institute for AI, Peking University)</i>	
Patch-Based 3D Natural Scene Generation From a Single Example	16762
<i>Weiyu Li (Shandong University), Xuelin Chen (Tencent AI Lab; Shandong University), Jue Wang (Tencent AI Lab), and Baoquan Chen (Peking University)</i>	
Consistent View Synthesis With Pose-Guided Diffusion Models	16773
<i>Hung-Yu Tseng (Meta), Qinbo Li (Meta), Changil Kim (Meta), Suhib Alsisan (Meta; University of Maryland, College Park), Jia-Bin Huang (Meta; University of Maryland, College Park), and Johannes Kopf (Meta)</i>	
Generalized Deep 3D Shape Prior via Part-Discretized Diffusion Process	16784
<i>Yuhan Li (Shanghai Jiao Tong University, Shanghai 200240, China), Yishun Dou (Huawei), Xuanhong Chen (Shanghai Jiao Tong University, Shanghai 200240, China), Bingbing Ni (Shanghai Jiao Tong University, Shanghai 200240, China; Huawei), Yilin Sun (Shanghai Jiao Tong University, Shanghai 200240, China), Yutian Liu (Shanghai Jiao Tong University, Shanghai 200240, China), and Fuzhen Wang (Shanghai Jiao Tong University, Shanghai 200240, China)</i>	
High Fidelity 3D Hand Shape Reconstruction via Scalable Graph Frequency Decomposition	16795
<i>Tianyu Luan (State University of New York at Buffalo), Yuanhao Zhai (State University of New York at Buffalo), Jingjing Meng (State University of New York at Buffalo), Zhong Li (OPPO US Research Center, InnoPeak Technology, Inc.), Zhang Chen (OPPO US Research Center, InnoPeak Technology, Inc.), Yi Xu (OPPO US Research Center, InnoPeak Technology, Inc.), and Junsong Yuan (State University of New York at Buffalo)</i>	
TAPS3D: Text-Guided 3D Textured Shape Generation From Pseudo Supervision	16805
<i>Jiacheng Wei (Nanyang Technological University, Singapore), Hao Wang (Nanyang Technological University, Singapore), Jiashi Feng (ByteDance), Guosheng Lin (Nanyang Technological University, Singapore), and Kim-Hui Yap (Nanyang Technological University, Singapore)</i>	

SECAD-Net: Self-Supervised CAD Reconstruction by Learning Sketch-Extrude Operations ...	16816
<i>Pu Li (MAIS, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Jianwei Guo (MAIS, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Xiaopeng Zhang (MAIS, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), and Dong-Ming Yan (MAIS, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences)</i>	
Interactive Cartoonization With Controllable Perceptual Factors	16827
<i>Namhyuk Ahn (NAVER WEBTOON AI), Patrick Kwon (NAVER WEBTOON AI), Jihye Back (NAVER WEBTOON AI), Kibeom Hong (Yonsei University), and Seungkwon Kim (NAVER WEBTOON AI)</i>	
High-Res Facial Appearance Capture From Polarized Smartphone Images	16836
<i>Dejan Azinović (Technical University of Munich), Olivier Maury (Meta Reality Labs), Christophe Hery (Meta Reality Labs), Matthias Nießner (Technical University of Munich), and Justus Thies (Max Planck Institute for Intelligent Systems)</i>	
GlassesGAN: Eyewear Personalization Using Synthetic Appearance Discovery and Targeted Subspace Modeling	16847
<i>Richard Plesh (Clarkson University, USA), Peter Peer (University of Ljubljana, Slovenia), and Vitomir Struc (University of Ljubljana, Slovenia)</i>	
Continuous Landmark Detection With 3D Queries	16858
<i>Prashanth Chandran (Disney Research Studios), Gaspard Zoss (Disney Research Studios), Paulo Gotardo (Disney Research Studios), and Derek Bradley (Disney Research Studios)</i>	
NeuFace: Realistic 3D Neural Face Rendering From Multi-View Images	16868
<i>Mingwu Zheng (Beihang University, China), Haiyu Zhang (Beihang University, China), Hongyu Yang (Beihang University, China), and Di Huang (Beihang University, China)</i>	
AVFace: Towards Detailed Audio-Visual 4D Face Reconstruction	16878
<i>Aggelina Chatziagapi (Stony Brook University) and Dimitris Samaras (Stony Brook University)</i>	
Learning Personalized High Quality Volumetric Head Avatars From Monocular RGB Videos .	16890
<i>Ziqian Bai (Google; Simon Fraser University), Feitong Tan (Google), Zeng Huang (Google), Kripasindhu Sarkar (Google), Danhang Tang (Google), Di Qiu (Google), Abhimitra Meka (Google), Ruofei Du (Google), Mingsong Dou (Google), Sergio Orts-Escolano (Google), Rohit Pandey (Google), Ping Tan (Simon Fraser University), Thabo Beeler (Google), Sean Fanello (Google), and Yinda Zhang (Google)</i>	
OTAvatar: One-Shot Talking Face Avatar With Controllable Tri-Plane Rendering	16901
<i>Zhiyuan Ma (The Hong Kong Polytechnic University; HKISI CAS), Xiangyu Zhu (CASIA), Guo-Jun Qi (OPPO Research), Zhen Lei (The Hong Kong Polytechnic University; HKISI CAS; CASIA), and Lei Zhang (The Hong Kong Polytechnic University)</i>	

X-Avatar: Expressive Human Avatars	16911
<i>Kaiyue Shen (ETH Zürich), Chen Guo (ETH Zürich), Manuel Kaufmann (ETH Zürich), Juan Jose Zarate (ETH Zürich), Julien Valentin (Microsoft), Jie Song (ETH Zürich), and Otmar Hilliges (ETH Zürich)</i>	
InstantAvatar: Learning Avatars From Monocular Video in 60 Seconds	16922
<i>Tianjian Jiang (ETH Zürich), Xu Chen (ETH Zürich; Max Planck Institute for Intelligent Systems, Tübingen), Jie Song (ETH Zürich), and Otmar Hilliges (ETH Zürich)</i>	
JAWS: Just a Wild Shot for Cinematic Transfer in Neural Radiance Fields	16933
<i>Xi Wang (Inria, IRISA, CNRS, Univ. Rennes), Robin Courant (Inria, IRISA, CNRS, Univ. Rennes; LIX, Ecole Polytechnique, IP Paris), Jinglei Shi (VCIP, CS, Nankai Univ.), Eric Marchand (Inria, IRISA, CNRS, Univ. Rennes), and Marc Christie (Inria, IRISA, CNRS, Univ. Rennes)</i>	
MonoHuman: Animatable Human Neural Field From Monocular Video	16943
<i>Zhengming Yu (SenseTime Research), Wei Cheng (SenseTime Research; Shanghai AI Laboratory), Xian Liu (The Chinese University of Hong Kong), Wayne Wu (Shanghai AI Laboratory), and Kwan-Yee Lin (Shanghai AI Laboratory; The Chinese University of Hong Kong)</i>	
Structured 3D Features for Reconstructing Controllable Avatars	16954
<i>Enric Corona (UPC, Barcelona; Google Research), Mihai Zanfir (Newton), Thiemo Alldieck (Google Research), Eduard Gabriel Bazavan (Google Research), Andrei Zanfir (Google Research), and Cristian Sminchisescu (Google Research)</i>	
HOOD: Hierarchical Graphs for Generalized Modelling of Clothing Dynamics	16965
<i>Artur Grigorev (ETH Zurich, Department of Computer Science), Michael J. Black (Max Planck Institute for Intelligent Systems, Tubingen), and Otmar Hilliges (ETH Zurich, Department of Computer Science; Max Planck Institute for Intelligent Systems, Tubingen)</i>	
Physically Realizable Natural-Looking Clothing Textures Evade Person Detectors via 3D Modeling	16975
<i>Zhanhao Hu (Tsinghua University, Beijing, China), Wenda Chu (Tsinghua University, China; Tsinghua University, Beijing, China), Xiaopei Zhu (Tsinghua University, China; Tsinghua University, Beijing, China), Hui Zhang (Beijing Institute of Fashion Technology, Beijing, China), Bo Zhang (Tsinghua University, Beijing, China), and Xiaolin Hu (Tsinghua University, China; Tsinghua University, China; Chinese Institute for Brain Research (CIBR), Beijing, China)</i>	
Learning Semantic-Aware Disentangled Representation for Flexible 3D Human Body Editing	16985
<i>Xiaokun Sun (Tianjin University, China), Qiao Feng (Tianjin University, China), Xiongzhen Li (Tianjin University, China), Jinsong Zhang (Tianjin University, China), Yu-Kun Lai (Cardiff University, United Kingdom), Jingyu Yang (Tianjin University, China), and Kun Li (Tianjin University, China)</i>	
Reconstructing Animatable Categories From Videos	16995
<i>Gengshan Yang (Carnegie Mellon University), Chaoyang Wang (Carnegie Mellon University), N. Dinesh Reddy (Carnegie Mellon University), and Deva Ramanan (Carnegie Mellon University)</i>	

Deformable Mesh Transformer for 3D Human Mesh Recovery	17006
<i>Yusuke Yoshiyasu (National Institute of Advanced Industrial Science and Technology (AIST), Japan)</i>	
Hi4D: 4D Instance Segmentation of Close Human Interaction	17016
<i>Yifei Yin (ETH Zurich), Chen Guo (ETH Zurich), Manuel Kaufmann (ETH Zurich), Juan Jose Zarate (ETH Zurich), Jie Song (ETH Zurich), and Otmar Hilliges (ETH Zurich)</i>	
Bringing Inputs to Shared Domains for 3D Interacting Hands Recovery in the Wild	17028
<i>Gyeongsik Moon (Meta Reality Labs)</i>	
Learning Human Mesh Recovery in 3D Scenes	17038
<i>Zehong Shen (Zhejiang University), Zhi Cen (Zhejiang University), Sida Peng (Zhejiang University), Qing Shuai (Zhejiang University), Hujun Bao (Zhejiang University), and Xiaowei Zhou (Zhejiang University)</i>	
H2ONet: Hand-Occlusion-and-Orientation-Aware Network for Real-Time 3D Hand Mesh Reconstruction	17048
<i>Hao Xu (Institute of Medical Intelligence and XR), Tianyu Wang (Department of Computer Science and Engineering), Xiao Tang (Department of Computer Science and Engineering), and Chi-Wing Fu (Department of Computer Science and Engineering; Institute of Medical Intelligence and XR; Shun Hing Institute of Advanced Engineering The Chinese University of Hong Kong)</i>	
What You Can Reconstruct From a Shadow	17059
<i>Ruoshi Liu (Columbia University), Sachit Menon (Columbia University), Chengzhi Mao (Columbia University), Dennis Park (Toyota Research Institute), Simon Stent (Woven Planet), and Carl Vondrick (Columbia University)</i>	
Autonomous Manipulation Learning for Similar Deformable Objects via Only One Demonstration.	17069
<i>Yu Ren (State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences; Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences; University of Chinese Academy of Sciences, Beijing, 100049, China), Ronghan Chen (State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences; Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences; University of Chinese Academy of Sciences, Beijing, 100049, China), and Yang Cong (State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences; Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences)</i>	
In-Hand 3D Object Scanning From an RGB Sequence	17079
<i>Shreyas Hampali (Reality Labs at Meta; Graz University of Technology, Austria), Tomas Hodan (Reality Labs at Meta), Luan Tran (Reality Labs at Meta), Lingni Ma (Reality Labs at Meta), Cem Keskin (Reality Labs at Meta), and Vincent Lepetit (Univ Gustave Eiffel, CNRS, France; Graz University of Technology, Austria)</i>	

Putting People in Their Place: Affordance-Aware Human Insertion Into Scenes	17089
<i>Sumith Kulal (Stanford University), Tim Brooks (UC Berkeley), Alex Aiken (Stanford University), Jiajun Wu (Stanford University), Jimei Yang (Adobe Research), Jingwan Lu (Adobe Research), Alexei A. Efros (UC Berkeley), and Krishna Kumar Singh (Stanford University)</i>	
Detecting Human-Object Contact in Images	17100
<i>Yixin Chen (Beijing Institute of General Artificial Intelligence, China; Work done while interning at MPI-IS2), Sai Kumar Dwivedi (Max Planck Institute for Intelligent Systems, Germany), Michael J. Black (Max Planck Institute for Intelligent Systems, Germany), and Dimitrios Tzionas (University of Amsterdam, The Netherlands)</i>	
What Happened 3 Seconds Ago? Inferring the Past With Thermal Imaging	17111
<i>Zitian Tang (Tsinghua University), Wenjie Ye (Tsinghua University), Wei-Chiu Ma (MIT), and Hang Zhao (Tsinghua University; Shanghai Qi Zhi Institute)</i>	
Trajectory-Aware Body Interaction Transformer for Multi-Person Pose Forecasting	17121
<i>Xiaogang Peng (Apple), Siyuan Mao (Apple), and Zizhao Wu (Apple)</i>	
Mutual Information-Based Temporal Difference Learning for Human Pose Estimation in Video	17131
<i>Runyang Feng (Jilin University; Ministry of Education, China), Yixing Gao (Jilin University; Ministry of Education, China), Xueqing Ma (Jilin University; Ministry of Education, China), Tze Ho Elden Tse (University of Birmingham), and Hyung Jin Chang (University of Birmingham)</i>	
Ego-Body Pose Estimation via Ego-Head Pose Estimation	17142
<i>Jiaman Li (Stanford University), Karen Liu (Stanford University), and Jiajun Wu (Stanford University)</i>	
ViPLO: Vision Transformer Based Pose-Conditioned Self-Loop Graph for Human-Object Interaction Detection	17152
<i>Jeeseung Park (mAy-I Inc., Korea), Jin-Woo Park (mAy-I Inc., Korea; Yonsei University, Korea), and Jong-Seok Lee (Yonsei University, Korea)</i>	
HS-Pose: Hybrid Scope Feature Extraction for Category-Level Object Pose Estimation	17163
<i>Linfang Zheng (Southern University of Science and Technology; University of Birmingham), Chen Wang (Southern University of Science and Technology; University of Hong Kong), Yinghan Sun (Southern University of Science and Technology), Esha Dasgupta (University of Birmingham), Hua Chen (Southern University of Science and Technology), Aleš Leonardis (University of Birmingham), Wei Zhang (Southern University of Science and Technology; Peng Cheng Laboratory, China), and Hyung Jin Chang (University of Birmingham)</i>	
ScarceNet: Animal Pose Estimation With Scarce Annotations	17174
<i>Chen Li (National University of Singapore) and Gim Hee Lee (National University of Singapore)</i>	
Cross-Domain 3D Hand Pose Estimation With Dual Modalities	17184
<i>Qiuxia Lin (National University of Singapore), Linlin Yang (National University of Singapore), and Angela Yao (National University of Singapore)</i>	

Linking Garment With Person via Semantically Associated Landmarks for Virtual Try-On	17194
<i>Keyu Yan (Alibaba Group; Hefei Institute of Physical Science, Chinese Academy of Sciences, China), Tingwei Gao (Alibaba Group), Hui Zhang (Hefei Institute of Physical Science, Chinese Academy of Sciences, China; University of Science and Technology of China, China), and Chengjun Xie (Hefei Institute of Physical Science, Chinese Academy of Sciences, China)</i>	
Level-S2fM: Structure From Motion on Neural Level Set of Implicit Surfaces	17205
<i>Yuxi Xiao (Wuhan University), Nan Xue (Wuhan University), Tianfu Wu (NC State University), and Gui-Song Xia (Wuhan University)</i>	
Revisiting Rotation Averaging: Uncertainties and Robust Losses	17215
<i>Ganlin Zhang (Lund University), Viktor Larsson (Lund University), and Daniel Barath (Lund University)</i>	
SliceMatch: Geometry-Guided Aggregation for Cross-View Pose Estimation	17225
<i>Ted Lentsch (Delft University of Technology, The Netherlands), Zimin Xia (Delft University of Technology, The Netherlands), Holger Caesar (Delft University of Technology, The Netherlands), and Julian F. P. Kooij (Delft University of Technology, The Netherlands)</i>	
Learning the Distribution of Errors in Stereo Matching for Joint Disparity and Uncertainty Estimation	17235
<i>Liyan Chen (Stevens Institute of Technology), Weihang Wang (Stevens Institute of Technology), and Philippos Mordohai (Stevens Institute of Technology)</i>	
Long-Term Visual Localization With Mobile Sensors	17245
<i>Shen Yan (National University of Defense Technology; Zhejiang University), Yu Liu (National University of Defense Technology), Long Wang (SenseTime Research), Zehong Shen (Zhejiang University), Zhen Peng (SenseTime Research), Haomin Liu (SenseTime Research), Maojun Zhang (National University of Defense Technology), Guofeng Zhang (Zhejiang University), and Xiaowei Zhou (Zhejiang University)</i>	
Learning To Predict Scene-Level Implicit 3D From Posed RGBD Data	17256
<i>Nilesh Kulkarni (University of Michigan), Linyi Jin (University of Michigan), Justin Johnson (University of Michigan), and David F. Fouhey (University of Michigan)</i>	
Paired-Point Lifting for Enhanced Privacy-Preserving Visual Localization	17266
<i>Chunghwan Lee (Hanyang University), Jaihoon Kim (Samsung Electronics; KAIST), Chanhyuk Yun (Hanyang University), and Je Hyeong Hong (Hanyang University)</i>	
The ObjectFolder Benchmark: Multisensory Learning With Neural and Real Objects	17276
<i>Ruohan Gao (Stanford Univeristy), Yiming Dou (Stanford Univeristy), Hao Li (Stanford Univeristy), Tanmay Agarwal (Stanford Univeristy), Jeannette Bohg (Stanford Univeristy), Yunzhu Li (Stanford Univeristy), Li Fei-Fei (Stanford Univeristy), and Jiajun Wu (Stanford Univeristy)</i>	

Learning Accurate 3D Shape Based on Stereo Polarimetric Imaging	17287
<i>Tianyu Huang (The Chinese University of Hong Kong, China), Haoang Li (The Chinese University of Hong Kong, China; Technical University of Munich, Germany), Kejing He (The Chinese University of Hong Kong, China), Congying Sui (The Chinese University of Hong Kong, China), Bin Li (The Chinese University of Hong Kong, China), and Yun-Hui Liu (The Chinese University of Hong Kong, China)</i>	
RUST: Latent Neural Scene Representations From Unposed Imagery	17297
<i>Mehdi S. M. Sajjadi (Google Research, Brain Team), Aravindh Mahendran (Google Research, Brain Team), Thomas Kipf (Google Research, Brain Team), Etienne Pot (Google Research, Brain Team), Daniel Duckworth (Google Research, Brain Team), Mario Lučić (Google Research, Brain Team), and Klaus Greff (Google Research, Brain Team)</i>	
Perspective Fields for Single Image Camera Calibration	17307
<i>Linyi Jin (University of Michigan), Jianming Zhang (Adobe Research), Yannick Hold-Geoffroy (Adobe Research), Oliver Wang (Adobe Research), Kevin Blackburn-Matzen (Adobe Research), Matthew Sticha (University of Michigan), and David F. Fouhey (University of Michigan)</i>	
VisFusion: Visibility-Aware Online 3D Scene Reconstruction From Videos	17317
<i>Huiyu Gao (Australian National University), Wei Mao (Australian National University), and Miaomiao Liu (Australian National University)</i>	
DeepLSD: Line Segment Detection and Refinement With Deep Image Gradients	17327
<i>Rémi Pautrat (Department of Computer Science, ETH Zurich), Daniel Barath (Department of Computer Science, ETH Zurich), Viktor Larsson (Lund University), Martin R. Oswald (Department of Computer Science, ETH Zurich; University of Amsterdam), and Marc Pollefeys (Department of Computer Science, ETH Zurich; Microsoft Mixed Reality and AI Zurich Lab)</i>	
Disentangling Orthogonal Planes for Indoor Panoramic Room Layout Estimation With Cross-Scale Distortion Awareness	17337
<i>Zhijie Shen (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology), Zishuo Zheng (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology), Chunyu Lin (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology), Lang Nie (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology), Kang Liao (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology), Shuai Zheng (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology), and Yao Zhao (Beijing Jiaotong University, China; Beijing Key Laboratory of Advanced Information Science and Network Technology)</i>	
Single Image Depth Prediction Made Better: A Multivariate Gaussian Take	17346
<i>Ce Liu (CVLETH Zurich), Suryansh Kumar (CVLETH Zurich), Shuhang Gu (UESTC China), Radu Timofte (CVLETH Zurich; University of Wurzburg), and Luc Van Gool (CVLETH Zurich; KU Leuven)</i>	

Wide-Angle Rectification via Content-Aware Conformal Mapping	17357
<i>Qi Zhang (Tencent AI Lab), Hongdong Li (Australian National University), and Qing Wang (Northwestern Polytechnical University)</i>	
All-in-Focus Imaging From Event Focal Stack	17366
<i>Hanyue Lou (Peking University), Minggui Teng (Peking University), Yixin Yang (Peking University), and Boxin Shi (Peking University)</i>	
Multi-View Stereo Representation Revist: Region-Aware MVSNet	17376
<i>Yisu Zhang (Zhejiang University), Jianke Zhu (Zhejiang University; Alibaba-Zhejiang University Joint Research Institute of Frontier Technologies), and Lixiang Lin (Zhejiang University)</i>	
Semantic Ray: Learning a Generalizable Semantic Field With Cross-Reprojection Attention .	17386
<i>Fangfu Liu (MEGVII Technology), Chubin Zhang (MEGVII Technology), Yu Zheng (MEGVII Technology), and Yueqi Duan (MEGVII Technology)</i>	
OmniCity: Omnipotent City Understanding With Multi-Level and Multi-View Images	17397
<i>Weijia Li (Sun Yat-Sen University), Yawen Lai (SenseTime Research), Linning Xu (The Chinese University of Hong Kong), Yuanbo Xiangli (The Chinese University of Hong Kong), Jinhua Yu (Sun Yat-Sen University), Conghui He (SenseTime Research; Shanghai Artificial Intelligence Laboratory), Gui-Song Xia (Wuhan University), and Dahua Lin (The Chinese University of Hong Kong; Shanghai Artificial Intelligence Laboratory)</i>	
ESLAM: Efficient Dense SLAM System Based on Hybrid Representation of Signed Distance Fields	17408
<i>Mohammad Mahdi Johari (Idiap Research Institute, EPFL), Camilla Carta (Idiap Research Institute, EPFL), and François Fleuret (Idiap Research Institute, EPFL)</i>	
Non-Line-of-Sight Imaging With Signal Superresolution Network	17420
<i>Jianyu Wang (Tsinghua University), Xintong Liu (Tsinghua University), Leping Xiao (Tsinghua University), Zuoqiang Shi (Tsinghua University; Yanqi Lake Beijing Institute of Mathematical Sciences and Applications), Lingyun Qiu (Tsinghua University; Yanqi Lake Beijing Institute of Mathematical Sciences and Applications), and Xing Fu (Tsinghua University)</i>	
Look, Radiate, and Learn: Self-Supervised Localisation via Radio-Visual Correspondence	17430
<i>Mohammed Alloulah (Nokia Bell Labs) and Maximilian Arnold (Nokia Bell Labs)</i>	
Learning Transformations To Reduce the Geometric Shift in Object Detection	17441
<i>Vidit Vidit (EPFL), Martin Engilberge (EPFL), and Mathieu Salzmann (ClearSpace SA)</i>	
Anchor3DLane: Learning To Regress 3D Anchors for Monocular 3D Lane Detection	17451
<i>Shaofei Huang (Institute of Information Engineering, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Zhenwei Shen (TuSimple), Zehao Huang (TuSimple), Zi-han Ding (Beihang University), Jiao Dai (Institute of Information Engineering, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Jizhong Han (Institute of Information Engineering, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Naiyan Wang (TuSimple), and Si Liu (Beihang University)</i>	

BEV-SAN: Accurate BEV 3D Object Detection via Slice Attention Networks	17461
<i>Xiaowei Chi (Peking University; The Chinese University of Hong Kong), Jiaming Liu (Peking University), Ming Lu (Peking University), Rongyu Zhang (Peking University), Zhaoqing Wang (The University of Sydney), Yandong Guo (AI2Robotics), and Shanghang Zhang (Peking University)</i>	
Semi-Supervised Stereo-Based 3D Object Detection via Cross-View Consensus	17471
<i>Wenhao Wu (City University of Hong Kong), Hau San Wong (City University of Hong Kong), and Si Wu (South China University of Technology)</i>	
Weakly Supervised Monocular 3D Object Detection Using Multi-View Projection and Direction Consistency	17482
<i>Runzhou Tao (Beijing Institute of Technology; QCraft), Wencheng Han (University of Macau), Zhongying Qiu (QCraft), Cheng-Zhong Xu (University of Macau), and Jianbing Shen (University of Macau)</i>	
MonoATT: Online Monocular 3D Object Detection With Adaptive Token Transformer	17493
<i>Yunsong Zhou (Shanghai Jiao Tong University), Hongzi Zhu (Shanghai Jiao Tong University), Quan Liu (Shanghai Jiao Tong University), Shan Chang (Donghua University), and Minyi Guo (Shanghai Jiao Tong University)</i>	
Azimuth Super-Resolution for FMCW Radar in Autonomous Driving	17504
<i>Yu-Jhe Li (Carnegie Mellon University), Shawn Hunt (DENSO International America, Inc.), Jinhung Park (Carnegie Mellon University), Matthew O'Toole (Carnegie Mellon University), and Kris Kitani (Carnegie Mellon University)</i>	
Pix2map: Cross-Modal Retrieval for Inferring Street Maps From Images	17514
<i>Xindi Wu (Carnegie Mellon University), KwunFung Lau (Argo AI), Francesco Ferroni (Argo AI), Aljoša Ošep (Carnegie Mellon University), and Deva Ramanan (Carnegie Mellon University; Argo AI)</i>	
LoGoNet: Towards Accurate 3D Object Detection With Local-to-Global Cross-Modal Fusion	17524
<i>Xin Li (East China Normal University), Tao Ma (The Chinese University of Hong Kong), Yuenan Hou (Shanghai AI Laboratory), Botian Shi (Shanghai AI Laboratory), Yuchen Yang (Fudan University), Youquan Liu (Hochschule Bremerhaven), Xingjiao Wu (Fudan University), Qin Chen (East China Normal University), Yikang Li (Shanghai AI Laboratory), Yu Qiao (Shanghai AI Laboratory), and Liang He (East China Normal University; Shanghai Key Laboratory of Multidimensional Information Processing)</i>	
Neural Map Prior for Autonomous Driving	17535
<i>Xuan Xiong (Shanghai Qi Zhi Institute), Yicheng Liu (Shanghai Qi Zhi Institute), Tianyuan Yuan (Tsinghua University), Yue Wang (MIT), Yilun Wang (Tsinghua University), and Hang Zhao (Shanghai Qi Zhi Institute; Tsinghua University)</i>	
Spherical Transformer for LiDAR-Based 3D Recognition	17545
<i>Xin Lai (The Chinese University of Hong Kong), Yukang Chen (The Chinese University of Hong Kong), Fanbin Lu (The Chinese University of Hong Kong), Jianhui Liu (The University of Hong Kong), and Jiaya Jia (The Chinese University of Hong Kong; SmartMore)</i>	

Density-Insensitive Unsupervised Domain Adaption on 3D Object Detection	17556
<i>Qianjiang Hu (Wangxuan Institute of Computer Technology, Peking University, China), Daizong Liu (Wangxuan Institute of Computer Technology, Peking University, China), and Wei Hu (Wangxuan Institute of Computer Technology, Peking University, China)</i>	
PillarNeXt: Rethinking Network Designs for 3D Object Detection in LiDAR Point Clouds	17567
<i>Jinyu Li (QCraft), Chenxu Luo (QCraft), and Xiaodong Yang (QCraft)</i>	
PeakConv: Learning Peak Receptive Field for Radar Semantic Segmentation	17577
<i>Liwen Zhang (Intelligent Science & Technology Academy of CASIC, Beijing 100144, China), Xinyan Zhang (Faculty of Computing, Harbin Institute of Technology, Heilongjiang 150001, China), Youcheng Zhang (Intelligent Science & Technology Academy of CASIC, Beijing 100144, China), Yufei Guo (Intelligent Science & Technology Academy of CASIC, Beijing 100144, China), Yuanpei Chen (Intelligent Science & Technology Academy of CASIC, Beijing 100144, China), Xuhui Huang (Intelligent Science & Technology Academy of CASIC, Beijing 100144, China), and Zhe Ma (Intelligent Science & Technology Academy of CASIC, Beijing 100144, China)</i>	
Single Domain Generalization for LiDAR Semantic Segmentation	17587
<i>Hyeonseong Kim (Visual Intelligence Lab., KAIST, Korea), Yoonsu Kang (Visual Intelligence Lab., KAIST, Korea), Changgyoon Oh (Visual Intelligence Lab., KAIST, Korea), and Kuk-Jin Yoon (Visual Intelligence Lab., KAIST, Korea)</i>	
Weakly Supervised Class-Agnostic Motion Prediction for Autonomous Driving	17599
<i>Ruibo Li (Nanyang Technological University), Hanyu Shi (Nanyang Technological University), Ziang Fu (SenseTime Research), Zhe Wang (SenseTime Research), and Guosheng Lin (Nanyang Technological University)</i>	
MethaneMapper: Spectral Absorption Aware Hyperspectral Transformer for Methane Detection ..	17609
<i>Satish Kumar (University of California Santa Barbara), Ivan Arevalo (University of California Santa Barbara), ASM Iftekhar (University of California Santa Barbara), and B S Manjunath (University of California Santa Barbara)</i>	
GrowSP: Unsupervised Semantic Segmentation of 3D Point Clouds	17619
<i>Zihui Zhang (Shenzhen Research Institute, The Hong Kong Polytechnic University; vLAR Group, The Hong Kong Polytechnic University), Bo Yang (Shenzhen Research Institute, The Hong Kong Polytechnic University; vLAR Group, The Hong Kong Polytechnic University), Bing Wang (Shenzhen Research Institute, The Hong Kong Polytechnic University; vLAR Group, The Hong Kong Polytechnic University), and Bo Li (Shenzhen Research Institute, The Hong Kong Polytechnic University; vLAR Group, The Hong Kong Polytechnic University)</i>	

SCoDA: Domain Adaptive Shape Completion for Real Scans	17630
<i>Yushuang Wu (FNii, CUHKSZ; SSE, CUHKSZ; SRIBD), Zizheng Yan (FNii, CUHKSZ; SSE, CUHKSZ; SRIBD), Ce Chen (FNii, CUHKSZ; SSE, CUHKSZ), Lai Wei (SDS, CUHKSZ), Xiao Li (Microsoft Research Asia), Guanbin Li (Sun Yat-sen University, Shenzhen), Yihao Li (FNii, CUHKSZ; SSE, CUHKSZ), Shuguang Cui (SSE, CUHKSZ; FNii, CUHKSZ), and Xiaoguang Han (SSE, CUHKSZ; FNii, CUHKSZ)</i>	
SCPNet: Semantic Scene Completion on Point Cloud	17642
<i>Zhaoyang Xia (Shanghai AI Laboratory), Youquan Liu (Shanghai AI Laboratory), Xin Li (East China Normal University), Xinge Zhu (The Chinese University of Hong Kong), Yuexin Ma (ShanghaiTech University), Yikang Li (Shanghai AI Laboratory), Yuenan Hou (Shanghai AI Laboratory), and Yu Qiao (Shanghai AI Laboratory)</i>	
ViewNet: A Novel Projection-Based Backbone With View Pooling for Few-Shot Point Cloud Classification	17652
<i>Jiajing Chen (Syracuse University, USA), Minmin Yang (Syracuse University, USA), and Senem Velipasalar (Syracuse University, USA)</i>	
Complete-to-Partial 4D Distillation for Self-Supervised Point Cloud Sequence Representation Learning	17661
<i>Zhuoyang Zhang (IIIS, Tsinghua University), Yuhao Dong (IIIS, Tsinghua University), Yunze Liu (IIIS, Tsinghua University; Shanghai Qi Zhi Institute), and Li Yi (IIIS, Tsinghua University; Shanghai Artificial Intelligence Laboratory; Shanghai Qi Zhi Institute)</i>	
Learnable Skeleton-Aware 3D Point Cloud Sampling	17671
<i>Cheng Wen (The University of Sydney, Australia), Baosheng Yu (The University of Sydney, Australia), and Dacheng Tao (The University of Sydney, Australia)</i>	
Meta Architecture for Point Cloud Analysis	17682
<i>Haojia Lin (Xiamen University, P.R. China), Xiawu Zheng (Peng Cheng Laboratory), Lijiang Li (Xiamen University, P.R. China), Fei Chao (Xiamen University, P.R. China), Shanshan Wang (Peng Cheng Laboratory; Chinese Academy of Sciences), Yan Wang (Samsara Inc.), Yonghong Tian (Peng Cheng Laboratory; National Engineering Research Center for Visual Technology, Peking University), and Rongrong Ji (Xiamen University, P.R. China; Peng Cheng Laboratory; Shenzhen Research Institute of Xiamen University)</i>	
PointListNet: Deep Learning on 3D Point Lists	17692
<i>Hehe Fan (Zhejiang University; National University of Singapore), Linchao Zhu (Zhejiang University), Yi Yang (Zhejiang University), and Mohan Kankanhalli (National University of Singapore)</i>	
PEAL: Prior-Embedded Explicit Attention Learning for Low-Overlap Point Cloud Registration	17702
<i>Junle Yu (Hangzhou Dianzi University), Luwei Ren (Hangzhou Dianzi University), Yu Zhang (Shanghai Jiaotong University), Wenhui Zhou (Hangzhou Dianzi University), Lili Lin (Zhejiang Gongshang University), and Guojun Dai (Hangzhou Dianzi University)</i>	

Unsupervised Inference of Signed Distance Functions From Single Sparse Point Clouds Without Learning Priors	17712
<i>Chao Chen (Tsinghua University, Beijing, China), Yu-Shen Liu (Tsinghua University, Beijing, China), and Zhizhong Han (Wayne State University, Detroit, USA)</i>	
Towards Better Gradient Consistency for Neural Signed Distance Functions via Level Set Alignment	17724
<i>Baorui Ma (Tsinghua University, Beijing, China), Junsheng Zhou (Tsinghua University, Beijing, China), Yu-Shen Liu (Tsinghua University, Beijing, China), and Zhizhong Han (Wayne State University, Detroit, USA)</i>	
Self-Supervised Learning for Multimodal Non-Rigid 3D Shape Matching	17735
<i>Dongliang Cao (University of Bonn) and Florian Bernard (University of Bonn)</i>	
3D Registration With Maximal Cliques	17745
<i>Xiyu Zhang (Northwestern Polytechnical University, China), Jiaqi Yang (Northwestern Polytechnical University, China), Shikun Zhang (Northwestern Polytechnical University, China), and Yanning Zhang (Northwestern Polytechnical University, China)</i>	
PanoSwin: A Pano-Style Swin Transformer for Panorama Understanding	17755
<i>Zhixin Ling (Fudan University), Zhen Xing (Fudan University), Xiangdong Zhou (Fudan University), Manliang Cao (Fudan University), and Guichun Zhou (Fudan University)</i>	
DKM: Dense Kernelized Feature Matching for Geometry Estimation	17765
<i>Johan Edstedt (Computer Vision Laboratory Linköping University), Ioannis Athanasiadis (Computer Vision Laboratory Linköping University), Mårten Wadenbäck (Computer Vision Laboratory Linköping University), and Michael Felsberg (Computer Vision Laboratory Linköping University)</i>	
PATS: Patch Area Transportation With Subdivision for Local Feature Matching	17776
<i>Junjie Ni (Zhejiang University; ZJU-SenseTime Joint Lab of 3D Vision), Yijin Li (Zhejiang University; ZJU-SenseTime Joint Lab of 3D Vision), Zhaoyang Huang (The Chinese University of Hong Kong), Hongsheng Li (The Chinese University of Hong Kong), Hujun Bao (Zhejiang University; ZJU-SenseTime Joint Lab of 3D Vision), Zhaopeng Cui (Zhejiang University), and Guofeng Zhang (Zhejiang University; ZJU-SenseTime Joint Lab of 3D Vision)</i>	
Correspondence Transformers With Asymmetric Feature Learning and Matching Flow Super-Resolution	17787
<i>Yixuan Sun (Academy of Engineering & Technology, Fudan University, China), Dongyang Zhao (Fudan University, China), Zhangyue Yin (Fudan University, China), Yiwen Huang (Fudan University, China), Tao Gui (Fudan University, China), Wenqiang Zhang (Academy of Engineering & Technology, Fudan University, China; Fudan University, China), and Weifeng Ge (Fudan University, China)</i>	
Learning Adaptive Dense Event Stereo From the Image Domain	17797
<i>Hoonhee Cho (Visual Intelligence Lab., KAIST, Korea), Jegyeong Cho (Visual Intelligence Lab., KAIST, Korea), and Kuk-Jin Yoon (Visual Intelligence Lab., KAIST, Korea)</i>	

On the Convergence of IRLS and Its Variants in Outlier-Robust Estimation	17808
<i>Liangzu Peng (Johns Hopkins University), Christian Kümmerle (UNC Charlotte), and René Vidal (University of Pennsylvania)</i>	
You Only Segment Once: Towards Real-Time Panoptic Segmentation	17819
<i>Jie Hu (Xiamen University), Linyan Huang (Xiamen University), Tianhe Ren (Xiamen University), Shengchuan Zhang (Xiamen University), Rongrong Ji (Xiamen University), and Liujuan Cao (Xiamen University)</i>	
BEVFormer v2: Adapting Modern Image Backbones to Bird's-Eye-View Recognition via Perspective Supervision	17830
<i>Chenyu Yang (Tsinghua University), Yuntao Chen (Centre for Artificial Intelligence and Robotics, HKISIXAS), Hao Tian (SenseTime Research), Chenxin Tao (Institute of Automation, Chinese Academy of Science (CASIA)), Xizhou Zhu (SenseTime Research), Zhaoxiang Zhang (Centre for Artificial Intelligence and Robotics, HKISIXAS; Institute of Automation, Chinese Academy of Science (CASIA)), Gao Huang (Tsinghua University), Hongyang Li (Shanghai Artificial Intelligence Laboratory), Yu Qiao (Shanghai Artificial Intelligence Laboratory), Lewei Lu (SenseTime Research), Jie Zhou (Tsinghua University), and Jifeng Dai (Tsinghua University; Shanghai Artificial Intelligence Laboratory)</i>	
UniHCP: A Unified Model for Human-Centric Perceptions	17840
<i>Yuanzheng Ci (The University of Sydney), Yizhou Wang (Zhejiang University; Shanghai AI Laboratory), Meilin Chen (Zhejiang University), Shixiang Tang (The University of Sydney), Lei Bai (Shanghai AI Laboratory), Feng Zhu (SenseTime Research), Rui Zhao (SenseTime Research; Shanghai Jiao Tong University, Shanghai, China), Fengwei Yu (SenseTime Research), Donglian Qi (Zhejiang University), and Wanli Ouyang (Shanghai AI Laboratory)</i>	
Planning-Oriented Autonomous Driving	17853
<i>Yihan Hu (OpenDriveLab and OpenGVLab, Shanghai AI Laboratory; Wuhan University), Jiazhi Yang (OpenDriveLab and OpenGVLab, Shanghai AI Laboratory), Li Chen (OpenDriveLab and OpenGVLab, Shanghai AI Laboratory), Keyu Li (OpenDriveLab and OpenGVLab, Shanghai AI Laboratory), Chonghao Sima (OpenDriveLab and OpenGVLab, Shanghai AI Laboratory), Xizhou Zhu (SenseTime Research; OpenDriveLab and OpenGVLab, Shanghai AI Laboratory), Siqi Chai (Wuhan University), Senyao Du (Wuhan University), Tianwei Lin (Wuhan University), Wenhai Wang (OpenDriveLab and OpenGVLab, Shanghai AI Laboratory), Lewei Lu (SenseTime Research), Xiaosong Jia (OpenDriveLab and OpenGVLab, Shanghai AI Laboratory), Qiang Liu (Wuhan University), Jifeng Dai (OpenDriveLab and OpenGVLab, Shanghai AI Laboratory), Yu Qiao (OpenDriveLab and OpenGVLab, Shanghai AI Laboratory), and Hongyang Li (OpenDriveLab and OpenGVLab, Shanghai AI Laboratory)</i>	
Query-Centric Trajectory Prediction	17863
<i>Zikang Zhou (City University of Hong Kong; City University of Hong Kong Shenzhen Research Institute), Jianping Wang (City University of Hong Kong; City University of Hong Kong Shenzhen Research Institute), Yung-Hui Li (Hon Hai Research Institute), and Yu-Kai Huang (Camegie Mellon University)</i>	

Unsupervised Sampling Promoting for Stochastic Human Trajectory Prediction	17874
<i>Guangyi Chen (Mohamed bin Zayed University of Artificial Intelligence; Carnegie Mellon University), Zhenhao Chen (Mohamed bin Zayed University of Artificial Intelligence), Shunxing Fan (Mohamed bin Zayed University of Artificial Intelligence), and Kun Zhang (Mohamed bin Zayed University of Artificial Intelligence; Carnegie Mellon University)</i>	
AdamsFormer for Spatial Action Localization in the Future	17885
<i>Hyung-gun Chi (Honda Research Institute USA; Purdue University), Kwonjoon Lee (Honda Research Institute USA), Nakul Agarwal (Honda Research Institute USA), Yi Xu (Honda Research Institute USA; Northeastern University), Karthik Ramani (Purdue University), and Chiho Choi (Samsung Semiconductor US)</i>	
PIRLNav: Pretraining With Imitation and RL Finetuning for ObjectNav	17896
<i>Ram Ramrakhya (Georgia Institute of Technology), Dhruv Batra (Georgia Institute of Technology; FAIR, Meta AI), Erik Wijmans (Georgia Institute of Technology), and Abhishek Das (FAIR, Meta AI)</i>	
NeRF in the Palm of Your Hand: Corrective Augmentation for Robotics via Novel-View Synthesis	17907
<i>Allan Zhou (Stanford), Moo Jin Kim (Stanford), Lirui Wang (MIT CSAIL), Pete Florence (Google), and Chelsea Finn (Stanford)</i>	
Camouflaged Instance Segmentation via Explicit De-Camouflaging	17918
<i>Naisong Luo (University of Science and Technology of China), Yuwen Pan (University of Science and Technology of China), Rui Sun (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center; Deep Space Exploration Lab), Zhiwei Xiong (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), and Feng Wu (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center)</i>	
Standing Between Past and Future: Spatio-Temporal Modeling for Multi-Camera 3D Multi-Object Tracking	17928
<i>Ziqi Pang (University of Illinois Urbana-Champaign), Jie Li (Toyota Research Institute), Pavel Tokmakov (Toyota Research Institute), Dian Chen (Toyota Research Institute), Sergey Zagoruyko (Woven Planet Level-5), and Yu-Xiong Wang (University of Illinois Urbana-Champaign)</i>	
MotionTrack: Learning Robust Short-Term and Long-Term Motions for Multi-Object Tracking	17939
<i>Zheng Qin (Xi'an Jiaotong University), Sanping Zhou (Xi'an Jiaotong University), Le Wang (Xi'an Jiaotong University), Jinghai Duan (Xi'an Jiaotong University), Gang Hua (Wormpex AI Research), and Wei Tang (University of Illinois at Chicago)</i>	
Multi-Modal Gait Recognition via Effective Spatial-Temporal Feature Fusion	17949
<i>Yufeng Cui (Beihang University, China) and Yimei Kang (Beihang University, China)</i>	

Rethinking the Learning Paradigm for Dynamic Facial Expression Recognition	17958
<i>Hanyang Wang (East China Normal University), Bo Li (Youtu Lab, Tencent), Shuang Wu (Youtu Lab, Tencent), Siyuan Shen (East China Normal University), Feng Liu (East China Normal University), Shouhong Ding (Youtu Lab, Tencent), and Aimin Zhou (East China Normal University)</i>	
One-Shot High-Fidelity Talking-Head Synthesis With Deformable Neural Radiance Field	17969
<i>Weichuang Li (Shanghai AI Laboratory), Longhao Zhang (DAMO Academy, Alibaba Group), Dong Wang (Shanghai AI Laboratory), Bin Zhao (Shanghai AI Laboratory; Northwestern Polytechnical University), Zhigang Wang (Shanghai AI Laboratory), Mulin Chen (Shanghai AI Laboratory; Northwestern Polytechnical University), Bang Zhang (DAMO Academy, Alibaba Group), Zhongjian Wang (DAMO Academy, Alibaba Group), Liefeng Bo (DAMO Academy, Alibaba Group), and Xuelong Li (Shanghai AI Laboratory; Northwestern Polytechnical University)</i>	
Progressive Disentangled Representation Learning for Fine-Grained Controllable Talking Head Synthesis	17979
<i>Duomin Wang (Xiaobing.AI), Yu Deng (Xiaobing.AI), Zixin Yin (Xiaobing.AI), Heung-Yeung Shum (Xiaobing.AI), and Baoyuan Wang (Xiaobing.AI)</i>	
Event-Guided Person Re-Identification via Sparse-Dense Complementary Learning	17990
<i>Chengzhi Cao (University of Science and Technology of China, China), Xueyang Fu (University of Science and Technology of China, China), Hongjian Liu (University of Science and Technology of China, China), Yukun Huang (University of Science and Technology of China, China), Kunyu Wang (University of Science and Technology of China, China), Jiebo Luo (University of Rochester, USA), and Zheng-Jun Zha (University of Science and Technology of China, China)</i>	
Executing Your Commands via Motion Diffusion in Latent Space	18000
<i>Xin Chen (Tencent PCG), Biao Jiang (Fudan University), Wen Liu (Tencent PCG), Zilong Huang (Tencent PCG), Bin Fu (Tencent PCG), Tao Chen (Fudan University), and Gang Yu (Tencent PCG)</i>	
MoLo: Motion-Augmented Long-Short Contrastive Learning for Few-Shot Action Recognition	18011
<i>Xiang Wang (Huazhong University of Science and Technology; Alibaba Group), Shiwei Zhang (Alibaba Group), Zhiwu Qing (Huazhong University of Science and Technology), Changxin Gao (Huazhong University of Science and Technology), Yingya Zhang (Alibaba Group), Deli Zhao (Alibaba Group), and Nong Sang (Huazhong University of Science and Technology)</i>	
“Seeing” Electric Network Frequency From Events	18022
<i>Lexuan Xu (Wuhan University, China), Guang Hua (Institute for Infocomm Research (I2R), A*STAR, Singapore), Haijian Zhang (Wuhan University, China), Lei Yu (Wuhan University, China), and Ning Qiao (SynSense Tech. Co. Ltd., Chengdu, China)</i>	

Event-Based Video Frame Interpolation With Cross-Modal Asymmetric Bidirectional Motion Fields	18032
<i>Taewoo Kim (Korea Advanced Institute of Science and Technology), Yujeong Chae (Korea Advanced Institute of Science and Technology), Hyun-Kurl Jang (Korea Advanced Institute of Science and Technology), and Kuk-Jin Yoon (Korea Advanced Institute of Science and Technology)</i>	
Event-Based Frame Interpolation With Ad-Hoc Deblurring	18043
<i>Lei Sun (Zhejiang University; ETH Zurich), Christos Sakaridis (ETH Zurich), Jingyun Liang (ETH Zurich), Peng Sun (Zhejiang University), Jiezhong Cao (ETH Zurich), Kai Zhang (ETH Zurich), Qi Jiang (Zhejiang University), Kaiwei Wang (Zhejiang University), and Luc Van Gool (ETH Zurich; KU Leuven)</i>	
Video Dehazing via a Multi-Range Temporal Alignment Network With Physical Prior	18053
<i>Jiaqi Xu (The Chinese University of Hong Kong; Shanghai Artificial Intelligence Laboratory), Xiaowei Hu (Shanghai Artificial Intelligence Laboratory), Lei Zhu (The Hong Kong University of Science and Technology (Guangzhou); The Hong Kong University of Science and Technology), Qi Dou (The Chinese University of Hong Kong), Jifeng Dai (Tsinghua University; Shanghai Artificial Intelligence Laboratory), Yu Qiao (Shanghai Artificial Intelligence Laboratory), and Pheng-Ann Heng (The Chinese University of Hong Kong)</i>	
TransFlow: Transformer As Flow Learner	18063
<i>Yawen Lu (Purdue University; Rochester Institute of Technology), Qifan Wang (Meta AI), Siqi Ma (Westlake University), Tong Geng (University of Rochester), Yingjie Victor Chen (Purdue University), Huaijin Chen (Vayu Robotics), and Dongfang Liu (Rochester Institute of Technology)</i>	
MP-Former: Mask-Piloted Transformer for Image Segmentation	18074
<i>Hao Zhang (The Hong Kong University of Science and Technology; International Digital Economy Academy (IDEA)), Feng Li (International Digital Economy Academy (IDEA)), Huaizhe Xu (The Hong Kong University of Science and Technology; International Digital Economy Academy (IDEA)), Shijia Huang (The Chinese University of Hong Kong; International Digital Economy Academy (IDEA)), Shilong Liu (Tsinghua University; International Digital Economy Academy (IDEA)), Lionel M. Ni (The Hong Kong University of Science and Technology, (Guangzhou)), and Lei Zhang (International Digital Economy Academy (IDEA))</i>	
GradICON: Approximate Diffeomorphisms via Gradient Inverse Consistency	18084
<i>Lin Tian (UNC Chapel Hill), Hastings Greer (UNC Chapel Hill), François-Xavier Vialard (LIGM, Université Gustave Eiffel; MOKAPLAN, INRIA Paris), Roland Kwitt (University of Salzburg), Raúl San José Estépar (Harvard Medical School), Richard Jarrett Rushmore (Boston University), Nikolaos Makris (Harvard Medical School), Sylvain Bouix (ÉTS Montréal), and Marc Niethammer (UNC Chapel Hill)</i>	
Neural Texture Synthesis With Guided Correspondence	18095
<i>Yang Zhou (Shenzhen University), Kaijian Chen (Shenzhen University), Rongjun Xiao (Shenzhen University), and Hui Huang (Shenzhen University)</i>	

Self-Supervised Non-Uniform Kernel Estimation With Flow-Based Motion Prior for Blind Image Deblurring	18105
<i>Zhenxuan Fang (Xidian University), Fangfang Wu (Xidian University), Weisheng Dong (Xidian University), Xin Li (West Virginia University), Jinjian Wu (Xidian University), and Guangming Shi (Xidian University)</i>	
Decoupling-and-Aggregating for Image Exposure Correction	18115
<i>Yang Wang (University of Science and Technology of China, China), Long Peng (University of Science and Technology of China, China), Liang Li (Key Lab of Intell. Info. Process., Inst. of Comput. Tech., CAS, China), Yang Cao (University of Science and Technology of China, China), and Zheng-Jun Zha (University of Science and Technology of China, China)</i>	
You Do Not Need Additional Priors or Regularizers in Retinex-Based Low-Light Image Enhancement	18125
<i>Huiyuan Fu (Beijing University of Posts and Telecommunications), Wenkai Zheng (Beijing University of Posts and Telecommunications), Xiangyu Meng (Beijing University of Posts and Telecommunications), Xin Wang (Stony Brook University), Chuanming Wang (Beijing University of Posts and Telecommunications), and Huadong Ma (Beijing University of Posts and Telecommunications)</i>	
DNF: Decouple and Feedback Network for Seeing in the Dark	18135
<i>Xin Jin (Nankai University), Ling-Hao Han (Nankai University), Zhen Li (Nankai University), Chun-Le Guo (Nankai University), Zhi Chai (Hisilicon Technologies Co. Ltd.), and Chongyi Li (Nanyang Technological University)</i>	
Contrastive Semi-Supervised Learning for Underwater Image Restoration via Reliable Bank	18145
<i>Shirui Huang (Xidian University), Keyan Wang (Xidian University), Huan Liu (McMaster University), Jun Chen (McMaster University), and Yunsong Li (Xidian University)</i>	
LG-BPN: Local and Global Blind-Patch Network for Self-Supervised Real-World Denoising ...	18156
<i>Zichun Wang (Beijing Institute of Technology), Ying Fu (Beijing Institute of Technology), Ji Liu (Baidu Inc., China), and Yulun Zhang (ETH Zurich)</i>	
Spectral Bayesian Uncertainty for Image Super-Resolution	18166
<i>Tao Liu (Huazhong University of Science and Technology, China), Jun Cheng (Huazhong University of Science and Technology, China), and Shan Tan (Huazhong University of Science and Technology, China)</i>	
Deep Random Projector: Accelerated Deep Image Prior	18176
<i>Taihui Li (University of Minnesota, USA), Hengkang Wang (University of Minnesota, USA), Zhong Zhuang (University of Minnesota, USA), and Ju Sun (University of Minnesota, USA)</i>	
Context-Aware Pretraining for Efficient Blind Image Decomposition	18186
<i>Chao Wang (Zhejiang University), Zhedong Zheng (National University of Singapore), Ruijie Quan (Zhejiang University), Yifan Sun (Baidu Inc.), and Yi Yang (Zhejiang University)</i>	

Metadata-Based RAW Reconstruction via Implicit Neural Functions	18196
<i>Leyi Li (Zhejiang University), Huijie Qiao (Chinese Academy of Sciences), Qi Ye (Key Lab of CS&AUS of Zhejiang Province), and Qinmin Yang (Zhejiang University)</i>	
Raw Image Reconstruction With Learned Compact Metadata	18206
<i>Yufei Wang (Nanyang Technological University), Yi Yu (Nanyang Technological University), Wenhan Yang (Peng Cheng Laboratory), Lanqing Guo (Nanyang Technological University), Lap-Pui Chau (The Hong Kong Polytechnic University), Alex C. Kot (Nanyang Technological University), and Bihan Wen (Nanyang Technological University)</i>	
AccellR: Task-Aware Image Compression for Accelerating Neural Restoration	18216
<i>Juncheol Ye (Korea Advanced Institute of Science and Technology (KAIST)), Hyunho Yeo (Korea Advanced Institute of Science and Technology (KAIST)), Jinwoo Park (Korea Advanced Institute of Science and Technology (KAIST)), and Dongsu Han (Korea Advanced Institute of Science and Technology (KAIST))</i>	
AutoFocusFormer: Image Segmentation off the Grid	18227
<i>Chen Ziwen (Oregon State University), Kaushik Patnaik (Apple Inc.), Shuangfei Zhai (Apple Inc.), Alvin Wan (Apple Inc.), Zhile Ren (Apple Inc.), Alexander G. Schwing (Apple Inc.), Alex Colburn (Apple Inc.), and Li Fuxin (Oregon State University; Apple Inc.)</i>	
Guided Depth Super-Resolution by Deep Anisotropic Diffusion	18237
<i>Nando Metzger (ETH Zurich), Rodrigo Caye Daudt (ETH Zurich), and Konrad Schindler (ETH Zurich)</i>	
Super-Resolution Neural Operator	18247
<i>Min Wei (Beijing University of Posts and Telecommunications, China) and Xuesong Zhang (Beijing University of Posts and Telecommunications, China)</i>	
Cascaded Local Implicit Transformer for Arbitrary-Scale Super-Resolution	18257
<i>Hao-Wei Chen (National Tsing Hua University; MediaTek Inc.), Yu-Syuan Xu (MediaTek Inc.), Min-Fong Hong (MediaTek Inc.), Yi-Min Tsai (MediaTek Inc.), Hsien-Kai Kuo (MediaTek Inc.), and Chun-Yi Lee (National Tsing Hua University)</i>	
GamutMLP: A Lightweight MLP for Color Loss Recovery	18268
<i>Hoang M. Le (York University), Brian Price (Adobe Research), Scott Cohen (Adobe Research), and Michael S. Brown (York University)</i>	
Efficient and Explicit Modelling of Image Hierarchies for Image Restoration	18278
<i>Yawei Li (ETH Zurich), Yuchen Fan (Meta Reality Labs), Xiaoyu Xiang (Meta Reality Labs), Denis Demandolx (Meta Reality Labs), Rakesh Ranjan (Meta Reality Labs), Radu Timofte (ETH Zurich; University of Wurzburg), and Luc Van Gool (ETH Zurich; KU Leuven)</i>	
LEMART: Label-Efficient Masked Region Transform for Image Harmonization	18290
<i>Sheng Liu (Amazon Prime Video), Cong Phuoc Huynh (Amazon Prime Video), Cong Chen (Amazon Prime Video), Maxim Arap (Amazon Prime Video), and Raffay Hamid (Amazon Prime Video)</i>	
CAP-VSTNet: Content Affinity Preserved Versatile Style Transfer	18300
<i>Linfeng Wen (Sun Yat-Sen University), Chengying Gao (Sun Yat-Sen University), and Changqing Zou (Zhejiang University; Zhejiang Lab)</i>	

ObjectStitch: Object Compositing With Diffusion Model	18310
<i>Yizhi Song (Purdue University), Zhifei Zhang (Adobe Research), Zhe Lin (Adobe Research), Scott Cohen (Adobe Research), Brian Price (Adobe Research), Jianming Zhang (Adobe Research), Soo Ye Kim (Adobe Research), and Daniel Aliaga (Purdue University)</i>	
DeepVecFont-v2: Exploiting Transformers To Synthesize Vector Fonts With Higher Quality ..	18320
<i>Yuqing Wang (Peking University, China), Yizhi Wang (Peking University, China), Longhui Yu (Peking University, China), Yuesheng Zhu (Peking University, China), and Zhouhui Lian (Peking University, China)</i>	
Master: Meta Style Transformer for Controllable Zero-Shot and Few-Shot Artistic Style Transfer	18329
<i>Hao Tang (Center for Data Science, Peking University), Songhua Liu (National University of Singapore), Tianwei Lin (VIS, Baidu Inc.), Shaoli Huang (Tencent AI Lab), Fu Li (VIS, Baidu Inc.), Dongliang He (VIS, Baidu Inc.), and Xinchao Wang (National University of Singapore)</i>	
CLIP-Sculptor: Zero-Shot Generation of High-Fidelity and Diverse Shapes From Natural Language	18339
<i>Aditya Sanghi (Autodesk Research), Rao Fu (Brown University), Vivian Liu (Columbia University), Karl D.D. Willis (Autodesk Research), Hooman Shayani (Autodesk Research), Amir H. Khasahmadi (Autodesk Research), Srinath Sridhar (Brown University), and Daniel Ritchie (Brown University)</i>	
LayoutDM: Transformer-Based Diffusion Model for Layout Generation	18349
<i>Shang Chai (University of Science and Technology of China), Liansheng Zhuang (University of Science and Technology of China), and Fengying Yan (Tianjin University)</i>	
Imagen Editor and EditBench: Advancing and Evaluating Text-Guided Image Inpainting	18359
<i>Su Wang (Google Research), Chitwan Saharia (Google Research), Ceslee Montgomery (Google Research), Jordi Pont-Tuset (Google Research), Shai Noy (Google Research), Stefano Pellegrini (Google Research), Yasumasa Onoe (Google Research), Sarah Laszlo (Google Research), David J. Fleet (Google Research), Radu Soricut (Google Research), Jason Baldridge (Google Research), Mohammad Norouzi (Google Research), Peter Anderson (Google Research), and William Chan (Google Research)</i>	
SpaText: Spatio-Textual Representation for Controllable Image Generation	18370
<i>Omri Avrahami (Meta AI; he Hebrew University of Jerusalem), Thomas Hayes (Meta AI), Oran Gafni (Meta AI), Sonal Gupta (Meta AI), Yaniv Taigman (Meta AI), Devi Parikh (Meta AI), Dani Lischinski (The Hebrew University of Jerusalem), Ohad Fried (Reichman University), and Xi Yin (Meta AI)</i>	
Paint by Example: Exemplar-Based Image Editing With Diffusion Models	18381
<i>Binxin Yang (University of Science and Technology of China), Shuyang Gu (Microsoft Research Asia), Bo Zhang (Microsoft Research Asia), Ting Zhang (Microsoft Research Asia), Xuejin Chen (University of Science and Technology of China), Xiaoyan Sun (University of Science and Technology of China), Dong Chen (Microsoft Research Asia), and Fang Wen (Microsoft Research Asia)</i>	

InstructPix2Pix: Learning To Follow Image Editing Instructions	18392
<i>Tim Brooks (University of California, Berkeley), Aleksander Holynski (University of California, Berkeley), and Alexei A. Efros (University of California, Berkeley)</i>	
LayoutFormer++: Conditional Graphic Layout Generation via Constraint Serialization and Decoding Space Restriction	18403
<i>Zhaoyun Jiang (Xi'an Jiaotong University), Jiaqi Guo (Microsoft Research Asia), Shizhao Sun (Microsoft Research Asia), Huayu Deng (Shanghai Jiaotong University), Zhongkai Wu (Beihang University), Vuksan Mijovic (Microsoft), Zijiang James Yang (Xi'an Jiaotong University), Jian-Guang Lou (Microsoft Research Asia), and Dongmei Zhang (Microsoft Research Asia)</i>	
Self-Guided Diffusion Models	18413
<i>Vincent Tao Hu (University of Amsterdam), David W. Zhang (University of Amsterdam), Yuki M. Asano (University of Amsterdam), Gertjan J. Burghouts (TNO), and Cees G. M. Snoek (University of Amsterdam)</i>	
HOLODIFFUSION: Training a 3D Diffusion Model Using 2D Images	18423
<i>Animesh Karnewar (UCL), Andrea Vedaldi (Meta AI), David Novotny (Meta AI), and Niloy J. Mitra (UCL)</i>	
Class-Balancing Diffusion Models	18434
<i>Yiming Qin (Shanghai Jiao-Tong University), Huangjie Zheng (University of Texas, Austin), Jiangchao Yao (Shanghai Jiao-Tong University; Shanghai AI Laboratory), Mingyuan Zhou (University of Texas, Austin), and Ya Zhang (Shanghai Jiao-Tong University; Shanghai AI Laboratory)</i>	
Conditional Image-to-Video Generation With Latent Flow Diffusion Models	18444
<i>Haomiao Ni (The Pennsylvania State University, USA), Changhao Shi (University of California, USA), Kai Li (NEC Laboratories America, USA), Sharon X. Huang (The Pennsylvania State University, USA), and Martin Renqiang Min (NEC Laboratories America, USA)</i>	
Video Probabilistic Diffusion Models in Projected Latent Space	18456
<i>Sihyun Yu (KAIST), Kihyuk Sohn (Google Research), Subin Kim (KAIST), and Jinwoo Shin (KAIST)</i>	
Regularized Vector Quantization for Tokenized Image Synthesis	18467
<i>Jiahui Zhang (Nanyang Technological University), Fangneng Zhan (Max Planck Institute for Informatics), Christian Theobalt (Max Planck Institute for Informatics), and Shijian Lu (Nanyang Technological University)</i>	
EfficientSCI: Densely Connected Network With Space-Time Factorization for Large-Scale Video Snapshot Compressive Imaging	18477
<i>Lishun Wang (Chengdu Institute of Computer Application Chinese Academy of Sciences), Miao Cao (University of Chinese Academy of Sciences), and Xin Yuan (Westlake University; Zhejiang University)</i>	
MMVC: Learned Multi-Mode Video Compression With Block-Based Prediction Mode Selection and Density-Adaptive Entropy Coding	18487
<i>Bowen Liu (University of Michigan, Ann Arbor), Yu Chen (University of Michigan, Ann Arbor), Rakesh Chowdary Machineni (University of Michigan, Ann Arbor), Shiyu Liu (University of Michigan, Ann Arbor), and Hun-Seok Kim (University of Michigan, Ann Arbor)</i>	

Video Compression With Entropy-Constrained Neural Representations	18497
<i>Carlos Gomes (ETH Zürich, Switzerland), Roberto Azevedo (DisneyResearch\Studios, Switzerland), and Christopher Schroers (DisneyResearch\Studios, Switzerland)</i>	
WIRE: Wavelet Implicit Neural Representations	18507
<i>Vishwanath Saragadam (Rice University), Daniel LeJeune (Rice University), Jasper Tan (Rice University), Guha Balakrishnan (Rice University), Ashok Veeraraghavan (Rice University), and Richard G. Baraniuk (Rice University)</i>	
TINC: Tree-Structured Implicit Neural Compression	18517
<i>Runzhao Yang (Tsinghua University, China)</i>	
CompletionFormer: Depth Completion With Convolutions and Vision Transformers	18527
<i>Youmin Zhang (University of Bologna), Xianda Guo (PhiGent Robotics), Matteo Poggi (University of Bologna), Zheng Zhu (PhiGent Robotics), Guan Huang (PhiGent Robotics), and Stefano Mattocchia (University of Bologna)</i>	
Lite-Mono: A Lightweight CNN and Transformer Architecture for Self-Supervised Monocular Depth Estimation	18537
<i>Ning Zhang (University of Twente), Francesco Nex (University of Twente), George Vosselman (University of Twente), and Norman Kerle (University of Twente)</i>	
Global Vision Transformer Pruning With Hessian-Aware Saliency	18547
<i>Huanrui Yang (NVIDIA; University of California, Berkeley), Hongxu Yin (NVIDIA), Maying Shen (NVIDIA), Pavlo Molchanov (NVIDIA), Hai Li (Duke University), and Jan Kautz (NVIDIA)</i>	
Lite DETR: An Interleaved Multi-Scale Encoder for Efficient DETR	18558
<i>Feng Li (The Hong Kong University of Science and Technology.), Ailing Zeng (International Digital Economy Academy (IDEA).), Shilong Liu (International Digital Economy Academy (IDEA).; BNRist Center, Institute for AI, Tsinghua University.), Hao Zhang (The Hong Kong University of Science and Technology.; International Digital Economy Academy (IDEA).), Hongyang Li (International Digital Economy Academy (IDEA); South China University of Science and Technology.), Lei Zhang (The Hong Kong University of Science and Technology.; International Digital Economy Academy (IDEA).), and Lionel M. Ni (International Digital Economy Academy (IDEA).)</i>	
PaCa-ViT: Learning Patch-to-Cluster Attention in Vision Transformers	18568
<i>Ryan Grainger (Department of ECE, NC State), Thomas Paniagua (Department of ECE, NC State), Xi Song (An Independent Researcher), Naresh Cuntoor (BlueHalo), Mun Wai Lee (BlueHalo), and Tianfu Wu (Department of ECE, NC State)</i>	
Visual Atoms: Pre-Training Vision Transformers With Sinusoidal Waves	18579
<i>Sora Takashima (National Institute of Advanced Industrial Science and Technology (AIST); Tokyo Institute of Technology), Ryo Hayamizu (National Institute of Advanced Industrial Science and Technology (AIST)), Nakamasa Inoue (National Institute of Advanced Industrial Science and Technology (AIST); Tokyo Institute of Technology), Hirokatsu Kataoka (National Institute of Advanced Industrial Science and Technology (AIST)), and Rio Yokota (Tokyo Institute of Technology)</i>	

Neuron Structure Modeling for Generalizable Remote Physiological Measurement	18589
<i>Hao Lu (The Hong Kong University of Science & Technology, Guangzhou), Zitong Yu (Great Bay University), Xuesong Niu (Kuaishou Technology), and Ying-Cong Chen (The Hong Kong University of Science & Technology, Guangzhou)</i>	
Explaining Image Classifiers With Multiscale Directional Image Representation	18600
<i>Stefan Kolek (Ludwig-Maximilians-Universität Munchen), Robert Windesheim (Ludwig-Maximilians-Universität Munchen), Hector Andrade-Loarca (Ludwig-Maximilians-Universität Munchen), Gitta Kutyniok (Ludwig-Maximilians-Universität Munchen; University of Tromsø), and Ron Levie (Technion-Israel Institute of Technology)</i>	
Integrally Pre-Trained Transformer Pyramid Networks	18610
<i>Yunjie Tian (UCAS), Lingxi Xie (Huawei Inc.), Zhaozhi Wang (UCAS), Longhui Wei (Huawei Inc), Xiaopeng Zhang (Huawei Inc), Jianbin Jiao (UCAS), Yaowei Wang (Pengcheng Lab.), Qi Tian (Huawei Inc.), and Qixiang Ye (UCAS; Pengcheng Lab.)</i>	
PartMix: Regularization Strategy To Learn Part Discovery for Visible-Infrared Person Re-Identification	18621
<i>Minsu Kim (Yonsei University), Seungryong Kim (Korea University), Jungin Park (Yonsei University), Seongheon Park (Yonsei University), and Kwanghoon Sohn (Yonsei University; Korea Institute of Science and Technology (KIST))</i>	
Knowledge Distillation for 6D Pose Estimation by Aligning Distributions of Local Predictions	18633
<i>Shuxuan Guo (CVLab, EPFL), Yinlin Hu (MagicLeap), Jose M. Alvarez (NVIDIA), and Mathieu Salzmann (CVLab, EPFL; ClearSpace)</i>	
Focused and Collaborative Feedback Integration for Interactive Image Segmentation	18643
<i>Qiaoqiao Wei (Tsinghua University, China), Hui Zhang (Tsinghua University, China), and Jun-Hai Yong (Tsinghua University, China)</i>	
PolyFormer: Referring Image Segmentation As Sequential Polygon Generation	18653
<i>Jiang Liu (Johns Hopkins University), Hui Ding (AWS AI Labs), Zhaowei Cai (AWS AI Labs), Yuting Zhang (AWS AI Labs), Ravi Kumar Satzoda (AWS AI Labs), Vijay Mahadevan (AWS AI Labs), and R. Manmatha (AWS AI Labs)</i>	
Devil's on the Edges: Selective Quad Attention for Scene Graph Generation	18664
<i>Deunsol Jung (Pohang University of Science and Technology (POSTECH), South Korea), Sanghyun Kim (Pohang University of Science and Technology (POSTECH), South Korea), Won Hwa Kim (Pohang University of Science and Technology (POSTECH), South Korea), and Minsu Cho (Pohang University of Science and Technology (POSTECH), South Korea)</i>	

Panoptic Video Scene Graph Generation	18675
<i>Jingkang Yang (Nanyang Technological University, Singapore), Wenxuan Peng (Nanyang Technological University, Singapore), Xiangtai Li (Nanyang Technological University, Singapore), Zujin Guo (Nanyang Technological University, Singapore), Liangyu Chen (Nanyang Technological University, Singapore), Bo Li (Nanyang Technological University, Singapore), Zheng Ma (SenseTime Research, China), Kaiyang Zhou (Nanyang Technological University, Singapore), Wayne Zhang (SenseTime Research, China), Chen Change Loy (Nanyang Technological University, Singapore), and Ziwei Liu (Nanyang Technological University, Singapore)</i>	
Generalized Relation Modeling for Transformer Tracking	18686
<i>Shenyuan Gao (The Hong Kong University of Science and Technology), Chunluan Zhou (Wormpex AI Research), and Jun Zhang (The Hong Kong University of Science and Technology)</i>	
Representation Learning for Visual Object Tracking by Masked Appearance Transfer	18696
<i>Haojie Zhao (Dalian University of Technology, China), Dong Wang (Dalian University of Technology, China), and Huchuan Lu (Dalian University of Technology, China; Peng Cheng Laboratory, China)</i>	
Unified Mask Embedding and Correspondence Learning for Self-Supervised Video Segmentation .	18706
<i>Liulei Li (Zhejiang University; Baidu VIS), Wenguan Wang (Zhejiang University), Tianfei Zhou (ETH Zurich), Jianwu Li (Beijing Institute of Technology), and Yi Yang (Zhejiang University)</i>	
EVAL: Explainable Video Anomaly Localization	18717
<i>Ashish Singh (CICS, University of Massachusetts Amherst; Mitsubishi Electric Research Labs), Michael J. Jones (Mitsubishi Electric Research Labs), and Erik G. Learned-Miller (CICS, University of Massachusetts Amherst; Mitsubishi Electric Research Labs)</i>	
MOSO: Decomposing MOtion, Scene and Object for Video Prediction	18727
<i>Mingzhen Sun (The Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS)), Weining Wang (The Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences (CASIA)), Xinxin Zhu (The Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences (CASIA)), and Jing Liu (The Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS))</i>	
TarViS: A Unified Approach for Target-Based Video Segmentation	18738
<i>Ali Athar (RWTH Aachen University, Germany), Alexander Hermans (RWTH Aachen University, Germany), Jonathon Luiten (RWTH Aachen University, Germany; Carnegie Mellon University, USA), Deva Ramanan (Carnegie Mellon University, USA), and Bastian Leibe (RWTH Aachen University, Germany)</i>	

Efficient Movie Scene Detection Using State-Space Transformers	18749
<i>Md Mohaiminul Islam (UNC Chapel Hill), Mahmudul Hasan (Comcast Labs), Kishan Shamsundar Athrey (Comcast Labs), Tony Braskich (Comcast Labs), and Gedas Bertasius (UNC Chapel Hill)</i>	
Latency Matters: Real-Time Action Forecasting Transformer	18759
<i>Harshayu Girase (Honda Research Institute, USA; UC Berkeley), Nakul Agarwal (Honda Research Institute, USA), Chiho Choi (Honda Research Institute, USA), and Karttikeya Mangalam (UC Berkeley)</i>	
Temporal Attention Unit: Towards Efficient Spatiotemporal Predictive Learning	18770
<i>Cheng Tan (Westlake University), Zhangyang Gao (Westlake University), Lirong Wu (Westlake University), Yongjie Xu (Westlake University), Jun Xia (Westlake University), Siyuan Li (Westlake University), and Stan Z. Li (Westlake University)</i>	
Watch or Listen: Robust Audio-Visual Speech Recognition With Visual Corruption Modeling and Reliability Scoring	18783
<i>Joanna Hong (Image and Video Systems Lab, KAIST), Minsu Kim (Image and Video Systems Lab, KAIST), Jeongsoo Choi (Image and Video Systems Lab, KAIST), and Yong Man Ro (Image and Video Systems Lab, KAIST)</i>	
ReVISE: Self-Supervised Speech Resynthesis With Visual Input for Universal and Generalized Speech Regeneration	18795
<i>Wei-Ning Hsu (FAIR, Meta AI Research), Tal Remez (FAIR, Meta AI Research), Bowen Shi (FAIR, Meta AI Research; Toyota Technological Institute at Chicago), Jacob Donley (Meta Reality Labs Research), and Yossi Adi (FAIR, Meta AI Research; The Hebrew University of Jerusalem)</i>	
SynthVSR: Scaling Up Visual Speech Recognition With Synthetic Supervision	18806
<i>Xubo Liu (University of Surrey), Egor Lakomkin (Meta AI), Konstantinos Vougioukas (Meta AI), Pingchuan Ma (Meta AI), Honglie Chen (Meta AI), Ruiming Xie (Meta AI), Morrie Doulaty (Meta AI), Niko Moritz (Meta AI), Jachym Kolar (Meta AI), Stavros Petridis (Meta AI), Maja Pantic (Meta AI), and Christian Fuegen (Meta AI)</i>	
SVFormer: Semi-Supervised Video Transformer for Action Recognition	18816
<i>Zhen Xing (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Qi Dai (Microsoft Research Asia), Han Hu (Microsoft Research Asia), Jingjing Chen (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Zuxuan Wu (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), and Yu-Gang Jiang (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing)</i>	
Collecting Cross-Modal Presence-Absence Evidence for Weakly-Supervised Audio-Visual Event Perception	18827
<i>Junyu Gao (Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS)), Mengyuan Chen (Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS)), and Changsheng Xu (Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences (UCAS); Peng Cheng Laboratory, China)</i>	

Post-Processing Temporal Action Detection	18837
<i>Sauradip Nag (CVSSP, University of Surrey, UK; iFlyTek-Surrey Joint Research Center on Artificial Intelligence, UK), Xiatian Zhu (CVSSP, University of Surrey, UK; Surrey Institute for People-Centred Artificial Intelligence, UK), Yi-Zhe Song (CVSSP, University of Surrey, UK; iFlyTek-Surrey Joint Research Center on Artificial Intelligence, UK), and Tao Xiang (CVSSP, University of Surrey, UK; iFlyTek-Surrey Joint Research Center on Artificial Intelligence, UK)</i>	
HaLP: Hallucinating Latent Positives for Skeleton-Based Self-Supervised Learning of Actions	18846
<i>Anshul Shah (Johns Hopkins University), Aniket Roy (Johns Hopkins University), Ketul Shah (Johns Hopkins University), Shlok Mishra (University of Maryland, College Park), David Jacobs (University of Maryland, College Park; Meta), Anoop Cherian (MERL), and Rama Chellappa (Johns Hopkins University)</i>	
TriDet: Temporal Action Detection With Relative Boundary Modeling	18857
<i>Dingfeng Shi (VRLab, Beihang University, China), Yujie Zhong (Meituan Inc.), Qiong Cao (JD Explore Academy), Lin Ma (Meituan Inc.), Jia Li (VRLab, Beihang University, China), and Dacheng Tao (JD Explore Academy)</i>	
Hybrid Active Learning via Deep Clustering for Video Action Detection	18867
<i>Aayush J. Rana (Center for Research in Computer Vision (CRCV), University of Central Florida) and Yogesh S. Rawat (Center for Research in Computer Vision (CRCV), University of Central Florida)</i>	
Two-Stream Networks for Weakly-Supervised Temporal Action Localization With Semantic-Aware Mechanisms	18878
<i>Yu Wang (Ant Group, China), Yadong Li (Ant Group, China), and Hongbin Wang (Ant Group, China)</i>	
Weakly Supervised Video Emotion Detection and Prediction via Cross-Modal Temporal Erasing Network	18888
<i>Zhicheng Zhang (Nankai University, China), Lijuan Wang (Nankai University, China), and Jufeng Yang (Nankai University, China)</i>	
Collaborative Noisy Label Cleaner: Learning Scene-Aware Trailers for Multi-Modal Highlight Detection in Movies	18898
<i>Bei Gan (Tencent YouTu Lab), Xiujun Shu (Tencent YouTu Lab), Ruizhi Qiao (Tencent YouTu Lab), Haoqian Wu (Tencent YouTu Lab), Keyu Chen (Tencent YouTu Lab), Hanjun Li (Tencent YouTu Lab), and Bo Ren (Tencent YouTu Lab)</i>	
Weakly Supervised Temporal Sentence Grounding With Uncertainty-Guided Self-Training ..	18908
<i>Yifei Huang (The University of Tokyo; Shanghai Artificial Intelligence Laboratory), Lijin Yang (The University of Tokyo), and Yoichi Sato (The University of Tokyo)</i>	
SViTT: Temporal Learning of Sparse Video-Text Transformers	18919
<i>Yi Li (UC San Diego), Kyle Min (Intel Labs), Subarna Tripathi (Intel Labs), and Nuno Vasconcelos (UC San Diego)</i>	

AutoAD: Movie Description in Context	18930
<i>Tengda Han (Visual Geometry Group, University of Oxford), Max Bain (Visual Geometry Group, University of Oxford), Arsha Nagrani (Visual Geometry Group, University of Oxford), Gül Varol (Visual Geometry Group, University of Oxford; Univ Gustave Eiffel, CNRS), Weidi Xie (Visual Geometry Group, University of Oxford; CMIC, Shanghai Jiao Tong University), and Andrew Zisserman (Visual Geometry Group, University of Oxford)</i>	
Text With Knowledge Graph Augmented Transformer for Video Captioning	18941
<i>Xin Gu (University of Chinese Academy of Sciences, China; Institute of Software Chinese Academy of Sciences, China), Guang Chen (ByteDance Inc., USA), Yufei Wang (ByteDance Inc., USA), Libo Zhang (Institute of Software Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Tiejian Luo (University of Chinese Academy of Sciences, China), and Longyin Wen (ByteDance Inc., USA)</i>	
StepFormer: Self-Supervised Step Discovery and Localization in Instructional Videos	18952
<i>Nikita Dvornik (Samsung AI Centre Toronto), Isma Hadji (Samsung AI Centre Toronto), Ran Zhang (Samsung AI Centre Toronto), Konstantinos G. Derpanis (Samsung AI Centre Toronto; York University), Richard P. Wildes (Samsung AI Centre Toronto; York University), and Allan D. Jepson (Samsung AI Centre Toronto)</i>	
Dual Alignment Unsupervised Domain Adaptation for Video-Text Retrieval	18962
<i>Xiaoshuai Hao (Chinese Academy of Sciences; University of Chinese Academy of Sciences), Wanqian Zhang (Chinese Academy of Sciences), Dayan Wu (Chinese Academy of Sciences), Fei Zhu (Chinese Academy of Sciences; University of Chinese Academy of Sciences), and Bo Li (Chinese Academy of Sciences; University of Chinese Academy of Sciences)</i>	
Hierarchical Semantic Correspondence Networks for Video Paragraph Grounding	18973
<i>Chaolei Tan (Sun Yat-sen University, China), Zihang Lin (Sun Yat-sen University, China), Jian-Fang Hu (Sun Yat-sen University, China; Guangdong Province Key Laboratory of Information Security Technology, China; Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, China), Wei-Shi Zheng (Sun Yat-sen University, China; Guangdong Province Key Laboratory of Information Security Technology, China; Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, China), and Jianhuang Lai (Sun Yat-sen University, China; Guangdong Province Key Laboratory of Information Security Technology, China; Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, China)</i>	
CLIPPING: Distilling CLIP-Based Models With a Student Base for Video-Language Retrieval ..	18983
<i>Renjing Pei (Huawei Noah's Ark Lab), Jianzhuang Liu (Huawei Noah's Ark Lab), Weimian Li (Huawei Noah's Ark Lab), Bin Shao (Huawei Noah's Ark Lab), Songcen Xu (Huawei Noah's Ark Lab), Peng Dai (Huawei Noah's Ark Lab), Juwei Lu (Huawei Noah's Ark Lab), and Youliang Yan (Huawei Noah's Ark Lab)</i>	

Learning Emotion Representations From Verbal and Nonverbal Communication	18993
<i>Sitao Zhang (The Pennsylvania State University, University Park, Pennsylvania, USA), Yimu Pan (The Pennsylvania State University, University Park, Pennsylvania, USA), and James Z. Wang (The Pennsylvania State University, University Park, Pennsylvania, USA)</i>	
Context De-Confounded Emotion Recognition	19005
<i>Dingkang Yang (Fudan University; Institute of Meta-Medical, IPASS), Zhaoyu Chen (Fudan University), Yuzheng Wang (Fudan University), Shunli Wang (Fudan University), Mingcheng Li (Fudan University), Siao Liu (Fudan University), Xiao Zhao (Fudan University), Shuai Huang (Fudan University), Zhiyan Dong (Fudan University), Peng Zhai (Fudan University), and Lihua Zhang (Fudan University; Institute of Meta-Medical, IPASS; Jilin Provincial Key Laboratory of Intelligence Science and Engineering, China; Engineering Research Center of AI and Robotics, Ministry of Education, China)</i>	
CiCo: Domain-Aware Sign Language Retrieval via Cross-Lingual Contrastive Learning	19016
<i>Yiting Cheng (Fudan University), Fangyun Wei (Microsoft Research Asia), Jianmin Bao (Microsoft Research Asia), Dong Chen (Microsoft Research Asia), and Wenqiang Zhang (Fudan University)</i>	
Discovering the Real Association: Multimodal Causal Reasoning in Video Question Answering	19027
<i>Chuanqi Zang (Beijing Institute of Technology), Hanqing Wang (Beijing Institute of Technology), Mingtao Pei (Beijing Institute of Technology), and Wei Liang (Beijing Institute of Technology; Yangtze Delta Region Academy of Beijing Institute of Technology, Jiaxing)</i>	
LEGO-Net: Learning Regular Rearrangements of Objects in Rooms	19037
<i>Qihong Anna Wei (Brown University), Sijie Ding (Brown University), Jeong Joon Park (Stanford University), Rahul Sajnani (Brown University), Adrien Poulenard (Stanford University), Srinath Sridhar (Brown University), and Leonidas Guibas (Stanford University)</i>	
LANA: A Language-Capable Navigator for Instruction Following and Generation	19048
<i>Xiaohan Wang (CCAI, Zhejiang University), Wenguan Wang (CCAI, Zhejiang University), Jiayi Shao (CCAI, Zhejiang University), and Yi Yang (CCAI, Zhejiang University)</i>	
Policy Adaptation From Foundation Model Feedback	19059
<i>Yuying Ge (University of Hong Kong), Annabella Macaluso (University of California, San Diego), Li Erran Li (AWS AI, Amazon), Ping Luo (University of Hong Kong), and Xiaolong Wang (University of California, San Diego)</i>	
Token Turing Machines	19070
<i>Michael S. Ryoo (Google Research), Keerthana Gopalakrishnan (Google Research), Kumara Kahatapitiya (Google Research), Ted Xiao (Google Research), Kanishka Rao (Google Research), Austin Stone (Google Research), Yao Lu (Google Research), Julian Ibarz (Google Research), and Anurag Arnab (Google Research)</i>	
Unicode Analogies: An Anti-Objectivist Visual Reasoning Challenge	19082
<i>Steven Spratley (The University of Melbourne), Krista A. Ehinger (The University of Melbourne), and Tim Miller (The University of Melbourne)</i>	

Exploring the Effect of Primitives for Compositional Generalization in Vision-and-Language. 19092	
<i>Chuanhao Li (Beijing Institute of Technology, China), Zhen Li (Beijing Institute of Technology, China), Chenchen Jing (Zhejiang University, Hangzhou, China), Yunde Jia (Shenzhen MSU-BIT University, China; Beijing Institute of Technology, China), and Yuwei Wu (Shenzhen MSU-BIT University, China; Beijing Institute of Technology, China)</i>	
VQACL: A Novel Visual Question Answering Continual Learning Setting 19102	
<i>Xi Zhang (School of Artificial Intelligence, University of Chinese Academy of Sciences), Feifei Zhang (Tianjin University of Technology), and Changsheng Xu (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; School of Artificial Intelligence, University of Chinese Academy of Sciences; Peng Cheng Laboratory)</i>	
MaPLe: Multi-Modal Prompt Learning 19113	
<i>Muhammad Uzair Khattak (Mohamed bin Zayed University of AI), Hanoona Rasheed (Mohamed bin Zayed University of AI), Muhammad Maaz (Mohamed bin Zayed University of AI), Salman Khan (Mohamed bin Zayed University of AI; Australian National University), and Fahad Shahbaz Khan (Mohamed bin Zayed University of AI; Linkoping University)</i>	
Meta-Personalizing Vision-Language Models To Find Named Instances in Video 19123	
<i>Chun-Hsiao Yeh (University of California, Berkeley; Adobe Research), Bryan Russell (University of California, Berkeley; Adobe Research), Josef Sivic (CIIRC CTU; Adobe Research), Fabian Caba Heilbron (Adobe Research), and Simon Jenni (University of California, Berkeley; Adobe Research)</i>	
Understanding and Improving Visual Prompting: A Label-Mapping Perspective 19133	
<i>Aochuan Chen (Michigan State University), Yuguang Yao (Michigan State University), Pin-Yu Chen (MIT-IBM Watson AI Lab, IBM Research), Yihua Zhang (Michigan State University), and Sijia Liu (Michigan State University; MIT-IBM Watson AI Lab, IBM Research)</i>	
RefTeacher: A Strong Baseline for Semi-Supervised Referring Expression Comprehension .. 19144	
<i>Jiamu Sun (Xiamen University), Gen Luo (Xiamen University), Yiyi Zhou (Xiamen University), Xiaoshuai Sun (Xiamen University), Guannan Jiang (Contemporary Ampere Technology), Zhiyu Wang (Contemporary Ampere Technology), and Rongrong Ji (Xiamen University; Shenzhen Research Institute of Xiamen University)</i>	
Leveraging per Image-Token Consistency for Vision-Language Pre-Training 19155	
<i>Yunhao Gou (Southern University of Science and Technology; Hong Kong University of Science and Technology), Tom Ko (ByteDance AI Lab), Hansi Yang (Hong Kong University of Science and Technology), James Kwok (Hong Kong University of Science and Technology), Yu Zhang (Southern University of Science and Technology; Peng Cheng Laboratory), and Mingxuan Wang (ByteDance AI Lab)</i>	
Improving Visual Grounding by Encouraging Consistent Gradient-Based Explanations 19165	
<i>Ziyan Yang (Rice University), Kushal Kafle (Adobe Research), Franck Dernoncourt (Adobe Research), and Vicente Ordonez (Rice University)</i>	

Image as a Foreign Language: BEiT Pretraining for Vision and Vision-Language Tasks	19175
<i>Wenhui Wang (Microsoft Corporation), Hangbo Bao (Microsoft Corporation), Li Dong (Microsoft Corporation), Johan Bjorck (Microsoft Corporation), Zhiliang Peng (Microsoft Corporation), Qiang Liu (Microsoft Corporation), Kriti Aggarwal (Microsoft Corporation), Owais Khan Mohammed (Microsoft Corporation), Saksham Singhal (Microsoft Corporation), Subhojit Som (Microsoft Corporation), and Furu Wei (Microsoft Corporation)</i>	
Language in a Bottle: Language Model Guided Concept Bottlenecks for Interpretable Image Classification	19187
<i>Yue Yang (University of Pennsylvania), Artemis Panagopoulou (University of Pennsylvania), Shenghao Zhou (University of Pennsylvania), Daniel Jin (University of Pennsylvania), Chris Callison-Burch (University of Pennsylvania), and Mark Yatskar (University of Pennsylvania)</i>	
Shepherding Slots to Objects: Towards Stable and Robust Object-Centric Learning	19198
<i>Jinwoo Kim (Yonsei University), Janghyuk Choi (KAIST), Ho-Jin Choi (KAIST), and Seon Joo Kim (Yonsei University)</i>	
Learning Visual Representations via Language-Guided Sampling	19208
<i>Mohamed El Banani (University of Michigan), Karan Desai (University of Michigan), and Justin Johnson (University of Michigan)</i>	
L-Colns: Language-Based Colorization With Instance Awareness	19221
<i>Zheng Chang (Beijing University of Posts and Telecommunications), Shuchen Weng (Peking University; Peking University), Peixuan Zhang (Beijing University of Posts and Telecommunications), Yu Li (International Digital Economy Academy), Si Li (Beijing University of Posts and Telecommunications), and Boxin Shi (Peking University; Peking University)</i>	
EDA: Explicit Text-Decoupling and Dense Alignment for 3D Visual Grounding	19231
<i>Yanmin Wu (Peking University, China), Xinhua Cheng (Peking University, China), Renrui Zhang (The Chinese University of Hong Kong, China; Shanghai AI Laboratory, China), Zesen Cheng (Peking University, China), and Jian Zhang (Peking University, China)</i>	
MSINet: Twins Contrastive Search of Multi-Scale Interaction for Object ReID	19243
<i>Jiayang Gu (Zhejiang University), Kai Wang (National University of Singapore), Hao Luo (Alibaba Group), Chen Chen (OPPO Research Institute), Wei Jiang (Zhejiang University), Yuqiang Fang (Space Engineering University), Shanghang Zhang (Peking University), Yang You (National University of Singapore), and Jian Zhao (Institute of North Electronic Equipment)</i>	

Unifying Vision, Text, and Layout for Universal Document Processing	19254
<i>Zineng Tang (University of North Carolina, Chapel Hill; Microsoft Azure Cognitive Services Research), Ziyi Yang (Microsoft Azure Cognitive Services Research), Guoxin Wang (Microsoft Azure Visual Document Intelligence), Yuwei Fang (Microsoft Azure Cognitive Services Research), Yang Liu (Microsoft Azure Cognitive Services Research), Chenguang Zhu (Microsoft Azure Cognitive Services Research), Michael Zeng (Microsoft Azure Cognitive Services Research), Cha Zhang (Microsoft Azure Visual Document Intelligence), and Mohit Bansal (University of North Carolina, Chapel Hill)</i>	
RA-CLIP: Retrieval Augmented Contrastive Language-Image Pre-Training	19265
<i>Chen-Wei Xie (Alibaba Group), Siyang Sun (Alibaba Group), Xiong Xiong (Alibaba Group), Yun Zheng (Alibaba Group), Deli Zhao (Alibaba Group), and Jingren Zhou (Alibaba Group)</i>	
Fine-Grained Image-Text Matching by Cross-Modal Hard Aligning Network	19275
<i>Zhengxin Pan (Zhejiang University), Fangyu Wu (Xian Jiaotong-liverpool University), and Bailing Zhang (NingboTech University)</i>	
Text-Guided Unsupervised Latent Transformation for Multi-Attribute Image Manipulation .	19285
<i>Xiwen Wei (South China University of Technology), Zhen Xu (South China University of Technology), Cheng Liu (Shantou University), Si Wu (South China University of Technology; Peng Cheng Laboratory), Zhiwen Yu (South China University of Technology), and Hau San Wong (City University of Hong Kong)</i>	
Improving Image Recognition by Retrieving From Web-Scale Image-Text Data	19295
<i>Ahmet Iscen (Google Research), Alireza Fathi (Google Research), and Cordelia Schmid (Google Research)</i>	
Pic2Word: Mapping Pictures to Words for Zero-Shot Composed Image Retrieval	19305
<i>Kuniaki Saito (Boston University; Google Cloud AI Research), Kihyuk Sohn (Google Research), Xiang Zhang (Google Cloud AI Research), Chun-Liang Li (Google Cloud AI Research), Chen-Yu Lee (Google Cloud AI Research), Kate Saenko (Boston University; MIT-IBM Watson AI Lab), and Tomas Pfister (Google Cloud AI Research)</i>	
DATE: Domain Adaptive Product Seeker for E-Commerce	19315
<i>Haoyuan Li (Zhejiang University), Hao Jiang (Alibaba Group), Tao Jin (Zhejiang University), Mengyan Li (Alibaba Group), Yan Chen (Alibaba Group), Zhijie Lin (Zhejiang University), Yang Zhao (Zhejiang University), and Zhou Zhao (Zhejiang University)</i>	
Multimodality Helps Unimodality: Cross-Modal Few-Shot Learning With Multimodal Models	19325
<i>Zhiqiu Lin (Carnegie Mellon University), Samuel Yu (Carnegie Mellon University), Zhiyi Kuang (Carnegie Mellon University), Deepak Pathak (Carnegie Mellon University), and Deva Ramanan (Carnegie Mellon University)</i>	
Finetune Like You Pretrain: Improved Finetuning of Zero-Shot Vision Models	19338
<i>Sachin Goyal (Carnegie Mellon University), Ananya Kumar (Stanford University), Sankalp Garg (Carnegie Mellon University), Zico Kolter (Carnegie Mellon University; Bosch Center for AI), and Aditi Raghunathan (Carnegie Mellon University)</i>	

DeepSolo: Let Transformer Decoder With Explicit Points Solo for Text Spotting	19348
<i>Maoyuan Ye (Wuhan University, China), Jing Zhang (The University of Sydney, Australia), Shanshan Zhao (JD Explore Academy, China), Juhua Liu (Wuhan University, China), Tongliang Liu (The University of Sydney, Australia), Bo Du (Wuhan University, China), and Dacheng Tao (The University of Sydney, Australia; JD Explore Academy, China)</i>	
EVA: Exploring the Limits of Masked Visual Representation Learning at Scale	19358
<i>Yuxin Fang (Huazhong University of Science and Technology; Beijing Academy of Artificial Intelligence), Wen Wang (Zhejiang University; Beijing Academy of Artificial Intelligence), Binhui Xie (Beijing Institute of Technology; Beijing Academy of Artificial Intelligence), Quan Sun (Beijing Academy of Artificial Intelligence), Ledell Wu (Beijing Academy of Artificial Intelligence), Xinggang Wang (Huazhong University of Science and Technology), Tiejun Huang (Beijing Academy of Artificial Intelligence), Xinlong Wang (Beijing Academy of Artificial Intelligence), and Yue Cao (Beijing Academy of Artificial Intelligence)</i>	
R2Former: Unified Retrieval and Reranking Transformer for Place Recognition	19370
<i>Sijie Zhu (University of Central Florida), Linjie Yang (University of Central Florida), Chen Chen (University of Central Florida), Mubarak Shah (University of Central Florida), Xiaohui Shen (University of Central Florida), and Heng Wang (University of Central Florida)</i>	
Open-Set Fine-Grained Retrieval via Prompting Vision-Language Evaluator	19381
<i>Shijie Wang (Dalian University of Technology, China), Jianlong Chang (Huawei Cloud & AI, China), Haojie Li (Dalian University of Technology, China; Shandong University of Science and Technology, China), Zhihui Wang (Dalian University of Technology, China), Wanli Ouyang (The University of Sydney, Australia), and Qi Tian (Huawei Cloud & AI, China)</i>	
Open-Category Human-Object Interaction Pre-Training via Language Modeling Framework	19392
<i>Sipeng Zheng (Renmin University of China), Boshen Xu (UESTC), and Qin Jin (Renmin University of China)</i>	
Neural Congealing: Aligning Images to a Joint Semantic Atlas	19403
<i>Dolev Ofri-Amar (Weizmann Institute of Science), Michal Geyer (Weizmann Institute of Science), Yoni Kasten (NVIDIA Research), and Tali Dekel (Weizmann Institute of Science)</i>	
Open Vocabulary Semantic Segmentation With Patch Aligned Contrastive Learning	19413
<i>Jishnu Mukhoti (University of Oxford; Meta AI), Tsung-Yu Lin (Meta AI), Omid Poursaeed (Meta AI), Rui Wang (Meta AI), Ashish Shah (Meta AI), Philip H.S. Torr (University of Oxford), and Ser-Nam Lim (Meta AI)</i>	
Semantic Human Parsing via Scalable Semantic Transfer Over Multiple Label Domains	19424
<i>Jie Yang (The Chinese University of Hong Kong, Shenzhen), Chaoqun Wang (The Chinese University of Hong Kong, Shenzhen), Zhen Li (The Chinese University of Hong Kong, Shenzhen), Junle Wang (Tencent), and Ruimao Zhang (The Chinese University of Hong Kong, Shenzhen)</i>	

Explicit Visual Prompting for Low-Level Structure Segmentations	19434
<i>Wei Huang Liu (University of Macau), Xi Shen (Tencent AI Lab), Chi-Man Pun (University of Macau), and Xiaodong Cun (Tencent AI Lab)</i>	
FreeSeg: Unified, Universal and Open-Vocabulary Image Segmentation	19446
<i>Jie Qin (Institute of Automation, Chinese Academy of Sciences; ByteDance Inc; University of Chinese Academy of Sciences), Jie Wu (ByteDance Inc), Pengxiang Yan (ByteDance Inc), Ming Li (ByteDance Inc), Ren Yuxi (ByteDance Inc), Xuefeng Xiao (ByteDance Inc), Yitong Wang (ByteDance Inc), Rui Wang (ByteDance Inc), Shilei Wen (ByteDance Inc), Xin Pan (ByteDance Inc), and Xingang Wang (Institute of Automation, Chinese Academy of Sciences)</i>	
Zero-Shot Referring Image Segmentation With Global-Local Context Features	19456
<i>Seonghoon Yu (AI Graduate School), Paul Hongsuck Seo (Google Research), and Jeany Son (AI Graduate School)</i>	
DejaVu: Conditional Regenerative Learning To Enhance Dense Prediction	19466
<i>Shubhankar Borse (Qualcomm AI Research), Debansmit Das (Qualcomm AI Research), Hyojin Park (Qualcomm AI Research), Hong Cai (Qualcomm AI Research), Rishiek Garrepalli (Qualcomm AI Research), and Fatih Porikli (Qualcomm AI Research)</i>	
Meta Compositional Referring Expression Segmentation	19478
<i>Li Xu (Singapore University of Technology and Design, Singapore), Mark He Huang (Singapore University of Technology and Design, Singapore; Institute for Infocomm Research (I2R) & Centre for Frontier AI Research (CFAR), A*STAR, Singapore), Xindi Shang (ByteDance), Zehuan Yuan (ByteDance), Ying Sun (Institute for Infocomm Research (I2R) & Centre for Frontier AI Research (CFAR), A*STAR, Singapore), and Jun Liu (Singapore University of Technology and Design, Singapore)</i>	
Interactive Segmentation As Gaussian Process Classification	19488
<i>Minghao Zhou (Xi'an Jiaotong University, China; Tencent Jarvis Lab, China), Hong Wang (Tencent Jarvis Lab, China), Qian Zhao (Xi'an Jiaotong University, China), Yuexiang Li (Tencent Jarvis Lab, China), Yawen Huang (Tencent Jarvis Lab, China), Deyu Meng (Xi'an Jiaotong University, China; Peng Cheng Laboratory, China; Macau University of Science and Technology, China), and Yefeng Zheng (Tencent Jarvis Lab, China)</i>	
Semantic-Promoted Debiasing and Background Disambiguation for Zero-Shot Instance Segmentation	19498
<i>Shuting He (Zhejiang University), Henghui Ding (Nanyang Technological University), and Wei Jiang (Zhejiang University)</i>	
Principles of Forgetting in Domain-Incremental Semantic Segmentation in Adverse Weather Conditions	19508
<i>Tobias Kalb (Porsche Engineering Group GmbH, Germany) and Jürgen Beyerer (Fraunhofer IOSB & Karlsruhe Institute of Technology, Germany)</i>	

AttentionShift: Iteratively Estimated Part-Based Attention Map for Pointly Supervised Instance Segmentation	19519
<i>Mingxiang Liao (University of Chinese Academy of Sciences), Zonghao Guo (University of Chinese Academy of Sciences), Yuze Wang (Huawei Noah's Ark Lab), Peng Yuan (Huawei Noah's Ark Lab), Bailan Feng (Huawei Noah's Ark Lab), and Fang Wan (University of Chinese Academy of Sciences)</i>	
PIDNet: A Real-Time Semantic Segmentation Network Inspired by PID Controllers	19529
<i>Jiacong Xu (Texas A&M University, College Station), Zixiang Xiong (Texas A&M University, College Station), and Shankar P. Bhattacharyya (Texas A&M University, College Station)</i>	
Leveraging Hidden Positives for Unsupervised Semantic Segmentation	19540
<i>Hyun Seok Seong (Sungkyunkwan University), WonJun Moon (Sungkyunkwan University), SuBeen Lee (Sungkyunkwan University), and Jae-Pil Heo (Sungkyunkwan University)</i>	
Understanding Imbalanced Semantic Segmentation Through Neural Collapse	19550
<i>Zhisheng Zhong (CUHK), Jiequan Cui (CUHK), Yibo Yang (JD Explore Academy), Xiaoyang Wu (HKU), Xiaojuan Qi (HKU), Xiangyu Zhang (MEGVII), and Jiaya Jia (CUHK; SmartMore)</i>	
Balancing Logit Variation for Long-Tailed Semantic Segmentation	19561
<i>Yuchao Wang (Shanghai Jiao Tong University), Jingjing Fei (SenseTime Research), Haochen Wang (Chinese Academy of Sciences), Wei Li (SenseTime Research), Tianpeng Bao (SenseTime Research), Liwei Wu (SenseTime Research), Rui Zhao (Shanghai Jiao Tong University; SenseTime Research), and Yujun Shen (CUHK)</i>	
Boundary-Enhanced Co-Training for Weakly Supervised Semantic Segmentation	19574
<i>Shenghai Rong (University of Science and Technology of China, China), Bohai Tu (University of Science and Technology of China, China), Zilei Wang (University of Science and Technology of China, China), and Junjie Li (University of Science and Technology of China, China)</i>	
Conflict-Based Cross-View Consistency for Semi-Supervised Semantic Segmentation	19585
<i>Zicheng Wang (University of Sydney; Samsung Research China-Beijing), Zhen Zhao (University of Sydney), Xiaoxia Xing (Samsung Research China-Beijing), Dong Xu (University of Hong Kong), Xiangyu Kong (Samsung Research China-Beijing), and Luping Zhou (University of Sydney)</i>	
Learning Multi-Modal Class-Specific Tokens for Weakly Supervised Dense Object Localization.....	19596
<i>Lian Xu (The University of Western Australia), Wanli Ouyang (Shanghai AI Laboratory), Mohammed Bennamoun (The University of Western Australia), Farid Boussaid (The University of Western Australia), and Dan Xu (Hong Kong University of Science and Technology)</i>	
WinCLIP: Zero-/Few-Shot Anomaly Classification and Segmentation	19606
<i>Jongheon Jeong (KAIST), Yang Zou (AWS AI Labs), Taewan Kim (AWS AI Labs), Dongqing Zhang (AWS AI Labs), Avinash Ravichandran (AWS AI Labs), and Onkar Dabeer (AWS AI Labs)</i>	

DualRel: Semi-Supervised Mitochondria Segmentation From a Prototype Perspective	19617
<i>Huayu Mai (University of Science and Technology of China), Rui Sun (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center; Deep Space Exploration Lab), Zhiwei Xiong (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), and Feng Wu (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center)</i>	
Distilling Self-Supervised Vision Transformers for Weakly-Supervised Few-Shot Classification & Segmentation	19627
<i>Dahyun Kang (Meta AI; POSTECH), Piotr Koniusz (Data61 CSIRO; Australian National University), Minsu Cho (POSTECH), and Naila Murray (Meta AI)</i>	
Co-Salient Object Detection With Uncertainty-Aware Group Exchange-Masking	19639
<i>Yang Wu (Nanjing University of Information Science and Technology, China), Huihui Song (Nanjing University of Information Science and Technology, China), Bo Liu (Walmart Global Tech, USA), Kaihua Zhang (Nanjing University of Information Science and Technology, Nanjing, China), and Dong Liu (Netflix Inc, USA)</i>	
Supervised Masked Knowledge Distillation for Few-Shot Transformers	19649
<i>Han Lin (Columbia University), Guangxing Han (Columbia University), Jiawei Ma (Columbia University), Shiyuan Huang (Columbia University), Xudong Lin (Columbia University), and Shih-Fu Chang (Columbia University)</i>	
Modeling the Distributional Uncertainty for Salient Object Detection Models	19660
<i>Xinyu Tian (Northwestern Polytechnical University, China), Jing Zhang (Australian National University, Australia), Mochu Xiang (Northwestern Polytechnical University, China), and Yuchao Dai (Northwestern Polytechnical University, China)</i>	
Weak-Shot Object Detection Through Mutual Knowledge Transfer	19671
<i>Xuanyi Du (WeChat, Tencent), Weitao Wan (WeChat, Tencent), Chong Sun (WeChat, Tencent), and Chen Li (WeChat, Tencent)</i>	
CAT: LoCalization and IdentificAtion Cascade Detection Transformer for Open-World Object Detection	19681
<i>Shuailei Ma (Northeast University, China), Yuefeng Wang (Northeast University, China), Ying Wei (Northeast University, China; Information Technology R&D Innovation Center of Peking University; Changsha Hisense Intelligent System Research Institute Co., Ltd), Jiaqi Fan (Northeast University, China), Thomas H. Li (Peking University Shenzhen Graduate School, China), Hongli Liu (Changsha Hisense Intelligent System Research Institute Co., Ltd), and Fanbing Lv (Changsha Hisense Intelligent System Research Institute Co., Ltd)</i>	
Adaptive Sparse Pairwise Loss for Object Re-Identification	19691
<i>Xiao Zhou (Tsinghua University, Beijing, 100084), Yujie Zhong (Meituan Inc.), Zhen Cheng (Tsinghua University, Beijing, 100084), Fan Liang (Meituan Inc.), and Lin Ma (Meituan Inc.)</i>	

DETRs With Hybrid Matching	19702
<i>Ding Jia (Peking University), Yuhui Yuan (Microsoft Research Asia), Haodi He (Stanford University), Xiaopei Wu (Zhejiang University), Haojun Yu (Peking University), Weihong Lin (Microsoft Research Asia), Lei Sun (Microsoft Research Asia), Chao Zhang (Peking University), and Han Hu (Microsoft Research Asia)</i>	
Generating Features With Increased Crop-Related Diversity for Few-Shot Object Detection .	19713
<i>Jingyi Xu (Stony Brook University), Hieu Le (EPFL), and Dimitris Samaras (Stony Brook University)</i>	
ScaleKD: Distilling Scale-Aware Knowledge in Small Object Detector	19723
<i>Yichen Zhu (Midea Group), Qiqi Zhou (Midea Group; Shanghai University of Electric Power), Ning Liu (Midea Group), Zhiyuan Xu (Midea Group), Zhikai Ou (Midea Group), Xiaofeng Mou (Midea Group), and Jian Tang (Midea Group)</i>	
Multiclass Confidence and Localization Calibration for Object Detection	19734
<i>Bimsara Pathiraja (Mohamed bin Zayed University of Artificial Intelligence, UAE), Malitha Gunawardhana (Mohamed bin Zayed University of Artificial Intelligence, UAE), and Muhammad Haris Khan (Mohamed bin Zayed University of Artificial Intelligence, UAE)</i>	
Open-Set Representation Learning Through Combinatorial Embedding	19744
<i>Geeho Kim (Computer Vision Laboratory, ECE & IPAI), Junoh Kang (Computer Vision Laboratory, ECE & IPAI), and Bohyung Han (Computer Vision Laboratory, ECE & IPAI; Seoul National University)</i>	
ProD: Prompting-To-Disentangle Domain Knowledge for Cross-Domain Few-Shot Image Classification	19754
<i>Tianyi Ma (University of Technology Sydney; Baidu Inc.), Yifan Sun (Baidu Inc.), Zongxin Yang (Zhejiang University), and Yi Yang (Zhejiang University)</i>	
Visual Language Pretrained Multiple Instance Zero-Shot Transfer for Histopathology Images	19764
<i>Ming Y. Lu (Massachusetts Institute of Technology; Harvard University; Mass General Brigham), Bowen Chen (Harvard University; Mass General Brigham), Andrew Zhang (Massachusetts Institute of Technology; Harvard University; Mass General Brigham), Drew F. K. Williamson (Harvard University; Mass General Brigham), Richard J. Chen (Harvard University; Mass General Brigham), Tong Ding (Harvard University; Mass General Brigham), Long Phi Le (Harvard University; Mass General Brigham), Yung-Sung Chuang (Massachusetts Institute of Technology), and Faisal Mahmood (Harvard University; Mass General Brigham)</i>	
FFF: Fragment-Guided Flexible Fitting for Building Complete Protein Structures	19776
<i>Weijie Chen (DP Technology, Ltd.; Peking University), Xinyan Wang (DP Technology, Ltd.), and Yuhang Wang (DP Technology, Ltd.)</i>	
Pseudo-Label Guided Contrastive Learning for Semi-Supervised Medical Image Segmentation	19786
<i>Hritam Basak (Stony Brook University, NY, USA) and Zhaozheng Yin (Stony Brook University, NY, USA)</i>	

Hierarchical Discriminative Learning Improves Visual Representations of Biomedical Microscopy	19798
<i>Cheng Jiang (University of Michigan), Xinhai Hou (University of Michigan), Akhil Kondepudi (University of Michigan), Asadur Chowdury (University of Michigan), Christian W. Freudiger (Invenio Imaging), Daniel A. Orringer (New York University), Honglak Lee (University of Michigan; LG AI Research), and Todd C. Hollon (University of Michigan)</i>	
KiUT: Knowledge-Injected U-Transformer for Radiology Report Generation	19809
<i>Zhongzhen Huang (Shanghai Jiao Tong University; Shanghai AI Laboratory), Xiaofan Zhang (Shanghai Jiao Tong University; Shanghai AI Laboratory), and Shaoting Zhang (Shanghai AI Laboratory; SenseTime Research)</i>	
Image Quality-Aware Diagnosis via Meta-Knowledge Co-Embedding	19819
<i>Haoxuan Che (The Hong Kong University of Science and Technology), Siyu Chen (The Hong Kong University of Science and Technology), and Hao Chen (The Hong Kong University of Science and Technology)</i>	
Interventional Bag Multi-Instance Learning on Whole-Slide Pathological Images	19830
<i>Tiancheng Lin (Shanghai Key Lab of Digital Media Processing and Transmission, Shanghai Jiao Tong University; MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University), Zhimiao Yu (Shanghai Key Lab of Digital Media Processing and Transmission, Shanghai Jiao Tong University; MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University), Hongyu Hu (Shanghai Key Lab of Digital Media Processing and Transmission, Shanghai Jiao Tong University; MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University), Yi Xu (Shanghai Key Lab of Digital Media Processing and Transmission, Shanghai Jiao Tong University; MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University), and Chang-Wen Chen (The Hong Kong Polytechnic University, Hong Kong, China)</i>	
Visual Prompt Tuning for Generative Transfer Learning	19840
<i>Kihyuk Sohn (Google Research), Huiwen Chang (Google Research), José Lezama (Google Research), Luisa Polania (Google Research), Han Zhang (Google Research), Yuan Hao (Google Research), Irfan Essa (Google Research), and Lu Jiang (Google Research)</i>	
LINE: Out-of-Distribution Detection by Leveraging Important Neurons	19852
<i>Yong Hyun Ahn (Kyung Hee University), Gyeong-Moon Park (Kyung Hee University), and Seong Tae Kim (Kyung Hee University)</i>	
GCFagg: Global and Cross-View Feature Aggregation for Multi-View Clustering	19863
<i>Weiqing Yan (Yantai University, China; Nanyang Technological University, Singapore), Yuanyang Zhang (Yantai University, China), Chenlei Lv (Shenzhen University, China), Chang Tang (China University of Geosciences, China), Guanghui Yue (Health Science Center, Shenzhen University, China), Liang Liao (Nanyang Technological University, Singapore), and Weisi Lin (Nanyang Technological University, Singapore)</i>	

Exploring and Exploiting Uncertainty for Incomplete Multi-View Classification	19873
<i>Mengyao Xie (Tianjin University), Zongbo Han (Tianjin University), Changqing Zhang (Tianjin University), Yichen Bai (Tianjin University), and Qinghua Hu (Tianjin University)</i>	
BiCro: Noisy Correspondence Rectification for Multi-Modality Data via Bi-Directional Cross-Modal Similarity Consistency	19883
<i>Shuo Yang (University of Technology Sydney), Zhaopan Xu (Harbin Institute of Technology), Kai Wang (National University Singapore), Yang You (National University Singapore), Hongxun Yao (Harbin Institute of Technology), Tongliang Liu (The University of Sydney), and Min Xu (University of Technology Sydney)</i>	
Bi-Directional Distribution Alignment for Transductive Zero-Shot Learning	19893
<i>Zhicai Wang (University of Science and Technology of China), Yanbin Hao (University of Science and Technology of China), Tingting Mu (The University of Manchester), Ouxiang Li (University of Science and Technology of China), Shuo Wang (University of Science and Technology of China), and Xiangnan He (University of Science and Technology of China)</i>	
HIER: Metric Learning Beyond Class Labels via Hierarchical Regularization	19903
<i>Sungyeon Kim (Dept. of CSE, POSTECH), Boseung Jeong (Dept. of CSE, POSTECH), and Suha Kwak (Graduate School of AI, POSTECH)</i>	
MaskCon: Masked Contrastive Learning for Coarse-Labelled Dataset	19913
<i>Chen Feng (Queen Mary University of London, UK) and Ioannis Patras (Queen Mary University of London, UK)</i>	
Class Prototypes Based Contrastive Learning for Classifying Multi-Label and Fine-Grained Educational Videos	19923
<i>Rohit Gupta (University of Central Florida), Anirban Roy (SRI International), Claire Christensen (SRI International), Sujeong Kim (SRI International), Sarah Gerard (SRI International), Madeline Cincebeaux (SRI International), Ajay Divakaran (SRI International), Todd Grindal (SRI International), and Mubarak Shah (University of Central Florida)</i>	
Learning From Noisy Labels With Decoupled Meta Label Purifier	19934
<i>Yuanpeng Tu (Dept. of Electronic and Information Engineering, Tongji Univeristy, Shanghai), Boshen Zhang (YouTu Lab, Tencent, Shanghai), Yuxi Li (YouTu Lab, Tencent, Shanghai), Liang Liu (YouTu Lab, Tencent, Shanghai), Jian Li (YouTu Lab, Tencent, Shanghai), Yabiao Wang (YouTu Lab, Tencent, Shanghai), Chengjie Wang (YouTu Lab, Tencent, Shanghai; Shanghai Jiao Tong University), and Cai Rong Zhao (Dept. of Electronic and Information Engineering, Tongji Univeristy, Shanghai)</i>	
SuperDisco: Super-Class Discovery Improves Visual Recognition for the Long-Tail	19944
<i>Yingjun Du (University of Amsterdam), Jiayi Shen (University of Amsterdam), Xiantong Zhen (University of Amsterdam; Inception Institute of Artificial Intelligence), and Cees G. M. Snoek (University of Amsterdam)</i>	

Matthias Eisenmann (Division of Intelligent Medical Systems, German Cancer Research Center (DKFZ), Germany; Helmholtz Imaging, German Cancer Research Center (DKFZ), Germany), Annika Reinke (Division of Intelligent Medical Systems, German Cancer Research Center (DKFZ), Germany; Helmholtz Imaging, German Cancer Research Center (DKFZ), Germany; Heidelberg University, Germany), Vivienn Weru (Division of Biostatistics, German Cancer Research Center (DKFZ), Germany), Minu D. Tizabi (Division of Intelligent Medical Systems, German Cancer Research Center (DKFZ), Germany; Helmholtz Imaging, German Cancer Research Center (DKFZ), Germany), Fabian Isensee (Division of Medical Image Computing, German Cancer Research Center (DKFZ), Germany; Helmholtz Imaging, German Cancer Research Center (DKFZ), Germany), Tim J. Adler (Division of Intelligent Medical Systems, German Cancer Research Center (DKFZ), Germany), Sharib Ali (University of Leeds, UK), Vincent Andrearczyk (HES-SO Valais-Wallis University of Applied Sciences and Arts Western Switzerland, Switzerland; Lausanne University Hospital, Switzerland), Marc Aubreville (Technische Hochschule Ingolstadt, Germany), Ujjwal Baid (University of Pennsylvania, USA; University of Pennsylvania, USA; University of Pennsylvania, USA), Spyridon Bakas (University of Pennsylvania, USA; University of Pennsylvania, USA; University of Pennsylvania, USA), Niranjana Balu (University of Washington, USA), Sophia Bano (University College London, UK), Jorge Bernal (Universitat Autònoma de Barcelona & Computer Vision Center, Spain), Sebastian Bodenstedt (Division of Translational Surgical Oncology, National Center for Tumor Diseases (NCT/UCC) Dresden, Germany), Alessandro Casella (Department of Advanced Robotics, Istituto Italiano di Tecnologia, Italy and Department of Electronics, Information and Bioengineering, Italy), Veronika Cheplygina (IT University of Copenhagen, Denmark), Marie Daum (Heidelberg University Hospital, Germany), Marleen de Bruijne (Biomedical Imaging Group Rotterdam, Department of Radiology and Nuclear Medicine, Erasmus MC, The Netherlands; University of Copenhagen, Denmark), Adrien Depeursinge (University of Applied Sciences Western Switzerland (HES-SO), Switzerland; Lausanne University Hospital, Switzerland), Reuben Dorent (Harvard Medical School, Brigham and Women's Hospital, USA; School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK), Jan Egger (University Hospital Essen (AöR), Germany), David G. Ellis (University of Nebraska Medical Center, USA), Sandy Engelhardt (Heidelberg University Hospital, Germany), Melanie Ganz (Copenhagen University Hospital, Denmark; University of Copenhagen, Denmark), Noha Ghatwary (Arab Academy of Science and Technology, Egypt), Gabriel Girard (CIBM Center for Biomedical Imaging, Switzerland; Centre Hospitalier Universitaire Vaudois (CHUV) and University of Lausanne (UNIL), Switzerland; Signal Processing Laboratory (LTS5), École Polytechnique Fédérale de Lausanne (EPFL), Switzerland), Patrick Godau (Division of Intelligent Medical Systems, German Cancer Research Center (DKFZ), Germany; Helmholtz Imaging, German Cancer Research Center (DKFZ), Germany; Heidelberg University, Germany; National Center for Tumor Diseases (NCT), Germany), Anubha Gupta (SBILab, Department of ECE, IIT-Delhi, India), Lasse Hansen (University of

Lübeck, Germany), Kanako Harada (The University of Tokyo, Japan), Mattias P. Heinrich (University of Lübeck, Germany), Nicholas Heller (University of Minnesota, USA), Alessa Hering (Radboud University Medical Center, The Netherlands; Fraunhofer MEVIS, Germany), Arnaud Huaulmé (Univ Rennes, INSERM, France), Pierre Jannin (Univ Rennes, INSERM, France), Ali Emre Kavur (Division of Intelligent Medical Systems, German Cancer Research Center (DKFZ), Germany; Helmholtz Imaging, German Cancer Research Center (DKFZ), Germany), Oldřich Kodým (Brno University of Technology, Czech Republic), Michal Kozubek (Masaryk University, Czech Republic), Jianning Li (University Hospital Essen (AöR), Germany), Hongwei Li (University of Zurich, Switzerland), Jun Ma (University of Toronto, Canada), Carlos Martín-Isla (Universitat de Barcelona, Spain), Bjoern Menze (University of Zurich, Switzerland), Alison Noble (University of Oxford, UK), Valentin Oreiller (University of Applied Sciences Western Switzerland (HES-SO), Switzerland; Lausanne University Hospital, Switzerland), Nicolas Padoy (University of Strasbourg, CNRS, France; IHU Strasbourg, France), Sarthak Pati (University of Pennsylvania, USA; University of Pennsylvania, USA; University of Pennsylvania, USA; Technical University of Munich, Germany), Kelly Payette (University Children's Hospital Zurich, University of Zurich, Switzerland; University of Zurich, Switzerland), Tim Rädtsch (Division of Intelligent Medical Systems, German Cancer Research Center (DKFZ), Germany; Helmholtz Imaging, German Cancer Research Center (DKFZ), Germany), Jonathan Rafael-Patiño (Centre Hospitalier Universitaire Vaudois (CHUV) and University of Lausanne (UNIL), Switzerland; Signal Processing Laboratory (LTS5), École Polytechnique Fédérale de Lausanne (EPFL), Switzerland), Vivek Singh Bawa (Oxford Brookes University, UK), Stefanie Speidel (Division of Translational Surgical Oncology, National Center for Tumor Diseases (NCT/UCC) Dresden, Germany; Centre for Tactile Internet with Human-in-the-Loop (CeTI), TU Dresden, Germany), Carole H. Sudre (University College London, UK; University College London, UK; School of Biomedical Engineering & Imaging Sciences, King's College London, UK; University College London, UK), Kimberlin van Wijnen (Biomedical Imaging Group Rotterdam, Department of Radiology and Nuclear Medicine, Erasmus MC, The Netherlands), Martin Wagner (Heidelberg University Hospital, Germany), Donglai Wei (Computer Science, Boston College, USA), Amine Yamlahi (Division of Intelligent Medical Systems, German Cancer Research Center (DKFZ), Germany), Moi Hoon Yap (Manchester Metropolitan University, UK), Chun Yuan (University of Washington, USA), Maximilian Zenk (Division of Medical Image Computing, German Cancer Research Center (DKFZ), Germany; Heidelberg University, Germany), Aneeq Zia (Intuitive Surgical, Inc., Sunnyvale, USA), David Zimmerer (Division of Medical Image Computing, German Cancer Research Center (DKFZ), Germany; Helmholtz Imaging, German Cancer Research Center (DKFZ), Germany), Dogu Baran Aydogan (University of Eastern Finland, Finland; Aalto University School of Science, Finland), Binod Bhattarai (University of Aberdeen, UK), Louise Bloch (University of Applied Sciences and Arts Dortmund, Germany; University Hospital Essen, Germany; University Hospital Essen, Germany), Raphael Brüngel (University of Applied Sciences and Arts Dortmund, Germany; University Hospital Essen,

Germany; University Hospital Essen, Germany), Jihoon Cho (School of Computing, Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea), Chanyeol Choi (Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, USA), Qi Dou (The Chinese University of Hong Kong, Hong Kong), Ivan Ezhov (Technical University of Munich, Germany), Christoph M. Friedrich (University of Applied Sciences and Arts Dortmund, Germany; University Hospital Essen, Germany), Clifton D. Fuller (The University of Texas MD Anderson Cancer Center, USA), Rebati Raman Gaire (Nepal Applied Mathematics and Informatics Institute for Research (NAAMII), Nepal), Adrian Galdran (Universidad Pompeu Fabra, Spain; University of Adelaide, Australia), Álvaro García Faura (XLAB d.o.o., Ljubljana, Slovenia), Maria Grammatikopoulou (Touch Surgery, Medtronic, UK), SeulGi Hong (CJ AI Center, Seoul, Republic of Korea), Mostafa Jahanifar (University of Warwick, UK), Ikbeom Jang (Hankuk University of Foreign Studies, Republic of Korea; Massachusetts General Hospital, USA; Harvard Medical School, USA), Abdolrahim Kadkhodamohammadi (Touch Surgery, Medtronic, UK), Inha Kang (School of Computing, Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea), Florian Kofler (Helmholtz AI, Helmholtz Zentrum München, Germany; Technical University Munich, Germany; Technical University of Munich, Germany; Technical University of Munich, Germany), Satoshi Kondo (Muroran Institute of Technology, Japan), Hugo Kuijff (UMC Utrecht, Utrecht, The Netherlands), Mingxing Li (University of Science and Technology of China, China), Minh Luu (Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea), Tomaž Martinčič (XLAB d.o.o., Ljubljana, Slovenia), Pedro Morais (2Ai, School of Technology, IPCA, Portugal), Mohamed A. Naser (The University of Texas MD Anderson Cancer Center, USA), Bruno Oliveira (2Ai, School of Technology, IPCA, Portugal; University of Minho, Portugal; University of Minho, Portugal), David Owen (Touch Surgery, Medtronic, UK), Subeen Pang (Massachusetts Institute of Technology, USA), Jinah Park (School of Computing, Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea), Sung-Hong Park (Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea), Szymon Plotka (Sano Centre for Computational Medicine, Poland; University of Amsterdam, The Netherlands; Amsterdam University Medical Center, University of Amsterdam, The Netherlands), Elodie Puybareau (LRE, EPITA, France), Nasir Rajpoot (University of Warwick, UK), Kanghyun Ryu (Artificial Intelligence and Robotics Institute, Korea Institute of Science and Technology, Republic of Korea), Numan Saeed (Mohamed Bin Zayed University of Artificial Intelligence, UAE), Adam Shephard (University of Warwick, UK), Pengcheng Shi (Mohamed Bin Zayed University of Artificial Intelligence, UAE), Dejan Štepec (XLAB d.o.o., Ljubljana, Slovenia; University of Ljubljana, Slovenia), Ronast Subedi (Nepal Applied Mathematics and Informatics Institute for Research (NAAMII), Nepal), Guillaume Tochon (LRE, EPITA, France), Helena R. Torres (2Ai, School of Technology, IPCA, Portugal; University of Minho, Portugal; University of Minho, Portugal), Helene Urien (ISEP, France), João L. Vilaça (2Ai, School of Technology, IPCA, Portugal), Kareem A. Wahid (The University of Texas MD Anderson Cancer Center, USA), Haojie Wang (Xiamen University, China), Jiacheng Wang

(Xiamen University, China), Liansheng Wang (Xiamen University, China), Xiyue Wang (Sichuan University, China), Benedikt Wiestler (Technical University of Munich, Germany), Marek Wodzinski (AGH UST, Department of Measurement and Electronics, Poland; University of Applied Sciences and Arts Western Switzerland (HES-SO), Switzerland), Fangfang Xia (Data Science and Learning Division, Argonne National Laboratory, USA; University of Chicago, USA), Juanying Xie (Shaanxi Normal University, China), Zhiwei Xiong (University of Science and Technology of China, China), Sen Yang (AI Lab, Tencent, China), Yanwu Yang (Harbin Institute of Technology, China), Zixuan Zhao (University of Chicago, USA), Klaus Maier-Hein (Division of Medical Image Computing, German Cancer Research Center (DKFZ), Germany; Helmholtz Imaging, German Cancer Research Center (DKFZ), Germany; Heidelberg University Hospital, Germany), Paul F. Jäger (Interactive Machine Learning Group, German Cancer Research Center (DKFZ), Germany; Helmholtz Imaging, German Cancer Research Center (DKFZ), Germany), Annette Kopp-Schneider (Division of Biostatistics, German Cancer Research Center (DKFZ), Germany), and Lena Maier-Hein (Division of Intelligent Medical Systems, German Cancer Research Center (DKFZ), Germany; Helmholtz Imaging, German Cancer Research Center (DKFZ), Germany; Heidelberg University, Germany; Heidelberg University, Germany; National Center for Tumor Diseases (NCT), Germany)

- Balanced Product of Calibrated Experts for Long-Tailed Recognition 19967
Emanuel Sanchez Aimar (Linköping University, Sweden), Arvi Jonnarth (Linköping University, Sweden), Michael Felsberg (Linköping University, Sweden), and Marco Kuhlmann (Linköping University, Sweden)
- Transfer Knowledge From Head to Tail: Uncertainty Calibration Under Long-Tailed Distribution 19978
Jiahao Chen (Renmin University of China) and Bing Su (Beijing Key Laboratory of Big Data Management and Analysis Methods)
- FREDOM: Fairness Domain Adaptation Approach to Semantic Scene Understanding 19988
Thanh-Dat Truong (University of Arkansas, USA), Ngan Le (University of Arkansas, USA), Bhiksha Raj (Carnegie Mellon University, USA), Jackson Cothren (University of Arkansas, USA), and Khoa Luu (University of Arkansas, USA)
- COT: Unsupervised Domain Adaptation With Clustering and Optimal Transport 19998
Yang Liu (Alibaba Group), Zhipeng Zhou (Alibaba Group), and Baigui Sun (Alibaba Group)
- MHPL: Minimum Happy Points Learning for Active Source Free Domain Adaptation 20008
Fan Wang (Shandong University, China), Zhongyi Han (Shandong University, China), Zhiyan Zhang (Shandong University, China), Rundong He (Shandong University, China), and Yilong Yin (Shandong University, China)
- Upcycling Models Under Domain and Category Shift 20019
Sanqing Qu (Tongji University), Tianpei Zou (Tongji University), Florian Röhrbein (Chemnitz University of Technology), Cewu Lu (Shanghai Jiao Tong University), Guang Chen (Tongji University), Dacheng Tao (JD Explore Academy; The University of Sydney), and Changjun Jiang (Tongji University)

PMR: Prototypical Modal Rebalance for Multimodal Learning	20029
<i>Yunfeng Fan (The Hong Kong Polytechnic University), Wenchao Xu (The Hong Kong Polytechnic University), Haozhao Wang (Huazhong University of Science and Technology), Junxiao Wang (KAUST; SDAIA-KAUST AI Center), and Song Guo (The Hong Kong Polytechnic University)</i>	
MMANet: Margin-Aware Distillation and Modality-Aware Regularization for Incomplete Multimodal Learning	20039
<i>Shicai Wei (University of Electronic Science and Technology of China), Chunbo Luo (University of Electronic Science and Technology of China), and Yang Luo (University of Electronic Science and Technology of China)</i>	
Feature Alignment and Uniformity for Test Time Adaptation	20050
<i>Shuai Wang (Tsmghua University), Daoan Zhang (Southern University of Science and Technology), Zipei Yan (Hongkong Polytechnic University), Jianguo Zhang (Southern University of Science and Technology), and Rui Li (Tsmghua University)</i>	
Revisiting Prototypical Network for Cross Domain Few-Shot Learning	20061
<i>Fei Zhou (Northwestern Polytechnical University), Peng Wang (University of Wollongong), Lei Zhang (Northwestern Polytechnical University), Wei Wei (Northwestern Polytechnical University), and Yanning Zhang (Northwestern Polytechnical University)</i>	
A Whac-a-Mole Dilemma: Shortcuts Come in Multiples Where Mitigating One Amplifies Others	20071
<i>Zhiheng Li (University of Rochester), Ivan Evtimov (Meta AI), Albert Gordo (Meta AI), Caner Hazirbas (Meta AI), Tal Hassner (Meta AI), Cristian Canton Ferrer (Meta AI), Chenliang Xu (University of Rochester), and Mark Ibrahim (Meta AI)</i>	
Independent Component Alignment for Multi-Task Learning	20083
<i>Dmitry Senushkin (Samsung Research), Nikolay Patakin (Samsung Research), Arseny Kuznetsov (Samsung Research), and Anton Konushin (Samsung Research)</i>	
MDL-NAS: A Joint Multi-Domain Learning Framework for Vision Transformer	20094
<i>Shiguang Wang (University of Electronic Science and Technology of China; SenseTime Research), Tao Xie (Harbin Institute of Technology; SenseTime Research), Jian Cheng (University of Electronic Science and Technology of China), Xingcheng Zhang (SenseTime Research), and Haijun Liu (Chongqing University)</i>	
MELTR: Meta Loss Transformer for Learning To Fine-Tune Video Foundation Models	20105
<i>Dohwan Ko (Korea University), Joonmyung Choi (Korea University), Hyeong Kyu Choi (Korea University), Kyoung-Woon On (Kakao Brain), Byungseok Roh (Kakao Brain), and Hyunwoo J. Kim (Kakao Brain)</i>	

1% VS 100%: Parameter-Efficient Low Rank Adapter for Dense Predictions	20116
<i>Dongshuo Yin (Chinese Academy of Sciences; University of Chinese Academy of Sciences), Yiran Yang (Chinese Academy of Sciences; University of Chinese Academy of Sciences), Zhechao Wang (Chinese Academy of Sciences; University of Chinese Academy of Sciences), Hongfeng Yu (Chinese Academy of Sciences), Kaiwen Wei (Chinese Academy of Sciences; University of Chinese Academy of Sciences), and Xian Sun (Chinese Academy of Sciences; University of Chinese Academy of Sciences)</i>	
Rebalancing Batch Normalization for Exemplar-Based Class-Incremental Learning	20127
<i>Sungmin Cha (Seoul National University), Sungjun Cho (LG AI Research), Dasol Hwang (LG AI Research), Sunwon Hong (Seoul National University), Moontae Lee (LG AI Research; University of Illinois Chicago), and Taesup Moon (Seoul National University)</i>	
Partial Network Cloning	20137
<i>Jingwen Ye (National University of Singapore), Songhua Liu (National University of Singapore), and Xinchao Wang (National University of Singapore)</i>	
ERM-KTP: Knowledge-Level Machine Unlearning via Knowledge Transfer	20147
<i>Shen Lin (Xidian University), Xiaoyu Zhang (Xidian University), Chenyang Chen (Xidian University), Xiaofeng Chen (Xidian University), and Willy Susilo (University of Wollongong)</i>	
Rethinking Feature-Based Knowledge Distillation for Face Recognition	20156
<i>Jingzhi Li (Samsung R&D Institute China Xi'an (SRCX)), Zidong Guo (Samsung R&D Institute China Xi'an (SRCX)), Hui Li (Samsung R&D Institute China Xi'an (SRCX)), Seungju Han (Samsung Advanced Institute of Technology (SAIT), South Korea), Ji-won Baek (Samsung Advanced Institute of Technology (SAIT), South Korea), Min Yang (Samsung R&D Institute China Xi'an (SRCX)), Ran Yang (Samsung R&D Institute China Xi'an (SRCX)), and Sungjoo Suh (Samsung Advanced Institute of Technology (SAIT), South Korea)</i>	
Regularizing Second-Order Influences for Continual Learning	20166
<i>Zhicheng Sun (Peking University), Yadong Mu (Peking University; Peng Cheng Laboratory), and Gang Hua (Wormpex AI Research)</i>	
Generalization Matters: Loss Minima Flattening via Parameter Hybridization for Efficient Online Knowledge Distillation	20176
<i>Tianli Zhang (Zhejiang University), Mengqi Xue (Hangzhou City University), Jiangtao Zhang (Zhejiang University), Haofei Zhang (Zhejiang University), Yu Wang (Zhejiang University), Lechao Cheng (Zhejiang Lab), Jie Song (Zhejiang University), and Mingli Song (Zhejiang University)</i>	
Decoupling Learning and Remembering: A Bilevel Memory Framework With Knowledge Projection for Task-Incremental Learning	20186
<i>Wenju Sun (Beijing Jiaotong University), Qingyong Li (Beijing Jiaotong University; Wuyi University), Jing Zhang (Beijing Jiaotong University), Wen Wang (Beijing Jiaotong University), and Yangli-ao Geng (Beijing Jiaotong University; Wuyi University)</i>	

On the Stability-Plasticity Dilemma of Class-Incremental Learning	20196
<i>Dongwan Kim (Computer Vision Laboratory, ECE1 & IPAI) and Bohyung Han (Computer Vision Laboratory, ECE1 & IPAI; Seoul National University)</i>	
Simulated Annealing in Early Layers Leads to Better Generalization	20205
<i>Amir M. Sarfi (Concordia University; Mila - Quebec AI Institute), Zahra Karimpour (Concordia University), Muawiz Chaudhary (Concordia University; Mila - Quebec AI Institute), Nasir M. Khalid (Concordia University; Mila - Quebec AI Institute), Mirco Ravanelli (Concordia University; Mila - Quebec AI Institute), Sudhir Mudur (Concordia University), and Eugene Belilovsky (Concordia University; Mila - Quebec AI Institute)</i>	
Frustratingly Easy Regularization on Representation Can Boost Deep Reinforcement Learning	20215
<i>Qiang He (Institute of Automation, Chinese Academy of Sciences, China), Huangyuan Su (Carnegie Mellon University, United States), Jieyu Zhang (University of Washington, United States), and Xinwen Hou (Institute of Automation, Chinese Academy of Sciences, China)</i>	
Tunable Convolutions With Parametric Multi-Loss Optimization	20226
<i>Matteo Maggioni (Huawei Noah's Ark Lab), Thomas Tanay (Huawei Noah's Ark Lab), Francesca Babiloni (Huawei Noah's Ark Lab), Steven McDonagh (Huawei Noah's Ark Lab), and Aleš Leonardis (Huawei Noah's Ark Lab)</i>	
Re-Basin via Implicit Sinkhorn Differentiation	20237
<i>Fidel A. Guerrero Peña (LIVIA, Dept. of Systems Engineering ETS Montreal, Canada), Heitor Rapela Medeiros (LIVIA, Dept. of Systems Engineering ETS Montreal, Canada), Thomas Dubail (LIVIA, Dept. of Systems Engineering ETS Montreal, Canada), Masih Aminbeidokhti (LIVIA, Dept. of Systems Engineering ETS Montreal, Canada), Eric Granger (LIVIA, Dept. of Systems Engineering ETS Montreal, Canada), and Marco Pedersoli (LIVIA, Dept. of Systems Engineering ETS Montreal, Canada)</i>	
Gradient Norm Aware Minimization Seeks First-Order Flatness and Improves Generalization	20247
<i>Xingxuan Zhang (Tsinghua University), Renzhe Xu (Tsinghua University), Han Yu (Tsinghua University), Hao Zou (Tsinghua University), and Peng Cui (Tsinghua University)</i>	
AstroNet: When Astrocyte Meets Artificial Neural Network	20258
<i>Mengqiao Han (Beijing Institute of Technology), Liyuan Pan (Beijing Institute of Technology), and Xiabi Liu (Beijing Institute of Technology)</i>	
Network Expansion for Practical Training Acceleration	20269
<i>Ning Ding (Peking University; Huawei Noah's Ark Lab), Yehui Tang (Peking University; Huawei Noah's Ark Lab), Kai Han (Huawei Noah's Ark Lab), Chao Xu (Peking University), and Yunhe Wang (Huawei Noah's Ark Lab)</i>	
Defining and Quantifying the Emergence of Sparse Concepts in DNNs	20280
<i>Jie Ren (Shanghai Jiao Tong University), Mingjie Li (Shanghai Jiao Tong University), Qirui Chen (Shanghai Jiao Tong University), Huiqi Deng (Shanghai Jiao Tong University), and Quanshi Zhang (Shanghai Jiao Tong University)</i>	

Samples With Low Loss Curvature Improve Data Efficiency	20290
<i>Isha Garg (Purdue University West Lafayette, USA) and Kaushik Roy (Purdue University West Lafayette, USA)</i>	
Masked Images Are Counterfactual Samples for Robust Fine-Tuning	20301
<i>Yao Xiao (Sun Yat-sen University), Ziyi Tang (Sun Yat-sen University), Pengxu Wei (Sun Yat-sen University), Cong Liu (Sun Yat-sen University), and Liang Lin (Sun Yat-sen University)</i>	
Bias Mimicking: A Simple Sampling Approach for Bias Mitigation	20311
<i>Maan Qraitem (Boston University), Kate Saenko (Boston University; MIT-IBM Watson AI Lab), and Bryan A. Plummer (Boston University)</i>	
NoisyQuant: Noisy Bias-Enhanced Post-Training Activation Quantization for Vision Transformers	20321
<i>Yijiang Liu (Nanjing University), Huanrui Yang (University of California, Berkeley), Zhen Dong (University of California, Berkeley), Kurt Keutzer (University of California, Berkeley), Li Du (Nanjing University), and Shanghang Zhang (Peking University)</i>	
Practical Network Acceleration With Tiny Sets	20331
<i>Guo-Hua Wang (State Key Laboratory for Novel Software Technology Nanjing University, China) and Jianxin Wu (State Key Laboratory for Novel Software Technology Nanjing University, China)</i>	
TeSLA: Test-Time Self-Learning With Automatic Adversarial Augmentation	20341
<i>Devavrat Tomar (EPFL), Guillaume Vray (EPFL), Behzad Bozorgtabar (CHUV), and Jean-Philippe Thiran (EPFL; CHUV)</i>	
Discriminator-Cooperated Feature Map Distillation for GAN Compression	20351
<i>Tie Hu (Xiamen University), Mingbao Lin (Tencent Youtu Lab), Lizhou You (Xiamen University), Fei Chao (Xiamen University), and Rongrong Ji (Xiamen University; Xiamen University; Shenzhen Research Institute of Xiamen University)</i>	
Private Image Generation With Dual-Purpose Auxiliary Classifier	20361
<i>Chen Chen (School of Computer Science, Faculty of Engineering, The University of Sydney), Daochang Liu (School of Computer Science, Faculty of Engineering, The University of Sydney), Siqi Ma (The University of New South Wales), Surya Nepal (CSIRO), and Chang Xu (School of Computer Science, Faculty of Engineering, The University of Sydney)</i>	
ImageNet-E: Benchmarking Neural Network Robustness via Attribute Editing	20371
<i>Xiaodan Li (Alibaba Group), Yuefeng Chen (Alibaba Group), Yao Zhu (Zhejiang University), Shuhui Wang (Inst. of Comput. Tech., CAS, China), Rong Zhang (Alibaba Group), and Hui Xue (Alibaba Group)</i>	
Masked Jigsaw Puzzle: A Versatile Position Embedding for Vision Transformers	20382
<i>Bin Ren (University of Pisa, Italy; University of Trento, Italy), Yahui Liu (University of Trento, Italy), Yue Song (University of Trento, Italy), Wei Bi (Tencent AI Lab, China), Rita Cucchiara (University of Modena and Reggio Emilia, Italy), Nicu Sebe (University of Trento, Italy), and Wei Wang (Beijing Jiaotong University, China)</i>	

A New Comprehensive Benchmark for Semi-Supervised Video Anomaly Detection and Anticipation.....	20392
<i>Congqi Cao (Northwestern Polytechnical University, China), Yue Lu (Northwestern Polytechnical University, China), Peng Wang (Northwestern Polytechnical University, China), and Yanning Zhang (Northwestern Polytechnical University, China)</i>	
SimpleNet: A Simple Network for Image Anomaly Detection and Localization	20402
<i>Zhikang Liu (University of Science and Technology of China), Yiming Zhou (Meka Technology Co., Ltd), Yuansheng Xu (Meka Technology Co., Ltd), and Zilei Wang (University of Science and Technology of China)</i>	
DaFKD: Domain-Aware Federated Knowledge Distillation	20412
<i>Haozhao Wang (Huazhong University of Science and Technology, China), Yichen Li (Huazhong University of Science and Technology, China; The Hong Kong Polytechnic University, Hong Kong), Wenchao Xu (Beijing Institute of Technology, China), Ruixuan Li (Huazhong University of Science and Technology, China), Yufeng Zhan (n/a), and Zhigang Zeng (Soochow University, China)</i>	
Reliable and Interpretable Personalized Federated Learning	20422
<i>Zixuan Qin (Tianjin University, China), Liu Yang (Tianjin University, China), Qilong Wang (Tianjin University, China), Yahong Han (Tianjin University, China), and Qinghua Hu (Tianjin University, China)</i>	
Adaptive Channel Sparsity for Federated Learning Under System Heterogeneity	20432
<i>Dongping Liao (CIS Dept, University of Macau, Macau SAR, China), Xitong Gao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China), Yiren Zhao (Imperial College London, UK), and Cheng-Zhong Xu (CIS Dept, University of Macau, Macau SAR, China)</i>	
Bias-Eliminating Augmentation Learning for Debaised Federated Learning	20442
<i>Yuan-Yi Xu (National Taiwan University, Taiwan), Ci-Siang Lin (National Taiwan University, Taiwan), and Yu-Chiang Frank Wang (National Taiwan University, Taiwan; NVIDIA, Taiwan)</i>	
Instance-Aware Domain Generalization for Face Anti-Spoofing	20453
<i>Qianyu Zhou (Shanghai Jiao Tong University; Shanghai Key Laboratory of Computer Software Evaluating and Testing), Ke-Yue Zhang (Youtu Lab, Tencent), Taiping Yao (Youtu Lab, Tencent), Xuequan Lu (Deakin University), Ran Yi (Shanghai Jiao Tong University), Shouhong Ding (Youtu Lab, Tencent; Deakin University), and Lizhuang Ma (Shanghai Jiao Tong University; Deakin University)</i>	
Adversarially Masking Synthetic To Mimic Real: Adaptive Noise Injection for Point Cloud Segmentation Adaptation	20464
<i>Guangrui Li (University of Technology Sydney; Baidu Inc.), Guoliang Kang (Beihang University), Xiaohan Wang (Zhejiang University), Yunchao Wei (Beijing Jiaotong University), and Yi Yang (Zhejiang University)</i>	

Model Barrier: A Compact Un-Transferable Isolation Domain for Model Intellectual Property Protection	20475
<i>Lianyu Wang (Nanjing University of Aeronautics and Astronautics, China; Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore), Meng Wang (Nanjing University of Aeronautics and Astronautics, China; Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore), Daoqiang Zhang (Nanjing University of Aeronautics and Astronautics, China; Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore), and Huazhu Fu (Nanjing University of Aeronautics and Astronautics, China; Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore)</i>	
MEDIC: Remove Model Backdoors via Importance Driven Cloning	20485
<i>Qiuling Xu (Purdue University), Guanhong Tao (Purdue University), Jean Honorio (Purdue University), Yingqi Liu (Purdue University), Shengwei An (Purdue University), Guangyu Shen (Purdue University), Siyuan Cheng (Purdue University), and Xiangyu Zhang (Purdue University)</i>	
Progressive Backdoor Erasing via Connecting Backdoor and Adversarial Attacks	20495
<i>Bingxu Mu (Xi'an Jiaotong University), Zhenxing Niu (Xidian University), Le Wang (IAIR, Xi'an Jiaotong University), Xue Wang (Alibaba Group), Qiguang Miao (Xidian University), Rong Jin (Alibaba Group), and Gang Hua (Wormpex AI Research)</i>	
Reinforcement Learning-Based Black-Box Model Inversion Attacks	20504
<i>Gyojin Han (KAIST), Jaehyun Choi (KAIST), Haeil Lee (KAIST), and Junmo Kim (KAIST)</i>	
T-SEA: Transfer-Based Self-Ensemble Attack on Object Detection	20514
<i>Hao Huang (Peking University Beijing, China), Ziyang Chen (Peking University Beijing, China), Huanran Chen (Peking University Beijing, China), Yongtao Wang (Peking University Beijing, China), and Kevin Zhang (Peking University Beijing, China)</i>	
Proximal Splitting Adversarial Attack for Semantic Segmentation	20524
<i>Jérôme Rony (ÉTS Montréal), Jean-Christophe Pesquet (Université Paris-Saclay, CentraleSupélec, Inria), and Ismail Ben Ayed (ÉTS Montréal)</i>	
Towards Transferable Targeted Adversarial Examples	20534
<i>Zhibo Wang (Zhejiang University, P. R. China; ZJU-Hangzhou Global Scientific and Technological Innovation Center), Hongshan Yang (Zhejiang University, P. R. China), Yunhe Feng (Zhejiang University, P. R. China), Peng Sun (Hunan University, P. R. China), Hengchang Guo (Wuhan University, P. R. China), Zhifei Zhang (Adobe Research), and Kui Ren (Zhejiang University, P. R. China)</i>	

AGAIN: Adversarial Training With Attribution Span Enlargement and Hybrid Feature Fusion	20544
<i>Shenglin Yin (Peking University, China), Kelu Yao (Zhejiang Laboratory, China; Institute of Computing Technology, Chinese Academy of Sciences, China), Sheng Shi (Northwest University, P. R. China; AI Lab, Lenovo Research, P. R. China), Yangzhou Du (AI Lab, Lenovo Research, P. R. China), and Zhen Xiao (Peking University, China)</i>	
Generalist: Decoupling Natural and Robust Generalization	20554
<i>Hongjun Wang (Peking University) and Yisen Wang (Peking University)</i>	
Cooperation or Competition: Avoiding Player Domination for Multi-Target Robustness via Adaptive Budgets	20564
<i>Yimu Wang (University of Waterloo, Canada), Dinghuai Zhang (University of Montreal Montreal, Canada), Yihan Wu (University of Pittsburgh, United States), Heng Huang (University of Pittsburgh, United States), and Hongyang Zhang (University of Waterloo, Canada)</i>	
Discrete Point-Wise Attack Is Not Enough: Generalized Manifold Adversarial Attack for Face Recognition	20575
<i>Qian Li (Eastern Institute for Advanced Study, China), Yuxiao Hu (Eastern Institute for Advanced Study, China), Ye Liu (Eastern Institute for Advanced Study, China), Dongxiao Zhang (Eastern Institute for Advanced Study, China), Xin Jin (Eastern Institute for Advanced Study, China), and Yuntian Chen (Eastern Institute for Advanced Study, China)</i>	
RIATIG: Reliable and Imperceptible Adversarial Text-to-Image Generation With Natural Prompts	20585
<i>Han Liu (Washington University in St. Louis), Yuhao Wu (Washington University in St. Louis), Shixuan Zhai (Washington University in St. Louis), Bo Yuan (Rutgers University), and Ning Zhang (Washington University in St. Louis)</i>	
CLIP2Protect: Protecting Facial Privacy Using Text-Guided Makeup via Adversarial Latent Search	20595
<i>Fahad Shamshad (Mohamed Bin Zayed University of AI, UAE), Muzammal Naseer (Mohamed Bin Zayed University of AI, UAE), and Karthik Nandakumar (Mohamed Bin Zayed University of AI, UAE)</i>	
TruFor: Leveraging All-Round Clues for Trustworthy Image Forgery Detection and Localization	20606
<i>Fabrizio Guillaro (University Federico II of Naples), Davide Cozzolino (University Federico II of Naples), Avneesh Sud (Google Research), Nicholas Dufour (Google Research), and Luisa Verdoliva (University Federico II of Naples)</i>	

Poster-Thu-PM

High-Fidelity Event-Radiance Recovery via Transient Event Frequency	20616
<i>Jin Han (The University of Tokyo; National Institute of Informatics), Yuta Asano (National Institute of Informatics), Boxin Shi (Peking University), Yinqiang Zheng (The University of Tokyo), and Imari Sato (The University of Tokyo; National Institute of Informatics)</i>	

RobustNeRF: Ignoring Distractors With Robust Losses	20626
<i>Sara Sabour (Google Research, Brain Team; University of Toronto), Suhani Vora (Google Research, Brain Team), Daniel Duckworth (Google Research, Brain Team), Ivan Krasin (Google Research, Brain Team), David J. Fleet (Google Research, Brain Team; University of Toronto), and Andrea Tagliasacchi (Google Research, Brain Team; University of Toronto; Simon Fraser University)</i>	
NeRDi: Single-View NeRF Synthesis With Language-Guided Diffusion As General Image Priors	20637
<i>Congyue Deng (Stanford University), Chiyu Max Jiang (Waymo), Charles R. Qi (Waymo), Xinchun Yan (Waymo), Yin Zhou (Waymo), Leonidas Guibas (Stanford University; Google Research), and Dragomir Anguelov (Waymo)</i>	
GM-NeRF: Learning Generalizable Model-Based Neural Radiance Fields From Multi-View Images ..	20648
<i>Jianchuan Chen (Dalian University of Technology, China), Wentao Yi (Dalian University of Technology, China), Liqian Ma (ZMO AI Inc.), Xu Jia (Dalian University of Technology, China), and Huchuan Lu (Dalian University of Technology, China)</i>	
MixNeRF: Modeling a Ray With Mixture Density for Novel View Synthesis From Sparse Inputs	20659
<i>Seunghyeon Seo (Seoul National Univeristy), Donghoon Han (Seoul National Univeristy), Yeonjin Chang (Seoul National Univeristy), and Nojun Kwak (Seoul National Univeristy)</i>	
SPIn-NeRF: Multiview Segmentation and Perceptual Inpainting With Neural Radiance Fields	20669
<i>Ashkan Mirzaei (Samsung AI Centre Toronto; University of Toronto), Tristan Aumentado-Armstrong (Samsung AI Centre Toronto; University of Toronto; Vector Institute for AI), Konstantinos G. Derpanis (Samsung AI Centre Toronto; York University; Vector Institute for AI), Jonathan Kelly (University of Toronto), Marcus A. Brubaker (Samsung AI Centre Toronto; York University; Vector Institute for AI), Igor Gilitschenski (University of Toronto), and Alex Levinshtein (Samsung AI Centre Toronto)</i>	
Masked Wavelet Representation for Compact Neural Radiance Fields	20680
<i>Daniel Rho (AI2XL, KT), Byeonghyeon Lee (Sungkyunkwan University), Seungtae Nam (Sungkyunkwan University), Joo Chan Lee (Sungkyunkwan University), Jong Hwan Ko (Sungkyunkwan University), and Eunbyung Park (Sungkyunkwan University)</i>	
PaletteNeRF: Palette-Based Appearance Editing of Neural Radiance Fields	20691
<i>Zhengfei Kuang (Stanford University), Fujun Luan (Adobe Research), Sai Bi (Adobe Research), Zhixin Shu (Adobe Research), Gordon Wetzstein (Stanford University), and Kalyan Sunkavalli (Stanford University)</i>	
SteerNeRF: Accelerating NeRF Rendering via Smooth Viewpoint Trajectory	20701
<i>Sicheng Li (Zhejiang University), Hao Li (Zhejiang University), Yue Wang (Zhejiang University), Yiyi Liao (Zhejiang University), and Lu Yu (Zhejiang University)</i>	

Transforming Radiance Field With Lipschitz Network for Photorealistic 3D Scene Stylization	20712
<i>Zicheng Zhang (University of Chinese Academy of Sciences), Yinglu Liu (JD AI Research), Congying Han (University of Chinese Academy of Sciences), Yingwei Pan (HiDream.ai Inc.), Tiande Guo (University of Chinese Academy of Sciences), and Ting Yao (HiDream.ai Inc.)</i>	
Occlusion-Free Scene Recovery via Neural Radiance Fields	20722
<i>Chengxuan Zhu (Peking University), Renjie Wan (Hong Kong Baptist University), Yunkai Tang (Peking University), and Boxin Shi (Peking University)</i>	
TriVol: Point Cloud Rendering via Triple Volumes	20732
<i>Tao Hu (The Chinese University of Hong Kong), Xiaogang Xu (The Chinese University of Hong Kong), Ruihang Chu (The Chinese University of Hong Kong), and Jiaya Jia (The Chinese University of Hong Kong; SmartMore)</i>	
DyNCA: Real-Time Dynamic Texture Synthesis Using Neural Cellular Automata	20742
<i>Ehsan Pajouheshgar (School of Computer and Communication Sciences, EPFL, Switzerland), Yitao Xu (School of Computer and Communication Sciences, EPFL, Switzerland), Tong Zhang (School of Computer and Communication Sciences, EPFL, Switzerland), and Sabine Ssstrunk (School of Computer and Communication Sciences, EPFL, Switzerland)</i>	
Neural Scene Chronology	20752
<i>Haotong Lin (Zhejiang University; Cornell University), Qianqian Wang (Cornell University), Ruojin Cai (Cornell University), Sida Peng (Zhejiang University), Hadar Averbuch-Elor (Tel Aviv University), Xiaowei Zhou (Zhejiang University), and Noah Snavely (Cornell University)</i>	
ReLight My NeRF: A Dataset for Novel View Synthesis and Relighting of Real World Objects .	20762
<i>Marco Toschi (Eyecan.ai), Riccardo De Matteo (Eyecan.ai), Riccardo Spezialetti (Eyecan.ai), Daniele De Gregorio (Eyecan.ai), Luigi Di Stefano (University of Bologna), and Samuele Salti (University of Bologna)</i>	
ORCa: Glossy Objects As Radiance-Field Cameras	20773
<i>Kushagra Tiwary (Massachusetts Institute ofTechnology), Akshat Dave (Rice University), Nikhil Behari (Massachusetts Institute ofTechnology), Tzofi Klinghoffer (Massachusetts Institute ofTechnology), Ashok Veeraraghavan (Rice University), and Ramesh Raskar (Massachusetts Institute ofTechnology)</i>	
Nighttime Smartphone Reflective Flare Removal Using Optical Center Symmetry Prior	20783
<i>Yuekun Dai (Nanyang Technological University), Yihang Luo (Nanyang Technological University), Shangchen Zhou (Nanyang Technological University), Chongyi Li (Nanyang Technological University), and Chen Change Loy (Nanyang Technological University)</i>	
SunStage: Portrait Reconstruction and Relighting Using the Sun as a Light Stage	20792
<i>Yifan Wang (University of Washington), Aleksander Holynski (University of Washington), Xiuming Zhang (Adobe Inc.), and Xuaner Zhang (Adobe Inc.)</i>	

The Differentiable Lens: Compound Lens Search Over Glass Surfaces and Materials for Object Detection	20803
<i>Geoffroi Côté (Universite Laval; Princeton University), Fahim Mannan (Algolux), Simon Thibault (Universite Laval), Jean-François Lalonde (Universite Laval), and Felix Heide (Princeton University; Algolux)</i>	
Teleidoscopic Imaging System for Microscale 3D Shape Reconstruction	20813
<i>Ryo Kawahara (Kyushu Institute of Technology), Meng-Yu Jennifer Kuo (University of Minnesota), and Shohei Nobuhara (Kyoto University)</i>	
Looking Through the Glass: Neural Surface Reconstruction Against High Specular Reflections.....	20823
<i>Jiaxiong Qiu (VCIP, CS, Nankai University), Peng-Tao Jiang (Zhejiang University), Yifan Zhu (VCIP, CS, Nankai University), Ze-Xin Yin (VCIP, CS, Nankai University), Ming-Ming Cheng (VCIP, CS, Nankai University), and Bo Ren (VCIP, CS, Nankai University)</i>	
NeuralUDF: Learning Unsigned Distance Fields for Multi-View Reconstruction of Surfaces With Arbitrary Topologies	20834
<i>Xiaoxiao Long (The University of Hong Kong; Tencent Games), Cheng Lin (Tencent Games), Lingjie Liu (Max Planck Institute for Informatics), Yuan Liu (The University of Hong Kong), Peng Wang (The University of Hong Kong), Christian Theobalt (Max Planck Institute for Informatics), Taku Komura (The University of Hong Kong), and Wenping Wang (Texas A&M University)</i>	
Sphere-Guided Training of Neural Implicit Surfaces	20844
<i>Andreea Dogaru (Skolkovo Institute of Science and Technology; Friedrich-Alexander-Universität Erlangen-Nürnberg), Andrei-Timotei Ardelean (Skolkovo Institute of Science and Technology; Friedrich-Alexander-Universität Erlangen-Nürnberg), Savva Ignatyev (Skolkovo Institute of Science and Technology), Egor Zakharov (Skolkovo Institute of Science and Technology), and Evgeny Burnaev (Skolkovo Institute of Science and Technology; Skolkovo Institute of Science and Technology; Artificial Intelligence Research Institute)</i>	
OReX: Object Reconstruction From Planar Cross-Sections Using Neural Fields	20854
<i>Haim Sawdayee (Tel Aviv University), Amir Vaxman (The University of Edinburgh), and Amit H. Bermano (Tel Aviv University)</i>	
Persistent Nature: A Generative Model of Unbounded 3D Worlds	20863
<i>Lucy Chai (MIT), Richard Tucker (Google Research), Zhengqi Li (Google Research), Phillip Isola (MIT), and Noah Snavely (Google Research; Cornell Tech)</i>	
3D Neural Field Generation Using Triplane Diffusion	20875
<i>J. Ryan Shue (Milton Academy), Eric Ryan Chan (Stanford University), Ryan Po (Stanford University), Zachary Ankner (Massachusetts Institute of Technology; MosaicML), Jiajun Wu (Stanford University), and Gordon Wetzstein (Stanford University)</i>	
Diffusion-Based Signed Distance Fields for 3D Shape Generation	20887
<i>Jaehyeok Shim (Artificial Intelligence Graduate School, UNIST), Changwoo Kang (Artificial Intelligence Graduate School, UNIST), and Kyungdon Joo (Artificial Intelligence Graduate School, UNIST)</i>	

Efficient View Synthesis and 3D-Based Multi-Frame Denoising With Multiplane Feature Representations	20898
<i>Thomas Tanay (Huawei Noah's Ark Lab), Aleš Leonardis (Huawei Noah's Ark Lab), and Matteo Maggioni (Huawei Noah's Ark Lab)</i>	
Dream3D: Zero-Shot Text-to-3D Synthesis Using 3D Shape Prior and Text-to-Image Diffusion Models	20908
<i>Jiale Xu (ARC Lab; ShanghaiTech University), Xintao Wang (ARC Lab), Weihao Cheng (ARC Lab), Yan-Pei Cao (ARC Lab), Ying Shan (ARC Lab), Xiaohu Qie (Tencent PCG), and Shenghua Gao (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging; Shanghai Engineering Research Center of Energy Efficient and Custom AI IC)</i>	
SINE: Semantic-Driven Image-Based NeRF Editing With Prior-Guided Editing Field	20919
<i>Chong Bao (Zhejiang University), Yinda Zhang (Google), Bangbang Yang (Zhejiang University), Tianxing Fan (Zhejiang University), Zesong Yang (Zhejiang University), Hujun Bao (Zhejiang University), Guofeng Zhang (Zhejiang University), and Zhaopeng Cui (Zhejiang University)</i>	
3D Highlighter: Localizing Regions on 3D Shapes via Text Descriptions	20930
<i>Dale Decatur (University of Chicago), Itai Lang (University of Chicago), and Rana Hanocka (University of Chicago)</i>	
Self-Supervised Geometry-Aware Encoder for Style-Based 3D GAN Inversion	20940
<i>Yushi Lan (Nanyang Technological University, Singapore), Xuyi Meng (Nanyang Technological University, Singapore), Shuai Yang (Nanyang Technological University, Singapore), Chen Change Loy (Nanyang Technological University, Singapore), and Bo Dai (Shanghai AI Laboratory)</i>	
PanoHead: Geometry-Aware 3D Full-Head Synthesis in 360°	20950
<i>Sizhe An (ByteDance Inc.; University of Wisconsin-Madison), Hongyi Xu (ByteDance Inc.), Yichun Shi (ByteDance Inc.), Guoxian Song (ByteDance Inc.), Umit Y. Ogras (University of Wisconsin-Madison), and Linjie Luo (ByteDance Inc.)</i>	
StyleGene: Crossover and Mutation of Region-Level Facial Genes for Kinship Face Synthesis	20960
<i>Hao Li (Computer Vision Institute, Shenzhen University), Xianxu Hou (National Engineering Laboratory for Big Data System Computing Technology, Shenzhen University; Xi'an Jiaotong-Liverpool University), Zepeng Huang (Computer Vision Institute, Shenzhen University), and Linlin Shen (Computer Vision Institute, Shenzhen University; National Engineering Laboratory for Big Data System Computing Technology, Shenzhen University; Shenzhen Institute of Artificial Intelligence and Robotics for Society; Guangdong Key Laboratory of Intelligent Information Processing, Shenzhen University)</i>	
Parameter Efficient Local Implicit Image Function Network for Face Segmentation	20970
<i>Mausoom Sarkar (MDSR Lab, Adobe), Nikitha SR (IIT Madras), Mayur Hemani (MDSR Lab, Adobe), Rishabh Jain (MDSR Lab, Adobe), and Balaji Krishnamurthy (MDSR Lab, Adobe)</i>	

Graphics Capsule: Learning Hierarchical 3D Face Representations From 2D Images	20981
<i>Chang Yu (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Xiangyu Zhu (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Xiaomei Zhang (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Zhaoxiang Zhang (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Centre for Artificial Intelligence and Robotics, Hong Kong Institute of Science & Innovation, Chinese Academy of Sciences), and Zhen Lei (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Centre for Artificial Intelligence and Robotics, Hong Kong Institute of Science & Innovation, Chinese Academy of Sciences)</i>	
Next3D: Generative Neural Texture Rasterization for 3D-Aware Head Avatars	20991
<i>Jingxiang Sun (Tsinghua University), Xuan Wang (Ant Group), Lizhen Wang (Tsinghua University; NNKosmos), Xiaoyu Li (Tencent AI Lab), Yong Zhang (Tencent AI Lab), Hongwen Zhang (Tsinghua University), and Yebin Liu (Tsinghua University)</i>	
Learning Neural Parametric Head Models	21003
<i>Simon Giebenhain (Technical University of Munich), Tobias Kirschstein (Technical University of Munich), Markos Georgopoulos (Synthesia), Martin Rünz (Synthesia), Lourdes Agapito (University College London), and Matthias Nießner (Synthesia)</i>	
Zero-Shot Text-to-Parameter Translation for Game Character Auto-Creation	21013
<i>Rui Zhao (Netease Fuxi AI Lab), Wei Li (Nankai University), Zhipeng Hu (Netease Fuxi AI Lab), Lincheng Li (Netease Fuxi AI Lab), Zhengxia Zou (Beihang University), Zhenwei Shi (Beihang University), and Changjie Fan (Netease Fuxi AI Lab)</i>	
Learning Locally Editable Virtual Humans	21024
<i>Hsuan-I Ho (Department of Computer Science, ETH Zürich), Lixin Xue (Department of Computer Science, ETH Zürich), Jie Song (Department of Computer Science, ETH Zürich), and Otmar Hilliges (Department of Computer Science, ETH Zürich)</i>	
Auto-CARD: Efficient and Robust Codec Avatar Driving for Real-Time Mobile Telepresence ..	21036
<i>Yonggan Fu (Georgia Institute of Technology), Yuecheng Li (Meta), Chenghui Li (Meta), Jason Saragih (Meta), Peizhao Zhang (Meta), Xiaoliang Dai (Meta), and Yingyan Celine Lin (Georgia Institute of Technology)</i>	
Ham2Pose: Animating Sign Language Notation Into Pose Sequences	21046
<i>Rotem Shalev Arkushin (Reichman University), Amit Moryossef (Bar-Ilan University), and Ohad Fried (Reichman University)</i>	

- PointAvatar: Deformable Point-Based Head Avatars From Videos 21057
Yufeng Zheng (ETH Zurich; Max Planck Institute for Intelligent Systems), Wang Yifan (Stanford University), Gordon Wetzstein (Stanford University), Michael J. Black (Max Planck Institute for Intelligent Systems), and Otmar Hilliges (ETH Zurich)
- PAniC-3D: Stylized Single-View 3D Reconstruction From Portraits of Anime Characters 21068
Shuhong Chen (University of Maryland - College Park, USA; ByteDance), Kevin Zhang (ByteDance), Yichun Shi (University of Maryland - College Park, USA), Heng Wang (University of Maryland - College Park, USA), Yiheng Zhu (University of Maryland - College Park, USA), Guoxian Song (University of Maryland - College Park, USA), Sizhe An (University of Maryland - College Park, USA), Janus Kristjansson (ByteDance), Xiao Yang (University of Maryland - College Park, USA), and Matthias Zwicker (ByteDance)
- HandNeRF: Neural Radiance Fields for Animatable Interacting Hands 21078
Zhiyang Guo (University of Science and Technology of China), Wengang Zhou (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Min Wang (Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Li Li (University of Science and Technology of China), and Houqiang Li (University of Science and Technology of China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center)
- VGFlow: Visibility Guided Flow Network for Human Reposing 21088
Rishabh Jain (MDSR Adobe), Krishna Kumar Singh (Adobe Research), Mayur Hemani (MDSR Adobe), Jingwan Lu (Adobe Research), Mausoom Sarkar (MDSR Adobe), Duygu Ceylan (Adobe Research), and Balaji Krishnamurthy (MDSR Adobe)
- Clothed Human Performance Capture With a Double-Layer Neural Radiance Fields 21098
Kangkan Wang (Key Lab of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education; Nanjing University of Science and Technology, China), Guofeng Zhang (Zhejiang University, China), Suxu Cong (Key Lab of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education), and Jian Yang (Key Lab of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education; Nanjing University of Science and Technology, China)
- POEM: Reconstructing Hand in a Point Embedded Multi-View Stereo 21108
Lixin Yang (Shanghai Jiao Tong University; Shanghai Qi Zhi Institute), Jian Xu (Nreal), Licheng Zhong (Shanghai Jiao Tong University), Xinyu Zhan (Shanghai Jiao Tong University), Zhicheng Wang (Nreal), Kejian Wu (Nreal), and Cewu Lu (Shanghai Jiao Tong University; Shanghai Qi Zhi Institute)
- FlexNeRF: Photorealistic Free-Viewpoint Rendering of Moving Humans From Sparse Views . 21118
Vinoj Jayasundara (University of Maryland, College Park), Amit Agrawal (Amazon.com, Inc.), Nicolas Heron (Amazon.com, Inc.), Abhinav Shrivastava (University of Maryland, College Park), and Larry S. Davis (University of Maryland, College Park; Amazon.com, Inc.)

Flow Supervision for Deformable NeRF	21128
<i>Chaoyang Wang (Carnegie Mellon University), Lachlan Ewen MacDonald (University of Adelaide), László A. Jeni (Carnegie Mellon University), and Simon Lucey (Carnegie Mellon University; University of Adelaide)</i>	
Building Rearticulable Models for Arbitrary 3D Objects From 4D Point Clouds	21138
<i>Shaowei Liu (University of Illinois Urbana-Champaign), Saurabh Gupta (University of Illinois Urbana-Champaign), and Shenlong Wang (University of Illinois Urbana-Champaign)</i>	
Implicit 3D Human Mesh Recovery Using Consistency With Pose and Shape From Unseen-View	21148
<i>Hanbyel Cho (Korea Advanced Institute of Science and Technology (KAIST), South Korea), Yooshin Cho (Korea Advanced Institute of Science and Technology (KAIST), South Korea), Jaesung Ahn (Korea Advanced Institute of Science and Technology (KAIST), South Korea), and Junmo Kim (Korea Advanced Institute of Science and Technology (KAIST), South Korea)</i>	
One-Stage 3D Whole-Body Mesh Recovery With Component Aware Transformer	21159
<i>Jing Lin (International Digital Economy Academy (IDEA); Shenzhen International Graduate School, Tsinghua University), Ailing Zeng (International Digital Economy Academy (IDEA)), Haoqian Wang (Shenzhen International Graduate School, Tsinghua University), Lei Zhang (International Digital Economy Academy (IDEA)), and Yu Li (International Digital Economy Academy (IDEA))</i>	
Im2Hands: Learning Attentive Implicit Representation of Interacting Two-Hand Shapes	21169
<i>Jihyun Lee (KAIST), Minhyuk Sung (KAIST), Honggyu Choi (KAIST), and Tae-Kyun Kim (KAIST; Imperial College London)</i>	
FLEX: Full-Body Grasping Without Full-Body Grasps	21179
<i>Purva Tendulkar (Columbia University), Dídac Surís (Columbia University), and Carl Vondrick (Columbia University)</i>	
DexArt: Benchmarking Generalizable Dexterous Manipulation With Articulated Objects	21190
<i>Chen Bao (Shanghai Jiao Tong University), Helin Xu (Tsinghua University), Yuzhe Qin (UC San Diego), and Xiaolong Wang (UC San Diego)</i>	
CARTO: Category and Joint Agnostic Reconstruction of ARTiculated Objects	21201
<i>Nick Heppert (University of Freiburg; Toyota Research Institute (TRI)), Muhammad Zubair Irshad (Georgia Institute of Technology), Sergey Zakharov (Toyota Research Institute (TRI)), Katherine Liu (Toyota Research Institute (TRI)), Rares Andrei Ambrus (Toyota Research Institute (TRI)), Jeannette Bohg (Stanford University), Abhinav Valada (University of Freiburg), and Thomas Kollar (Toyota Research Institute (TRI))</i>	
CIRCLE: Capture in Rich Contextual Environments	21211
<i>João Pedro Araújo (Stanford University), Jiaman Li (Stanford University), Karthik Vetrivel (Stanford University), Rishi Agarwal (Stanford University), Jiajun Wu (Stanford University), Deepak Gopinath (Meta AI), Alexander William Clegg (Meta AI), and Karen Liu (Stanford University)</i>	

Decoupling Human and Camera Motion From Videos in the Wild	21222
<i>Vickie Ye (University of California, Berkeley), Georgios Pavlakos (University of California, Berkeley), Jitendra Malik (University of California, Berkeley), and Angjoo Kanazawa (University of California, Berkeley)</i>	
GarmentTracking: Category-Level Garment Pose Tracking	21233
<i>Han Xue (Shanghai Qi Zhi Institute; Shanghai Jiao Tong University), Wenqiang Xu (Shanghai Jiao Tong University), Jieyi Zhang (Shanghai Jiao Tong University), Tutian Tang (Shanghai Jiao Tong University), Yutong Li (Shanghai Jiao Tong University), Wenxin Du (Shanghai Jiao Tong University), Ruolin Ye (Cornell University), and Cewu Lu (Shanghai Jiao Tong University; Shanghai Jiao Tong University)</i>	
Hierarchical Temporal Transformer for 3D Hand Pose Estimation and Action Recognition From Egocentric RGB Videos	21243
<i>Yilin Wen (The University of Hong Kong), Hao Pan (Microsoft Research Asia), Lei Yang (The University of Hong Kong; TransGP), Jia Pan (The University of Hong Kong), Taku Komura (The University of Hong Kong), and Wenping Wang (Texas A&M University)</i>	
PSVT: End-to-End Multi-Person 3D Pose and Shape Estimation With Progressive Video Transformers	21254
<i>Zhongwei Qiu (University of Science and Technology Beijing; University of Sydney; Beijing Engineering Research Center of Industrial Spectrum Imaging), Qiansheng Yang (Baidu), Jian Wang (Baidu), Haocheng Feng (Baidu), Junyu Han (Baidu), Errui Ding (Baidu), Chang Xu (University of Sydney), Dongmei Fu (University of Science and Technology Beijing; Beijing Engineering Research Center of Industrial Spectrum Imaging), and Jingdong Wang (Baidu)</i>	
Delving Into Discrete Normalizing Flows on SO(3) Manifold for Probabilistic Rotation Modeling	21264
<i>Yulin Liu (Peking University), Haoran Liu (Peking University), Yingda Yin (Peking University), Yang Wang (Peking University), Baoquan Chen (Peking University), and He Wang (Peking University)</i>	

3D-POP – An Automated Annotation Approach to Facilitate Markerless 2D-3D Tracking of Freely Moving Birds With Marker-Based Motion Capture	21274
<i>Hemal Naik (Dept. of Collective Behavior and Dept. of Ecology of Animal Societies, Max Planck Institute of Animal Behavior; Dept. of Biology, University of Konstanz; Centre for the Advanced Study of Collective Behaviour, University of Konstanz; Informatik Department, Technische Universität München), Alex Hoi Hang Chan (Dept. of Collective Behavior and Dept. of Ecology of Animal Societies, Max Planck Institute of Animal Behavior; Dept. of Biology, University of Konstanz), Junran Yang (Dept. of Biology, University of Konstanz), Mathilde Delacoux (Dept. of Collective Behavior and Dept. of Ecology of Animal Societies, Max Planck Institute of Animal Behavior; Dept. of Biology, University of Konstanz), Iain D. Couzin (Dept. of Collective Behavior and Dept. of Ecology of Animal Societies, Max Planck Institute of Animal Behavior; Dept. of Biology, University of Konstanz; Centre for the Advanced Study of Collective Behaviour, University of Konstanz), Fumihiro Kano (Dept. of Collective Behavior and Dept. of Ecology of Animal Societies, Max Planck Institute of Animal Behavior; Dept. of Biology, University of Konstanz), and Máté Nagy (Dept. of Collective Behavior and Dept. of Ecology of Animal Societies, Max Planck Institute of Animal Behavior; Dept. of Biology, University of Konstanz; Centre for the Advanced Study of Collective Behaviour, University of Konstanz; Dept. of Biological Physics, Eotvos Lorand University; MTA-ELTE 'Lendület' Collective Behaviour Research Group, Hungarian Academy of Sciences)</i>	
TTA-COPE: Test-Time Adaptation for Category-Level Object Pose Estimation	21285
<i>Taeyeop Lee (KAIST), Jonathan Tremblay (NVIDIA), Valts Blukis (NVIDIA), Bowen Wen (NVIDIA), Byeong-Uk Lee (KAIST), Inkyu Shin (KAIST), Stan Birchfield (KAIST), In So Kweon (KAIST), and Kuk-Jin Yoon (KAIST)</i>	
Markerless Camera-to-Robot Pose Estimation via Self-Supervised Sim-to-Real Transfer	21296
<i>Jingpei Lu (University of California, San Diego), Florian Richter (University of California, San Diego), and Michael C. Yip (University of California, San Diego)</i>	
SMOC-Net: Leveraging Camera Pose for Self-Supervised Monocular Object Pose Estimation	21307
<i>Tao Tan (School of Artificial Intelligence, UCAS; State Key Laboratory of Multimodal Artificial Intelligence Systems, CASIA) and Qiulei Dong (School of Artificial Intelligence, UCAS; State Key Laboratory of Multimodal Artificial Intelligence Systems, CASIA; Center for Excellence in Brain Science and Intelligence Technology, CAS)</i>	
IMP: Iterative Matching and Pose Estimation With Adaptive Pooling	21317
<i>Fei Xue (University of Cambridge), Ignas Budvytis (University of Cambridge), and Roberto Cipolla (University of Cambridge)</i>	
Self-Supervised Representation Learning for CAD	21327
<i>Benjamin T. Jones (University of Washington), Michael Hu (University of Washington), Milin Kodnongbua (University of Washington), Vladimir G. Kim (Adobe Research), and Adriana Schulz (University of Washington)</i>	

Few-Shot Geometry-Aware Keypoint Localization	21337
<i>Xingzhe He (Flawless AI), Gaurav Bharaj (University of British Columbia), David Ferman (Flawless AI), Helge Rhodin (Flawless AI), and Pablo Garrido (University of British Columbia)</i>	
SparsePose: Sparse-View Camera Pose Regression and Refinement	21349
<i>Samarth Sinha (University of Toronto), Jason Y. Zhang (Carnegie Mellon University), Andrea Tagliasacchi (University of Toronto; Simon Fraser University; Google), Igor Gilitschenski (University of Toronto), and David B. Lindell (University of Toronto; Vector Institute)</i>	
A Large-Scale Homography Benchmark	21360
<i>Daniel Barath (Computer Vision and Geometry Group, ETH Zurich, Switzerland), Dmytro Mishkin (VRG, Faculty of Electrical Engineering, CTU in Prague, Czech Republic; HOVER Inc.), Michal Polic (VRG, Faculty of Electrical Engineering, CTU in Prague, Czech Republic; CIIRC, CTU in Prague, Czech Republic), Wolfgang Förstner (University Bonn), and Jiri Matas (VRG, Faculty of Electrical Engineering, CTU in Prague, Czech Republic)</i>	
Learning Geometric-Aware Properties in 2D Representation Using Lightweight CAD Models, or Zero Real 3D Pairs	21371
<i>Pattaramanee Arsomngern (VISTEC, Thailand), Sarana Nutanong (VISTEC, Thailand), and Supasorn Suwajanakorn (VISTEC, Thailand)</i>	
AutoRecon: Automated 3D Object Discovery and Reconstruction	21382
<i>Yuang Wang (Zhejiang University), Xingyi He (Zhejiang University), Sida Peng (Zhejiang University), Haotong Lin (Zhejiang University), Hujun Bao (Zhejiang University), and Xiaowei Zhou (Zhejiang University)</i>	
Multi-Sensor Large-Scale Dataset for Multi-View 3D Reconstruction	21392
<i>Oleg Voynov (Skolkovo Institute of Science and Technology; Artificial Intelligence Research Institute), Gleb Bobrovskikh (Skolkovo Institute of Science and Technology), Pavel Karpyshev (Skolkovo Institute of Science and Technology), Saveliy Galochkin (Skolkovo Institute of Science and Technology), Andrei-Timotei Ardelean (Skolkovo Institute of Science and Technology), Arseniy Bozhenko (Skolkovo Institute of Science and Technology), Ekaterina Karmanova (Skolkovo Institute of Science and Technology), Pavel Kopanev (Skolkovo Institute of Science and Technology), Yaroslav Labutin-Rymsho (Moscow Engineering Physics Institute), Ruslan Rakhimov (Skolkovo Institute of Science and Technology), Aleksandr Safin (Skolkovo Institute of Science and Technology), Valerii Serpiva (Skolkovo Institute of Science and Technology), Alexey Artemov (Technical University of Munich), Evgeny Burnaev (Skolkovo Institute of Science and Technology; Artificial Intelligence Research Institute), Dzmitry Tsetserukou (Skolkovo Institute of Science and Technology), and Denis Zorin (New York University)</i>	
NeurOCS: Neural NOCS Supervision for Monocular 3D Object Localization	21404
<i>Zhixiang Min (Stevens Institute of Technology), Bingbing Zhuang (NEC Laboratories America), Samuel Schuster (NEC Laboratories America), Buyu Liu (NEC Laboratories America), Enrique Dunn (Stevens Institute of Technology), and Manmohan Chandraker (NEC Laboratories America)</i>	

Self-Supervised Super-Plane for Neural 3D Reconstruction	21415
<i>Botao Ye (University of Chinese Academy of Sciences), Sifei Liu (NVIDIA), Xueting Li (NVIDIA), and Ming-Hsuan Yang (University of California, Merced; Yonsei University)</i>	
PlaneDepth: Self-Supervised Depth Estimation via Orthogonal Planes	21425
<i>Ruoyu Wang (ShanghaiTech University), Zehao Yu (University of Tübingen), and Shenghua Gao (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging; Shanghai Engineering Research Center of Energy Efficient and Custom AI IC)</i>	
Single View Scene Scale Estimation Using Scale Field	21435
<i>Byeong-Uk Lee (KAIST, Republic of Korea), Jianming Zhang (Adobe Inc., USA), Yannick Hold-Geoffroy (Adobe Inc., USA), and In So Kweon (KAIST, Republic of Korea)</i>	
3D Line Mapping Revisited	21445
<i>Shaohui Liu (Department of Computer Science, ETH Zurich), Yifan Yu (Department of Computer Science, ETH Zurich), Rémi Pautrat (Department of Computer Science, ETH Zurich), Marc Pollefeys (Department of Computer Science, ETH Zurich; Microsoft), and Viktor Larsson (Lund University)</i>	
Inverting the Imaging Process by Learning an Implicit Camera Model	21456
<i>Xin Huang (Northwestern Polytechnical University, China), Qi Zhang (Tencent AI Lab), Ying Feng (Tencent AI Lab), Hongdong Li (Australian National University), and Qing Wang (Northwestern Polytechnical University, China)</i>	
SfM-TTR: Using Structure From Motion for Test-Time Refinement of Single-View Depth Networks	21466
<i>Sergio Izquierdo (University of Zaragoza, Spain) and Javier Civera (University of Zaragoza, Spain)</i>	
iDisc: Internal Discretization for Monocular Depth Estimation	21477
<i>Luigi Piccinelli (Computer Vision Lab, ETH Zurich), Christos Sakaridis (Computer Vision Lab, ETH Zurich), and Fisher Yu (Computer Vision Lab, ETH Zurich)</i>	
DC2: Dual-Camera Defocus Control by Learning To Refocus	21488
<i>Hadi Alzayer (Google), Abdullah Abuolaim (Google), Leung Chun Chan (Google), Yang Yang (Google), Ying Chen Lou (Google), Jia-Bin Huang (University of Maryland, College Park), and Abhishek Kar (Google)</i>	
A Practical Stereo Depth System for Smart Glasses	21498
<i>Jialiang Wang (Meta Platforms Inc.), Daniel Scharstein (Meta Platforms Inc.), Akash Bapat (Meta Platforms Inc.), Kevin Blackburn-Matzen (Adobe), Matthew Yu (Meta Platforms Inc.), Jonathan Lehman (Meta Platforms Inc.), Suhub Alsisan (Meta Platforms Inc.), Yanghan Wang (Meta Platforms Inc.), Sam Tsai (Meta Platforms Inc.), Jan-Michael Frahm (Meta Platforms Inc.), Zijian He (Meta Platforms Inc.), Peter Vajda (Meta Platforms Inc.), Michael F. Cohen (Meta Platforms Inc.), and Matt Uyttendaele (Meta Platforms Inc.)</i>	

GeoMVSNet: Learning Multi-View Stereo With Geometry Perception	21508
<i>Zhe Zhang (Peking University, China), Rui Peng (Peking University, China), Yuxi Hu (The Chinese University of Hong Kong, China), and Ronggang Wang (Peking University, China)</i>	
DINN360: Deformable Invertible Neural Network for Latitude-Aware 360° Image Rescaling	21519
<i>Yichen Guo (Beihang University, China), Mai Xu (Beihang University, China), Lai Jiang (University of British Columbia, Canada), Leonid Sigal (University of British Columbia, Canada), and Yunjin Chen (Beihang University, China)</i>	
OmniVidar: Omnidirectional Depth Estimation From Multi-Fisheye Images	21529
<i>Sheng Xie (Harbin Institute of Technology, Shenzhen), Daochuan Wang (Hong Kong Center for Logistics Robotics), and Yun-Hui Liu (Harbin Institute of Technology, Shenzhen; Hong Kong Center for Logistics Robotics; The Chinese University of Hong Kong)</i>	
Learning To Fuse Monocular and Multi-View Cues for Multi-Frame Depth Estimation in Dynamic Scenes	21539
<i>Rui Li (Northwestern Polytechnical University), Dong Gong (The University of New South Wales), Wei Yin (DJI), Hao Chen (Zhejiang University), Yu Zhu (Northwestern Polytechnical University), Kaixuan Wang (DJI), Xiaozhi Chen (DJI), Jinqiu Sun (Northwestern Polytechnical University), and Yanning Zhang (Northwestern Polytechnical University)</i>	
Modality-Invariant Visual Odometry for Embodied Vision	21549
<i>Marius Memmel (University of Washington), Roman Bachmann (Swiss Federal Institute of Technology (EPFL)), and Amir Zamir (Swiss Federal Institute of Technology (EPFL))</i>	
VL-SAT: Visual-Linguistic Semantics Assisted Training for 3D Semantic Scene Graph Prediction in Point Cloud	21560
<i>Ziqin Wang (Beihang University), Bowen Cheng (Beihang University), Lichen Zhao (Beihang University), Dong Xu (The University of Hong Kong), Yang Tang (East China University of Science and Technology), and Lu Sheng (Beihang University)</i>	
CAPE: Camera View Position Embedding for Multi-View 3D Object Detection	21570
<i>Kaixin Xiong (Huazhong University of Science and Technology), Shi Gong (Baidu Inc.), Xiaoqing Ye (Baidu Inc.), Xiao Tan (Baidu Inc.), Ji Wan (Baidu Inc.), Errui Ding (Baidu Inc.), Jingdong Wang (Baidu Inc.), and Xiang Bai (Huazhong University of Science and Technology)</i>	
AeDet: Azimuth-Invariant Multi-View 3D Object Detection	21580
<i>Chengjian Feng (Meituan Inc.), Zequn Jie (Meituan Inc.), Yujie Zhong (Meituan Inc.), Xiangxiang Chu (Meituan Inc.), and Lin Ma (Meituan Inc.)</i>	
Object Detection With Self-Supervised Scene Adaptation	21589
<i>Zekun Zhang (Stony Brook University, USA) and Minh Hoai (VinAI Research, Vietnam)</i>	

Understanding the Robustness of 3D Object Detection With Bird’s-Eye-View Representations in Autonomous Driving	21600
<i>Zijian Zhu (Shanghai Jiao Tong University), Yichi Zhang (BNRist Center, Tsinghua University; Zhongguancun Laboratory), Hai Chen (Anhui University), Yinpeng Dong (BNRist Center, Tsinghua University), Shu Zhao (Anhui University), Wenbo Ding (SAIC Motor AI Lab), Jiachen Zhong (SAIC Motor AI Lab), and Shibao Zheng (Shanghai Jiao Tong University)</i>	
BEVHeight: A Robust Framework for Vision-Based Roadside 3D Object Detection	21611
<i>Lei Yang (Tsinghua University), Kaicheng Yu (Autonomous Driving Lab, Alibaba Group), Tao Tang (Sun Yat-sen University), Jun Li (Tsinghua University), Kun Yuan (Center for Machine Learning Research, Peking University), Li Wang (Tsinghua University), Xinyu Zhang (Tsinghua University), and Peng Chen (Autonomous Driving Lab, Alibaba Group)</i>	
Uncertainty-Aware Vision-Based Metric Cross-View Geolocalization	21621
<i>Florian Fervers (Fraunhofer IOSB), Sebastian Bullinger (Fraunhofer IOSB), Christoph Bodensteiner (Fraunhofer IOSB), Michael Arens (Fraunhofer IOSB), and Rainer Stiefelhagen (Karlsruhe Institute of Technology)</i>	
OrienterNet: Visual Localization in 2D Public Maps With Neural Matching	21632
<i>Paul-Edouard Sarlin (ETH Zurich), Daniel DeTone (Meta Reality Labs), Tsun-Yi Yang (Meta Reality Labs), Armen Avetisyan (Meta Reality Labs), Julian Straub (Meta Reality Labs), Tomasz Malisiewicz (Meta Reality Labs), Samuel Rota Bulò (Meta Reality Labs), Richard Newcombe (Meta Reality Labs), Peter Kotschieder (Meta Reality Labs), and Vasileios Balntas (Meta Reality Labs)</i>	
MSMDFusion: Fusing LiDAR and Camera at Multiple Scales With Multi-Depth Seeds for 3D Object Detection	21643
<i>Yang Jiao (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Zequn Jie (Meituan), Shaoxiang Chen (Meituan), Jingjing Chen (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Lin Ma (Meituan), and Yu-Gang Jiang (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing)</i>	
Virtual Sparse Convolution for Multimodal 3D Object Detection	21653
<i>Hai Wu (Xiamen University), Chenglu Wen (Xiamen University), Shaoshuai Shi (Max-Planck Institute), Xin Li (Texas A&M University), and Cheng Wang (Xiamen University)</i>	
Optimal Transport Minimization: Crowd Localization on Density Maps for Semi-Supervised Counting	21663
<i>Wei Lin (City University of Hong Kong) and Antoni B. Chan (City University of Hong Kong)</i>	
VoxelNeXt: Fully Sparse VoxelNet for 3D Object Detection and Tracking	21674
<i>Yukang Chen (The Chinese University of Hong Kong), Jianhui Liu (The University of Hong Kong), Xiangyu Zhang (MEGVII), Xiaojuan Qi (The University of Hong Kong), and Jiaya Jia (The Chinese University of Hong Kong; SmartMore)</i>	
GraVoS: Voxel Selection for 3D Point-Cloud Detection	21684
<i>Oren Shrouf (Technion, Israel), Yizhak Ben-Shabat (Technion, Israel & ANU Australia), and Ayellet Tal (Technion, Israel)</i>	

MSeg3D: Multi-Modal 3D Semantic Segmentation for Autonomous Driving	21694
<i>Jiale Li (Zhejiang University), Hang Dai (University of Glasgow), Hao Han (Zhejiang University), and Yong Ding (Zhejiang University)</i>	
LaserMix for Semi-Supervised LiDAR Semantic Segmentation	21705
<i>Lingdong Kong (Nanyang Technological University; National University of Singapore; CNRS(c)CREATE), Jiawei Ren (Nanyang Technological University), Liang Pan (Nanyang Technological University), and Ziwei Liu (Nanyang Technological University)</i>	
Implicit Surface Contrastive Clustering for LiDAR Point Clouds	21716
<i>Zaiwei Zhang (Nuro Inc), Min Bai (AWS AI), and Erran Li (AWS AI)</i>	
Semi-Weakly Supervised Object Kinematic Motion Prediction	21726
<i>Gengxin Liu (Shenzhen University), Qian Sun (Shenzhen University), Haibin Huang (Kuaishou Technology), Chongyang Ma (Kuaishou Technology), Yulan Guo (Sun Yat-sen University), Li Yi (Tsinghua University), Hui Huang (Shenzhen University), and Ruizhen Hu (Shenzhen University)</i>	
PartSLIP: Low-Shot Part Segmentation for 3D Point Clouds via Pretrained Image-Language Models	21736
<i>Minghua Liu (UC San Diego), Yinhao Zhu (Qualcomm AI Research), Hong Cai (Qualcomm AI Research), Shizhong Han (Qualcomm AI Research), Zhan Ling (UC San Diego), Fatih Porikli (Qualcomm AI Research), and Hao Su (UC San Diego)</i>	
Learning Weather-General and Weather-Specific Features for Image Restoration Under Multiple Adverse Weather Conditions	21747
<i>Yurui Zhu (University of Science and Technology of China; Shanghai Artificial Intelligence Laboratory), Tianyu Wang (The Chinese University of Hong Kong), Xueyang Fu (University of Science and Technology of China), Xuanyu Yang (Shanghai Artificial Intelligence Laboratory), Xin Guo (University of Science and Technology of China), Jifeng Dai (Tsinghua University), Yu Qiao (Shanghai Artificial Intelligence Laboratory), and Xiaowei Hu (Shanghai Artificial Intelligence Laboratory)</i>	
Geometry and Uncertainty-Aware 3D Point Cloud Class-Incremental Semantic Segmentation	21759
<i>Yuwei Yang (Sichuan University), Munawar Hayat (Monash University), Zhao Jin (Sichuan University), Chao Ren (Sichuan University), and Yinjie Lei (Sichuan University)</i>	
Learning 3D Representations From 2D Pre-Trained Models via Image-to-Point Masked Autoencoders	21769
<i>Renrui Zhang (Shanghai Artificial Intelligence Laboratory; CUHK MMLab), Lihui Wang (Peking University), Yu Qiao (Shanghai Artificial Intelligence Laboratory), Peng Gao (Shanghai Artificial Intelligence Laboratory), and Hongsheng Li (CUHK MMLab; CPII under InnoHK)</i>	
ToThePoint: Efficient Contrastive Learning of 3D Point Clouds via Recycling	21781
<i>Xinglin Li (Hunan University, China; Syracuse University, USA), Jiajing Chen (Syracuse University, USA), Jinhui Ouyang (Hunan University, China), Hanhui Deng (Hunan University, China; Exponential Innovation, China), Senem Velipasalar (Syracuse University, USA), and Di Wu (Hunan University, China; Exponential Innovation, China)</i>	

PointDistiller: Structured Knowledge Distillation Towards Efficient and Compact 3D Detection	21791
<i>Linfeng Zhang (Tsinghua University), Runpei Dong (Xi'an Jiaotong University), Hung-Shuo Tai (DIDI), and Kaisheng Ma (Tsinghua University)</i>	
PointConvFormer: Revenge of the Point-Based Convolution	21802
<i>Wenxuan Wu (Oregon State University; CASIA), Li Fuxin (Oregon State University; Apple, Inc.), and Qi Shan (Apple, Inc.)</i>	
Self-Positioning Point-Based Transformer for Point Cloud Understanding	21814
<i>Jinyoung Park (Korea University), Sanghyeok Lee (Korea University), Sihyeon Kim (Korea University), Yunyang Xiong (Meta Reality Labs), and Hyunwoo J. Kim (Korea University)</i>	
PointClustering: Unsupervised Point Cloud Pre-Training Using Transformation Invariance in Clustering	21824
<i>Fuchen Long (JD Explore Academy, China), Ting Yao (JD Explore Academy, China), Zhaofan Qiu (JD Explore Academy, China), Lusong Li (JD Explore Academy, China), and Tao Mei (JD Explore Academy, China)</i>	
Neural Intrinsic Embedding for Non-Rigid Point Cloud Matching	21835
<i>Puhua Jiang (Tsinghua Shenzhen International Graduate School; Peng Cheng Laboratory), Mingze Sun (Tsinghua Shenzhen International Graduate School), and Ruqi Huang (Tsinghua Shenzhen International Graduate School)</i>	
HGNet: Learning Hierarchical Geometry From Points, Edges, and Surfaces	21846
<i>Ting Yao (HiDream.ai Inc.), Yehao Li (HiDream.ai Inc.), Yingwei Pan (HiDream.ai Inc.), and Tao Mei (HiDream.ai Inc.)</i>	
LP-DIF: Learning Local Pattern-Specific Deep Implicit Function for 3D Objects and Scenes ...	21856
<i>Meng Wang (Tsinghua University, China), Yu-Shen Liu (Tsinghua University, China), Yue Gao (Tsinghua University, China), Kanle Shi (Kuaishou Technology, China), Yi Fang (New York University Abu Dhabi, UAE), and Zhizhong Han (Wayne State University, USA)</i>	
Conjugate Product Graphs for Globally Optimal 2D-3D Shape Matching	21866
<i>Paul Roetzer (University of Bonn), Zorah Löhner (University of Siegen), and Florian Bernard (University of Bonn)</i>	
UTM: A Unified Multiple Object Tracking Model With Identity-Aware Feature Enhancement .	21876
<i>Sisi You (Nanjing University of Posts and Telecommunications), Hantao Yao (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences (CASIA)), Bing-Kun Bao (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences), and Changsheng Xu (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences (CASIA); University of Chinese Academy of Sciences)</i>	

Learning Rotation-Equivariant Features for Visual Correspondence	21887
<i>Jongmin Lee (Pohang University of Science and Technology (POSTECH), South Korea), Byungjin Kim (Pohang University of Science and Technology (POSTECH), South Korea), Seungwook Kim (Pohang University of Science and Technology (POSTECH), South Korea), and Minsu Cho (Pohang University of Science and Technology (POSTECH), South Korea)</i>	
Adaptive Spot-Guided Transformer for Consistent Local Feature Matching	21898
<i>Jiahuan Yu (University of Science and Technology of China), Jiahao Chang (University of Science and Technology of China), Jianfeng He (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China; Deep Space Exploration Lab), Jiyang Yu (China Academy of Space Technology), and Feng Wu (University of Science and Technology of China)</i>	
PMatch: Paired Masked Image Modeling for Dense Geometric Matching	21909
<i>Shengjie Zhu (Department of Computer Science and Engineering, Michigan State University) and Xiaoming Liu (Department of Computer Science and Engineering, Michigan State University)</i>	
Iterative Geometry Encoding Volume for Stereo Matching	21919
<i>Gangwei Xu (Huazhong University of Science and Technology), Xianqi Wang (Huazhong University of Science and Technology), Xiaohuan Ding (Huazhong University of Science and Technology), and Xin Yang (Huazhong University of Science and Technology)</i>	
Adaptive Annealing for Robust Geometric Estimation	21929
<i>Chitturi Sidhartha (Indian Institute of Science, Bengaluru, India-560012), Lalit Manam (Indian Institute of Science, Bengaluru, India-560012), and Venu Madhav Govindu (Indian Institute of Science, Bengaluru, India-560012)</i>	
Tangentially Elongated Gaussian Belief Propagation for Event-Based Incremental Optical Flow Estimation	21940
<i>Jun Nagata (DENSO IT LAB., INC., Japan) and Yusuke Sekikawa (DENSO IT LAB., INC., Japan)</i>	
Robust and Scalable Gaussian Process Regression and Its Applications	21950
<i>Yifan Lu (Wuhan University, China), Jiayi Ma (Wuhan University, China), Leyuan Fang (Hunan University, China), Xin Tian (Wuhan University, China), and Junjun Jiang (Harbin Institute of Technology, China)</i>	
BEV-Guided Multi-Modality Fusion for Driving Perception	21960
<i>Yunze Man (UIUC), Liang-Yan Gui (UIUC), and Yu-Xiong Wang (UIUC)</i>	
HumanBench: Towards General Human-Centric Perception With Projector Assisted Pretraining ...	21970
<i>Shixiang Tang (The University of Sydney; SenseTime Research), Cheng Chen (SenseTime Research), Qingsong Xie (SenseTime Research), Meilin Chen (Zhejiang University; SenseTime Research), Yizhou Wang (Zhejiang University; SenseTime Research), Yuanzheng Ci (The University of Sydney), Lei Bai (Shanghai AI Laboratory), Feng Zhu (SenseTime Research), Haiyang Yang (SenseTime Research), Li Yi (SenseTime Research), Rui Zhao (SenseTime Research; Shanghai Jiao Tong University, China), and Wanli Ouyang (Shanghai AI Laboratory)</i>	

Think Twice Before Driving: Towards Scalable Decoders for End-to-End Autonomous Driving	21983
<i>Xiaosong Jia (Shanghai Jiao Tong University; Shanghai AI Laboratory), Penghao Wu (Shanghai AI Laboratory; University of California at San Diego), Li Chen (Shanghai AI Laboratory), Jiangwei Xie (Shanghai AI Laboratory), Conghui He (Shanghai AI Laboratory), Junchi Yan (Shanghai Jiao Tong University; Shanghai AI Laboratory), and Hongyang Li (Shanghai AI Laboratory; Shanghai Jiao Tong University)</i>	
ProphNet: Efficient Agent-Centric Motion Forecasting With Anchor-Informed Proposals	21995
<i>Xishun Wang (QCraft), Tong Su (QCraft), Fang Da (QCraft), and Xiaodong Yang (QCraft)</i>	
StarCraftImage: A Dataset for Prototyping Spatial Reasoning Methods for Multi-Agent Environments	22004
<i>Sean Kulinski (Purdue University), Nicholas R. Waytowich (ARL), James Z. Hare (ARL), and David I. Inouye (Purdue University)</i>	
Stimulus Verification Is a Universal and Effective Sampler in Multi-Modal Human Trajectory Prediction	22014
<i>Jianhua Sun (Shanghai Jiao Tong University, China), Yuxuan Li (Shanghai Jiao Tong University, China), Liang Chai (Shanghai Jiao Tong University, China), and Cewu Lu (Shanghai Jiao Tong University, China)</i>	
PyPose: A Library for Robot Learning With Physics-Based Optimization	22024
<i>Chen Wang (Carnegie Mellon University, USA; State University of New York, USA), Dasong Gao (Carnegie Mellon University, USA; Massachusetts Institute of Technology, USA), Kuan Xu (Nanyang Technological University, Singapore), Junyi Geng (Carnegie Mellon University, USA), Yaoyu Hu (Carnegie Mellon University, USA), Yuheng Qiu (Carnegie Mellon University, USA), Bowen Li (Carnegie Mellon University, USA), Fan Yang (ETH Zürich, Switzerland), Brady Moon (Carnegie Mellon University, USA), Abhinav Pandey (Pennsylvania State University, USA), Aryan Aryan (Carnegie Mellon University, USA; Delhi Technological University, India), Jiahe Xu (Carnegie Mellon University, USA), Tianhao Wu (University of Virginia, USA), Haonan He (Carnegie Mellon University, USA), Daning Huang (Pennsylvania State University, USA), Zhongqiang Ren (Carnegie Mellon University, USA), Shibo Zhao (Carnegie Mellon University, USA), Taimeng Fu (The Chinese University of Hong Kong, China), Pranay Reddy (University of Massachusetts Amherst, USA), Xiao Lin (Georgia Institute of Technology, USA), Wenshan Wang (Carnegie Mellon University, USA), Jingnan Shi (Massachusetts Institute of Technology, USA), Rajat Talak (Massachusetts Institute of Technology, USA), Kun Cao (Nanyang Technological University, Singapore), Yi Du (State University of New York, USA), Han Wang (Nanyang Technological University, Singapore), Huai Yu (Wuhan University, China), Shanzhao Wang (University of Michigan, USA), Siyu Chen (Nanyang Technological University, Singapore), Ananth Kashyap (Fox Chapel Area High School, USA), Rohan Bandaru (Lexington High School, USA), Karthik Dantu (State University of New York, USA), Jiajun Wu (Stanford University, USA), Lihua Xie (Nanyang Technological University, Singapore), Luca Carlone (Massachusetts Institute of Technology, USA), Marco Hutter (ETH Zürich, Switzerland), and Sebastian Scherer (Carnegie Mellon University, USA)</i>	

Source-Free Adaptive Gaze Estimation by Uncertainty Reduction	22035
<i>Xin Cai (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Jiabei Zeng (Institute of Computing Technology, Chinese Academy of Sciences, China), Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China; Peng Cheng Laboratory, China), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China)</i>	
Camouflaged Object Detection With Feature Decomposition and Edge Reconstruction	22046
<i>Chunming He (Tsinghua University), Kai Li (NEC Laboratories America), Yachao Zhang (Tsinghua University), Longxiang Tang (Tsinghua University), Yulun Zhang (ETH Zurich), Zhenhua Guo (Tianyi Traffic Technology), and Xiu Li (Tsinghua University)</i>	
MOTRv2: Bootstrapping End-to-End Multi-Object Tracking by Pretrained Object Detectors ..	22056
<i>Yuang Zhang (Shanghai Jiao Tong University), Tiancai Wang (MEGVII Technology), and Xiangyu Zhang (MEGVII Technology; Beijing Academy of Artificial Intelligence)</i>	
Clothing-Change Feature Augmentation for Person Re-Identification	22066
<i>Ke Han (Institute of Automation, Chinese Academy of Sciences (CASIA)), Shaogang Gong (Institute of Automation, Chinese Academy of Sciences (CASIA)), Yan Huang (University of Chinese Academy of Sciences (UCAS)), Liang Wang (Queen Mary University of London (QMUL)), and Tieniu Tan (Nanjing University)</i>	
Dynamic Aggregated Network for Gait Recognition	22076
<i>Kang Ma (Beijing Institute of Technology), Ying Fu (Beijing Institute of Technology), Dezhi Zheng (Beijing Institute of Technology), Chunshui Cao (WATRIX.AI), Xuecai Hu (WATRIX.AI), and Yongzhen Huang (WATRIX.AI; Beijing Normal University)</i>	
Feature Representation Learning With Adaptive Displacement Generation and Transformer Fusion for Micro-Expression Recognition	22086
<i>Zhijun Zhai (Wuhan University, China), Jianhui Zhao (Wuhan University, China), Chengjiang Long (Meta Reality Labs, USA), Wenju Xu (OPPO US Research Center, InnoPeak Technology Inc, USA), Shuangjiang He (FiberHome Telecommunication Technologies Co., Ltd, China), and Huijuan Zhao (FiberHome Telecommunication Technologies Co., Ltd, China)</i>	
MetaPortrait: Identity-Preserving Talking Head Generation With Fast Personalized Adaptation	22096
<i>Bowen Zhang (USTC), Chenyang Qi (HKUST), Pan Zhang (USTC), Bo Zhang (Microsoft), HsiangTao Wu (Microsoft), Dong Chen (HKUST), Qifeng Chen (HKUST), Yong Wang (USTC), and Fang Wen (Microsoft)</i>	
FLAG3D: A 3D Fitness Activity Dataset With Language Instruction	22106
<i>Yansong Tang (Tsinghua University), Jinpeng Liu (Tsinghua University), Aoyang Liu (Tsinghua University), Bin Yang (Tsinghua University), Wenxun Dai (Tsinghua University), Yongming Rao (Tsinghua University), Jiwen Lu (Tsinghua University), Jie Zhou (Tsinghua University), and Xiu Li (Tsinghua University)</i>	

TranSG: Transformer-Based Skeleton Graph Prototype Contrastive Learning With Structure-Trajectory Prompted Reconstruction for Person Re-Identification	22118
<i>Haocong Rao (Nanyang Technological University, Singapore) and Chunyan Miao (Nanyang Technological University, Singapore)</i>	
NeMo: Learning 3D Neural Motion Fields From Multiple Video Instances of the Same Action	22129
<i>Kuan-Chieh Wang (Stanford University), Zhenzhen Weng (Stanford University), Maria Xenochristou (Stanford University), João Pedro Araújo (Stanford University), Jeffrey Gu (Stanford University), Karen Liu (Stanford University), and Serena Yeung (Stanford University)</i>	
Unsupervised Space-Time Network for Temporally-Consistent Segmentation of Multiple Motions..	22139
<i>Etienne Meunier (Inna, France) and Patrick Bouthemy (Inna, France)</i>	
Deep Polarization Reconstruction With PDAVIS Events	22149
<i>Haiyang Mei (Dalian University of Technology, China; University of Zurich and ETH Zurich, Switzerland), Zuowen Wang (University of Zurich and ETH Zurich, Switzerland), Xin Yang (Dalian University of Technology, China), Xiaopeng Wei (Dalian University of Technology, China), and Tobi Delbruck (University of Zurich and ETH Zurich, Switzerland)</i>	
Range-Nullspace Video Frame Interpolation With Focalized Motion Estimation	22159
<i>Zhiyang Yu (Harbin Institute of Technology, China), Yu Zhang (SenseTime Research and Tetras.AI, China), Dongqing Zou (SenseTime Research and Tetras.AI, China; Shanghai Jiao Tong University, China), Xijun Chen (Harbin Institute of Technology, China), Jimmy S. Ren (SenseTime Research and Tetras.AI, China; Shanghai Jiao Tong University, China), and Shunqing Ren (Harbin Institute of Technology, China)</i>	
Exploring Motion Ambiguity and Alignment for High-Quality Video Frame Interpolation	22169
<i>Kun Zhou (SSE, CUHK-Shenzhen; SmartMore Corporation), Wenbo Li (CUHK), Xiaoguang Han (SSE, CUHK-Shenzhen), and Jiangbo Lu (SmartMore Corporation)</i>	
1000 FPS HDR Video With a Spike-RGB Hybrid Camera	22180
<i>Yakun Chang (Peking University), Chu Zhou (Peking University), Yuchen Hong (Peking University), Liwen Hu (Peking University), Chao Xu (Peking University), Tiejun Huang (Peking University), and Boxin Shi (Peking University)</i>	
Deep Discriminative Spatial and Temporal Network for Efficient Video Deblurring	22191
<i>Jinshan Pan (Nanjing University of Science and Technology), Boming Xu (Nanjing University of Science and Technology), Jiangxin Dong (Nanjing University of Science and Technology), Jianjun Ge (China Electronics Technology Group Corporation), and Jinhui Tang (Nanjing University of Science and Technology)</i>	

Gated Multi-Resolution Transfer Network for Burst Restoration and Enhancement	22201
<i>Nancy Mehta (CVPR Lab, Indian Institute of Technology Ropar), Akshay Dudhane (Mohamed bin Zayed University of AI), Subrahmanyam Murala (CVPR Lab, Indian Institute of Technology Ropar), Syed Waqas Zamir (Inception Institute of AI), Salman Khan (Mohamed bin Zayed University of AI; Australian National University), and Fahad Shahbaz Khan (Mohamed bin Zayed University of AI; Linkoping University)</i>	
A Unified HDR Imaging Method With Pixel and Patch Level	22211
<i>Qingsen Yan (Northwestern Polytechnical University), Weiye Chen (Xidian University), Song Zhang (Xidian University), Yu Zhu (Northwestern Polytechnical University), Jinqiu Sun (Northwestern Polytechnical University), and Yanning Zhang (Northwestern Polytechnical University)</i>	
BiasBed – Rigorous Texture Bias Evaluation	22221
<i>Nikolai Kalischek (Photogrammetry and Remote Sensing, ETH Zurich), Rodrigo Caye Daudt (Photogrammetry and Remote Sensing, ETH Zurich), Torben Peters (Photogrammetry and Remote Sensing, ETH Zurich), Reinhard Furrer (Institute for Mathematics, University of Zurich), Jan D. Wegner (Institute for Computational Science, University of Zurich), and Konrad Schindler (Photogrammetry and Remote Sensing, ETH Zurich)</i>	
Learning a Practical SDR-to-HDRTV Up-Conversion Using New Dataset and Degradation Models ...	22231
<i>Cheng Guo (Communication University of China; Peng Cheng Laboratory), Leidong Fan (Peng Cheng Laboratory; Peking University), Ziyu Xue (Academy of Broadcasting Science, National Radio and Television Administration; Communication University of China), and Xiuhua Jiang (Peng Cheng Laboratory; Peking University)</i>	
Learning a Deep Color Difference Metric for Photographic Images	22242
<i>Haoyu Chen (City University of Hong Kong), Zhihua Wang (City University of Hong Kong; Shenzhen MSU-BIT University), Yang Yang (Shenzhen Transsion Holdings Co., Ltd.), Qilin Sun (The Chinese University of Hong Kong (Shenzhen)), and Kede Ma (City University of Hong Kong)</i>	
Learning a Simple Low-Light Image Enhancer From Paired Low-Light Instances	22252
<i>Zhenqi Fu (Xiamen University, China), Yan Yang (Hangzhou Dianzi University, China), Xiaotong Tu (Xiamen University, China), Yue Huang (Xiamen University, China), Xinghao Ding (Xiamen University, China), and Kai-Kuang Ma (Nanyang Technological University, Singapore)</i>	
Residual Degradation Learning Unfolding Framework With Mixing Priors Across Spectral and Spatial for Compressive Spectral Imaging	22262
<i>Yubo Dong (Xidian University), Dahua Gao (Xidian University), Tian Qiu (Xidian University), Yuyan Li (Xidian University), Minxi Yang (Xidian University), and Guangming Shi (Xidian University)</i>	
Toward Stable, Interpretable, and Lightweight Hyperspectral Super-Resolution	22272
<i>Wen-jin Guo (Hunan University), Weiyang Xie (Xidian University), Kai Jiang (Xidian University), Yunsong Li (Xidian University), Jie Lei (Xidian University), and Leyuan Fang (Xidian University)</i>	

RIDCP: Revitalizing Real Image Dehazing via High-Quality Codebook Priors	22282
<i>Rui-Qi Wu (Nankai University), Zheng-Peng Duan (Nankai University), Chun-Le Guo (Nankai University), Zhi Chai (Hisilicon Technologies Co. Ltd.), and Chongyi Li (Nanyang Technological University)</i>	
Robust Unsupervised StyleGAN Image Restoration	22292
<i>Yohan Poirier-Ginter (Inria, Universite Cote d'Azur; Universite Laval) and Jean-François Lalonde (Universite Laval)</i>	
Quality-Aware Pre-Trained Models for Blind Image Quality Assessment	22302
<i>Kai Zhao (Kuaishou Technology), Kun Yuan (Kuaishou Technology), Ming Sun (Kuaishou Technology), Mading Li (Kuaishou Technology), and Xing Wen (Kuaishou Technology)</i>	
Learning To Exploit the Sequence-Specific Prior Knowledge for Image Processing Pipelines Optimization	22314
<i>Haina Qin (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Longfei Han (Beijing Technology and Business University), Weihua Xiong (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences), Juan Wang (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences), Wentao Ma (Zeku Technology), Bing Li (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences), and Weiming Hu (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; School of Artificial Intelligence, University of Chinese Academy of Sciences)</i>	
Multi-Realism Image Compression With a Conditional Generator	22324
<i>Eirikur Agustsson (Google Research), David Minnen (Google Research), George Toderici (Google Research), and Fabian Mentzer (Google Research)</i>	
RGB No More: Minimally-Decoded JPEG Vision Transformers	22334
<i>Jeongsoo Park (University of Michigan) and Justin Johnson (University of Michigan)</i>	
Kernel Aware Resampler	22347
<i>Michael Bernasconi (ETH Zurich; DisneyResearch), Abdelaziz Djelouah (DisneyResearch), Farnood Salehi (DisneyResearch), Markus Gross (ETH Zurich; DisneyResearch), and Christopher Schroers (DisneyResearch)</i>	
Spatial-Frequency Mutual Learning for Face Super-Resolution	22356
<i>Chenyang Wang (School of Computer Science and Technology, Harbin Institute of Technology, China), Junjun Jiang (School of Computer Science and Technology, Harbin Institute of Technology, China), Zhiwei Zhong (School of Computer Science and Technology, Harbin Institute of Technology, China), and Xianming Liu (School of Computer Science and Technology, Harbin Institute of Technology, China)</i>	

Activating More Pixels in Image Super-Resolution Transformer	22367
<i>Xiangyu Chen (University of Macau; Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences; Shanghai Artificial Intelligence Laboratory), Xintao Wang (ARC Lab, Tencent PCG), Jiantao Zhou (University of Macau), Yu Qiao (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences; Shanghai Artificial Intelligence Laboratory), and Chao Dong (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences; Shanghai Artificial Intelligence Laboratory)</i>	
Omni Aggregation Networks for Lightweight Image Super-Resolution	22378
<i>Hang Wang (Huawei), Xuanhong Chen (Shanghai Jiao Tong University, China), Bingbing Ni (Shanghai Jiao Tong University, China; Huawei), Yutian Liu (Shanghai Jiao Tong University, China), and Jinfan Liu (Shanghai Jiao Tong University, China)</i>	
Towards Artistic Image Aesthetics Assessment: A Large-Scale Dataset and a New Method ..	22388
<i>Ran Yi (Shanghai Jiao Tong University), Haoyuan Tian (Shanghai Jiao Tong University), Zhihao Gu (Shanghai Jiao Tong University), Yu-Kun Lai (Cardiff University), and Paul L. Rosin (Cardiff University)</i>	
RWSC-Fusion: Region-Wise Style-Controlled Fusion Network for the Prohibited X-Ray Security Image Synthesis	22398
<i>Luwen Duan (Zhejiang Dahua Technology Co., Ltd.; Zhejiang University), Min Wu (Zhejiang Dahua Technology Co., Ltd.), Lijian Mao (Zhejiang Dahua Technology Co., Ltd.), Jun Yin (Zhejiang Dahua Technology Co., Ltd.), Jianping Xiong (Zhejiang Dahua Technology Co., Ltd.), and Xi Li (Zhejiang University)</i>	
Efficient Scale-Invariant Generator With Column-Row Entangled Pixel Synthesis	22408
<i>Thuan Hoang Nguyen (VinAI Research, Vietnam), Thanh Van Le (VinAI Research, Vietnam), and Anh Tran (VinAI Research, Vietnam)</i>	
Masked and Adaptive Transformer for Exemplar Based Image Translation	22418
<i>Chang Jiang (Hangzhou Dianzi University), Fei Gao (Hangzhou Dianzi University; Xidian University), Biao Ma (Hangzhou Dianzi University), Yuhao Lin (Hangzhou Dianzi University), Nannan Wang (Xidian University), and Gang Xu (Hangzhou Dianzi University)</i>	
SmartBrush: Text and Shape Guided Object Inpainting With Diffusion Model	22428
<i>Shaoan Xie (Carnegie Mellon University), Zhifei Zhang (Adobe Research), Zhe Lin (Adobe Research), Tobias Hinz (Adobe Research), and Kun Zhang (Carnegie Mellon University; Mohamed bin Zayed University of Artificial Intelligence)</i>	
Neural Transformation Fields for Arbitrary-Styled Font Generation	22438
<i>Bin Fu (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), Junjun He (Shanghai Artificial Intelligence Laboratory), Jianjun Wang (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), and Yu Qiao (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences; Shanghai Artificial Intelligence Laboratory)</i>	
Referring Image Matting	22448
<i>Jizhizi Li (The University of Sydney, Australia), Jing Zhang (The University of Sydney, Australia), and Dacheng Tao (The University of Sydney, Australia)</i>	

Handwritten Text Generation From Visual Archetypes	22458
<i>Vittorio Pippi (University of Modena and Reggio Emilia, Via Pietro Vivarelli, Modena (Italy)), Silvia Cascianelli (University of Modena and Reggio Emilia, Via Pietro Vivarelli, Modena (Italy)), and Rita Cucchiara (University of Modena and Reggio Emilia, Via Pietro Vivarelli, Modena (Italy))</i>	
SceneComposer: Any-Level Semantic Image Synthesis	22468
<i>Yu Zeng (Johns Hopkins University), Zhe Lin (Adobe Research), Jianming Zhang (Adobe Research), Qing Liu (Adobe Research), John Collomosse (Adobe Research), Jason Kuen (Adobe Research), and Vishal M. Patel (Johns Hopkins University)</i>	
Affordance Diffusion: Synthesizing Hand-Object Interactions	22479
<i>Yufei Ye (Carnegie Mellon University), Xueting Li (NVIDIA), Abhinav Gupta (Carnegie Mellon University), Shalini De Mello (NVIDIA), Stan Birchfield (NVIDIA), Jiaming Song (NVIDIA), Shubham Tulsiani (Carnegie Mellon University), and Sifei Liu (NVIDIA)</i>	
LayoutDiffusion: Controllable Diffusion Model for Layout-to-Image Generation	22490
<i>Guangcong Zheng (Zhejiang University), Xianpan Zhou (Zhejiang University), Xuwei Li (Zhejiang University), Zhongang Qi (ARC Lab, Tencent PCG), Ying Shan (ARC Lab, Tencent PCG), and Xi Li (Zhejiang University; Shanghai Institute for Advanced Study of Zhejiang University; Shanghai AI Lab; Zhejiang - Singapore Innovation and AI Joint Research Lab)</i>	
DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation ...	22500
<i>Nataniel Ruiz (Google Research; Boston University), Yuanzhen Li (Google Research), Varun Jampani (Google Research), Yael Pritch (Google Research), Michael Rubinstein (Google Research), and Kfir Aberman (Google Research)</i>	
GLIGEN: Open-Set Grounded Text-to-Image Generation	22511
<i>Yuheng Li (University of Wisconsin-Madison), Haotian Liu (University of Wisconsin-Madison), Qingyang Wu (Columbia University), Fangzhou Mu (University of Wisconsin-Madison), Jianwei Yang (Microsoft), Jianfeng Gao (Microsoft), Chunyuan Li (Microsoft), and Yong Jae Lee (University of Wisconsin-Madison)</i>	
Safe Latent Diffusion: Mitigating Inappropriate Degeneration in Diffusion Models	22522
<i>Patrick Schramowski (DFKI; Hessian.AI; Computer Science Department, TU Darmstadt; LAION), Manuel Brack (Computer Science Department, TU Darmstadt), Björn Deiseroth (Hessian.AI; Computer Science Department, TU Darmstadt; Aleph Alpha), and Kristian Kersting (DFKI; Hessian.AI; Computer Science Department, TU Darmstadt)</i>	
EDICT: Exact Diffusion Inversion via Coupled Transformations	22532
<i>Bram Wallace (Salesforce Research), Akash Gokul (Salesforce Research), and Nikhil Naik (Salesforce Research)</i>	
Solving 3D Inverse Problems Using Pre-Trained 2D Diffusion Models	22542
<i>Hyungjin Chung (Korea Advanced Institute of Science & Technology; Los Alamos National Laboratory), Dohoon Ryu (Korea Advanced Institute of Science & Technology), Michael T. McCann (Los Alamos National Laboratory), Marc L. Klasky (Los Alamos National Laboratory), and Jong Chul Ye (Korea Advanced Institute of Science & Technology)</i>	

Diffusion Probabilistic Model Made Slim	22552
<i>Xingyi Yang (National University of Singapore), Daquan Zhou (ByteDance Inc.), Jiashi Feng (ByteDance Inc.), and Xinchao Wang (National University of Singapore)</i>	
Align Your Latents: High-Resolution Video Synthesis With Latent Diffusion Models	22563
<i>Andreas Blattmann (LMU Munich), Robin Rombach (LMU Munich), Huan Ling (NVIDIA; Vector Institute; University of Toronto), Tim Dockhorn (NVIDIA; Vector Institute; University of Waterloo), Seung Wook Kim (NVIDIA; Vector Institute; University of Toronto), Sanja Fidler (NVIDIA; Vector Institute; University of Toronto), and Karsten Kreis (NVIDIA)</i>	
Binary Latent Diffusion	22576
<i>Ze Wang (Purdue University), Jiang Wang (Microsoft Corporation), Zicheng Liu (Microsoft Corporation), and Qiang Qiu (Purdue University)</i>	
Semi-Supervised Video Inpainting With Cycle Consistency Constraints	22586
<i>Zhiliang Wu (Nanjing University of Science and Technology, China), Hanyu Xuan (Anhui University, China), Changchang Sun (Illinois Institute of Technology, USA), Weili Guan (Monash University, Australia; Peng Cheng Laboratory, China), Kang Zhang (Nanjing University of Science and Technology, China), and Yan Yan (Illinois Institute of Technology, USA)</i>	
Towards Accurate Image Coding: Improved Autoregressive Image Generation With Dynamic Vector Quantization	22596
<i>Mengqi Huang (University of Science and Technology of China, China), Zhendong Mao (University of Science and Technology of China, China; Institute of Artificial intelligence, Hefei Comprehensive National Science Center, China), Zhuowei Chen (University of Science and Technology of China, China), and Yongdong Zhang (University of Science and Technology of China, China; Institute of Artificial intelligence, Hefei Comprehensive National Science Center, China)</i>	
Large-Capacity and Flexible Video Steganography via Invertible Neural Network	22606
<i>Chong Mou (Peking University Shenzhen Graduate School, China), Youmin Xu (Peking University Shenzhen Graduate School, China), Jiechong Song (Peking University Shenzhen Graduate School, China), Chen Zhao (King Abdullah University of Science and Technology (KAUST), Saudi Arabia), Bernard Ghanem (King Abdullah University of Science and Technology (KAUST), Saudi Arabia), and Jian Zhang (Peking University Shenzhen Graduate School, China; Peng Cheng Laboratory, China)</i>	
Neural Video Compression With Diverse Contexts	22616
<i>Jiahao Li (Microsoft Research Asia), Bin Li (Microsoft Research Asia), and Yan Lu (Microsoft Research Asia)</i>	
Efficient Semantic Segmentation by Altering Resolutions for Compressed Videos	22627
<i>Yubin Hu (Tsinghua University), Yuze He (Tsinghua University), Yanghao Li (Tsinghua University), Jisheng Li (Tsinghua University), Yuxing Han (Tsinghua University), Jiangtao Wen (Eastern Institute for Advanced Study), and Yong-Jin Liu (Tsinghua University)</i>	

Structured Sparsity Learning for Efficient Video Super-Resolution	22638
<i>Bin Xia (Tsinghua University), Jingwen He (Shanghai AI Laboratory), Yulun Zhang (ETH Zurich), Yitong Wang (ByteDance Inc), Yapeng Tian (University of Texas at Dallas), Wenming Yang (Tsinghua University), and Luc Van Gool (ETH Zurich)</i>	
DisCo-CLIP: A Distributed Contrastive Loss for Memory Efficient CLIP Training	22648
<i>Yihao Chen (International Digital Economy Academy (IDEA), China), Xianbiao Qi (International Digital Economy Academy (IDEA), China), Jianan Wang (International Digital Economy Academy (IDEA), China), and Lei Zhang (International Digital Economy Academy (IDEA), China)</i>	
Boost Vision Transformer With GPU-Friendly Sparsity and Quantization	22658
<i>Chong Yu (Fudan University; NVIDIA Corporation), Tao Chen (Fudan University), Zhongxue Gan (Fudan University), and Jiayuan Fan (Fudan University)</i>	
All Are Worth Words: A ViT Backbone for Diffusion Models	22669
<i>Fan Bao (Tsinghua University, China), Shen Nie (Renmin University of China; Beijing Key Laboratory of Big Data Management and Analysis Methods, China), Kaiwen Xue (Renmin University of China; Beijing Key Laboratory of Big Data Management and Analysis Methods, China), Yue Cao (Beijing Academy of Artificial Intelligence), Chongxuan Li (Renmin University of China; Beijing Key Laboratory of Big Data Management and Analysis Methods, China), Hang Su (Tsinghua University, China), and Jun Zhu (Tsinghua University, China)</i>	
Sparsifiner: Learning Sparse Instance-Dependent Attention for Efficient Vision Transformers	22680
<i>Cong Wei (University of Toronto; Modiface, Inc.), Brendan Duke (University of Toronto; Modiface, Inc.), Ruowei Jiang (Modiface, Inc.), Parham Aarabi (University of Toronto; Modiface, Inc.), Graham W. Taylor (University of Guelph; Vector Institute), and Florian Shkurti (University of Toronto; Vector Institute)</i>	
Vision Transformer With Super Token Sampling	22690
<i>Huaibo Huang (MAIS&CRIPAC, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Xiaoqiang Zhou (MAIS&CRIPAC, Institute of Automation, Chinese Academy of Sciences, China; University of Science and Technology of China, China), Jie Cao (MAIS&CRIPAC, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Ran He (MAIS&CRIPAC, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China; ShanghaiTech University, China), and Tieniu Tan (MAIS&CRIPAC, Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China; University of Science and Technology of China, China; Nanjing University, Nanjing, China)</i>	
DropKey for Vision Transformer	22700
<i>Bonan Li (University of Chinese Academy of Sciences), Yinhan Hu (University of Chinese Academy of Sciences), Xuecheng Nie (MT Lab, Meitu Inc.), Congying Han (University of Chinese Academy of Sciences), Xiangjian Jiang (University of Cambridge), Tiande Guo (University of Chinese Academy of Sciences), and Luoqi Liu (MT Lab, Meitu Inc.)</i>	

Seeing Beyond the Brain: Conditional Diffusion Model With Sparse Masked Modeling for Vision Decoding	22710
<i>Zijiao Chen (National University of Singapore), Jiaxin Qing (The Chinese University of Hong Kong), Tiange Xiang (Stanford University), Wan Lin Yue (National University of Singapore), and Juan Helen Zhou (National University of Singapore)</i>	
ResFormer: Scaling ViTs With Multi-Resolution Training	22721
<i>Rui Tian (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Zuxuan Wu (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Qi Dai (Microsoft Research Asia), Han Hu (Microsoft Research Asia), Yu Qiao (Shanghai AI Laboratory), and Yu-Gang Jiang (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing)</i>	
Stare at What You See: Masked Image Modeling Without Reconstruction	22732
<i>Hongwei Xue (University of Science and Technology of China; Shanghai Artificial Intelligence Laboratory), Peng Gao (Shanghai Artificial Intelligence Laboratory; Shenzhen Institutes of Advanced Technology, Chinese Academy of Science), Hongyang Li (Shanghai Artificial Intelligence Laboratory), Yu Qiao (Shanghai Artificial Intelligence Laboratory), Hao Sun (China Telecom Corporation Ltd. Data&AI Technology Company), Houqiang Li (University of Science and Technology of China), and Jiebo Luo (University of Rochester)</i>	
Mixed Autoencoder for Self-Supervised Visual Representation Learning	22742
<i>Kai Chen (Hong Kong University of Science and Technology), Zhili Liu (Hong Kong University of Science and Technology; Huawei Noah's Ark Lab), Lanqing Hong (Huawei Noah's Ark Lab), Hang Xu (Huawei Noah's Ark Lab), Zhenguo Li (Huawei Noah's Ark Lab), and Dit-Yan Yeung (Hong Kong University of Science and Technology)</i>	
Shape-Erased Feature Learning for Visible-Infrared Person Re-Identification	22752
<i>Jiawei Feng (Sun Yat-sen University, China), Ancong Wu (Sun Yat-sen University, China), and Wei-Shi Zheng (Sun Yat-sen University, China; Ministry of Education, China; Guangdong Key Laboratory of Information Security Technology, China)</i>	
G-MSM: Unsupervised Multi-Shape Matching With Graph-Based Affinity Priors	22762
<i>Marvin Eisenberger (TU Munich), Aysim Toker (TU Munich), Laura Leal-Taixé (TU Munich), and Daniel Cremers (TU Munich; Munich Center for Machine Learning; University of Oxford)</i>	
Efficient Mask Correction for Click-Based Interactive Image Segmentation	22773
<i>Fei Du (Alibaba Group), Jianlong Yuan (Alibaba Group), Zhibin Wang (Alibaba Group), and Fan Wang (Alibaba Group)</i>	
Prototype-Based Embedding Network for Scene Graph Generation	22783
<i>Chaofan Zheng (University of Electronic Science and Technology of China, China), Xinyu Lyu (University of Electronic Science and Technology of China, China), Lianli Gao (University of Electronic Science and Technology of China, China), Bo Dai (University of Electronic Science and Technology of China, China), and Jingkuan Song (University of Electronic Science and Technology of China, China)</i>	

Graph Representation for Order-Aware Visual Transformation	22793
<i>Yue Qiu (National Institute of Advanced Industrial Science and Technology (AIST)), Yanjun Sun (National Institute of Advanced Industrial Science and Technology (AIST)), Fumiya Matsuzawa (National Institute of Advanced Industrial Science and Technology (AIST)), Kenji Iwata (National Institute of Advanced Industrial Science and Technology (AIST)), and Hirokatsu Kataoka (National Institute of Advanced Industrial Science and Technology (AIST))</i>	
Unbiased Scene Graph Generation in Videos	22803
<i>Sayak Nag (University of California, USA), Kyle Min (Intel Corporation, USA), Subarna Tripathi (Intel Corporation, USA), and Amit K. Roy-Chowdhury (University of California, USA)</i>	
Recurrence Without Recurrence: Stable Video Landmark Detection With Deep Equilibrium Models	22814
<i>Paul Micaelli (University of Edinburgh), Arash Vahdat (NVIDIA), Hongxu Yin (NVIDIA), Jan Kautz (NVIDIA), and Pavlo Molchanov (NVIDIA)</i>	
VideoTrack: Learning To Track Objects via Video Transformer	22826
<i>Fei Xie (Shanghai Jiao Tong University), Lei Chu (Microsoft Research Asia), Jiahao Li (Microsoft Research Asia), Yan Lu (Microsoft Research Asia), and Chao Ma (Shanghai Jiao Tong University)</i>	
Breaking the “Object” in Video Object Segmentation	22836
<i>Pavel Tokmakov (Toyota Research Institute), Jie Li (Toyota Research Institute), and Adrien Gaidon (Toyota Research Institute)</i>	
Hierarchical Semantic Contrast for Scene-Aware Video Anomaly Detection	22846
<i>Shengyang Sun (Zhejiang University, China) and Xiaojin Gong (Zhejiang University, China)</i>	
Mask-Free Video Instance Segmentation	22857
<i>Lei Ke (ETH Zurich; HKUST), Martin Danelljan (ETH Zurich), Henghui Ding (ETH Zurich), Yu-Wing Tai (HKUST), Chi-Keung Tang (HKUST), and Fisher Yu (ETH Zurich)</i>	
Hierarchical Neural Memory Network for Low Latency Event Processing	22867
<i>Ryuhei Hamaguchi (National Institute of Advanced Industrial Science and Technology (AIST)), Yasutaka Furukawa (Simon Fraser University), Masaki Onishi (National Institute of Advanced Industrial Science and Technology (AIST)), and Ken Sakurada (National Institute of Advanced Industrial Science and Technology (AIST))</i>	
Unifying Short and Long-Term Tracking With Graph Hierarchies	22877
<i>Orcun Cetintas (Technical University of Munich), Guillem Brasó (Technical University of Munich; Munich Center for Machine Learning), and Laura Leal-Taixé (Technical University of Munich)</i>	
Towards End-to-End Generative Modeling of Long Videos With Memory-Efficient Bidirectional Transformers	22888
<i>Jaehoon Yoo (KAIST), Semin Kim (KAIST), Doyup Lee (Kakao Brain), Chiheon Kim (Kakao Brain), and Seunghoon Hong (KAIST)</i>	

An Empirical Study of End-to-End Video-Language Transformers With Masked Visual Modeling	22898
<i>Tsu-Jui Fu (UC Santa Barbara), Linjie Li (UC Santa Barbara), Zhe Gan (UC Santa Barbara), Kevin Lin (Microsoft), William Yang Wang (UC Santa Barbara), Lijuan Wang (Microsoft), and Zicheng Liu (Microsoft)</i>	
Egocentric Audio-Visual Object Localization	22910
<i>Chao Huang (University of Rochester), Yapeng Tian (University of Rochester), Anurag Kumar (Meta Reality Labs Research), and Chenliang Xu (University of Rochester)</i>	
AVFormer: Injecting Vision Into Frozen Speech Models for Zero-Shot AV-ASR	22922
<i>Paul Hongsuck Seo (Google Research), Arsha Nagrani (Google Research), and Cordelia Schmid (Google Research)</i>	
A Light Weight Model for Active Speaker Detection	22932
<i>Junhua Liao (Sichuan University, Chengdu, China), Haihan Duan (The Chinese University of Hong Kong, Shenzhen, China.), Kanghui Feng (Sichuan University, Chengdu, China), Wanbing Zhao (Sichuan University, Chengdu, China), Yanbing Yang (The Chinese University of Hong Kong, China.; Sichuan University, Chengdu, China.), and Liangyin Chen (The Chinese University of Hong Kong, China.; Sichuan University, Chengdu, China.)</i>	
Dense-Localizing Audio-Visual Events in Untrimmed Videos: A Large-Scale Benchmark and Baseline	22942
<i>Tiantian Geng (Southern University of Science and Technology; University of Birmingham), Teng Wang (Southern University of Science and Technology; The University of Hong Kong), Jinming Duan (University of Birmingham), Runmin Cong (Shandong University), and Feng Zheng (Southern University of Science and Technology; Peng Cheng Laboratory)</i>	
Video Test-Time Adaptation for Action Recognition	22952
<i>Wei Lin (Graz University of Technology, Austria; Christian Doppler Laboratory for Semantic 3D Computer Vision), Muhammad Jehanzeb Mirza (Graz University of Technology, Austria; Christian Doppler Laboratory for Embedded Machine Learning), Mateusz Kozinski (Graz University of Technology, Austria), Horst Possegger (Graz University of Technology, Austria), Hilde Kuehne (Goethe University Frankfurt, Germany; MIT-IBM Watson AI Lab), and Horst Bischof (Christian Doppler Laboratory for Embedded Machine Learning)</i>	
Unified Keypoint-Based Action Recognition Framework via Structured Keypoint Pooling	22962
<i>Ryo Hachiuma (Konica Minolta, Inc.), Fumiaki Sato (Konica Minolta, Inc.), and Taiki Sekii (Konica Minolta, Inc.)</i>	
Object Discovery From Motion-Guided Tokens	22972
<i>Zhipeng Bao (CMU), Pavel Tokmakov (Toyota Research Institute), Yu-Xiong Wang (UIUC), Adrien Gaidon (Toyota Research Institute), and Martial Hebert (CMU)</i>	
Open Set Action Recognition via Multi-Label Evidential Learning	22982
<i>Chen Zhao (na), Dawei Du (na), Anthony Hoogs (na), and Christopher Funk (na)</i>	

PivoTAL: Prior-Driven Supervision for Weakly-Supervised Temporal Action Localization	22992
<i>Mamshad Nayeem Rizve (University of Central Florida), Gaurav Mittal (Microsoft), Ye Yu (Microsoft), Matthew Hall (Microsoft), Sandra Sajeev (Microsoft), Mubarak Shah (University of Central Florida), and Mei Chen (Microsoft)</i>	
Improving Weakly Supervised Temporal Action Localization by Bridging Train-Test Gap in Pseudo Labels	23003
<i>Jingqiu Zhou (The Chinese University of Hong Kong), Linjiang Huang (The Chinese University of Hong Kong; Centre for Perceptual and Interactive Intelligence, Hong Kong), Liang Wang (Institute of Automation Chinese Academy of Science), Si Liu (Beihang University), and Hongsheng Li (The Chinese University of Hong Kong; Centre for Perceptual and Interactive Intelligence, Hong Kong; Xidian University)</i>	
Are Binary Annotations Sufficient? Video Moment Retrieval via Hierarchical Uncertainty-Based Active Learning	23013
<i>Wei Ji (National University of Singapore), Renjie Liang (National University of Singapore), Zhedong Zheng (National University of Singapore), Wenqiao Zhang (Zhejiang University), Shengyu Zhang (Zhejiang University), Juncheng Li (Zhejiang University), Mengze Li (Zhejiang University), and Tat-seng Chua (National University of Singapore)</i>	
Query-Dependent Video Representation for Moment Retrieval and Highlight Detection	23023
<i>WonJun Moon (Sungkyunkwan University), Sangeek Hyun (Sungkyunkwan University), SangUk Park (Pylar), Dongchan Park (Pylar), and Jae-Pil Heo (Sungkyunkwan University)</i>	
Vita-CLIP: Video and Text Adaptive CLIP via Multimodal Prompting	23034
<i>Syed Talal Wasim (Mohamed bin Zayed University of AI), Muzammal Naseer (Mohamed bin Zayed University of AI), Salman Khan (Mohamed bin Zayed University of AI; Australian National University), Fahad Shahbaz Khan (Mohamed bin Zayed University of AI; Linköping University), and Mubarak Shah (University of Central Florida)</i>	
Towards Generalisable Video Moment Retrieval: Visual-Dynamic Injection to Image-Text Pre-Training	23045
<i>Dezhao Luo (Queen Mary University of London), Jiabo Huang (Queen Mary University of London), Shaogang Gong (Queen Mary University of London), Hailin Jin (Queen Mary University of London), and Yang Liu (Queen Mary University of London)</i>	
Hierarchical Video-Moment Retrieval and Step-Captioning	23056
<i>Abhay Zala (UNC Chapel Hill), Jaemin Cho (UNC Chapel Hill), Satwik Kottur (Meta AI2), Xilun Chen (Meta AI2), Barlas Oguz (Meta AI2), Yashar Mehdad (Meta AI2), and Mohit Bansal (UNC Chapel Hill)</i>	
HierVL: Learning Hierarchical Video-Language Embeddings	23066
<i>Kumar Ashutosh (UT Austin), Rohit Girdhar (FAIR, Meta AI), Lorenzo Torresani (FAIR, Meta AI), and Kristen Grauman (UT Austin; FAIR, Meta AI)</i>	

Learning Transferable Spatiotemporal Representations From Natural Script Knowledge	23079
<i>Ziyun Zeng (Tsinghua University; Applied Research Center (ARC), Tencent PCG), Yuying Ge (The University of Hong Kong), Xihui Liu (The University of Hong Kong), Bin Chen (Harbin Institute of Technology, Shenzhen), Ping Luo (The University of Hong Kong), Shu-Tao Xia (Tsinghua University), and Yixiao Ge (Applied Research Center (ARC), Tencent PCG)</i>	
WINNER: Weakly-Supervised Hierarchical decomposition and alignment for Spatio-temporal Video Grounding	23090
<i>Mengze Li (Zhejiang University), Han Wang (Zhejiang University), Wenqiao Zhang (National university of Singapore), Jiayu Miao (Zhejiang University), Zhou Zhao (Zhejiang University; Shanghai Institute for Advanced Study of Zhejiang University; Shanghai AI Laboratory), Shengyu Zhang (Zhejiang University), Wei Ji (National university of Singapore), and Fei Wu (Shanghai Institute for Advanced Study of Zhejiang University; Shanghai AI Laboratory)</i>	
Collaborative Static and Dynamic Vision-Language Streams for Spatio-Temporal Video Grounding	23100
<i>Zihang Lin (Sun Yat-sen University, China), Chaolei Tan (Sun Yat-sen University, China), Jian-Fang Hu (Sun Yat-sen University, China; Guangdong Province Key Laboratory of Information Security Technology, China; Ministry of Education, China), Zhi Jin (Sun Yat-sen University, China), Tiancai Ye (Tencent, China), and Wei-Shi Zheng (Sun Yat-sen University, China; Guangdong Province Key Laboratory of Information Security Technology, China; Ministry of Education, China)</i>	
Learning Action Changes by Measuring Verb-Adverb Textual Relationships	23110
<i>Davide Moltisanti (The University of Edinburgh, United Kingdom), Frank Keller (The University of Edinburgh, United Kingdom), Hakan Bilen (The University of Edinburgh, United Kingdom), and Laura Sevilla-Lara (The University of Edinburgh, United Kingdom)</i>	
LAVENDER: Unifying Video-Language Understanding As Masked Language Modeling	23119
<i>Linjie Li (Microsoft), Zhe Gan (Microsoft), Kevin Lin (Microsoft), Chung-Ching Lin (Microsoft), Zicheng Liu (Microsoft), Ce Liu (Microsoft), and Lijuan Wang (Microsoft)</i>	
DeCo: Decomposition and Reconstruction for Compositional Temporal Grounding via Coarse-To-Fine Contrastive Ranking	23130
<i>Lijin Yang (The University of Tokyo), Quan Kong (Woven by Toyota), Hsuan-Kung Yang (Woven by Toyota), Wadim Kehl (Woven by Toyota), Yoichi Sato (The University of Tokyo), and Norimasa Kobori (Woven by Toyota)</i>	

CVT-SLR: Contrastive Visual-Textual Transformation for Sign Language Recognition With Variational Alignment	23141
<i>Jiangbin Zheng (Research Center for Industries of the Future, Westlake University), Yile Wang (Research Center for Industries of the Future, Westlake University; Institute for AI Industry Research (AIR), Tsinghua University), Cheng Tan (Research Center for Industries of the Future, Westlake University), Siyuan Li (Research Center for Industries of the Future, Westlake University), Ge Wang (Research Center for Industries of the Future, Westlake University), Jun Xia (Research Center for Industries of the Future, Westlake University), Yidong Chen (Xiamen University), and Stan Z. Li (Research Center for Industries of the Future, Westlake University)</i>	
Joint Visual Grounding and Tracking With Natural Language Specification	23151
<i>Li Zhou (Harbin Institute of Technology, Shenzhen), Zikun Zhou (Peng Cheng Laboratory; Harbin Institute of Technology, Shenzhen), Kaige Mao (Harbin Institute of Technology, Shenzhen), and Zhenyu He (Harbin Institute of Technology, Shenzhen)</i>	
Accelerating Vision-Language Pretraining With Free Language Modeling	23161
<i>Teng Wang (Southern University of Science and Technology; The University of Hong Kong), Yixiao Ge (ARC Lab), Feng Zheng (Southern University of Science and Technology; Peng Cheng Laboratory), Ran Cheng (Southern University of Science and Technology), Ying Shan (ARC Lab), Xiaohu Qie (Tencent PCG), and Ping Luo (The University of Hong Kong; Shanghai AI Laboratory)</i>	
CoWs on Pasture: Baselines and Benchmarks for Language-Driven Zero-Shot Object Navigation ...	23171
<i>Samir Yitzhak Gadre (Columbia University), Mitchell Wortsman (University of Washington), Gabriel Ilharco (University of Washington), Ludwig Schmidt (University of Washington), and Shuran Song (Columbia University)</i>	
Where We Are and What We're Looking At: Query Based Worldwide Image Geo-Localization Using Hierarchies and Scenes	23182
<i>Brandon Clark (University of Central Florida, USA), Alec Kerrigan (University of Central Florida, USA), Parth Parag Kulkarni (University of Central Florida, USA), Vicente Vivanco Cepeda (University of Central Florida, USA), and Mubarak Shah (University of Central Florida, USA)</i>	
ANetQA: A Large-Scale Benchmark for Fine-Grained Compositional Reasoning Over Untrimmed Videos	23191
<i>Zhou Yu (Hangzhou Dianzi University, China), Lixiang Zheng (Hangzhou Dianzi University, China), Zhou Zhao (Zhejiang University, China), Fei Wu (Zhejiang University, China), Jianping Fan (Hangzhou Dianzi University, China; AI Lab at Lenovo Research, China), Kui Ren (Zhejiang University, China), and Jun Yu (Hangzhou Dianzi University, China)</i>	

MetaCLUE: Towards Comprehensive Visual Metaphors Research	23201
<i>Arjun R. Akula (Google), Brendan Driscoll (Google), Pradyumna Narayana (Google), Soravit Changpinyo (Google), Zhiwei Jia (Google), Suyash Damle (Google), Garima Pruthi (Google), Sugato Basu (Google), Leonidas Guibas (Google), William T. Freeman (Google), Yuanzhen Li (Google), and Varun Jampani (Google)</i>	
GeoVLN: Learning Geometry-Enhanced Visual Representation With Slot Attention for Vision-and-Language Navigation	23212
<i>Jingyang Huo (Fudan University), Qiang Sun (Fudan University), Boyan Jiang (Fudan University), Haitao Lin (Fudan University), and Yanwei Fu (Fudan University)</i>	
Being Comes From Not-Being: Open-Vocabulary Text-to-Motion Generation With Wordless Training	23222
<i>Junfan Lin (Sun Yat-sen University The Hong Kong Polytechnic University), Jianlong Chang (Huawei Cloud), Lingbo Liu (Kong Polytechnic University), Guanbin Li (Sun Yat-sen University The Hong Kong Polytechnic University), Liang Lin (Sun Yat-sen University The Hong Kong Polytechnic University), Qi Tian (Huawei Cloud), and Chang-Wen Chen (Kong Polytechnic University)</i>	
LASP: Text-to-Text Optimization for Language-Aware Soft Prompting of Vision & Language Models	23232
<i>Adrian Bulat (Samsung AI Cambridge; Technical University of Iasi) and Georgios Tzimiropoulos (Samsung AI Cambridge; Queen Mary University of London)</i>	
Position-Guided Text Prompt for Vision-Language Pre-Training	23242
<i>Jinpeng Wang (National University of Singapore), Pan Zhou (Sea AI Lab), Mike Zheng Shou (National University of Singapore), and Shuicheng Yan (Sea AI Lab)</i>	
Intrinsic Physical Concepts Discovery With Object-Centric Predictive Models	23252
<i>Qu Tang (University of Chinese Academy of Sciences; MAIS, Institute of Automation, Chinese Academy of Sciences), Xiangyu Zhu (University of Chinese Academy of Sciences; MAIS, Institute of Automation, Chinese Academy of Sciences), Zhen Lei (University of Chinese Academy of Sciences; MAIS, Institute of Automation, Chinese Academy of Sciences; Centre for Artificial Intelligence and Robotics, Hong Kong Institute of Science & Innovation), and Zhaoxiang Zhang (University of Chinese Academy of Sciences; MAIS, Institute of Automation, Chinese Academy of Sciences; Centre for Artificial Intelligence and Robotics, Hong Kong Institute of Science & Innovation)</i>	
MAP: Multimodal Uncertainty-Aware Vision-Language Pre-Training Model	23262
<i>Yatai Ji (Tsinghua University), Junjie Wang (Waseda University), Yuan Gong (Tsinghua University), Lin Zhang (IDEA), Yanru Zhu (Tsinghua University), Hongfa Wang (Tencent TEG), Jiaying Zhang (IDEA), Tetsuya Sakai (Waseda University), and Yujie Yang (Tsinghua University)</i>	
CLAMP: Prompt-Based Contrastive Learning for Connecting Language and Animal Pose	23272
<i>Xu Zhang (The University of Sydney, Australia), Wen Wang (Zhejiang University, China), Zhe Chen (The University of Sydney, Australia), Yufei Xu (The University of Sydney, Australia), Jing Zhang (The University of Sydney, Australia), and Dacheng Tao (The University of Sydney, Australia)</i>	

Teacher-Generated Spatial-Attention Labels Boost Robustness and Accuracy of Contrastive Models	23282
<i>Yushi Yao (Waymo), Chang Ye (Google), Junfeng He (Google), and Gamaleldin F. Elsayed (Google)</i>	
DegAE: A New Pretraining Paradigm for Low-Level Vision	23292
<i>Yihao Liu (Shanghai Artificial Intelligence Laboratory; Chinese Academy of Sciences; University of Chinese Academy of Sciences), Jingwen He (Shanghai Artificial Intelligence Laboratory), Jinjin Gu (Shanghai Artificial Intelligence Laboratory; The University of Sydney), Xiangtao Kong (Shanghai Artificial Intelligence Laboratory; Chinese Academy of Sciences; University of Chinese Academy of Sciences), Yu Qiao (Shanghai Artificial Intelligence Laboratory; Chinese Academy of Sciences), and Chao Dong (Shanghai Artificial Intelligence Laboratory; Chinese Academy of Sciences)</i>	
RILS: Masked Visual Reconstruction in Language Semantic Space	23304
<i>Shusheng Yang (Huazhong University of Science & Technology), Yixiao Ge (ARC Lab), Kun Yi (ARC Lab), Dian Li (Foundation Technology Center), Ying Shan (ARC Lab), Xiaohu Qie (Tencent PCG), and Xinggang Wang (Huazhong University of Science & Technology)</i>	
Learning Geometry-Aware Representations by Sketching	23315
<i>Hyundo Lee (AI Institute, Seoul National University), Inwoo Hwang (AI Institute, Seoul National University), Hyunsung Go (AI Institute, Seoul National University), Won-Seok Choi (AI Institute, Seoul National University), Kibeom Kim (AI Institute, Seoul National University), and Byoung-Tak Zhang (AI Institute, Seoul National University)</i>	
SketchXAI: A First Look at Explainability for Human Sketches	23327
<i>Zhiyu Qu (SkctchX, CVSSP, University of Surrey; iFlyTek-Surrey Joint Research Centre on Artificial Intelligence), Yulia Gryaditskaya (SkctchX, CVSSP, University of Surrey), Ke Li (SkctchX, CVSSP, University of Surrey; Beijing University of Posts and Telecommunications), Kaiyue Pang (SkctchX, CVSSP, University of Surrey), Tao Xiang (SkctchX, CVSSP, University of Surrey; iFlyTek-Surrey Joint Research Centre on Artificial Intelligence), and Yi-Zhe Song (SkctchX, CVSSP, University of Surrey; iFlyTek-Surrey Joint Research Centre on Artificial Intelligence)</i>	
MAGVLT: Masked Generative Vision-and-Language Transformer	23338
<i>Sungwoong Kim (Korea University, South Korea), Daejin Jo (Kakao Brain, South Korea), Donghoon Lee (Kakao Brain, South Korea), and Jongmin Kim (Kakao Brain, South Korea)</i>	
Zero-Shot Everything Sketch-Based Image Retrieval, and in Explainable Style	23349
<i>Fengyin Lin (Beijing University of Posts and Telecommunications), Mingkang Li (Beijing University of Posts and Telecommunications), Da Li (Samsung AI Centre, Cambridge), Timothy Hospedales (Samsung AI Centre, Cambridge; University of Edinburgh), Yi-Zhe Song (CVSSP, University of Surrey), and Yonggang Qi (Beijing University of Posts and Telecommunications)</i>	

Semantic-Conditional Diffusion Networks for Image Captioning	23359
<i>Jianjie Luo (Sun Yat-sen University, China; HiDream.ai Inc.), Yehao Li (Sun Yat-sen University, China; HiDream.ai Inc.), Yingwei Pan (Sun Yat-sen University, China; HiDream.ai Inc.), Ting Yao (Sun Yat-sen University, China; HiDream.ai Inc.), Jianlin Feng (Sun Yat-sen University, China; HiDream.ai Inc.), Hongyang Chao (Sun Yat-sen University, China; HiDream.ai Inc.), and Tao Mei (Sun Yat-sen University, China; HiDream.ai Inc.)</i>	
REVEAL: Retrieval-Augmented Visual-Language Pre-Training With Multi-Source Multimodal Knowledge Memory	23369
<i>Ziniu Hu (University of California, Los Angeles), Ahmet Iscen (Google Research), Chen Sun (Google Research), Zirui Wang (Google Research), Kai-Wei Chang (University of California, Los Angeles), Yizhou Sun (University of California, Los Angeles), Cordelia Schmid (Google Research), David A. Ross (Google Research), and Alireza Fathi (Google Research)</i>	
Variational Distribution Learning for Unsupervised Text-to-Image Generation	23380
<i>Minsoo Kang (ECE), Doyup Lee (Seoul National University), Jiseob Kim (Seoul National University), Saehoon Kim (Seoul National University), and Bohyung Han (ECE; Kakao Brain)</i>	
Scaling Language-Image Pre-Training via Masking	23390
<i>Yanghao Li (Meta AI, FAIR), Haoqi Fan (Meta AI, FAIR), Ronghang Hu (Meta AI, FAIR), Christoph Feichtenhofer (Meta AI, FAIR), and Kaiming He (Meta AI, FAIR)</i>	
LANIT: Language-Driven Image-to-Image Translation for Unlabeled Data	23401
<i>Jihye Park (Korea University, Korea), Sunwoo Kim (Korea University, Korea), Soohyun Kim (Korea University, Korea), Seokju Cho (Korea University, Korea), Jaejun Yoo (UNIST, Korea), Youngjung Uh (Korea University, Korea), and Seungryong Kim (Korea University, Korea)</i>	
Revisiting Self-Similarity: Structural Embedding for Image Retrieval	23412
<i>Seongwon Lee (Yonsei University, Seoul, Korea), Suhyeon Lee (Yonsei University, Seoul, Korea), Hongje Seong (Yonsei University, Seoul, Korea), and Euntai Kim (Yonsei University, Seoul, Korea)</i>	
Improving Cross-Modal Retrieval With Set of Diverse Embeddings	23422
<i>Dongwon Kim (Dept. of CSE, POSTECH), Namyup Kim (Dept. of CSE, POSTECH), and Suha Kwak (Graduate School of AI, POSTECH)</i>	
Masked Autoencoding Does Not Help Natural Language Supervision at Scale	23432
<i>Floris Weers (NA), Vaishaal Shankar (NA), Angelos Katharopoulos (NA), Yinfei Yang (NA), and Tom Gunter (NA)</i>	
Few-Shot Learning With Visual Distribution Calibration and Cross-Modal Distribution Alignment	23445
<i>Runqi Wang (Beihang University; Huawei Noah's Ark Lab), Hao Zheng (Huawei Noah's Ark Lab; Tokyo Institute of Technology), Xiaoyue Duan (Beihang University), Jianzhuang Liu (Huawei Noah's Ark Lab), Yuning Lu (Huawei Noah's Ark Lab; University of Science and Technology of China), Tian Wang (Beihang University), Songcen Xu (Huawei Noah's Ark Lab), and Baochang Zhang (Beihang University; Zhongguancun Laboratory)</i>	

Deep Hashing With Minimal-Distance-Separated Hash Centers	23455
<i>Liangdao Wang (Sun Yat-Sen University), Yan Pan (Sun Yat-Sen University), Cong Liu (Sun Yat-Sen University), Hanjiang Lai (Sun Yat-Sen University), Jian Yin (Sun Yat-Sen University), and Ye Liu (Sun Yat-Sen University; Big Data Department, Lizhi Inc.)</i>	
ConZIC: Controllable Zero-Shot Image Captioning by Sampling-Based Polishing	23465
<i>Zequn Zeng (Xidian University, China), Hao Zhang (Xidian University, China), Ruiying Lu (Xidian University, China), Dongsheng Wang (Xidian University, China), Bo Chen (Xidian University, China), and Zhengjue Wang (Xidian University, China)</i>	
Learning To Name Classes for Vision and Language Models	23477
<i>Sarah Parisot (Huawei Noah's Ark Lab), Yongxin Yang (Huawei Noah's Ark Lab), and Steven McDonagh (Huawei Noah's Ark Lab)</i>	
Data-Efficient Large Scale Place Recognition With Graded Similarity Supervision	23487
<i>María Leyva-Vallina (University of Groningen; University of Twente), Nicola Strisciuglio (University of Groningen), and Nicolai Petkov (University of Groningen; University of Twente)</i>	
DetCLIPv2: Scalable Open-Vocabulary Object Detection Pre-Training via Word-Region Alignment	23497
<i>Lewei Yao (Hong Kong University of Science and Technology; Huawei Noah's Ark Lab), Jianhua Han (Huawei Noah's Ark Lab), Xiaodan Liang (Shenzhen Campus of Sun Yat-Sen University), Dan Xu (Hong Kong University of Science and Technology), Wei Zhang (Hong Kong University of Science and Technology), Zhenguo Li (Hong Kong University of Science and Technology), and Hang Xu (Hong Kong University of Science and Technology)</i>	
HOICLIP: Efficient Knowledge Transfer for HOI Detection With Vision-Language Models	23507
<i>Shan Ning (ShanghaiTech University, China), Longtian Qiu (ShanghaiTech University, China), Yongfei Liu (ByteDance Inc.), and Xuming He (ShanghaiTech University, China; Shanghai Engineering Research Center of Intelligent Vision and Imaging)</i>	
OvarNet: Towards Open-Vocabulary Object Attribute Recognition	23518
<i>Keyan Chen (Beihang University), Xiaolong Jiang (Xiaohongshu Inc), Yao Hu (Xiaohongshu Inc), Xu Tang (Xiaohongshu Inc), Yan Gao (Xiaohongshu Inc), Jianqi Chen (Beihang University), and Weidi Xie (CMIC, Shanghai Jiao Tong University; Shanghai AI Laboratory)</i>	
NeRF-RPN: A General Framework for Object Detection in NeRFs	23528
<i>Benran Hu (The Hong Kong University of Science and Technology), Junkai Huang (The Hong Kong University of Science and Technology), Yichen Liu (The Hong Kong University of Science and Technology), Yu-Wing Tai (The Hong Kong University of Science and Technology; Kuaishou Technology), and Chi-Keung Tang (The Hong Kong University of Science and Technology)</i>	

Mask-Free OVIS: Open-Vocabulary Instance Segmentation Without Manual Mask Annotations	23539
<i>Vibashan VS (Johns Hopkins University), Ning Yu (Salesforce Research), Chen Xing (Salesforce Research), Can Qin (Northeastern University), Mingfei Gao (Salesforce Research), Juan Carlos Niebles (Salesforce Research), Vishal M. Patel (Johns Hopkins University), and Ran Xu (Salesforce Research)</i>	
GP-VTON: Towards General Purpose Virtual Try-On via Collaborative Local-Flow Global-Parsing Learning	23550
<i>Zhenyu Xie (Shenzhen Campus of Sun Yat-Sen University), Zaiyu Huang (Shenzhen Campus of Sun Yat-Sen University), Xin Dong (ByteDance), Fuwei Zhao (Shenzhen Campus of Sun Yat-Sen University), Haoye Dong (Carnegie Mellon University), Xijin Zhang (ByteDance), Feida Zhu (ByteDance), and Xiaodan Liang (Shenzhen Campus of Sun Yat-Sen University; Peng Cheng Laboratory)</i>	
Decomposed Soft Prompt Guided Fusion Enhancing for Compositional Zero-Shot Learning	23560
<i>Xiaocheng Lu (The Hong Kong Polytechnic University), Song Guo (The Hong Kong Polytechnic University; The Hong Kong Polytechnic University Shenzhen Research Institute), Ziming Liu (The Hong Kong Polytechnic University), and Jingcai Guo (The Hong Kong Polytechnic University; The Hong Kong Polytechnic University Shenzhen Research Institute)</i>	
Contrastive Grouping With Transformer for Referring Image Segmentation	23570
<i>Jiajin Tang (ShanghaiTech University), Ge Zheng (ShanghaiTech University), Cheng Shi (ShanghaiTech University), and Sibe Yang (ShanghaiTech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging)</i>	
Semantic Prompt for Few-Shot Image Recognition	N/A
<i>Wentao Chen (University of Science and Technology of China; Center for Research on Intelligent Perception and Computing, NLPR, CASIA), Chenyang Si (Nanyang Technological University, Singapore), Zhang Zhang (Center for Research on Intelligent Perception and Computing, NLPR, CASIA; University of Chinese Academy of Sciences), Liang Wang (Center for Research on Intelligent Perception and Computing, NLPR, CASIA; University of Chinese Academy of Sciences), Zilei Wang (University of Science and Technology of China), and Tieniu Tan (University of Science and Technology of China; Center for Research on Intelligent Perception and Computing, NLPR, CASIA; University of Chinese Academy of Sciences)</i>	
GRES: Generalized Referring Expression Segmentation	23592
<i>Chang Liu (Nanyang Technological University, Singapore), Henghui Ding (Nanyang Technological University, Singapore), and Xudong Jiang (Nanyang Technological University, Singapore)</i>	
Network-Free, Unsupervised Semantic Segmentation With Synthetic Images	23602
<i>Qianli Feng (Amazon), Raghudeep Gadde (Amazon), Wentong Liao (Amazon), Eduard Ramon (Amazon), and Alex Martinez (Amazon)</i>	
Few-Shot Semantic Image Synthesis With Class Affinity Transfer	23611
<i>Marlène Careil (LTCI, Telecom Paris, IP Paris; Meta AI), Jakob Verbeek (Meta AI), and Stéphane Lathuilière (LTCI, Telecom Paris, IP Paris)</i>	

Ultra-High Resolution Segmentation With Ultra-Rich Context: A Novel Benchmark	23621
<i>Deyi Ji (University of Science and Technology of China; Alibaba Group), Feng Zhao (University of Science and Technology of China), Hongtao Lu (Shanghai Jiao Tong University), Mingyuan Tao (Alibaba Group), and Jieping Ye (Alibaba Group)</i>	
Content-Aware Token Sharing for Efficient Semantic Segmentation With Vision Transformers	23631
<i>Chenyang Lu (Eindhoven University of Technology), Daan de Geus (Eindhoven University of Technology), and Gijs Dubbelman (Eindhoven University of Technology)</i>	
Hierarchical Dense Correlation Distillation for Few-Shot Segmentation	23641
<i>Bohao Peng (The Chinese University of Hong Kong), Zhuotao Tian (SmartMore), Xiaoyang Wu (The University of Hong Kong), Chengyao Wang (The Chinese University of Hong Kong), Shu Liu (SmartMore), Jingyong Su (Harbin Institute of Technology, Shenzhen), and Jiaya Jia (The Chinese University of Hong Kong; SmartMore)</i>	
On Calibrating Semantic Segmentation Models: Analyses and an Algorithm	23652
<i>Dongdong Wang (University of Central Florida), Boqing Gong (Google Research), and Liqiang Wang (University of Central Florida)</i>	
FastInst: A Simple Query-Based Model for Real-Time Instance Segmentation	23663
<i>Junjie He (Alibaba Group), Pengyu Li (Alibaba Group), Yifeng Geng (Alibaba Group), and Xuansong Xie (Alibaba Group)</i>	
Out-of-Candidate Rectification for Weakly Supervised Semantic Segmentation	23673
<i>Zesen Cheng (Peking University, China; Peking University Shenzhen Graduate School, China), Pengchong Qiao (Peking University, China; Peking University Shenzhen Graduate School, China; Peng Cheng Laboratory, China), Kehan Li (Peking University, China; Peking University Shenzhen Graduate School, China), Siheng Li (Tsinghua University), Pengxu Wei (Sun Yat-Sen University), Xiangyang Ji (Tsinghua University), Li Yuan (Peking University, China; Peng Cheng Laboratory, China), Chang Liu (Tsinghua University), and Jie Chen (Peking University, China; Peking University Shenzhen Graduate School, China; Peng Cheng Laboratory, China)</i>	
Foundation Model Drives Weakly Incremental Learning for Semantic Segmentation	23685
<i>Chaohui Yu (Alibaba Group), Qiang Zhou (Alibaba Group), Jingliang Li (University of the Chinese Academy of Sciences), Jianlong Yuan (Alibaba Group), Zhibin Wang (Alibaba Group), and Fan Wang (Alibaba Group)</i>	
Long-Tailed Visual Recognition via Self-Heterogeneous Integration With Knowledge Excavation	23695
<i>Yan Jin (Xiamen University, China), Mengke Li (Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ), China), Yang Lu (Xiamen University, China), Yiu-ming Cheung (Hong Kong Baptist University, China), and Hanzi Wang (Xiamen University, China)</i>	
Instance-Specific and Model-Adaptive Supervision for Semi-Supervised Semantic Segmentation....	23705
<i>Zhen Zhao (University of Sydney; Baidu VIS), Sifan Long (Baidu VIS; Jilin University), Jimin Pi (Baidu VIS), Jingdong Wang (Baidu VIS), and Luping Zhou (University of Sydney)</i>	

Active Finetuning: Exploiting Annotation Budget in the Pretraining-Finetuning Paradigm	23715
<i>Yichen Xie (University of California, Berkeley), Han Lu (Shanghai Jiao Tong University), Junchi Yan (Shanghai Jiao Tong University), Xiaokang Yang (Shanghai Jiao Tong University), Masayoshi Tomizuka (University of California, Berkeley), and Wei Zhan (University of California, Berkeley)</i>	
IDG: A Framework To Eliminate Explanation Noise From Integrated Gradients	23725
<i>Ruo Yang (Department of Computer Science, Illinois Institute of Technology, USA), Binghui Wang (Department of Computer Science, Illinois Institute of Technology, USA), and Mustafa Bilgic (Department of Computer Science, Illinois Institute of Technology, USA)</i>	
Weakly Supervised Posture Mining for Fine-Grained Classification	23735
<i>Zhenchao Tang (Sun Yat-sen University), Hualin Yang (Sun Yat-sen University), and Calvin Yu-Chian Chen (Sun Yat-sen University; China Medical University Hospital; Asia University)</i>	
Vision Transformers Are Good Mask Auto-Labelers	23745
<i>Shiyi Lan (NVIDIA), Xitong Yang (Meta AI, FAIR), Zhiding Yu (NVIDIA), Zuxuan Wu (Fudan University), Jose M. Alvarez (NVIDIA), and Anima Anandkumar (NVIDIA; Caltech)</i>	
Enhanced Training of Query-Based Object Detection via Selective Query Recollection	23756
<i>Fangyi Chen (Carnegie Mellon University), Han Zhang (Carnegie Mellon University), Kai Hu (Carnegie Mellon University), Yu-Kai Huang (Carnegie Mellon University), Chenchen Zhu (Meta AI), and Marios Savvides (Meta AI)</i>	
Box-Level Active Detection	23766
<i>Mengyao Lyu (Tsinghua University; BNRist; Hangzhou Zhuoxi Institute of Brain and Intelligence), Jundong Zhou (Tsinghua University; BNRist; Hangzhou Zhuoxi Institute of Brain and Intelligence), Hui Chen (Tsinghua University; BNRist), Yijie Huang (OPPO Research Institute), Dongdong Yu (OPPO Research Institute), Yaqian Li (OPPO Research Institute), Yandong Guo (OPPO Research Institute), Yuchen Guo (Tsinghua University; BNRist), Liuyu Xiang (Beijing University of Posts and Telecommunications), and Guiguang Ding (Tsinghua University; BNRist)</i>	
CIGAR: Cross-Modality Graph Reasoning for Domain Adaptive Object Detection	23776
<i>Yabo Liu (Harbin Institute of Technology, Shenzhen; Peng Cheng Laboratory), Jinghua Wang (Harbin Institute of Technology, Shenzhen), Chao Huang (Shenzhen Campus of Sun Yat-sen University), Yaowei Wang (Peng Cheng Laboratory), and Yong Xu (Harbin Institute of Technology, Shenzhen; Peng Cheng Laboratory)</i>	
DA-DETR: Domain Adaptive Detection Transformer With Information Fusion	23787
<i>Jingyi Zhang (Nanyang Technological University), Jiaying Huang (Nanyang Technological University), Zhipeng Luo (Nanyang Technological University; SenseTime Research), Gongjie Zhang (Nanyang Technological University; Black Sesame Technologies), Xiaoqin Zhang (Wenzhou University), and Shijian Lu (Nanyang Technological University)</i>	

Continual Detection Transformer for Incremental Object Detection	23799
<i>Yaoyao Liu (Max Planck Institute for Informatics, Saarland Informatics Campus), Bernt Schiele (Max Planck Institute for Informatics, Saarland Informatics Campus), Andrea Vedaldi (University of Oxford), and Christian Rupprecht (University of Oxford)</i>	
Semi-DETR: Semi-Supervised Object Detection With Detection Transformers	23809
<i>Jiacheng Zhang (Sun Yat-sen University, China; department of Computer Vision Technology (VIS), Baidu Inc., China), Xiangru Lin (Department of Computer Vision Technology (VIS), Baidu Inc., China), Wei Zhang (Department of Computer Vision Technology (VIS), Baidu Inc., China), Kuo Wang (Sun Yat-sen University, China), Xiao Tan (Department of Computer Vision Technology (VIS), Baidu Inc., China), Junyu Han (Department of Computer Vision Technology (VIS), Baidu Inc., China), Errui Ding (Department of Computer Vision Technology (VIS), Baidu Inc., China), Jingdong Wang (Department of Computer Vision Technology (VIS), Baidu Inc., China), and Guanbin Li (Sun Yat-sen University, China)</i>	
Hierarchical Supervision and Shuffle Data Augmentation for 3D Semi-Supervised Object Detection	23819
<i>Chuangdong Liu (Chongqing University of Posts and Telecommunications, China; Chongqing Key Laboratory of Signal and Information Processing, China), Chenqiang Gao (Chongqing University of Posts and Telecommunications, China; Chongqing Key Laboratory of Signal and Information Processing, China), Fangcen Liu (Chongqing University of Posts and Telecommunications, China; Chongqing Key Laboratory of Signal and Information Processing, China), Pengcheng Li (Chongqing University of Posts and Telecommunications, China; Chongqing Key Laboratory of Signal and Information Processing, China), Deyu Meng (Xi'an Jiaotong University, China; Macau University of Science and Technology, Macau), and Xinbo Gao (Chongqing University of Posts and Telecommunications, China)</i>	
Harmonious Teacher for Cross-Domain Object Detection	23829
<i>Jinhong Deng (University of Electronic Science and Technology of China), Dongli Xu (University of Sydney), Wen Li (Shenzhen Institute for Advanced Study, UESTC), and Lixin Duan (Shenzhen Institute for Advanced Study, UESTC; Sichuan Provincial People's Hospital, UESTC)</i>	
Contrastive Mean Teacher for Domain Adaptive Object Detectors	23839
<i>Shengcao Cao (University of Illinois at Urbana-Champaign), Dhiraj Joshi (IBM Research), Liang-Yan Gui (University of Illinois at Urbana-Champaign), and Yu-Xiong Wang (University of Illinois at Urbana-Champaign)</i>	
Out-of-Distributed Semantic Pruning for Robust Semi-Supervised Learning	23849
<i>Yu Wang (Peking University, China; Peking University Shenzhen), Pengchong Qiao (Peking University, China; Peng Cheng Laboratory, China; Peking University Shenzhen), Chang Liu (Tsinghua University, China), Guoli Song (Peng Cheng Laboratory, China; Peking University Shenzhen), Xiawu Zheng (Peng Cheng Laboratory, China; Peking University Shenzhen), and Jie Chen (Peking University, China; Peng Cheng Laboratory, China; Peking University Shenzhen)</i>	

(ML)2P-Encoder: On Exploration of Channel-Class Correlation for Multi-Label Zero-Shot Learning	23859
<i>Ziming Liu (The Hong Kong Polytechnic University, China), Song Guo (The Hong Kong Polytechnic University, China; The Hong Kong Polytechnic University Shenzhen Research Institute, China), Xiaocheng Lu (The Hong Kong Polytechnic University, China), Jingcai Guo (The Hong Kong Polytechnic University, China; The Hong Kong Polytechnic University Shenzhen Research Institute, China), Jiewei Zhang (The Hong Kong Polytechnic University, China), Yue Zeng (The Hong Kong Polytechnic University, China), and Fushuo Huo (The Hong Kong Polytechnic University, China)</i>	
MagicNet: Semi-Supervised Multi-Organ Segmentation via Magic-Cube Partition and Recovery	23869
<i>Duowen Chen (East China Normal University), Yunhao Bai (East China Normal University), Wei Shen (Shanghai Jiao Tong University), Qingli Li (East China Normal University), Lequan Yu (The University of Hong Kong), and Yan Wang (East China Normal University)</i>	
Devil Is in the Queries: Advancing Mask Transformers for Real-World Medical Image Segmentation and Out-of-Distribution Localization	23879
<i>Mingze Yuan (Alibaba Group; Peking University), Yingda Xia (Alibaba Group), Hexin Dong (Alibaba Group; Peking University), Zifan Chen (Peking University), Jiawen Yao (Alibaba Group), Mingyan Qiu (Alibaba Group), Ke Yan (Alibaba Group), Xiaoli Yin (Shengjing Hospital), Yu Shi (Shengjing Hospital), Xin Chen (Guangdong Province People's Hospital), Zaiyi Liu (Guangdong Province People's Hospital), Bin Dong (Peking University; Peking University Changsha Institute for Computing and Digital Economy), Jingren Zhou (Alibaba Group), Le Lu (Alibaba Group), Ling Zhang (Alibaba Group), and Li Zhang (Peking University)</i>	
SQUID: Deep Feature In-Painting for Unsupervised Anomaly Detection	23890
<i>Tiange Xiang (University of Sydney), Yixiao Zhang (Johns Hopkins University), Yongyi Lu (Johns Hopkins University), Alan L. Yuille (Johns Hopkins University), Chaoyi Zhang (University of Sydney), Weidong Cai (University of Sydney), and Zongwei Zhou (Johns Hopkins University)</i>	
OCELOT: Overlapped Cell on Tissue Dataset for Histopathology	23902
<i>Jeongun Ryu (Lunit Inc.), Aaron Valero Puche (Lunit Inc.), JaeWoong Shin (Lunit Inc.), Seonwook Park (Lunit Inc.), Biagio Brattoli (Lunit Inc.), Jinhee Lee (Lunit Inc.), Wonkyung Jung (Lunit Inc.), Soo Ick Cho (Lunit Inc.), Kyunghyun Paeng (Lunit Inc.), Chan-Young Ock (Lunit Inc.), Donggeun Yoo (Lunit Inc.), and Sérgio Pereira (Lunit Inc.)</i>	
DeGPR: Deep Guided Posterior Regularization for Multi-Class Cell Detection and Counting .	23913
<i>Aayush Kumar Tyagi (IIT Delhi), Chirag Mohapatra (IIT Delhi), Prasenjit Das (AIIMS, New Delhi), Govind Makharia (AIIMS, New Delhi), Lalita Mehra (AIIMS, New Delhi), Prathosh AP (IISc, Bangalore), and Mausam Mausam (IIT Delhi)</i>	

Best of Both Worlds: Multimodal Contrastive Learning With Tabular and Imaging Data	23924
<i>Paul Hager (Technical University of Munich; Klinikum Rechts der Isar), Martin J. Menten (Technical University of Munich; Klinikum Rechts der Isar; Imperial College London), and Daniel Rueckert (Technical University of Munich; Klinikum Rechts der Isar; Imperial College London)</i>	
RankMix: Data Augmentation for Weakly Supervised Learning of Classifying Whole Slide Images With Diverse Sizes and Imbalanced Categories	23936
<i>Yuan-Chih Chen (Apple) and Chun-Shien Lu (Apple)</i>	
GEN: Pushing the Limits of Softmax-Based Out-of-Distribution Detection	23946
<i>Xixi Liu (Chalmers University of Technology), Yaroslava Lochman (Chalmers University of Technology), and Christopher Zach (Chalmers University of Technology)</i>	
Discriminating Known From Unknown Objects via Structure-Enhanced Recurrent Variational AutoEncoder	23956
<i>Aming Wu (Xidian University, China) and Cheng Deng (Xidian University, China)</i>	
Sample-Level Multi-View Graph Clustering	23966
<i>Yuze Tan (Sichuan University), Yixi Liu (Sichuan University), Shudong Huang (Sichuan University), Wentao Feng (Sichuan University), and Jiancheng Lv (Sichuan University)</i>	
On the Effects of Self-Supervision and Contrastive Alignment in Deep Multi-View Clustering	23976
<i>Daniel J. Trosten (UiT The Arctic University of Norway), Sigurd Løkse (UiT The Arctic University of Norway), Robert Jenssen (UiT The Arctic University of Norway), and Michael C. Kampffmeyer (UiT The Arctic University of Norway)</i>	
Deep Fair Clustering via Maximizing and Minimizing Mutual Information: Theory, Algorithm and Metric	23986
<i>Pengxin Zeng (Sichuan University, China), Yunfan Li (Sichuan University, China), Peng Hu (Sichuan University, China), Dezhong Peng (Sichuan University, China), Jiancheng Lv (Sichuan University, China), and Xi Peng (Sichuan University, China)</i>	
Transductive Few-Shot Learning With Prototype-Based Label Propagation by Iterative Graph Refinement	23996
<i>Hao Zhu (Australian National University) and Piotr Koniusz (Data61*CSIRO)</i>	
Open-Set Likelihood Maximization for Few-Shot Learning	24007
<i>Malik Boudiaf (ÉTS Montreal), Etienne Bennequin (Sicara - MICS), Myriam Tami (MICS), Antoine Toubhans (Sicara), Pablo Piantanida (CNRS), Celine Hudelot (MICS), and Ismail Ben Ayed (ÉTS Montreal)</i>	
HyperMatch: Noise-Tolerant Semi-Supervised Learning via Relaxed Contrastive Constraint .	24017
<i>Beitong Zhou (Hikvision Research Institute), Jing Lu (Hikvision Research Institute), Kerui Liu (Hikvision Research Institute), Yunlu Xu (Hikvision Research Institute), Zhazhan Cheng (Hikvision Research Institute), and Yi Niu (Hikvision Research Institute)</i>	

Token Boosting for Robust Self-Supervised Visual Transformer Pre-Training	24027
<i>Tianjiao Li (Singapore University of Technology and Design), Lin Geng Foo (Singapore University of Technology and Design), Ping Hu (Boston University), Xindi Shang (ByteDance), Hossein Rahmani (Lancaster University), Zehuan Yuan (ByteDance), and Jun Liu (Singapore University of Technology and Design)</i>	
Difficulty-Based Sampling for Debaised Contrastive Representation Learning	24039
<i>Taeuk Jang (Purdue University) and Xiaoqian Wang (Purdue University)</i>	
Improving Selective Visual Question Answering by Learning From Your Peers	24049
<i>Corentin Dancette (FAIR, Meta AI; Sorbonne Universite), Spencer Whitehead (FAIR, Meta AI), Rishabh Maheshwary (FAIR, Meta AI), Ramakrishna Vedantam (FAIR, Meta AI), Stefan Scherer (Reality Labs Research, Meta), Xinlei Chen (Sorbonne Universite; Valeo.ai), Matthieu Cord (Sorbonne Universite), and Marcus Rohrbach (FAIR, Meta AI)</i>	
Superclass Learning With Representation Enhancement	24060
<i>Zeyu Gan (Key Lab of Data Engineering and Knowledge Engineering of MOE Renmin University of China; Renmin University of China, China), Suyun Zhao (Key Lab of Data Engineering and Knowledge Engineering of MOE Renmin University of China; Renmin University of China, China), Jinlong Kang (Renmin University of China, China), Liyuan Shang (Renmin University of China, China), Hong Chen (Key Lab of Data Engineering and Knowledge Engineering of MOE Renmin University of China; Renmin University of China, China), and Cuiping Li (Key Lab of Data Engineering and Knowledge Engineering of MOE Renmin University of China; Renmin University of China, China)</i>	
DISC: Learning From Noisy Labels via Dynamic Instance-Specific Selection and Correction ..	24070
<i>Yifan Li (Institute of Computing Technology, CAS, China; University of the Chinese Academy of Sciences, China), Hu Han (Institute of Computing Technology, CAS, China; University of the Chinese Academy of Sciences, China; Peng Cheng Laboratory, China), Shiguang Shan (Institute of Computing Technology, CAS, China; University of the Chinese Academy of Sciences, China; Peng Cheng Laboratory, China), and Xilin Chen (Institute of Computing Technology, CAS, China; University of the Chinese Academy of Sciences, China)</i>	
FCC: Feature Clusters Compression for Long-Tailed Visual Recognition	24080
<i>Jian Li (Jilin University), Ziyao Meng (University of Minho), Daqian Shi (University of Trento), Rui Song (Jilin University), Xiaolei Diao (University of Trento), Jingwen Wang (Jilin University), and Hao Xu (Jilin University)</i>	
Dynamically Instance-Guided Adaptation: A Backward-Free Approach for Test-Time Domain Adaptive Semantic Segmentation	24090
<i>Wei Wang (Western University), Zhun Zhong (University of Trento), Weijie Wang (University of Trento), Xi Chen (Huawei Noah's Ark Lab), Charles Ling (Western University), Boyu Wang (Western University), and Nicu Sebe (University of Trento)</i>	
Semi-Supervised Domain Adaptation With Source Label Adaptation	24100
<i>Yu-Chu Yu (National Taiwan University) and Hsuan-Tien Lin (National Taiwan University)</i>	

Adjustment and Alignment for Unbiased Open Set Domain Adaptation	24110
<i>Wuyang Li (City University of Hong Kong), Jie Liu (City University of Hong Kong), Bo Han (Hong Kong Baptist University), and Yixuan Yuan (The Chinese University of Hong Kong)</i>	
C-SFDA: A Curriculum Learning Aided Self-Training Framework for Efficient Source Free Domain Adaptation	24120
<i>Nazmul Karim (Department of ECE, UCF, USA), Niluthpol Chowdhury Mithun (SRI International, Princeton, USA), Abhinav Rajvanshi (SRI International, Princeton, USA), Han-pang Chiu (SRI International, Princeton, USA), Supun Samarasekera (SRI International, Princeton, USA), and Nazanin Rahnavard (Department of ECE, UCF, USA)</i>	
ALOFT: A Lightweight MLP-Like Architecture With Dynamic Low-Frequency Transform for Domain Generalization	24132
<i>Jintao Guo (Nanjing University), Na Wang (Nanjing University; Nanjing University), Lei Qi (Southeast University), and Yinghuan Shi (Nanjing University; Nanjing University)</i>	
Modality-Agnostic Debiasing for Single Domain Generalization	24142
<i>Sanqing Qu (Tongji University, HiDream.ai Inc.), Yingwei Pan (Tongji University, HiDream.ai Inc.), Guang Chen (Tongji University, HiDream.ai Inc.), Ting Yao (Tongji University, HiDream.ai Inc.), Changjun Jiang (Tongji University, HiDream.ai Inc.), and Tao Mei (Tongji University, HiDream.ai Inc.)</i>	
ActMAD: Activation Matching To Align Distributions for Test-Time-Training	24152
<i>Muhammad Jehanzeb Mirza (Institute for Computer Graphics and Vision, TU Graz, Austria; Christian Doppler Laboratory for Embedded Machine Learning), Pol Jané Soneira (Institute of Control Systems, KIT, Germany), Wei Lin (Institute for Computer Graphics and Vision, TU Graz, Austria; Christian Doppler Laboratory for Semantic 3D Computer Vision), Mateusz Kozinski (Institute for Computer Graphics and Vision, TU Graz, Austria), Horst Possegger (Institute for Computer Graphics and Vision, TU Graz, Austria), and Horst Bischof (Institute for Computer Graphics and Vision, TU Graz, Austria; Christian Doppler Laboratory for Embedded Machine Learning)</i>	
TIPI: Test Time Adaptation With Transformation Invariance	24162
<i>A. Tuan Nguyen (University of Oxford), Thanh Nguyen-Tang (Johns Hopkins University), Ser-Nam Lim (Meta AI), and Philip H.S. Torr (University of Oxford)</i>	
Improved Test-Time Adaptation for Domain Generalization	24172
<i>Liang Chen (The University of Adelaide), Yong Zhang (Tencent AI Lab), Yibing Song (AI3 Institute, Fudan University), Ying Shan (Tencent AI Lab), and Lingqiao Liu (The University of Adelaide)</i>	
Learning With Fantasy: Semantic-Aware Virtual Contrastive Constraint for Few-Shot Class-Incremental Learning	24183
<i>Zeyin Song (Peking University), Yifan Zhao (Peking University), Yujun Shi (National University of Singapore), Peixi Peng (Peking University; Peng Cheng Laboratory), Li Yuan (Peking University; Peng Cheng Laboratory), and Yonghong Tian (Peking University; Peng Cheng Laboratory)</i>	

NIFF: Alleviating Forgetting in Generalized Few-Shot Object Detection via Neural Instance Feature Forging	24193
<i>Karim Guirguis (Robert Bosch GmbH; Karlsruhe Institute of Technology), Johannes Meier (Robert Bosch GmbH), George Eskandar (University of Stuttgart), Matthias Kayser (Robert Bosch GmbH), Bin Yang (University of Stuttgart), and Jürgen Beyerer (Karlsruhe Institute of Technology; Fraunhofer IOSB)</i>	
MixPHM: Redundancy-Aware Parameter-Efficient Tuning for Low-Resource Visual Question Answering	24203
<i>Jingjing Jiang (Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University) and Nanning Zheng (Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University)</i>	
PIVOT: Prompting for Video Continual Learning	24214
<i>Andrés Villa (Pontificia Universidad Católica de Chile; King Abdullah University of Science and Technology (KAUST)), Juan León Alcázar (King Abdullah University of Science and Technology (KAUST)), Motasem Alfarra (King Abdullah University of Science and Technology (KAUST)), Kumail Alhamoud (King Abdullah University of Science and Technology (KAUST)), Julio Hurtado (University of Pisa), Fabian Caba Heilbron (Adobe Research), Alvaro Soto (Pontificia Universidad Católica de Chile), and Bernard Ghanem (King Abdullah University of Science and Technology (KAUST))</i>	
BlackVIP: Black-Box Visual Prompting for Robust Transfer Learning	24224
<i>Changdae Oh (University of Seoul), Hyeji Hwang (University of Seoul), Hee-young Lee (Sungkyunkwan University), YongTaek Lim (University of Seoul), Geunyoung Jung (University of Seoul), Jiyong Jung (University of Seoul), Hosik Choi (University of Seoul), and Kyungwoo Song (Yonsei University)</i>	
DKT: Diverse Knowledge Transfer Transformer for Class Incremental Learning	24236
<i>Xinyuan Gao (Xi'an Jiaotong University), Yuhang He (Xi'an Jiaotong University), Songlin Dong (Xi'an Jiaotong University), Jie Cheng (Huawei Technologies, China), Xing Wei (Xi'an Jiaotong University), and Yihong Gong (Xi'an Jiaotong University)</i>	
PCR: Proxy-Based Contrastive Replay for Online Class-Incremental Continual Learning	24246
<i>Huiwei Lin (Harbin Institute of Technology, Shenzhen), Baoquan Zhang (Harbin Institute of Technology, Shenzhen), Shanshan Feng (Harbin Institute of Technology, Shenzhen), Xutao Li (Harbin Institute of Technology, Shenzhen), and Yunming Ye (Harbin Institute of Technology, Shenzhen)</i>	
Masked Autoencoders Enable Efficient Knowledge Distillers	24256
<i>Yutong Bai (Johns Hopkins University), Zeyu Wang (University of California, Santa Cruz), Junfei Xiao (Johns Hopkins University), Chen Wei (Johns Hopkins University), Huiyu Wang (Johns Hopkins University), Alan L. Yuille (Johns Hopkins University), Yuyin Zhou (University of California, Santa Cruz), and Cihang Xie (University of California, Santa Cruz)</i>	

Data-Free Knowledge Distillation via Feature Exchange and Activation Region Constraint	24266
<i>Shikang Yu (Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, China; University of Chinese Academy of Sciences, China), Jiachen Chen (Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, China; University of Chinese Academy of Sciences, China), Hu Han (University of Chinese Academy of Sciences, China; Peng Cheng Laboratory, China), and Shuqiang Jiang (Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, China; University of Chinese Academy of Sciences, China)</i>	
Multi-Level Logit Distillation	24276
<i>Ying Jin (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong), Jiaqi Wang (Shanghai AI Laboratory), and Dahua Lin (CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong; Shanghai AI Laboratory)</i>	
Preserving Linear Separability in Continual Learning by Backward Feature Projection	24286
<i>Qiao Gu (University of Toronto), Dongsub Shim (LG AI Research), and Florian Shkurti (University of Toronto)</i>	
Critical Learning Periods for Multisensory Integration in Deep Networks	24296
<i>Michael Kleinman (University of California, Los Angeles), Alessandro Achille (AWS AI Labs), and Stefano Soatto (AWS AI Labs)</i>	
SLACK: Stable Learning of Augmentations With Cold-Start and KL Regularization	24306
<i>Juliette Marrie (Univ. Grenoble Alpes, Inria, LJK; NAVER LABS Europe), Michael Arbel (Univ. Grenoble Alpes, Inria, LJK), Diane Larlus (NAVER LABS Europe), and Julien Mairal (Univ. Grenoble Alpes, Inria, LJK)</i>	
Improving Generalization With Domain Convex Game	24315
<i>Fangrui Lv (Beijing Institute of Technology, China), Jian Liang (Beijing Institute of Technology, China), Shuang Li (Beijing Institute of Technology, China), Jinming Zhang (Beijing Institute of Technology, China), and Di Liu (Beijing Institute of Technology, China)</i>	
Exploring Data Geometry for Continual Learning	24325
<i>Zhi Gao (Beijing Institute of Technology, China), Chen Xu (Shenzhen MSU-BIT University, China), Feng Li (n/a), Yunde Jia (Beijing Institute of Technology, China; Shenzhen MSU-BIT University, China), Mehrtash Harandi (Monash University, Australia), and Yuwei Wu (Beijing Institute of Technology, China; Shenzhen MSU-BIT University, China)</i>	
FlowGrad: Controlling the Output of Generative ODEs With Gradients	24335
<i>Xingchao Liu (University of Texas at Austin), Lemeng Wu (University of Texas at Austin), Shujian Zhang (University of Texas at Austin), Chengyue Gong (University of Texas at Austin), Wei Ping (NVIDIA), and Qiang Liu (University of Texas at Austin)</i>	
Deep Graph Reprogramming	24345
<i>Yongcheng Jing (The University of Sydney, Australia), Chongbin Yuan (National University of Singapore, Singapore), Li Ju (National University of Singapore, Singapore), Yiding Yang (National University of Singapore, Singapore), Xinchao Wang (National University of Singapore, Singapore), and Dacheng Tao (The University of Sydney, Australia)</i>	

X-Pruner: eXplainable Pruning for Vision Transformers	24355
<i>Lu Yu (James Cook University) and Wei Xiang (La Trobe University)</i>	
Bias in Pruned Vision Models: In-Depth Analysis and Countermeasures	24364
<i>Eugenia Iofinova (IST Austria), Alexandra Peste (IST Austria), and Dan Alistarh (IST Austria & Neural Magic)</i>	
Compacting Binary Neural Networks by Sparse Kernel Selection	24374
<i>Yikai Wang (BNRist Center, State Key Lab on Intelligent Technology and Systems, Tsinghua University), Wenbing Huang (Gaoling School of Artificial Intelligence, Renmin University of China), Yinpeng Dong (BNRist Center, State Key Lab on Intelligent Technology and Systems, Tsinghua University; RealAI), Fuchun Sun (BNRist Center, State Key Lab on Intelligent Technology and Systems, Tsinghua University), and Anbang Yao (Intel Labs China)</i>	
Deep Deterministic Uncertainty: A New Simple Baseline	24384
<i>Jishnu Mukhoti (University of Oxford), Andreas Kirsch (University of Oxford), Joost van Amersfoort (University of Oxford), Philip H.S. Torr (University of Oxford), and Yarin Gal (University of Oxford)</i>	
Understanding Deep Generative Models With Generalized Empirical Likelihoods	24395
<i>Suman Ravuri (DeepMind London, UK), Mélanie Rey (DeepMind London, UK), Shakir Mohamed (DeepMind London, UK), and Marc Peter Deisenroth (University College London London, UK)</i>	
Fair Scratch Tickets: Finding Fair Sparse Networks Without Weight Training	24406
<i>Pengwei Tang (Renmin University of China, Beijing; Beijing Key Laboratory of Big Data Management and Analysis Methods, Beijing), Wei Yao (Renmin University of China, Beijing; Beijing Key Laboratory of Big Data Management and Analysis Methods, Beijing), Zhicong Li (Renmin University of China, Beijing; Beijing Key Laboratory of Big Data Management and Analysis Methods, Beijing), and Yong Liu (Renmin University of China, Beijing; Beijing Key Laboratory of Big Data Management and Analysis Methods, Beijing)</i>	
Hard Sample Matters a Lot in Zero-Shot Quantization	24417
<i>Huantong Li (South China University of Technology; Information Technology R&D Innovation Center of Peking University), Xiangmiao Wu (South China University of Technology), Fanbing Lv (Changsha Hisense Intelligent System Research Institute Co., Ltd), Daihai Liao (Changsha Hisense Intelligent System Research Institute Co., Ltd), Thomas H. Li (Peking University Shenzhen Graduate School, China), Yonggang Zhang (Hong Kong Baptist University), Bo Han (Hong Kong Baptist University), and Mingkui Tan (South China University of Technology; Key Laboratory of Big Data and Intelligent Robot, Ministry of Education; PengCheng Laboratory)</i>	
PD-Quant: Post-Training Quantization Based on Prediction Difference Metric	24427
<i>Jiawei Liu (Huazhong University of Science & Technology), Lin Niu (Huazhong University of Science & Technology), Zhihang Yuan (Houmo AI), Dawei Yang (Houmo AI), Xinggong Wang (Huazhong University of Science & Technology), and Wenyu Liu (Huazhong University of Science & Technology)</i>	

Vector Quantization With Self-Attention for Quality-Independent Representation Learning .	24438
<i>Zhou Yang (Xidian University), Weisheng Dong (Xidian University), Xin Li (West Virginia University), Mengluan Huang (Xidian University), Yulin Sun (Xidian University), and Guangming Shi (Xidian University)</i>	
Masked Auto-Encoders Meet Generative Adversarial Networks and Beyond	24449
<i>Zhengcong Fei (NA), Mingyuan Fan (NA), Li Zhu (NA), Junshi Huang (NA), Xiaoming Wei (NA), and Xiaolin Wei (NA)</i>	
Sequential Training of GANs Against GAN-Classifiers Reveals Correlated “Knowledge Gaps” Present Among Independently Trained GAN Instances	24460
<i>Arkanath Pathak (Google Research) and Nicholas Dufour (Google Research)</i>	
Edges to Shapes to Concepts: Adversarial Augmentation for Robust Vision	24470
<i>Aditay Tripathi (CDS, Indian Institute of Science), Rishubh Singh (Google Research India), Anirban Chakraborty (CDS, Indian Institute of Science), and Pradeep Shenoy (Google Research India)</i>	
Towards Universal Fake Image Detectors That Generalize Across Generative Models	24480
<i>Utkarsh Ojha (University of Wisconsin-Madison), Yuheng Li (University of Wisconsin-Madison), and Yong Jae Lee (University of Wisconsin-Madison)</i>	
Explicit Boundary Guided Semi-Push-Pull Contrastive Learning for Supervised Anomaly Detection	24490
<i>Xincheng Yao (Shanghai Jiao Tong University), Ruoqi Li (Shanghai Jiao Tong University), Jing Zhang (Academy Military Science, Beijing, China), Jun Sun (Shanghai Jiao Tong University), and Chongyang Zhang (Shanghai Jiao Tong University; Shanghai Jiao Tong University)</i>	
Generating Anomalies for Video Anomaly Detection With Prompt-Based Feature Mapping .	24500
<i>Zuhao Liu (Sun Yat-sen University, China; Ministry of Education, China), Xiao-Ming Wu (Sun Yat-sen University, China; Ministry of Education, China), Dian Zheng (Sun Yat-sen University, China; Ministry of Education, China), Kun-Yu Lin (Sun Yat-sen University, China; Ministry of Education, China), and Wei-Shi Zheng (Sun Yat-sen University, China; Ministry of Education, China)</i>	
Revisiting Reverse Distillation for Anomaly Detection	24511
<i>Tran Dinh Tien (VinBrain JSC, Vietnam), Anh Tuan Nguyen (VinBrain JSC, Vietnam), Nguyen Hoang Tran (VinBrain JSC, Vietnam), Ta Duc Huy (VinBrain JSC, Vietnam), Soan T.M. Duong (VinBrain JSC, Vietnam; Le Quy Don Technical University, Vietnam), Chanh D. Tr. Nguyen (VinBrain JSC, Vietnam; VinUniversity, Vietnam), and Steven Q. H. Truong (VinBrain JSC, Vietnam; VinUniversity, Vietnam)</i>	
MetaMix: Towards Corruption-Robust Continual Learning With Temporally Self-Adaptive Data Transformation	24521
<i>Zhenyi Wang (State University of New York at Buffalo, USA), Li Shen (JD Explore Academy, China), Donglin Zhan (Columbia University, USA), Qiuling Suo (State University of New York at Buffalo, USA), Yanjun Zhu (State University of New York at Buffalo, USA), Tiehang Duan (Meta, USA), and Mingchen Gao (State University of New York at Buffalo, USA)</i>	

ScaleFL: Resource-Adaptive Federated Learning With Heterogeneous Clients	24532
<i>Fatih Ilhan (Georgia Institute of Technology Atlanta, GA), Gong Su (IBM Research Yorktown Heights, NY), and Ling Liu (Georgia Institute of Technology Atlanta, GA)</i>	
Confidence-Aware Personalized Federated Learning via Variational Expectation Maximization	24542
<i>Junyi Zhu (ESAT-PSI, KU Leuven), Xingchen Ma (Amazon Web Services), and Matthew B. Blaschko (ESAT-PSI, KU Leuven)</i>	
Make Landscape Flatter in Differentially Private Federated Learning	24552
<i>Yifan Shi (Tsinghua University, China), Yingqi Liu (Nanjing University of Science and Technology, China), Kang Wei (Nanjing University of Science and Technology, China), Li Shen (JD Explore Academy, China), Xueqian Wang (Tsinghua University, China), and Dacheng Tao (JD Explore Academy, China)</i>	
Rethinking Domain Generalization for Face Anti-Spoofing: Separability and Alignment	24563
<i>Yiyao Sun (University of Wisconsin-Madison), Yaojie Liu (Google Research), Xiaoming Liu (Google Research; Michigan State University), Yixuan Li (University of Wisconsin-Madison), and Wen-Sheng Chu (Google Research)</i>	
StyleAdv: Meta Style Adversarial Training for Cross-Domain Few-Shot Learning	24575
<i>Yuqian Fu (Shanghai Key Lab of Intelligent Information Processing, School of Computer Science, Fudan University), Yu Xie (Purple Mountain Laboratories, China), Yanwei Fu (School of Data Science, Fudan University), and Yu-Gang Jiang (Shanghai Key Lab of Intelligent Information Processing, School of Computer Science, Fudan University)</i>	
The Dark Side of Dynamic Routing Neural Networks: Towards Efficiency Backdoor Injection	24585
<i>Simin Chen (University of Texas at Dallas), Hanlin Chen (Purdue University), Mirazul Haque (University of Texas at Dallas), Cong Liu (University of California, Riverside), and Wei Yang (University of Texas at Dallas)</i>	
Architectural Backdoors in Neural Networks	24595
<i>Mikel Bober-Irizar (University of Cambridge, UK), Ilia Shumailov (Vector Institute, CA; University of Oxford, UK), Yiren Zhao (Imperial College London, UK; University of Cambridge, UK), Robert Mullins (University of Cambridge, UK), and Nicolas Papernot (University of Toronto, CA; Vector Institute, CA)</i>	
You Are Catching My Attention: Are Vision Transformers Bad Learners Under Backdoor Attacks?	24605
<i>Zenghui Yuan (Huazhong University of Science and Technology), Pan Zhou (Huazhong University of Science and Technology), Kai Zou (Protagolabs Inc), and Yu Cheng (Microsoft Research)</i>	
A Practical Upper Bound for the Worst-Case Attribution Deviations	24616
<i>Fan Wang (Nanyang Technological University) and Adams Wai-Kin Kong (Nanyang Technological University)</i>	

Sibling-Attack: Rethinking Transferable Adversarial Attacks Against Face Recognition	24626
<i>Zexin Li (University of California, Riverside), Bangjie Yin (Tencent), Taiping Yao (Tencent), Junfeng Guo (The University of Texas at Dallas), Shouhong Ding (Tencent), Simin Chen (The University of Texas at Dallas), and Cong Liu (University of California, Riverside)</i>	
Angelic Patches for Improving Third-Party Object Detector Performance	24638
<i>Wenwen Si (University of Pennsylvania), Shuo Li (University of Pennsylvania), Sangdon Park (School of Cybersecurity & Privacy, Georgia Institute of Technology), Insup Lee (University of Pennsylvania), and Osbert Bastani (University of Pennsylvania)</i>	
Introducing Competition To Boost the Transferability of Targeted Adversarial Examples Through Clean Feature Mixup	24648
<i>Junyoung Byun (Korea Advanced Institute of Science and Technology (KAIST)), Myung-Joon Kwon (Korea Advanced Institute of Science and Technology (KAIST)), Seungju Cho (Korea Advanced Institute of Science and Technology (KAIST)), Yoonji Kim (Korea Advanced Institute of Science and Technology (KAIST)), and Changick Kim (Korea Advanced Institute of Science and Technology (KAIST))</i>	
Towards Compositional Adversarial Robustness: Generalizing Adversarial Training to Composite Semantic Perturbations	24658
<i>Lei Hsiung (National Tsing Hua University; The Chinese University of Hong Kong), Yun-Yun Tsai (Columbia University), Pin-Yu Chen (IBM Research), and Tsung-Yi Ho (National Tsing Hua University; The Chinese University of Hong Kong)</i>	
Boosting Accuracy and Robustness of Student Models via Adaptive Adversarial Distillation .	24668
<i>Bo Huang (The Hong Kong University of Science and Technology, China), Mingyang Chen (The Hong Kong University of Science and Technology, China), Yi Wang (Dongguan University of Technology, China), Junda Lu (Macquarie University, Australia), Minhao Cheng (The Hong Kong University of Science and Technology, China), and Wei Wang (The Hong Kong University of Science and Technology, China)</i>	
The Enemy of My Enemy Is My Friend: Exploring Inverse Adversaries for Improving Adversarial Training	24678
<i>Junhao Dong (Sun Yat-Sen University, China), Seyed-Mohsen Moosavi-Dezfooli (Imperial College London, UK), Jianhuang Lai (Guangdong Province Key Laboratory of Information Security Technology, China; Ministry of Education, China), and Xiaohua Xie (Guangdong Province Key Laboratory of Information Security Technology, China; Ministry of Education, China)</i>	
Robust Single Image Reflection Removal Against Adversarial Attacks	24688
<i>Zhenbo Song (Nanjing University of Science and Technology), Zhenyuan Zhang (Nanjing University of Science and Technology), Kaihao Zhang (Australian National University), Wenhan Luo (Shenzhen Campus of Sun Yat-sen University), Zhaoxin Fan (Renmin University of China), Wenqi Ren (Shenzhen Campus of Sun Yat-sen University), and Jianfeng Lu (Nanjing University of Science and Technology)</i>	

Physical-World Optical Adversarial Attacks on 3D Face Recognition	24699
<i>Yanjie Li (The Hong Kong Polytechnic University), Yiquan Li (The Hong Kong Polytechnic University), Xuelong Dai (The Hong Kong Polytechnic University), Songtao Guo (Chongqing University), and Bin Xiao (The Hong Kong Polytechnic University)</i>	
AUNet: Learning Relations Between Action Units for Face Forgery Detection	24709
<i>Weiming Bai (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Yufan Liu (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Zhipeng Zhang (DiDiChuxing), Bing Li (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; People AI, Inc.), and Weiming Hu (State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; CAS Center for Excellence in Brain Science and Intelligence Technology)</i>	

Author Index