2021 IEEE/CVF International Conference on Computer Vision (ICCV 2021)

Virtual Conference
11-17 October 2021

Pages 1-589

IEEE Catalog Number: CFP21198-POD
ISBN: 978-1-6654-2813-2

Copyright © 2021 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:	CFP21198-POD
ISBN (Print-On-Demand):	978-1-6654-2813-2
ISBN (Online):	978-1-6654-2812-5
ISSN:	1550-5499

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

2021 IEEE/CVF International Conference on Computer Vision (ICCV)

ICCV 2021

Table of Contents

Message from the General and Program Chairs ..... **clxvi**
Conference Organizers ..... **clxviii**

Segmentation and Scene Analysis

MVTN: Multi-View Transformation Network for 3D Shape Recognition 1
Abdullah Hamdi (KAUST), Silvio Giancola (KAUST), and Bernard Ghanem (KAUST)

GLiT: Neural Architecture Search for Global and Local Image Transformer 12
Boyu Chen (The University of Sydney), Peixia Li (The University of Sydney), Chuming Li (sensetime), Baopu Li (BAIDU USA LLC), Lei Bai (University of Sydney), Chen Lin (University of Oxford), Ming Sun (SenseTime Group Limited), Junjie Yan (Sensetime Group Limited), and Wanli Ouyang (The University of Sydney)

CvT: Introducing Convolutions to Vision Transformers 22
Haiping Wu (McGill University), Bin Xiao (Microsoft), Noel Codella (Microsoft), Mengchen Liu (Microsoft), Xiyang Dai (Microsoft), Lu Yuan (Microsoft), and Lei Zhang (International Digital Economy Academy, (IDEA))

Going Deeper With Image Transformers 32
Hugo Touvron (Facebook AI Research), Matthieu Cord (Sorbonne University), Alexandre Sablayrolles (Facebook AI Research), Gabriel Synnaeve (Facebook), and Hervé Jégou (Facebook AI Research)

DTMNet: A Discrete Tchebichef Moments-Based Deep Neural Network for Multi-Focus Image Fusion 43
Bin Xiao (Chongqing University of Posts and Telecommunications), Haifeng Wu (Chongqing University of Posts and Telecommunications), and Xiuli Bi (Chongqing University of Posts and Telecommunications)

CrossNorm and SelfNorm for Generalization Under Distribution Shifts 52
Zhiqiang Tang (Rutgers), Yunhe Gao (Rutgers University), Yi Zhu (Amazon), Zhi Zhang (Amazon), Mu Li (Amazon), and Dimitris N. Metaxas (Rutgers)

NGC: A Unified Framework for Learning With Open-World Noisy Data 62
Zhi-Fan Wu (Nanjing University), Tong Wei (Nanjing University), Jianwen Jiang (Alibaba Group), Chaojie Mao (Alibaba Group), Mingqian Tang (Alibaba Group), and Yu-Feng Li (Nanjing University)

Learning With Noisy Labels via Sparse Regularization 72
Xiong Zhou (Harbin Institute of Technology), Xianming Liu (Harbin Institute of Technology), Chenyang Wang (Harbin Institute of Technology), Deming Zhai (Harbin Institute of Technology), Junjun Jiang (Harbin Institute of Technology), and Xiangyang Ji (Tsinghua University)

Asymmetric Loss for Multi-Label Classification 82
Tal Ridnik (Alibaba), Emanuel Ben-Baruch (Alibaba), Nadav Zamir (Alibaba), Asaf Noy (Alibaba), Itamar Friedman (Alibaba), Matan Protter (Alibaba group), and Lihi Zelnik-Manor (Alibaba)

Procrustean Training for Imbalanced Deep Learning 92
Han-Jia Ye (Nanjing University), De-Chuan Zhan (Nanjing University), and Wei-Lun Chao (The Ohio State University)

Conditional Variational Capsule Network for Open Set Recognition 103
Yunrui Guo (National University of Defense Technology), Guglielmo Camporese (Padova University), Wenjing Yang (National University of Defense Technology), Alessandro Sperduti (Università di Padova, (IT)), and Lamberto Ballan (University of Padova)

ACE: Ally Complementary Experts for Solving Long-Tailed Recognition in One-Shot 112
Jiarui Cai (University of Washington), Yizhou Wang (University of Washington), and Jenq-Neng Hwang (University of Washington)

FREE: Feature Refinement for Generalized Zero-Shot Learning 122
Shiming Chen (Huazhong University of Science and Technology), Wenjie Wang (Huazhong University of Science and Technology), Beihao Xia (Huazhong University of Science and Technology), Qinmu Peng (Huazhong University of Science and Technology), Xinge You (Huazhong University of Science and Technology), Feng Zheng (SUSTech), and Ling Shao (Inception Institute of Artificial Intelligence)

Online Refinement of Low-Level Feature Based Activation Map for Weakly Supervised Object Localization 132
Jinheng Xie (Shenzhen University), Cheng Luo (Shenzhen University), Xiangping Zhu (Computer Vision Institute, School of Computer Science \& Software Engineering, Shenzhen University), Ziqi Jin (ShenZhen University), Weizeng Lu (Shenzhen University), and Linlin Shen (Shenzhen University)

Z-Score Normalization, Hubness, and Few-Shot Learning 142
Nanyi Fei (Renmin University of China), Yizhao Gao (Renmin University of China), Zhiwu Lu (Renmin University of China), and Tao Xiang (University of Surrey)

Spatio-Temporal Representation Factorization for Video-Based Person Re-Identification 152
Abhishek Aich (University of California, Riverside), Meng Zheng (United Imaging Intelligence), Srikrishna Karanam (United Imaging Intelligence), Terrence Chen (United Imaging Intelligence), Amit K. Roy-Chowdhury (University of California, Riverside), and Ziyan Wu (United Imaging Intelligence)

Transformer-Based Dual Relation Graph for Multi-Label Image Recognition 163
Jiawei Zhao (Beihang University), Ke Yan (Tencent), Yifan Zhao (Beihang University), Xiaowei Guo (Tencent Youtu Lab), Feiyue Huang (Tencent), and Jia Li (Beihang University)

Dance With Self-Attention: A New Look of Conditional Random Fields on Anomaly Detection in Videos 173
Didik Purwanto (National Taiwan University of Science and Technology), Yie-Tarng Chen (National Taiwan University of Science and Technology), and Wen-Hsien Fang (National Taiwan University of Science and Technology)

Residual Attention: A Simple but Effective Method for Multi-Label Recognition 184
Ke Zhu (Nanjing University) and Jianxin Wu (Nanjing University)

Self-Supervised Geometric Features Discovery via Interpretable Attention for Vehicle Re-Identification and Beyond 194
Ming Li (Worcester Polytechnic Institute), Xinming Huang (Worcester Polytechnic Institute), and Ziming Zhang (Worcester Polytechnic Institute)

Heterogeneous Relational Complement for Vehicle Re-Identification 205
Jiajian Zhao (Beihang University), Yifan Zhao (Beihang University), Jia Li (Beihang University), Ke Yan (Tencent), and Yonghong Tian (Peking University)

Attack-Guided Perceptual Data Generation for Real-World Re-Identification 215
Yukun Huang (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), and Zheng-Jun Zha (University of Science and Technology of China)

Syncretic Modality Collaborative Learning for Visible Infrared Person Re-Identification 225
Ziyu Wei (Xidian University), Xi Yang (Xidian University), Nannan Wang (Xidian University), and Xinbo Gao (Chongqing University of Posts and Telecommunications)

Distilling Virtual Examples for Long-Tailed Recognition 235
Yin-Yin He (Nanjing University), Jianxin Wu (Nanjing University), and Xiu-Shen Wei (Nanjing University of Science and Technology)

Neural Photofit: Gaze-Based Mental Image Reconstruction 245
Florian Strohm (University of Stuttgart), Ekta Sood (University of Stuttgart), Sven Mayer (LMU Munich), Philipp Müller (DFKI GmbH), Mihai Bâce (University of Stuttgart), and Andreas Bulling (University of Stuttgart)

When Pigs Fly: Contextual Reasoning in Synthetic and Natural Scenes 255
Philipp Bomatter (ETH Zürich), Mengmi Zhang (BCH; HMS), Dimitar Karev (Harvard College), Spandan Madan (Harvard), Claire Tseng (Harvard), and Gabriel Kreiman (Harvard)

MAAS: Multi-Modal Assignation for Active Speaker Detection 265
Juan Léon Alcázar (KAUST), Fabian Caba (Adobe Research), Ali K. Thabet (Facebook), and Bernard Ghanem (KAUST)

Move2Hear: Active Audio-Visual Source Separation 275
Sagnik Majumder (University of Texas at Austin), Ziad Al-Halah (UT Austin), and Kristen Grauman (Facebook AI Research & UT Austin)

Image2Reverb: Cross-Modal Reverb Impulse Response Synthesis 286
Nikhil Singh (MIT Media Lab), Jeff Mentch (Harvard University), Jerry Ng (MIT), Matthew Beveridge (Massachusetts Institute of Technology), and Iddo Drori (MIT)

Multi-Modality Associative Bridging Through Memory: Speech Sound Recollected From Face Video 296
Minsu Kim (KAIST), Joanna Hong (KAIST), Se Jin Park (KAIST), and Yong Man Ro (KAIST)

BN-NAS: Neural Architecture Search With Batch Normalization 307
Boyu Chen (The University of Sydney), Peixia Li (The University of Sydney), Baopu Li (BAIDU USA LLC), Chen Lin (University of Oxford), Chuming Li (sensetime), Ming Sun (SenseTime Group Limited), Junjie Yan (Sensetime Group Limited), and Wanli Ouyang (The University of Sydney)

Differentiable Dynamic Wirings for Neural Networks 317
Kun Yuan (SenseTime Research), Quanquan Li (SenseTime Research), Shaopeng Guo (SenseTime), Dapeng Chen (The Chinese University of Hong Kong), Aojun Zhou (Sensetime&tetras.ai), Fengwei Yu (Sensetime Research), and Ziwei Liu (Nanyang Technological University)

AutoSpace: Neural Architecture Search With Less Human Interference 327
Daquan Zhou (National University of Singapore), Xiaojie Jin (NUS), Xiaochen Lian (ByteDance), Linjie Yang (ByteDance AI Lab), Yujing Xue (National University of Singapore), Qibin Hou (National University of Singapore), and Jiashi Feng (NUS)

Zen-NAS: A Zero-Shot NAS for High-Performance Image Recognition 337
Ming Lin (Alibaba Group), Pichao Wang (Alibaba Group), Zhenhong Sun (Alibaba Group), Hesen Chen (Alibaba Group), Xiuyu Sun (Alibaba Group), Qi Qian (Alibaba Group), Hao Li (Alibaba Group), and Rong Jin (alibaba group)

CrossViT: Cross-Attention Multi-Scale Vision Transformer for Image Classification 347
Chun-Fu Richard Chen (MIT-IBM Watson AI Lab, IBM Research AI), Quanfu Fan (IBM Research), and Rameswar Panda (MIT-IBM Watson AI Lab, IBM Research)

Conformer: Local Features Coupling Global Representations for Visual Recognition 357
Zhiliang Peng (University of Chinese Academy of Sciences), Wei Huang (University of Chinese Academy of Sciences), Shanzhi Gu (PengCheng Laboratory), Lingxi Xie (Huawei Inc.), Yaowei Wang (PengCheng Laboratory), Jianbin Jiao (University of Chinese Academy of Sciences), and Qixiang Ye (University of Chinese Academy of Sciences, China)

Scalable Vision Transformers With Hierarchical Pooling 367
Zizheng Pan (Monash University), Bohan Zhuang (Monash University), Jing Liu (Monash University), Haoyu He (Monash University), and Jianfei Cai (Monash University)

Vision Transformer With Progressive Sampling 377
Xiaoyu Yue (Centre for Perceptual and Interactive Intelligence), Shuyang Sun (University of Oxford, UK), Zhanghui Kuang (Sensetime Ltd.), Meng Wei (Tsinghua University, China), Philip H.S. Torr (University of Oxford, UK), Wayne Zhang (SenseTime Research), and Dahua Lin (The Chinese University of Hong Kong)

Generic Attention-Model Explainability for Interpreting Bi-Modal and Encoder-Decoder Transformers 387
Hila Chefer (Tel Aviv University), Shir Gur (Tel Aviv University), and Lior Wolf (Tel Aviv University, Israel)

Learning Canonical View Representation for 3D Shape Recognition With Arbitrary Views 397
Xin Wei (Xi'an Jiaotong University), Yifei Gong (Tencent), Fudong Wang (Wuhan University), Xing Sun (Tencent), and Jian Sun (Xi'an Jiaotong University)

MosaicOS: A Simple and Effective Use of Object-Centric Images for Long-Tailed Object Detection 407
Cheng Zhang (The Ohio State University), Tai-Yu Pan (The Ohio State University), Yandong Li (University of Central Florida), Hexiang Hu (University of Southern California), Dong Xuan (Ohio State University), Soravit Changpinyo (Google Research), Boqing Gong (Google), and Wei-Lun Chao (The Ohio State University)

Towers of Babel: Combining Images, Language, and 3D Geometry for Learning Multimodal Vision 418
Xiaoshi Wu (Tsinghua University), Hadar Averbuch-Elor (Cornell University), Jin Sun (Cornell University), and Noah Snavely (Cornell University and Google AI)

Virtual Multi-Modality Self-Supervised Foreground Matting for Human-Object Interaction 428
Bo Xu (OPPO Research Institute), Han Huang (OPPO Research Institute), Cheng Lu (XPENG), Ziwen Li (OPPO Research Institute), and Yandong Guo (MSR)

An Asynchronous Kalman Filter for Hybrid Event Cameras 438
Ziwei Wang (Australian National University), Yonhon Ng (Australian National University), Cedric Scheerlinck (The Australian National University), and Robert Mahony (Australian National University)

Amplitude-Phase Recombination: Rethinking Robustness of Convolutional Neural Networks in Frequency Domain 448
Guangyao Chen (Peking University), Peixi Peng (Peking University), Li Ma (Peking University), Jia Li (Beihang University), Lin Du (AI Application Research Center, Huawei Technologies Co., Ltd), and Yonghong Tian (Peking University)

MicroNet: Improving Image Recognition With Extremely Low FLOPs 458
Yunsheng Li (UCSD), Yinpeng Chen (Microsoft), Xiyang Dai (Microsoft), Dongdong Chen (Microsoft Cloud AI), Mengchen Liu (Microsoft), Lu Yuan (Microsoft), Zicheng Liu (Microsoft), Lei Zhang (International Digital Economy Academy, (IDEA)), and Nuno Vasconcelos (UC San Diego)

Group-Wise Inhibition Based Feature Regularization for Robust Classification 468
Haozhe Liu (Shenzhen University), Haoqian Wu (ShenZhen University), Weicheng Xie (Shenzhen University), Feng Liu (Shenzhen University), and Linlin Shen (Shenzhen University)

Exploration and Estimation for Model Compression 477
Yanfu Zhang (University of Pittsburgh), Shangqian Gao (University of Pittsburgh), and Heng Huang (University of Pittsburgh & JD Explore Academy)

Learning To Resize Images for Computer Vision Tasks 487
Hossein Talebi (Google) and Peyman Milanfar (Google)

Learning Meta-Class Memory for Few-Shot Semantic Segmentation 497
Zhonghua Wu (Nanyang Technological University), Xiangxi Shi (Oregon State University), Guosheng Lin (Nanyang Technological University), and Jianfei Cai (Monash University)

Aggregation With Feature Detection 507
Shuyang Sun (University of Oxford), Xiaoyu Yue (Centre for Perceptual and Interactive Intelligence), Xiaojuan Qi (The University of Hong Kong), Wanli Ouyang (The University of Sydney), Victor Adrian Prisacariu (University of Oxford), and Philip H.S. Torr (University of Oxford)

Continual Learning on Noisy Data Streams via Self-Purified Replay 517
Chris Dongjoo Kim (Seoul National University), Jinseo Jeong (Seoul National University), Sangwoo Moon (Seoul National University), and Gunhee Kim (Seoul National University)

Point Cloud Augmentation With Weighted Local Transformations 528
Sihyeon Kim (Korea University), Sanghyeok Lee (Korea University), Dasol Hwang (Korea University), Jaewon Lee (Korea University), Seong Jae Hwang (University of Pittsburgh), and Hyunwoo J. Kim (Korea University)

Tokens-to-Token ViT: Training Vision Transformers From Scratch on ImageNet 538
Li Yuan (National University of Singapore), Yunpeng Chen (National University of Singapore), Tao Wang (National University of Singapore), Weihao Yu (National University of Singapore), Yujun Shi (National University of Singapore), Zi-Hang Jiang (National University of Singapore), Francis E.H. Tay (National University of Singapore), Jiashi Feng (National University of Singapore), and Shuicheng Yan (National University of Singapore)

Pyramid Vision Transformer: A Versatile Backbone for Dense Prediction Without Convolutions.... 548
Wenhai Wang (Nanjing University), Enze Xie (The University of Hong Kong), Xiang Li (Nanjing University of Science and Technology), Deng-Ping Fan (Inception Institute of Artificial Intelligence), Kaitao Song (Nanjing University of Science and Technology), Ding Liang (Sensetime Group Limited), Tong Lu (Nanjing University), Ping Luo (The University of Hong Kong), and Ling Shao (Inception Institute of Artificial Intelligence)

Incorporating Convolution Designs Into Visual Transformers 559
Kun Yuan (SenseTime Research), Shaopeng Guo (SenseTime), Ziwei Liu (Nanyang Technological University), Aojun Zhou (Sensetime&tetras.ai), Fengwei Yu (Sensetime Research), and Wei Wu (SenseTime Group Limited)

Visformer: The Vision-Friendly Transformer 569
Zhengsu Chen (Beihang University, China), Lingxi Xie (Huawei Inc.), Jianwei Niu (Beihang University, China), Xuefeng Liu (Beihang University, China), Longhui Wei (University of Science and Technology of China), and Qi Tian (Xidian University, China)

Visual Transformers: Where Do Transformers Really Belong in Vision Models? 579
Bichen Wu (Facebook Research), Chenfeng Xu (University of California, Berkeley, USA), Xiaoliang Dai (Facebook), Alvin Wan (University of California, Berkeley, USA), Peizhao Zhang (Facebook), Zhicheng Yan (Facebook AI), Masayoshi Tomizuka (University of California, Berkeley, USA), Joseph E. Gonzalez (University of California, Berkeley, USA), Kurt Keutzer (University of California, Berkeley, USA), and Peter Vajda (Facebook)

Joint Representation Learning and Novel Category Discovery on Single- and Multi-Modal Data.... 590
Xuhui Jia (google), Kai Han (University of Bristol), Yukun Zhu (Google Inc.), and Bradley Green (Google Inc.)

Stochastic Partial Swap: Enhanced Model Generalization and Interpretability for Fine-Grained Recognition 600
Shaoli Huang (University of Sydney), Xinchao Wang (National University of Singapore), and Dacheng Tao (The University of Sydney)

Self Supervision to Distillation for Long-Tailed Visual Recognition 610
Tianhao Li (Nanjing University), Limin Wang (Nanjing University), and Gangshan Wu (Nanjing University)

Semantic Diversity Learning for Zero-Shot Multi-Label Classification 620
Avi Ben-Cohen (Alibaba group), Nadav Zamir (Alibaba), Emanuel Ben-Baruch (Alibaba), Itamar Friedman (Alibaba), and Lihi Zelnik-Manor (Alibaba)

Shallow Bayesian Meta Learning for Real-World Few-Shot Recognition 631
Xueting Zhang (University of Edinburgh), Debin Meng (SIAT_MMLab), Henry Gouk (University of Edinburgh), and Timothy M. Hospedales (Edinburgh University)

Adversarial Attacks Are Reversible With Natural Supervision 641
Chengzhi Mao (Columbia University), Mia Chiquier (Columbia University), Hao Wang (Rutgers University), Junfeng Yang (Columbia University), and Carl Vondrick (Columbia University)

Architecture Disentanglement for Deep Neural Networks 652
Jie Hu (Xiamen University), Liujuan Cao (Xiamen University), Tong Tong (Xiamen university), Qixiang Ye (University of Chinese Academy of Sciences, China), Shengchuan Zhang (Xiamen University), Ke Li (Tencent), Feiyue Huang (Tencent), Ling Shao (Inception Institute of Artificial Intelligence), and Rongrong Ji (Xiamen University, China)

Exploiting Explanations for Model Inversion Attacks 662
Xuejun Zhao (National University of Singapore), Wencan Zhang (National University of Singapore), Xiaokui Xiao (National University of Singapore), and Brian Lim (National University of Singapore)

Explaining in Style: Training a GAN To Explain a Classifier in StyleSpace 673
Oran Lang (Google), Yossi Gandelsman (Google), Michal Yarom (Google), Yoav Wald (Google), Gal Elidan (Google), Avinatan Hassidim (MIT), William T. Freeman (Weizmann Institute, Israel), Phillip Isola (Google), Amir Globerson (Google & Hebrew University), Michal Irani (Hebrew University), and Inbar Mosseri (Google)

Ground-Truth or DAER: Selective Re-Query of Secondary Information 683
Stephan J. Lemmer (University of Michigan) and Jason J. Corso (University of Michigan)

Parametric Contrastive Learning 695
Jiequan Cui (Chinese University of Hong Kong), Zhisheng Zhong (Peking University), Shu Liu (SmartMore), Bei Yu (CUHK), and Jiaya Jia (Chinese University of Hong Kong)

Learning Fast Sample Re-Weighting Without Reward Data 705
Zizhao Zhang (Google) and Tomas Pfister (Google)

Influence-Balanced Loss for Imbalanced Visual Classification 715
Seulki Park (Seoul National University), Jongin Lim (Seoul National University), Younghan Jeon (Seoul National University), and Jin Young Choi (Seoul National University)

Statistically Consistent Saliency Estimation 725
Shunyan Luo (George Washington University), Emre Barut (Amazon), and Fang Jin (George Washington University)

Contrastive Multimodal Fusion With TupleInfoNCE 734
Yunze Liu (Tsinghua University), Qingnan Fan (Tencent AI Lab), Shanghang Zhang (UC Berkeley), Hao Dong (Peking University), Thomas Funkhouser (Google Research), and Li Yi (Tsinghua University)

Recursively Conditional Gaussian for Ordinal Unsupervised Domain Adaptation 744
Xiaofeng Liu (CMU), Site Li (CMU), Yubin Ge (University of Illinois at Urbana-Champaign), Pengyi Ye (Harvard), Jane You (PolyU), and Jun Lu (Harvard)

TrivialAugment: Tuning-Free Yet State-of-the-Art Data Augmentation 754
Samuel G. Müller (University of Freiburg) and Frank Hutter (University of Freiburg)

FcaNet: Frequency Channel Attention Networks 763
Zequn Qin (Zhejiang University), Pengyi Zhang (Zhejiang University), Fei Wu (Zhejiang University), and Xi Li (Zhejiang University)

Global Pooling, More Than Meets the Eye: Position Information Is Encoded Channel-Wise in CNNs 773
Md Amirul Islam (Ryerson University), Matthew Kowal (Ryerson University), Sen Jia (University of Waterloo), Konstantinos G. Derpanis (Ryerson University), and Neil D. B. Bruce (University of Guelph)

Neural Video Portrait Relighting in Real-Time via Consistency Modeling 782
Longwen Zhang (ShanghaiTech University), Qixuan Zhang (ShanghaiTech University), Minye Wu (Shanghaitech University), Jingyi Yu (Shanghai Tech University), and Lan Xu (HKUST)

OpenGAN: Open-Set Recognition via Open Data Generation 793
Shu Kong (Carnegie Mellon University) and Deva Ramanan (Carnegie Mellon University)

MixMo: Mixing Multiple Inputs for Multiple Outputs via Deep Subnetworks 803
Alexandre Ramé (Heuritech), Rémy Sun (Laboratoire d'informatique de Paris 6), and Matthieu Cord (Sorbonne University)

Learning To Diversify for Single Domain Generalization 814
Zijian Wang (University of Queensland), Yadan Luo (University of Queensland), Ruihong Qiu (The University of Queensland), Zi Huang (University of Queensland), and Mahsa Baktashmotlagh (University of Queensland)

SS-IL: Separated Softmax for Incremental Learning 824
Hongjoon Ahn (SKKU), Jihwan Kwak (Sungkyunkwan University), Subin Lim (Sungkyunkwan University), Hyeonsu Bang (Sungkyunkwan University), Hyojun Kim (Sungkyunkwan University), and Taesup Moon (Sungkyunkwan University)

Multimodal Knowledge Expansion 834
Zihui Xue (University of Texas at Austin), Sucheng Ren (South China University of Technology), Zhengqi Gao (MIT), and Hang Zhao (Tsinghua University)

FaPN: Feature-Aligned Pyramid Network for Dense Image Prediction 844
Shihua Huang (Southern University of Science and Technology), Zhichao Lu (Southern University of Science and Technology), Ran Cheng (Southern University of Science and Technology), and Cheng He (Southern University of Science and Technology)

Grafit: Learning Fine-Grained Image Representations With Coarse Labels 854
Hugo Touvron (Facebook AI Research), Alexandre Sablayrolles (Facebook AI Research), Matthijs Douze (Facebook AI Research), Matthieu Cord (Sorbonne University), and Hervé Jégou (Facebook AI Research)

Attentional Pyramid Pooling of Salient Visual Residuals for Place Recognition 865
Guohao Peng (Nanyang Technological University), Jun Zhang (Nanyang Technological University), Heshan Li (Nanyang Technological University), and Danwei Wang (Nanyang Technological University)

Interpretable Image Recognition by Constructing Transparent Embedding Space 875
Jiaqi Wang (Beijing Jiaotong University), Huafeng Liu (Beijing Jiaotong University), Xinyue Wang (Beijing Jiaotong University), and Liping Jing (Beijing Jiaotong University)

Generating Attribution Maps With Disentangled Masked Backpropagation 885
Adria Ruiz (Institut de Robotica i Informatica Industrial, CSIC-UPC), Antonio Agudo (Institut de Robotica i Informatica Industrial, CSIC-UPC), and Francesc Moreno-Noguer (IRI)

Walk in the Cloud: Learning Curves for Point Clouds Shape Analysis 895
Tiange Xiang (University of Sydney), Chaoyi Zhang (University of Sydney), Yang Song (University of New South Wales), Jianhui Yu (University of Sydney), and Weidong Cai (University of Sydney)

End-to-End Trainable Trident Person Search Network Using Adaptive Gradient Propagation 905
Byeong-Ju Han (Ulsan National Institute of Science and Technology), Kuhyeun Ko (Ulsan National Institute of Science and Technology), and Jae-Young Sim (Ulsan National Institute of Science and Technology)

Graph-Based Asynchronous Event Processing for Rapid Object Recognition 914
Yijin Li (Zhejiang University), Han Zhou (Zhejiang University), Bangbang Yang (Zhejiang University), Ye Zhang (Zhejiang University), Zhaopeng Cui (Zhejiang University), Hujun Bao (Zhejiang University), and Guofeng Zhang (Zhejiang University)

Parsing Table Structures in the Wild 924
Rujiao Long (Alibaba Group), Wen Wang (Wuhan University), Nan Xue (Wuhan University), Feiyu Gao (Alibaba Group), Zhibo Yang (Alibaba Group), Yongpan Wang (Alibaba Group), and Gui-Song Xia (Wuhan University)

SketchLattice: Latticed Representation for Sketch Manipulation 933
Yonggang Qi (Beijing University of Posts and Telecommunications), Guoyao Su (Beijing University of Posts and Telecommunications), Pinaki Nath Chowdhury (University of Surrey), Mingkang Li (Beijing University of Posts and Telecommunications), and Yi-Zhe Song (University of Surrey)

Spatial and Semantic Consistency Regularizations for Pedestrian Attribute Recognition 942
Jian Jia (Institute of Automation, Chinese Academy of Sciences), Xiaotang Chen (Institute of Automation, Chinese Academy of Sciences), and Kaiqi Huang (Institute of Automation, Chinese Academy of Sciences)

Detecting Persuasive Atypicality by Modeling Contextual Compatibility 952
Meiqi Guo (University of Pittsburgh), Rebecca Hwa (University of Pittsburgh), and Adriana Kovashka (University of Pittsburgh)

Text Is Text, No Matter What: Unifying Text Recognition Using Knowledge Distillation 963
Ayan Kumar Bhunia (University of Surrey), Aneeshan Sain (University of Surrey), Pinaki Nath Chowdhury (University of Surrey), and Yi-Zhe Song (University of Surrey)

DocFormer: End-to-End Transformer for Document Understanding 973
Srikar Appalaraju (Amazon), Bhavan Jasani (Amazon AWS AI), Bhargava Urala Kota (Amazon Web Services), Yusheng Xie (Amazon), and R. R. Manmatha (Amazon)

LayoutTransformer: Layout Generation and Completion With Self-Attention 984
Kamal Gupta (University of Maryland), Justin Lazarow (UC San Diego), Alessandro Achille (Amazon Web Services), Larry S. Davis (University of Maryland), Vijay Mahadevan (Amazon Inc.), and Abhinav Shrivastava (University of Maryland)

Effectively Leveraging Attributes for Visual Similarity 995
Samarth Mishra (Boston University), Zhongping Zhang (Boston University), Yuan Shen (University of Illinois at Urbana-Champaign, USA), Ranjitha Kumar (University of Illinois at Urbana-Champaign, USA), Venkatesh Saligrama (Boston University), and Bryan A. Plummer (Boston University)

Counterfactual Attention Learning for Fine-Grained Visual Categorization and Re-Identification 1005
Yongming Rao (Tsinghua University), Guangyi Chen (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)

Learning Canonical 3D Object Representation for Fine-Grained Recognition 1015
Sunghun Joung (Yonsei Univ.), Seungryong Kim (Korea University), Minsu Kim (Yonsei Univ.), Ig-Jae Kim (KIST), and Kwanghoon Sohn (Yonsei Univ.)

SCOUTER: Slot Attention-Based Classifier for Explainable Image Recognition 1026
Liangzhi Li (Osaka University), Bowen Wang (Osaka University), Manisha Verma (Osaka University), Yuta Nakashima (Osaka University), Ryo Kawasaki (Osaka University), and Hajime Nagahara (Osaka University)

Beyond Trivial Counterfactual Explanations With Diverse Valuable Explanations 1036
Pau Rodríguez (ElementAI / ServiceNow), Massimo Caccia (MILA), Alexandre Lacoste (ServiceNow), Lee Zamparo (ServiceNow), Issam Laradji (McGill), Laurent Charlin (HEC Montreal), and David Vazquez (Element AI)

From Culture to Clothing: Discovering the World Events Behind a Century of Fashion Images 1046
Wei-Lin Hsiao (UT-Austin) and Kristen Grauman (Facebook AI Research & UT Austin)

De-Rendering Stylized Texts 1056
Wataru Shimoda (CyberAgent.Inc), Daichi Haraguchi (Kyushu University), Seiichi Uchida (Kyushu University), and Kota Yamaguchi (Cyberagent)

Handwriting Transformers 1066
Ankan Kumar Bhunia (MBZUAI), Salman Khan (Australian National University, (ANU)), Hisham Cholakkal (MBZUAI), Rao Muhammad Anwer (MBZUAI), Fahad Shahbaz Khan (MBZUAI), and Mubarak Shah (University of Central Florida)

Interpreting Attributions and Interactions of Adversarial Attacks 1075
Xin Wang (Shanghai Jiao Tong University), Shuyun Lin (Shanghai Jiao Tong Universarity), Hao Zhang (Shanghai Jiao Tong University), Yufei Zhu (shanghai jiaotong university), and Quanshi Zhang (Shanghai Jiao Tong University)

The Right To Talk: An Audio-Visual Transformer Approach 1085
Thanh-Dat Truong (University of Arkansas), Chi Nhan Duong (Concordia University), The De Vu (VinAI), Hoang Anh Pham (VinAI Research), Bhiksha Raj (Carnegie Mellon University), Ngan Le (University of Arkansas), and Khoa Luu (University of Arkansas)

Why Approximate Matrix Square Root Outperforms Accurate SVD in Global Covariance Pooling? 1095
Yue Song (University of Trento), Nicu Sebe (University of Trento), and Wei Wang (EPFL)

Striking a Balance Between Stability and Plasticity for Class-Incremental Learning 1104
Guile Wu (Queen Mary University of London), Shaogang Gong (Queen Mary University of London), and Pan Li (Queen Mary University of London)

Predicting With Confidence on Unseen Distributions 1114
Devin Guillory (UC Berkeley), Vaishaal Shankar (UC Berkeley), Sayna Ebrahimi (UC Berkeley), Trevor Darrell (UC Berkeley), and Ludwig Schmidt (Toyota Research Institute)

Transforms Based Tensor Robust PCA: Corrupted Low-Rank Tensors Recovery via Convex Optimization 1125
Canyi Lu (CMU)

CODEs: Chamfer Out-of-Distribution Examples Against Overconfidence Issue 1133
Keke Tang (The University of Hong Kong; Guangzhou University), Dingruibo Miao (Guangzhou University), Weilong Peng (Tianjin University; Guangzhou University), Jianpeng Wu (Guangzhou University), Yawen Shi (Guangzhou University), Zhaoquan Gu (Guangzhou University), Zhihong Tian (Guangzhou University), and Wenping Wang (The University of Hong Kong)

IDARTS: Interactive Differentiable Architecture Search 1143
Song Xue (Beihang Unversity), Runqi Wang (Beihang University), Baochang Zhang (Beihang University), Tian Wang (Beihang University), Guodong Guo (Baidu), and David Doermann (University at Buffalo)

MeshTalk: 3D Face Animation From Speech Using Cross-Modality Disentanglement 1153
Alexander Richard (Facebook Reality Labs), Michael Zollhöfer (Facebook Reality Labs), Yandong Wen (Carnegie Mellon University), Fernando de la Torre (Facebook), and Yaser Sheikh (Facebook Reality Labs)

Audio-Visual Floorplan Reconstruction 1163
Senthil Purushwalkam (Carnegie Mellon University), Sebastià Vicenc Amengual Garí (Facebook Reality Labs), Vamsi Krishna Ithapu (Facebook Reality Labs), Carl Schissler (Facebook Reality Labs), Philip Robinson (Facebook Reality Labs), Abhinav Gupta (CMU/FAIR), and Kristen Grauman (Facebook AI Research & UT Austin)

How To Design a Three-Stage Architecture for Audio-Visual Active Speaker Detection in the Wild 1173
Okan Köpüklü (Technical University of Munich), Maja Taseska (Microsoft), and Gerhard Rigoll (Institute for Human-Machine Communication, TU Munich, Germany)

Visual Scene Graphs for Audio Source Separation 1184
Moitreya Chatterjee (University of Illinois at Urbana Champaign), Jonathan Le Roux (Mitsubishi Electric Research Laboratories, (MERL)), Narendra Ahuja (University of Illinois at Urbana-Champaign, USA), and Anoop Cherian (MERL)

Better Aggregation in Test-Time Augmentation 1194
Divya Shanmugam (MIT), Davis Blalock (MIT), Guha Balakrishnan (Amazon Web Services, (AWS)), and John Guttag (MIT)

Explaining Local, Global, and Higher-Order Interactions in Deep Learning 1204
Samuel Lerman (University of Rochester), Charles Venuto (University of Rochester), Henry Kautz (University of Rochester), and Chenliang Xu (University of Rochester)

Explanations for Occluded Images 1214
Hana Chockler (King's College London), Daniel Kroening (Amazon.com, Inc.), and Youcheng Sun (Queen's University Belfast)

E-ViL: A Dataset and Benchmark for Natural Language Explanations in Vision-Language Tasks .. 1224
Maxime Kayser (University of Oxford), Oana-Maria Camburu (University of Oxford), Leonard Salewski (University of Tübingen), Cornelius Emde (University of Oxford), Virginie Do (Université Paris Dauphine, Facebook AI Research), Zeynep Akata (University of Tübingen), and Thomas Lukasiewicz (University of Oxford)

Broaden Your Views for Self-Supervised Video Learning 1235
Adrià Recasens (DeepMind), Pauline Luc (Deepmind), Jean-Baptiste Alayrac (DeepMind), Luyu Wang (DeepMind), Florian Strub (DeepMind), Corentin Tallec (Deepmind), Mateusz Malinowski (DeepMind), Viorica Pătrăucean (DeepMind), Florent Altché (DeepMind), Michal Valko (DeepMind), Jean-Bastien Grill (DeepMind), Aäron van den Oord (DeepMind), and Andrew Zisserman (University of Oxford)

Hypergraph Neural Networks for Hypergraph Matching 1246
Xiaowei Liao (South China University of Technology), Yong Xu (South China University of Technology), and Haibin Ling (Stony Brook University)

Embed Me if You Can: A Geometric Perceptron 1256
Pavlo Melnyk (Linköping University), Michael Felsberg (Linköping University), and Mårten Wadenbäck (Linköping University)

Learning To Discover Reflection Symmetry via Polar Matching Convolution 1265
Ahyun Seo (POSTECH), Woohyeon Shim (POSTECH), and Minsu Cho (POSTECH)

TGRNet: A Table Graph Reconstruction Network for Table Structure Recognition 1275
Wenyuan Xue (Beijing Jiaotong University), Baosheng Yu (The University of Sydney), Wen Wang (Beijing Jiaotong University), Dacheng Tao (The University of Sydney), and Qingyong Li (Beijing Jiaotong University)

Adaptive Boundary Proposal Network for Arbitrary Shape Text Detection 1285
Shi-Xue Zhang (University of Science and Technology Beijing), Xiaobin Zhu (University of Science and Technology Beijing), Chun Yang (University of Science and Technology Beijing), Hongfa Wang (Tencent Computer System Co., Ltd.), and Xu-Cheng Yin (University of Science and Technology Beijing)

Shape-Biased Domain Generalization via Shock Graph Embeddings 1295
Maruthi Narayanan (JHU Applied Physics Laboratory), Vickram Rajendran (The Johns Hopkins University Applied Physics Laboratory), and Benjamin Kimia (USA)

Towards Learning Spatially Discriminative Feature Representations 1306
Chaofei Wang (Tsinghua University), Jiayu Xiao (Tsinghua University), Yizeng Han (Tsinghua University), Qisen Yang (tsinghua university), Shiji Song (Tsinghua), and Gao Huang (Tsinghua)

Towards Better Explanations of Class Activation Mapping 1316
Hyungsik Jung (Samsung SDS) and Youngrock Oh (Mobilint)

Finding Representative Interpretations on Convolutional Neural Networks 1325
Peter Cho-Ho Lam (Huawei Technologies Canada Co., Ltd.), Lingyang Chu (McMaster University), Maxim Torgonskiy (Huawei Canada), Jian Pei (Simon Fraser University), Yong Zhang (Huawei Technologies Canada Co., Ltd.), and Lanjun Wang (Huawei Technologies Canada Co., Ltd.)

LFI-CAM: Learning Feature Importance for Better Visual Explanation 1335
Kwang Hee Lee (Boeing Korea Engineering and Technology Center, (BKETC)), Chaewon Park (Boeing Korea Engineering and Technology Center, (BKETC)), Junghyun Oh (Boeing Korea Engineering and Technology Center, (BKETC)), and Nojun Kwak (Seoul National University)

Vision and Language Understanding

Panoptic Narrative Grounding 1344
Cristina González (Universidad de los Andes), Nicolás Ayobi (University of Los Andes), Isabela Hernández (Universidad de los Andes), José Hernández (Universidad de los Andes, Colombia), Jordi Pont-Tuset (Google), and Pablo Arbeláez (Universidad de los Andes)

Who's Waldo? Linking People Across Text and Images 1354
Yuqing Cui (Cornell University), Apoorv Khandelwal (Cornell University), Yoav Artzi (Cornell University), Noah Snavely (Cornell University and Google AI), and Hadar Averbuch-Elor (Cornell University)

YouRefIt: Embodied Reference Understanding With Language and Gesture 1365
Yixin Chen (UCLA), Qing Li (UCLA), Deqian Kong (University of California, Los Angeles), Yik Lun Kei (UCLA), Song-Chun Zhu (UCLA), Tao Gao (UCLA), Yixin Zhu (UCLA), and Siyuan Huang (UCLA)

Synthesis of Compositional Animations From Textual Descriptions 1376
Anindita Ghosh (DFKI), Noshaba Cheema (DFKI), Cennet Oguz (DFKI), Christian Theobalt (MPI Informatik), and Philipp Slusallek (German Research Center for Artificial Intelligence, (DFKI) & Saarland University)

In Defense of Scene Graphs for Image Captioning 1387
Kien Nguyen (UC San Diego), Subarna Tripathi (Intel Labs), Bang Du (UC San Diego), Tanaya Guha (University of Warwick), and Truong Q. Nguyen (UC San Diego)

Unshuffling Data for Improved Generalization in Visual Question Answering 1397
Damien Teney (University of Adelaide), Ehsan Abbasnejad (The University of Adelaide), and Anton van den Hengel (University of Adelaide)

Compressing Visual-Linguistic Model via Knowledge Distillation 1408
Zhiyuan Fang (Arizona State University), Jianfeng Wang (Microsoft), Xiaowei Hu (Microsoft), Lijuan Wang (Microsoft), Yezhou Yang (Arizona State University), and Zicheng Liu (Microsoft)

UniT: Multimodal Multitask Learning With a Unified Transformer 1419
Ronghang Hu (Facebook) and Amanpreet Singh (Facebook)

CrossCLR: Cross-Modal Contrastive Learning for Multi-Modal Video Representations 1430
Mohammadreza Zolfaghari (University of Freiburg), Yi Zhu (Amazon), Peter Gehler (Amazon), and Thomas Brox (University of Freiburg)

Graph Constrained Data Representation Learning for Human Motion Segmentation 1440
Mariella Dimiccoli (CSIC-UPC), Lluís Garrido (UB), Guillem Rodriguez-Corominas (UPC), and Herwig Wendt (CNRS, University of Toulouse)

Zero-Shot Natural Language Video Localization 1450
Jinwoo Nam (GIST), Daechul Ahn (GIST), Dongyeop Kang (University of Minnesota), Seong Jong Ha (NCSOFT), and Jonghyun Choi (GIST)

Learning Temporal Dynamics From Cycles in Narrated Video 1460
Dave Epstein (UC Berkeley), Jiajun Wu (Stanford University), Cordelia Schmid (Google), and Chen Sun (Google)

Dense Interaction Learning for Video-Based Person Re-Identification 1470
Tianyu He (Alibaba Group), Xin Jin (University of Science and Technology of China), Xu Shen (Alibaba Group), Jianqiang Huang (Damo Academy, Alibaba Group), Zhibo Chen (University of Science and Technology of China), and Xian-Sheng Hua (Damo Academy, Alibaba Group)

Vi2CLR: Video and Image for Visual Contrastive Learning of Representation 1482
Ali Diba (KU Leuven), Vivek Sharma (Harvard, MIT), Reza Safdari (Sensifai), Dariush Lotfi (Sensifai), Saquib Sarfraz (Karlsruhe Institute of Technology), Rainer Stiefelhagen (Karlsruhe Institute of Technology), and Luc Van Gool (ETH Zurich)

MGSampler: An Explainable Sampling Strategy for Video Action Recognition 1493
Yuan Zhi (Nanjing University), Zhan Tong (Nanjing University), Limin Wang (Nanjing University), and Gangshan Wu (Nanjing University)

Fast Video Moment Retrieval 1503
Junyu Gao (CASIA) and Changsheng Xu (CASIA)

STVGBert: A Visual-Linguistic Transformer Based Framework for Spatio-Temporal Video Grounding 1513
Rui Su (the University of Sydney), Qian Yu (Beihang University), and Dong Xu (University of Sydney)

Motion Guided Region Message Passing for Video Captioning 1523
Shaoxiang Chen (Fudan University) and Yu-Gang Jiang (Fudan University)

Dynamic Context-Sensitive Filtering Network for Video Salient Object Detection 1533
Miao Zhang (Dalian University of Technology, China), Jie Liu (Dalian University of Technology, China), Yifei Wang (Dalian University of Technology, China), Yongri Piao (Dalian University of Technology, China), Shunyu Yao (Dalian University of Technology, China), Wei Ji (University of Alberta, Canada), Jingjing Li (University of Alberta, Canada), Huchuan Lu (Dalian University of Technology, China), and Zhongxuan Luo (Dalian University of Technology, China)

Learning Motion-Appearance Co-Attention for Zero-Shot Video Object Segmentation 1544
Shu Yang (Dalian University of Technology), Lu Zhang (Dalian University of Technology), Jinqing Qi (Dalian University of Technology), Huchuan Lu (Dalian University of Technology), Shuo Wang (Meituan), and Xiaoxing Zhang (Meituan)

Beyond Question-Based Biases: Assessing Multimodal Shortcut Learning in Visual Question Answering 1554
Corentin Dancette (Sorbonne Universite, CNRS, LIP6, Paris), Rémi Cadène (Sorbonne Universite, CNRS, LIP6, Paris; Brown University, USA), Damien Teney (Idiap Research Institute; University of Adelaide), and Matthieu Cord (Sorbonne Universite, CNRS, LIP6, Paris; Valeo.ai)

Greedy Gradient Ensemble for Robust Visual Question Answering 1564
Xinzhe Han (University of Chinese Academy of Sciences), Shuhui Wang (VIPL, ICT, Chinese academic of science), Chi Su (Kingsoft Cloud), Qingming Huang (University of Chinese Academy of Sciences), and Qi Tian (Huawei Cloud & AI)

Self-Motivated Communication Agent for Real-World Vision-Dialog Navigation 1574
Yi Zhu (Noah's Ark Lab, Huawei Technologies), Yue Weng (Sun Yat-sen University), Fengda Zhu (Monash), Xiaodan Liang (Sun Yat-sen University), Qixiang Ye (University of Chinese Academy of Sciences, China), Yutong Lu (Sun Yat-sen University), and Jianbin Jiao (University of Chinese Academy of Sciences)

Contrast and Classify: Training Robust VQA Models 1584
Yash Kant (Georgia Institute of Technology), Abhinav Moudgil (Georgia Institute of Technology), Dhruv Batra (Georgia Tech & Facebook AI Research), Devi Parikh (Georgia Tech & Facebook AI Research), and Harsh Agrawal (Georgia Institute of Technology)

Linguistically Routing Capsule Network for Out-of-Distribution Visual Question Answering 1594
Qingxing Cao (Sun Yat-sen University, China), Wentao Wan (Sun Yat-sen University, China), Keze Wang (Sun Yat-sen University, China), Xiaodan Liang (Sun Yat-sen University, China), and Liang Lin (Sun Yat-sen University, China)

LapsCore: Language-Guided Person Search via Color Reasoning 1604
Yushuang Wu (CUHKSZ), Zizheng Yan (CUHK, Shenzhen), Xiaoguang Han (Shenzhen Research Institute of Big Data, the Chinese University of Hong Kong, (Shenzhen)), Guanbin Li (Sun Yat-sen University), Changqing Zou (Huawei Technologies Canada), and Shuguang Cui (The Chinese University of Hong Kong, Shenzhen)

Airbert: In-Domain Pretraining for Vision-and-Language Navigation 1614
Pierre-Louis Guhur (INRIA), Makarand Tapaswi (Wadhwani AI, IIIT Hyderbad), Shizhe Chen (INRIA), Ivan Laptev (Inria/Google), and Cordelia Schmid (INRIA Paris)

Vision-Language Navigation With Random Environmental Mixup 1624
Chong Liu (University of Chinese Academy of Sciences), Fengda Zhu (Monash University, Australia), Xiaojun Chang (RMIT University, Australia), Xiaodan Liang (Sun Yat-sen University, China), Zongyuan Ge (Monash University, Australia), and Yi-Dong Shen (Chinese Academy of Sciences)

The Road To Know-Where: An Object-and-Room Informed Sequential BERT for Indoor Vision-Language Navigation 1635
Yuankai Qi (The University of Adelaide), Zizheng Pan (Monash University), Yicong Hong (Australian National University), Ming-Hsuan Yang (University of California at Merced), Anton van den Hengel (University of Adelaide), and Qi Wu (University of Adelaide)

VLGrammar: Grounded Grammar Induction of Vision and Language 1645
Yining Hong (University of California, Los Angeles), Qing Li (UCLA), Song-Chun Zhu (UCLA), and Siyuan Huang (UCLA)

Env-QA: A Video Question Answering Benchmark for Comprehensive Understanding of Dynamic Environments 1655
Difei Gao (ICT, CAS), Ruiping Wang (Institute of Computing Technology, Chinese Academy of Sciences), Ziyi Bai (Institute of Computing Technology, Chinese Academy of Sciences), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences)

Just Ask: Learning To Answer Questions From Millions of Narrated Videos 1666
Antoine Yang (Inria), Antoine Miech (DeepMind), Josef Sivic (Czech Technical University), Ivan Laptev (INRIA Paris), and Cordelia Schmid (Inria/Google)

HAIR: Hierarchical Visual-Semantic Relational Reasoning for Video Question Answering 1678
Fei Liu (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences and University of Chinese Academy of Sciences), Jing Liu (National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences), Weining Wang (Institute of Automation, Chinese Academy of Sciences), and Hanqing Lu (NLPR, Institute of Automation, CAS)

Video Question Answering Using Language-Guided Deep Compressed-Domain Video Feature .. 1688
Nayoung Kim (Ewha Womans University), Seong Jong Ha (NCSOFT), and Je-Won Kang (Ewha Womans University)

Multiple Pairwise Ranking Networks for Personalized Video Summarization 1698
Yassir Saquil (University of Bath), Da Chen (Alibaba Group), Yuan He (Alibaba Group), Chuan Li (Lambda Labs), and Yong-Liang Yang (University of Bath)

Frozen in Time: A Joint Video and Image Encoder for End-to-End Retrieval 1708
Max Bain (University of Oxford, UK), Arsha Nagrani (University of Oxford, UK), Gül Varol (Ecole des Ponts ParisTech), and Andrew Zisserman (University of Oxford, UK)

Video Instance Segmentation With a Propose-Reduce Paradigm 1719
Huaijia Lin (The Chinese University of Hong Kong), Ruizheng Wu (The Chinese University of Hong Kong), Shu Liu (SmartMore), Jiangbo Lu (SmartMore Corporation), and Jiaya Jia (Chinese University of Hong Kong)

Deep 3D Mask Volume for View Synthesis of Dynamic Scenes 1729
Kai-En Lin (University of California San Diego), Lei Xiao (Facebook Reality Labs), Feng Liu (Portland State University, Facebook Reality Labs), Guowei Yang (University of California, San Diego), and Ravi Ramamoorthi (University of California San Diego)

Unsupervised Deep Video Denoising 1739
Dev Yashpal Sheth (Indian Institute of Technology Madras), Sreyas Mohan (NYU), Joshua L. Vincent (ASU), Ramon Manzorro (ASU), Peter A. Crozier (ASU), Mitesh M. Khapra (Indian Institute of Technology Madras), Eero P. Simoncelli (New York University), and Carlos Fernandez-Granda (NYU)

TransVG: End-to-End Visual Grounding With Transformers 1749
Jiajun Deng (University of Science and Technology of China), Zhengyuan Yang (University of Rochester), Tianlang Chen (University of Rochester), Wengang Zhou (University of Science and Technology of China), and Houqiang Li (University of Science and Technology of China)

MDETR – Modulated Detection for End-to-End Multi-Modal Understanding 1760
Aishwarya Kamath (New York University), Mannat Singh (Facebook AI Research), Yann LeCun (New York University), Gabriel Synnaeve (Facebook), Ishan Misra (Facebook AI Research), and Nicolas Carion (NYU)

InstanceRefer: Cooperative Holistic Understanding for Visual Grounding on Point Clouds Through Instance Multi-Level Contextual Referring 1771
Zhihao Yuan (The Chinese University of Hong Kong, Shenzhen), Xu Yan (The Chinese University of Hong Kong, Shenzhen), Yinghong Liao (The Chinese University of Hong Kong, Shenzhen), Ruimao Zhang (The Chinese University of Hong Kong, Shenzhen), Sheng Wang (Tencent), Zhen Li (Chinese University of Hong Kong, Shenzhen), and Shuguang Cui (The Chinese University of Hong Kong, Shenzhen)

Detector-Free Weakly Supervised Grounding by Separation 1781
Assaf Arbelle (IBM Research AI), Sivan Doveh (IBM-Research), Amit Alfassy (IBM-Research), Joseph Shtok (IBM-Reseach), Guy Lev (Haifa Research Lab), Eli Schwartz (IBM-Research), Hilde Kuehne (University of Frankfurt), Hila Barak Levi (Weizmann Institute of Science), Prasanna Sattigeri (IBM Research), Rameswar Panda (MIT-IBM Watson AI Lab, IBM Research), Chun-Fu Richard Chen (MIT-IBM Watson AI Lab, IBM Research AI), Alex Bronstein (Tel Aviv University, Israel), Kate Saenko (Boston University), Shimon Ullman (Weizmann institute of science), Raja Giryes (Tel Aviv University), Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research), and Leonid Karlinsky (IBM-Research)

Wasserstein Coupled Graph Learning for Cross-Modal Retrieval 1793
Yun Wang (NanJing University of Science & Technology), Tong Zhang (Nanjing University of Science and Technology), Xueya Zhang (Nanjing University of Science and Technology), Zhen Cui (Nanjing University of Science and Technology), Yuge Huang (Tencent YouTu), Pengcheng Shen (Tencent), Shaoxin Li (Tencent), and Jian Yang (Nanjing University of Science and Technology)

Learning To Generate Scene Graph From Natural Language Supervision 1803
Yiwu Zhong (University of Wisconsin-Madison), Jing Shi (University of Rochester), Jianwei Yang (Microsoft Research), Chenliang Xu (University of Rochester), and Yin Li (University of Wisconsin-Madison)

Ask&Confirm: Active Detail Enriching for Cross-Modal Retrieval With Partial Query 1815
Guanyu Cai (Tongji University), Jun Zhang (Tencent), Xinyang Jiang (Tencent), Yifei Gong (Tencent), Lianghua He (), Fufu Yu (Tencent Youtu), Pai Peng (Tencent Youtu Lab), Xiaowei Guo (Tencent Youtu Lab), Feiyue Huang (Tencent), and Xing Sun (Tencent)

Weakly Supervised Human-Object Interaction Detection in Video via Contrastive Spatiotemporal Regions 1825
Shuang Li (MIT), Yilun Du (MIT), Antonio Torralba (MIT), Josef Sivic (Czech Technical University), and Bryan Russell (Adobe Research)

SAT: 2D Semantics Assisted Training for 3D Visual Grounding 1836
Zhengyuan Yang (University of Rochester), Songyang Zhang (University of Rochester), Liwei Wang (CUHK), and Jiebo Luo (U. Rochester)

Adaptive Hierarchical Graph Reasoning With Semantic Coherence for Video-and-Language Inference 1847
Juncheng Li (Zhejiang University), Siliang Tang (Zhejiang University), Linchao Zhu (University of Technology, Sydney), Haochen Shi (Université de Montréal), Xuanwen Huang (Zhejiang University), Fei Wu (Zhejiang University, China), Yi Yang (UTS), and Yueting Zhuang (Zhejiang University)

Interpretable Visual Reasoning via Induced Symbolic Space 1858
Zhonghao Wang (UIUC), Kai Wang (SHI Lab @ UOregon), Mo Yu (IBM T. J. Watson), Jinjun Xiong (IBM Thomas J. Watson Research Center), Wen-mei Hwu (University of Illinois at Urbana-Champaign), Mark Hasegawa-Johnson (University of Illinois), and Humphrey Shi (U of Oregon; UIUC)

Factorizing Perception and Policy for Interactive Instruction Following 1868
Kunal Pratap Singh (Allen Institute for AI), Suvaansh Bhambri (GIST), Byeonghwi Kim (GIST), Roozbeh Mottaghi (Allen Institute for AI), and Jonghyun Choi (GIST)

Unified Questioner Transformer for Descriptive Question Generation in Goal-Oriented Visual Dialogue 1878
Shoya Matsumori (Keio University), Kosuke Shingyouchi (Keio University), Yuki Abe (Keio University), Yosuke Fukuchi (Keio University), Komei Sugiura (Keio University), and Michita Imai (Keio University)

Weakly Supervised Relative Spatial Reasoning for Visual Question Answering 1888
Pratyay Banerjee (Arizona State University), Tejas Gokhale (Arizona State University), Yezhou Yang (Arizona State University), and Chitta Baral (Arizona State University)

Mixed SIGNals: Sign Language Production via a Mixture of Motion Primitives 1899
Ben Saunders (University of Surrey), Necati Cihan Camgoz (University of Surrey), and Richard Bowden (University of Surrey)

Localize to Binauralize: Audio Spatialization From Visual Sound Source Localization 1910
Kranthi Kumar Rachavarapu (Indian Institute of Technology Madras), Aakanksha Aakanksha (Indian Institute of Technology Madras), Vignesh Sundaresha (Indian Institute of Technology Madras), and A. N. Rajagopalan (Indian Institute of Technology Madras)

Spatial-Temporal Consistency Network for Low-Latency Trajectory Forecasting 1920
Shijie Li (Bonn University), Yanying Zhou (University of Bonn), Jinhui Yi (University of Bonn), and Juergen Gall (University of Bonn)

T-Net: Effective Permutation-Equivariant Network for Two-View Correspondence Learning 1930
Zhen Zhong (Minjiang University), Guobao Xiao (Minjiang University), Linxin Zheng (Fuzhou University), Yan Lu (Fuzhou University), and Jiayi Ma (Wuhan University)

IntraTomo: Self-Supervised Learning-Based Tomography via Sinogram Synthesis and Prediction. 1940
Guangming Zang (KAUST), Ramzi Idoughi (KAUST), Rui Li (KAUST), Peter Wonka (KAUST), and Wolfgang Heidrich (KAUST)

Describing and Localizing Multiple Changes With Transformers 1951
Yue Qiu (National Institute of Advanced Industrial Science and Technology, (AIST)), Shintaro Yamamoto (National Institute of Advanced Industrial Science and Technology, (AIST) / Waseda University), Kodai Nakashima (National Institute of Advanced Industrial Science and Technology, (AIST)), Ryota Suzuki (National Institute of Advanced Industrial Science and Technology, (AIST)), Kenji Iwata (National Institute of Advanced Industrial Science and Technology, (AIST)), Hirokatsu Kataoka (National Institute of Advanced Industrial Science and Technology, (AIST)), and Yutaka Satoh (National Institute of Advanced Industrial Science and Technology, (AIST))

Cross-Camera Convolutional Color Constancy 1961
Mahmoud Afifi (York University), Jonathan T. Barron (Google Research), Chloe LeGendre (Google Inc.), Yun-Ta Tsai (GOOGLE INC), and Francois Bleibel (Google, Inc.)

IICNet: A Generic Framework for Reversible Image Conversion 1971
Ka Leong Cheng (The Hong Kong University of Science and Technology), Yueqi Xie (The Hong Kong University of Science and Technology), and Qifeng Chen (HKUST)

Dual-Camera Super-Resolution With Aligned Attention Modules 1981
Tengfei Wang (HKUST), Jiaxin Xie (The Hong Kong university of science and technology), Wenxiu Sun (SenseTime Research and Tetras.AI), Qiong Yan (SenseTime Group Limited), and Qifeng Chen (HKUST)

Let's See Clearly: Contaminant Artifact Removal for Moving Cameras 1991
Xiaoyu Li (Hong Kong University of Science and Technology), Bo Zhang (Microsoft Research Asia), Jing Liao (City University of Hong Kong), and Pedro V. Sander (HKUST)

Explainable Video Entailment With Grounded Visual Evidence 2001
Junwen Chen (Rochester Institute of Technology) and Yu Kong (Rochester Institute of Technology)

Pano-AVQA: Grounded Audio-Visual Question Answering on 360° Videos 2011
Heeseung Yun (Seoul National University), Youngjae Yu (Seoul National University Vision and Learning Lab), Wonsuk Yang (Vision and Learning Lab, Seoul National University), Kangil Lee (Hyundai Motor Company), and Gunhee Kim (Seoul National University)

Adversarial VQA: A New Benchmark for Evaluating the Robustness of VQA Models 2022
Linjie Li (Microsoft), Jie Lei (UNC Chapel Hill), Zhe Gan (Microsoft), and Jingjing Liu (Tsinghua University)

AESOP: Abstract Encoding of Stories, Objects, and Pictures 2032
Hareesh Ravi (Rutgers University), Kushal Kafle (Adobe Research), Scott Cohen (Adobe Research), Jonathan Brandt (-), and Mubbasir Kapadia (Rutgers University)

On the Hidden Treasure of Dialog in Video Question Answering 2044
Deniz Engin (Inria), François Schnitzler (InterDigital), Ngoc Q. K. Duong (Interdigital), and Yannis Avrithis (Inria)

TRAR: Routing the Attention Spans in Transformer for Visual Question Answering 2054
Yiyi Zhou (Xiamen University), Tianhe Ren (XMU), Chaoyang Zhu (Xiamen University), Xiaoshuai Sun (Xiamen University), Jianzhuang Liu (Noah's Ark Lab, Huawei Technologies), Xinghao Ding (Xiamen University), Mingliang Xu (Zhengzhou University), and Rongrong Ji (Xiamen University, China)

StyleCLIP: Text-Driven Manipulation of StyleGAN Imagery 2065
Or Patashnik (Tel Aviv University), Zongze Wu (Hebrew University of Jerusalem), Eli Shechtman (The Hebrew University of Jerusalem), Daniel Cohen-Or (Adobe Research, US), and Dani Lischinski (Tel Aviv University)

Viewpoint-Agnostic Change Captioning With Cycle Consistency 2075
Hoeseong Kim (Seoul National University), Jongseok Kim (Violet), Hyungseok Lee (Hyundai Motor Group AIRS Company), Hyunsung Park (Hyundai Motor Group AIRS Company), and Gunhee Kim (Seoul National University)

Motion-Focused Contrastive Learning of Video Representations 2085
Rui Li (University of Science and Technology of China), Yiheng Zhang (JD AI Research), Zhaofan Qiu (JD.com), Ting Yao (JD AI Research), Dong Liu (University of Science and Technology of China), and Tao Mei (AI Research of JD.com)

Language-Guided Global Image Editing via Cross-Modal Cyclic Mechanism 2095
Wentao Jiang (Beihang University), Ning Xu (Adobe Research), Jiayun Wang (Beihang University), Chen Gao (Beihang University), Jing Shi (University of Rochester), Zhe Lin (Adobe Research), and Si Liu (Beihang University)

Image Retrieval on Real-Life Images With Pre-Trained Vision-and-Language Models 2105
Zheyuan Liu (The Australian National University), Cristian Rodriguez-Opazo (University of Adelaide), Damien Teney (University of Adelaide), and Stephen Gould (Australian National University, Australia)

Dual Transfer Learning for Event-Based End-Task Prediction via Pluggable Event to Image Translation 2115
Lin Wang (KAIST), Yujeong Chae (KAIST), and Kuk-Jin Yoon (KAIST)

N-ImageNet: Towards Robust, Fine-Grained Object Recognition With Event Cameras 2126
Junho Kim (Seoul National University), Jaehyeok Bae (Seoul National University), Gangin Park (Seoul National University), Dongsu Zhang (Seoul National University), and Young Min Kim (Seoul National University)

Patch Craft: Video Denoising by Deep Modeling and Patch Matching 2137
Gregory Vaksman (Technion), Michael Elad (Google), and Peyman Milanfar (Google)

LocTex: Learning Data-Efficient Visual Representations From Localized Textual Supervision 2147
Zhijian Liu (MIT), Simon Stent (Toyota Research Institute), Jie Li (Toyota Research Institute), John Gideon (Toyota Research Institute), and Song Han (MIT)

Hierarchical Graph Attention Network for Few-Shot Visual-Semantic Learning 2157
Chengxiang Yin (Syracuse University), Kun Wu (Syracuse University), Zhengping Che (Didi Chuxing), Bo Jiang (DiDi AI Labs, Didi Chuxing), Zhiyuan Xu (Syracuse University), and Jian Tang (Syracuse University)

Partial Off-Policy Learning: Balance Accuracy and Diversity for Human-Oriented Image Captioning 2167
Jiahe Shi (Tsinghua University), Yali Li (Tsinghua University), and Shengjin Wang (Tsinghua University)

Auto-Parsing Network for Image Captioning and Visual Question Answering 2177
Xu Yang (Nanyang Technological University), Chongyang Gao (Dartmouth College), Hanwang Zhang (Nanyang Technological University), and Jianfei Cai (Monash University)

COOKIE: Contrastive Cross-Modal Knowledge Sharing Pre-Training for Vision-Language Representation 2188
Keyu Wen (Bytedance), Jin Xia (ByteDance), Yuanyuan Huang (Bytedance Co.), Linyang Li (Fudan University), Jiayan Xu (ByteDance), and Jie Shao (Bytedance)

Adversarial Attack on Deep Cross-Modal Hamming Retrieval 2198
Chao Li (Xidian University), Shangqian Gao (University of Pittsburgh), Cheng Deng (Xidian University), Wei Liu (Tencent), and Heng Huang (University of Pittsburgh & JD Explore Academy)

Defocus Map Estimation and Deblurring From a Single Dual-Pixel Image 2208
Shumian Xin (Carnegie Mellon University), Neal Wadhwa (Google), Tianfan Xue (Google), Jonathan T. Barron (Google Research), Pratul P. Srinivasan (Google Research), Jiawen Chen (Google), Ioannis Gkioulekas (Carnegie Mellon University), and Rahul Garg (Google)

How To Train Neural Networks for Flare Removal 2219
Yicheng Wu (Rice University), Qiurui He (Google Research), Tianfan Xue (Google), Rahul Garg (Google), Jiawen Chen (Adobe), Ashok Veeraraghavan (Rice University), and Jonathan T. Barron (Google Research)

Hyperspectral Image Denoising With Realistic Data 2228
Tao Zhang (Beijing Institute of Technology), Ying Fu (Beijing Institute of Technology), and Cheng Li (Huawei Noah's Ark Lab)

Dynamic CT Reconstruction From Limited Views With Implicit Neural Representations and Parametric Motion Fields 2238
Albert W. Reed (Arizona State University), Hyojin Kim (LLNL), Rushil Anirudh (Lawrence Livermore National Laboratory), K. Aditya Mohan (Lawrence Livermore National Laboratory), Kyle Champley (LLNL), Jingu Kang (Lawrence Livermore National Laboratory), and Suren Jayasuriya (Arizona State University)

High Quality Disparity Remapping With Two-Stage Warping 2249
Bing Li (KAUST), Chia-Wen Lin (National Tsing Hua University), Cheng Zheng (KAUST), Shan Liu (Tencent America), Junsong Yuan (State University of New York at Buffalo, USA), Bernard Ghanem (KAUST), and C.-C. Jay Kuo (USC)

Semantic-Embedded Unsupervised Spectral Reconstruction From Single RGB Images in the Wild 2259
Zhiyu Zhu (City University of Hong Kong), Hui Liu (City University of Hong Kong), Junhui Hou (City University of Hong Kong, Hong Kong), Huanqiang Zeng (Huaqiao University), and Qingfu Zhang (City University of Hong Kong)

Learning To Reduce Defocus Blur by Realistically Modeling Dual-Pixel Data 2269
Abdullah Abuolaim (York University), Mauricio Delbracio (Google Research), Damien Kelly (Google Inc.), Michael S. Brown (York University), and Peyman Milanfar (Google)

Hybrid Neural Fusion for Full-Frame Video Stabilization 2279
Yu-Lun Liu (National Taiwan University; MediaTek), Wei-Sheng Lai (Google), Ming-Hsuan Yang (University of California at Merced), Yung-Yu Chuang (National Taiwan University), and Jia-Bin Huang (Virginia Tech)

Spatially-Adaptive Image Restoration Using Distortion-Guided Networks 2289
Kuldeep Purohit (Michigan State University), Maitreya Suin (Indian Institute of Technology Madras), A. N. Rajagopalan (Indian Institute of Technology Madras), and Vishnu Naresh Boddeti (Michigan State University)

Anonymizing Egocentric Videos 2300
Daksh Thapar (Indian Institute of Technology, Mandi), Aditya Nigam (IIT mandi), and Chetan Arora (Indian Institute of Technology Delhi)

What You Can Learn by Staring at a Blank Wall 2310
Prafull Sharma (MIT), Miika Aittala (NVIDIA), Yoav Y. Schechner (Technion), Antonio Torralba (MIT), Gregory W. Wornell (MIT), William T. Freeman (MIT), and Frédo Durand (MIT)

Inference of Black Hole Fluid-Dynamics From Sparse Interferometric Measurements 2320
Aviad Levis (Caltech), Daeyoung Lee (University of Illinois at Urbana-Champaign), Joel A. Tropp (Caltech), Charles F. Gammie (University of Illinois at Urbana-Champaign), and Katherine L. Bouman (Caltech)

C2N: Practical Generative Noise Modeling for Real-World Denoising 2330
Geonwoon Jang (Seoul National University), Wooseok Lee (Seoul National University), Sanghyun Son (Seoul National University), and Kyoung Mu Lee (Seoul National University)

Fourier Space Losses for Efficient Perceptual Image Super-Resolution 2340
Dario Fuoli (ETH Zurich), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)

Lucas-Kanade Reloaded: End-to-End Super-Resolution From Raw Image Bursts 2350
Bruno Lecouat (Inria), Jean Ponce (Inria), and Julien Mairal (INRIA)

Variable-Rate Deep Image Compression Through Spatially-Adaptive Feature Transform 2360
Myungseo Song (Seoul National University), Jinyoung Choi (Seoul National University), and Bohyung Han (Seoul National University)

V-DESIRR: Very Fast Deep Embedded Single Image Reflection Removal 2370
B H Pawan Prasad (Samsung Research), Green Rosh K K S (Samsung Research Institute Bangalore), Lokesh R. Boregowda (Samsung Research Institute Bangalore), Kaushik Mitra (IIT Madras), and Sanjoy Chowdhury (Samsung R&D Institute Bangalore)

NeuSpike-Net: High Speed Video Reconstruction via Bio-Inspired Neuromorphic Cameras 2380
Lin Zhu (Peking University), Jianing Li (Peking University), Xiao Wang (Peng Cheng Laboratory), Tiejun Huang (Peking University), and Yonghong Tian (Peking University)

Large Scale Multi-Illuminant (LSMI) Dataset for Developing White Balance Algorithm Under Mixed Illumination 2390
Dongyoung Kim (Yonsei University), Jinwoo Kim (Yonsei University), Seonghyeon Nam (York University), Dongwoo Lee (Yonsei University), Yeonkyung Lee (Samsung Advanced Institute of Technology), Nahyup Kang (Samsung Advanced Institute of Technology), Hyong-Euk Lee (Samsung Advanced Institute of Technology), ByungIn Yoo (Samsung Advanced Institute of Technology), Jae-Joon Han (Samsung), and Seon Joo Kim (Yonsei University)

A Light Stage on Every Desk 2400
Soumyadip Sengupta (University of Washington), Brian Curless (University of Washington), Ira Kemelmacher-Shlizerman (University of Washington), and Steven M. Seitz (University of Washington)

A Dark Flash Normal Camera 2410
Zhihao Xia (Washington University in St. Louis), Jason Lawrence (Google), and Supreeth Achar (Google, Inc.)

Virtual Light Transport Matrices for Non-Line-of-Sight Imaging 2420
Julio Marco (Universidad de Zaragoza), Adrian Jarabo (University of Zaragoza), Ji Hyun Nam (University of Wisconsin Madison), Xiaochun Liu (University of Wisconsin - Madison), Miguel Ángel Cosculluela (Universidad de Zaragoza), Andreas Velten (University of Wisconsin - Madison), and Diego Gutierrez (University of Zaragoza)

Learning Dynamic Interpolation for Extremely Sparse Light Fields With Wide Baselines 2430
Mantang Guo (City University of Hong Kong), Jing Jin (City University of Hong Kong), Hui Liu (City University of Hong Kong), and Junhui Hou (City University of Hong Kong, Hong Kong)

Deep Reparametrization of Multi-Frame Super-Resolution and Denoising 2440
Goutam Bhat (ETH Zurich), Martin Danelljan (ETH Zurich), Fisher Yu (ETH Zurich), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)

Real-Time Image Enhancer via Learnable Spatial-Aware 3D Lookup Tables 2451
Tao Wang (Huawei Noah's Ark Lab), Yong Li (Huawei Noah's Ark Lab), Jingyang Peng (Huawei Noah's Ark Lab), Yipeng Ma (Huawei Noah's Ark Lab), Xian Wang (Huawei Noah's Ark Lab), Fenglong Song (Huawei Noah's Ark Lab), and Youliang Yan (Huawei)

Distillation-Guided Image Inpainting 2461
Maitreya Suin (Indian Institute of Technology Madras), Kuldeep Purohit (Michigan State University), and A. N. Rajagopalan (Indian Institute of Technology Madras)

SeLFVi: Self-Supervised Light-Field Video Reconstruction From Stereo Video 2471
Prasan Shedligeri (Indian Institute of Technology Madras), Florian Schiffers (Northwestern University, Department of Radiology), Sushobhan Ghosh (Northwestern University), Oliver Cossairt (Northwestern University, USA), and Kaushik Mitra (IIT Madras)

HDR Video Reconstruction: A Coarse-To-Fine Network and a Real-World Benchmark Dataset 2482
Guanying Chen (The University of Hong Kong), Chaofeng Chen (The University of Hong Kong), Shi Guo (The Hong Kong Polytechnic University), Zhetong Liang (The Hong Kong Polytechnic University), Kwan-Yee K. K. Wong (The University of Hong Kong), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)

Photon-Starved Scene Inference Using Single Photon Cameras 2492
Bhavya Goyal (University of Wisconsin-Madison) and Mohit Gupta (University of Wisconsin-Madison, USA)

Unsupervised Non-Rigid Image Distortion Removal via Grid Deformation 2502
Nianyi Li (Louisiana State University), Simron Thapa (Louisiana State University), Cameron Whyte (Arizona State University), Albert W. Reed (Arizona State University), Suren Jayasuriya (Arizona State University), and Jinwei Ye (Louisiana State University)

Super Resolve Dynamic Scene From Continuous Spike Streams 2513
Jing Zhao (Peking University), Jiyu Xie (University of Science and Technology of China), Ruiqin Xiong (Peking University), Jian Zhang (Peking University Shenzhen Graduate School), Zhaofei Yu (Peking University), and Tiejun Huang (Peking University)

COMISR: Compression-Informed Video Super-Resolution 2523
Yinxiao Li (Google), Pengchong Jin (Google Research), Feng Yang (Google Research), Ce Liu (Google), Ming-Hsuan Yang (Google Research), and Peyman Milanfar (Google)

Multitask AET With Orthogonal Tangent Regularity for Dark Object Detection 2533
Ziteng Cui (Shanghai Jiaotong University), Guo-Jun Qi (Futurewei Technologies), Lin Gu (RIKENAIP / The University of Tokyo), Shaodi You (University of Amsterdam), Zenghui Zhang (Shanghai Jiao Tong university), and Tatsuya Harada (The University of Tokyo / RIKEN)

Event-Based Video Reconstruction Using Transformer 2543
Wenming Weng (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)

Learning Privacy-Preserving Optics for Human Pose Estimation 2553
Carlos Hinojosa (Universidad Industrial de Santander), Juan Carlos Niebles (Stanford University), and Henry Arguello (Universidad Industrial Santander)

Motion Deblurring With Real Events 2563
Fang Xu (Wuhan University), Lei Yu (Wuhan University), Bishan Wang (Wuhan University), Wen Yang (Wuhan University), Gui-Song Xia (Wuhan University), Xu Jia (Dalian University of Technology), Zhendong Qiao (Huawei, Multimedia Technology Dept), and Jianzhuang Liu (Noah's Ark Lab, Huawei Technologies)

Objects As Cameras: Estimating High-Frequency Illumination From Shadows 2573
Tristan Swedish (Massachusetts Institute of Technology), Connor Henley (Massachusetts Institute of Technology), and Ramesh Raskar (Massachusetts Institute of Technology)

A Simple Framework for 3D Lensless Imaging With Programmable Masks 2583
Yucheng Zheng (University of California, Riverside), Yi Hua (Carnegie Mellon University), Aswin C. Sankaranarayanan (Carnegie Mellon University), and M. Salman Asif (University of California, Riverside)

Universal and Flexible Optical Aberration Correction Using Deep-Prior Based Deconvolution 2593
Xiu Li (Tsinghua University), Jinli Suo (Tsinghua University), Weihang Zhang (Tsinghua University), Xin Yuan (Bell Labs), and Qionghai Dai (Tsinghua University)

Self-Supervised Neural Networks for Spectral Snapshot Compressive Imaging 2602
Ziyi Meng (Beijing University of Posts and Telecommunications), Zhenming Yu (Beijing University of Posts and Telecommunications), Kun Xu (Beijing University of Posts and Telecommunications), and Xin Yuan (Bell Labs)

Extreme-Quality Computational Imaging via Degradation Framework 2612
Shiqi Chen (Zhejiang University, China), Huajun Feng (Zhejiang University, China), Keming Gao (Zhejiang University, China), Zhihai Xu (Zhejiang University, China), and Yueting Chen (Zhejiang University, China)

Single Image Defocus Deblurring Using Kernel-Sharing Parallel Atrous Convolutions 2622
Hyeongseok Son (POSTECH), Junyong Lee (POSTECH), Sunghyun Cho (POSTECH), and Seungyong Lee (POSTECH)

Single-Shot Hyperspectral-Depth Imaging With Learned Diffractive Optics 2631
Seung-Hwan Baek (Princeton University), Hayato Ikoma (Stanford University), Daniel S. Jeon (KAIST), Yuqi Li (King Abdullah University of Science and Technology), Wolfgang Heidrich (KAUST), Gordon Wetzstein (Stanford University), and Min H. Kim (KAIST)

Incorporating Learnable Membrane Time Constant To Enhance Learning of Spiking Neural Networks 2641
Wei Fang (Peking University), Zhaofei Yu (Peking University), Yanqi Chen (Peking University), Timothée Masquelier (Centre de Recherche Cerveau et Cognition, (CERCO)), Tiejun Huang (Peking University), and Yonghong Tian (Peking University)

Multispectral Illumination Estimation Using Deep Unrolling Network 2652
Yuqi Li (King Abdullah University of Science and Technology), Qiang Fu (KAUST), and Wolfgang Heidrich (KAUST)

A Hybrid Frequency-Spatial Domain Model for Sparse Image Reconstruction in Scanning Transmission Electron Microscopy 2662
Bintao He (Shandong University), Fa Zhang (Institute of Computing Technology, Chinese Academy of Sciences), Huanshui Zhang (Shandong University), and Renmin Han (Research Center for Mathematics and Interdisciplinary Sciences, Shandong University)

Time-Multiplexed Coded Aperture Imaging: Learned Coded Aperture and Pixel Exposures for Compressive Imaging Systems 2672
Edwin Vargas (Universidad Industrial de Santander), Julien N. P. Martel (Stanford University), Gordon Wetzstein (Stanford University), and Henry Arguello (Universidad Industrial Santander)

Video Analysis and Understanding

Dual Bipartite Graph Learning: A General Approach for Domain Adaptive Object Detection 2683
Chaoqi Chen (The University of Hong Kong), Jiongcheng Li (Xiamen University), Zebiao Zheng (Xiamen University), Yue Huang (Xiamen University), Xinghao Ding (Xiamen University), and Yizhou Yu (The University of Hong Kong)

The Devil Is in the Task: Exploiting Reciprocal Appearance-Localization Features for Monocular 3D Object Detection 2693
Zhikang Zou (Huazhong University of Science and Technology, China), Xiaoqing Ye (Baidu Inc), Liang Du (Fudan University, China), Xianhui Cheng (Fudan University, China), Xiao Tan (Baidu Inc.), Li Zhang (Fudan University, China), Jianfeng Feng (Fudan University, China), Xiangyang Xue (Fudan University, China), and Errui Ding (Baidu Inc.)

Pyramid R-CNN: Towards Better Performance and Adaptability for 3D Object Detection 2703
Jiageng Mao (The Chinese University of Hong Kong), Minzhe Niu (Huawei Noah's Ark Lab), Haoyue Bai (The Hong Kong University of Science and Technology), Xiaodan Liang (Sun Yat-sen University), Hang Xu (Huawei Noah's Ark Lab), and Chunjing Xu (Huawei Noah's Ark Lab)

Learning Multi-Scene Absolute Pose Regression With Transformers 2713
Yoli Shavit (Bar Ilan University), Ron Ferens (Bar Ilan University), and Yosi Keller (Bar Ilan University)

Improving 3D Object Detection With Channel-Wise Transformer 2723
Hualian Sheng (College of Information Science and Electronic Engineering, Zhejiang University; DAMO Academy, Alibaba Group), Sijia Cai (DAMO Academy, Alibaba Group), Yuan Liu (Damo Academy, Alibaba Group), Bing Deng (Damo Academy, Alibaba Group), Jianqiang Huang (Damo Academy, Alibaba Group), Xian-Sheng Hua (Damo Academy, Alibaba Group), and Min-Jian Zhao (Zhejiang University)

HPNet: Deep Primitive Segmentation Using Hybrid Representations 2733
Siming Yan (The University of Texas at Austin), Zhenpei Yang (The University of Texas at Austin), Chongyang Ma (Kuaishou Technology), Haibin Huang (Kuaishou Technology), Etienne Vouga (The University of Texas at Austin), and Qixing Huang (The University of Texas at Austin)

GraphFPN: Graph Feature Pyramid Network for Object Detection 2743
Gangming Zhao (The University of Hong Kong), Weifeng Ge (Fudan University), and Yizhou Yu (The University of Hong Kong)

SGPA: Structure-Guided Prior Adaptation for Category-Level 6D Object Pose Estimation 2753
Kai Chen (The Chinese University of Hong Kong) and Qi Dou (The Chinese University of Hong Kong)

Instance Segmentation in 3D Scenes Using Semantic Superpoint Tree Networks 2763
Zhihao Liang (South China University of Technology), Zhihao Li (Noah's Ark Lab, Huawei Technologies Co., Ltd.), Songcen Xu (Noah's Ark Lab, Huawei Technologies Co., Ltd.), Mingkui Tan (South China University of Technology), and Kui Jia (South China University of Technology)

PR-GCN: A Deep Graph Convolutional Network With Point Refinement for 6D Pose Estimation . 2773
Guangyuan Zhou (Beihang University), Huiqun Wang (Beihang University), Jiaxin Chen (Beihang University), and Di Huang (Beihang University, China)

Contrastive Attention Maps for Self-Supervised Co-Localization 2783
Minsong Ki (LG Uplus), Youngjung Uh (Yonsei University), Junsuk Choe (Sogang University), and Hyeran Byun (Yonsei University)

Sparse-Shot Learning With Exclusive Cross-Entropy for Extremely Many Localisations 2793
Andreas Panteli (Netherlands Cancer institute; University of Amsterdam,), Jonas Teuwen (Netherlands Cancer Institute; University of Amsterdam; Radboud University Medical Center), Hugo Horlings (Netherlands Cancer Institute), and Efstratios Gavves (University of Amsterdam, The Netherlands; Ellogon.AI)

Prior to Segment: Foreground Cues for Weakly Annotated Classes in Partially Supervised Instance Segmentation 2804
David Biertimpel (TomTom), Sindi Shkodrani (TomTom), Anil S. Baslamisli (University of Amsterdam), and Nóra Baka (TomTom)

Weakly Supervised 3D Semantic Segmentation Using Cross-Image Consensus and Inter-Voxel Affinity Relations 2814
Xiaoyu Zhu (Carnegie Mellon University), Jeffrey Chen (Carnegie Mellon University), Xiangrui Zeng (Carnegie Mellon University), Junwei Liang (Carnegie Mellon University), Chengqi Li (UC San Diego), Sinuo Liu (Carnegie Mellon University), Sima Behpour (Carnegie Mellon University), and Min Xu (Carnegie Mellon University)

Self-Supervised Image Prior Learning With GMM From a Single Noisy Image 2825
Haosen Liu (Huazhong University of Science and Technology), Xuan Liu (Huazhong University of Science and Technology), Jiangbo Lu (SmartMore Corporation), and Shan Tan (Huazhong University of Science and Technology)

Human Detection and Segmentation via Multi-View Consensus 2835
Isinsu Katircioglu (EPFL), Helge Rhodin (UBC), Jörg Spörri (Sports Medical Research Group, Department of Orthopedics, Balgrist University Hospital, University of Zurich), Mathieu Salzmann (EPFL), and Pascal Fua (EPFL, Switzerland)

PreDet: Large-Scale Weakly Supervised Pre-Training for Detection 2845
Vignesh Ramanathan (Facebook), Rui Wang (Facebook), and Dhruv Mahajan (Facebook)

Boosting Weakly Supervised Object Detection via Learning Bounding Box Adjusters 2856
Bowen Dong (Harbin Institute of Technology), Zitong Huang (Harbin Institute of Technology), Yuelin Guo (Harbin Institute of Technology), Qilong Wang (Tianjin University), Zhenxing Niu (Alibaba Group), and Wangmeng Zuo (Harbin Institute of Technology, China)

TS-CAM: Token Semantic Coupled Attention Map for Weakly Supervised Object Localization 2866
Wei Gao (University of Chinese Academy of Science), Fang Wan (University of Chinese Academy of Sciences), Xingjia Pan (Tencent), Zhiliang Peng (University of Chinese Academy of Sciences), Qi Tian (Huawei Cloud & AI), Zhenjun Han (University of Chinese Academy of Sciences), Bolei Zhou (CUHK), and Qixiang Ye (University of Chinese Academy of Sciences, China)

Semi-Supervised Active Learning for Semi-Supervised Models: Exploit Adversarial Examples With Graph-Based Virtual Labels 2876
Jiannan Guo (Zhejiang University), Haochen Shi (Université de Montréal), Yangyang Kang (Alibaba Group), Kun Kuang (Zhejiang University), Siliang Tang (Zhejiang University), Zhuoren Jiang (Zhejiang University), Changlong Sun (Alibaba Group), Fei Wu (Zhejiang University, China), and Yueting Zhuang (Zhejiang University)

An End-to-End Transformer Model for 3D Object Detection 2886
Ishan Misra (Facebook AI Research), Rohit Girdhar (Facebook AI Research), and Armand Joulin (Facebook AI Research)

RangeDet: In Defense of Range View for LiDAR-Based 3D Object Detection 2898
Lue Fan (Institute of Automation, Chinese Academy of Sciences), Xuan Xiong (Tusimple), Feng Wang (TuSimple), Naiyan Wang (TuSimple), and ZhaoXiang Zhang (Chinese Academy of Sciences, China)

3DVG-Transformer: Relation Modeling for Visual Grounding on Point Clouds 2908
Lichen Zhao (Beihang University), Daigang Cai (Beihang University), Lu Sheng (Beihang University), and Dong Xu (University of Sydney)

Gated3D: Monocular 3D Object Detection From Temporal Illumination Cues 2918
Frank Julca-Aguilar (Algolux), Jason Taylor (Algolux), Mario Bijelic (Mercedes-Benz AG), Fahim Mannan (Algolux), Ethan Tseng (Princeton University), and Felix Heide (Princeton / Algolux)

Group-Free 3D Object Detection via Transformers 2929
Ze Liu (USTC), Zheng Zhang (MSRA, Huazhong University of Science and Technolog), Yue Cao (Microsoft Research), Han Hu (Microsoft Research Asia), and Xin Tong (Microsoft)

Body-Face Joint Detection via Embedding and Head Hook 2939
Junfeng Wan (Aibee), Jiangfan Deng (Aibee), Xiaosong Qiu (Aibee), and Feng Zhou (Aibee)

ELSD: Efficient Line Segment Detector and Descriptor 2949
Haotian Zhang (Megvii), Yicheng Luo (Megvii), Fangbo Qin (Institute of Automation, Chinese Academy of Sciences), Yijia He (Megvii Technology Limited.), and Xiao Liu (Megvii Inc)

WB-DETR: Transformer-Based Detector Without Backbone 2959
Fanfan Liu (University of Chinese Academy of Sciences), Haoran Wei (University of Chinese Academy of Sciences), Wenzhe Zhao (Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China), Guozhen Li (Dalian University of Technology), Jingquan Peng (University of Chinese Academy of Sciences), and Zihao Li (University of Chinese Academy of Sciences)

Dynamic DETR: End-to-End Object Detection With Dynamic Attention 2968
Xiyang Dai (Microsoft), Yinpeng Chen (Microsoft), Jianwei Yang (Microsoft Research), Pengchuan Zhang (Microsoft Research AI), Lu Yuan (Microsoft), and Lei Zhang (International Digital Economy Academy, (IDEA))

Multi-Scale Vision Longformer: A New Vision Transformer for High-Resolution Image Encoding 2978
Pengchuan Zhang (Microsoft Research AI), Xiyang Dai (Microsoft), Jianwei Yang (Microsoft Research), Bin Xiao (Microsoft), Lu Yuan (Microsoft), Lei Zhang (International Digital Economy Academy, (IDEA)), and Jianfeng Gao (Microsoft Research)

Rank & Sort Loss for Object Detection and Instance Segmentation 2989
Kemal Oksuz (Middle East Technical University), Baris Can Cam (Middle East Technical University), Emre Akbas (Middle East Technical University), and Sinan Kalkan (Middle East Technical University)

Switchable K-Class Hyperplanes for Noise-Robust Representation Learning 2999
Boxiao Liu (Institute of Computing Technology, Chinese Academy of Sciences), Guanglu Song (Sensetime), Manyuan Zhang (Sensetime), Haihang You (ICT, Chinese Academy of Sciences), and Yu Liu (The Chinese University of Hong Kong)

DecentLaM: Decentralized Momentum SGD for Large-Batch Deep Training 3009
Kun Yuan (Alibaba Group, (US)), Yiming Chen (Alibaba Group), Xinmeng Huang (University of Pennsylvania), Yingya Zhang (Alibaba Group), Pan Pan (Alibaba Group), Yinghui Xu (Alibaba DAMO Academy), and Wotao Yin (Alibaba US, DAMO Academy)

Large-Scale Robust Deep AUC Maximization: A New Surrogate Loss and Empirical Studies on Medical Image Classification 3020
Zhuoning Yuan (University of Iowa), Yan Yan (Washington State University), Milan Sonka (U Iowa, USA), and Tianbao Yang (University of Iowa)

Robust Small-Scale Pedestrian Detection With Cued Recall via Memory Learning 3030
Jung Uk Kim (KAIST), Sungjune Park (KAIST), and Yong Man Ro (KAIST)

End-to-End Semi-Supervised Object Detection With Soft Teacher 3040
Mengde Xu (Huazhong University of Science and Technology), Zheng Zhang (MSRA, Huazhong University of Science and Technolog), Han Hu (Microsoft Research Asia), Jianfeng Wang (Microsoft), Lijuan Wang (Microsoft), Fangyun Wei (Microsoft Research Asia), Xiang Bai (Huazhong University of Science and Technology), and Zicheng Liu (Microsoft)

CaT: Weakly Supervised Object Detection With Category Transfer 3050
Tianyue Cao (Shanghai Jiao Tong University), Lianyu Du (Shanghai JiaoTong University), Xiaoyun Zhang (Shanghai Jiao Tong University), Siheng Chen (Shanghai Jiao Tong University), Ya Zhang (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University), and Yan-Feng Wang (Cooperative medianet innovation center of Shanghai Jiao Tong University)

ADNet: Leveraging Error-Bias Towards Normal Direction in Face Alignment 3060
Yangyu Huang (Microsoft Research Asia), Hao Yang (Microsoft Research Asia), Chong Li (Microsoft Research Asia), Jongyoo Kim (Microsoft Research Asia), and Fangyun Wei (Microsoft Research Asia)

Causal Attention for Unbiased Visual Recognition 3071
Tan Wang (Nanyang Technological University), Chang Zhou (Alibaba), Qianru Sun (Singapore Management University), and Hanwang Zhang (Nanyang Technological University)

MLVSNet: Multi-Level Voting Siamese Network for 3D Visual Tracking 3081
Zhoutao Wang (Nanjing University of Aeronautics and Astronautics), Qian Xie (Nanjing University of Aeronautics and Astronautics), Yu-Kun Lai (Cardiff University), Jing Wu (Cardiff), Kun Long (Nanjing University of Aeronautics and Astronautics), and Jun Wang (Nanjing University of Aeronautics and Astronautics)

Geometry Uncertainty Projection Network for Monocular 3D Object Detection 3091
Yan Lu (University of Sydney), Xinzhu Ma (The University of Sydney), Lei Yang (Sensetime Group Limited), Tianzhu Zhang (University of Science and Technology of China), Yating Liu (USTC), Qi Chu (University of Science and Technology of China), Junjie Yan (Sensetime Group Limited), and Wanli Ouyang (The University of Sydney)

Multi-Instance Pose Networks: Rethinking Top-Down Pose Estimation 3102
Rawal Khirodkar (Carnegie Mellon University), Visesh Chari (Amazon), Amit Agrawal (Amazon), and Ambrish Tyagi (Amazon)

OMNet: Learning Overlapping Mask for Partial-to-Partial Point Cloud Registration 3112
Hao Xu (UESTC), Shuaicheng Liu (UESTC; Megvii), Guangfu Wang (Megvii), Guanghui Liu (UESTC), and Bing Zeng (University of Electronic Science and Technology of China)

Is Pseudo-Lidar Needed for Monocular 3D Object Detection? 3122
Dennis Park (Toyota Research Institute), Rares Ambrus (Toyota Research Institute), Vitor Guizilini (Toyota Research Institute), Jie Li (Toyota Research Institute), and Adrien Gaidon (Toyota Research Institute)

LIGA-Stereo: Learning LiDAR Geometry Aware Representations for Stereo-Based 3D Detector ... 3133
Xiaoyang Guo (The Chinese University of Hong Kong), Shaoshuai Shi (The Chinese University of Hong Kong), Xiaogang Wang (Chinese University of Hong Kong, Hong Kong), and Hongsheng Li (Chinese University of Hong Kong)

Voxel Transformer for 3D Object Detection 3144
Jiageng Mao (The Chinese University of Hong Kong), Yujing Xue (National University of Singapore), Minzhe Niu (Huawei Noah's Ark Lab), Haoyue Bai (The Hong Kong University of Science and Technology), Jiashi Feng (NUS), Xiaodan Liang (Sun Yat-sen University), Hang Xu (Huawei Noah's Ark Lab), and Chunjing Xu (Huawei Noah's Ark Lab)

Detecting Invisible People 3154
Tarasha Khurana (Carnegie Mellon University), Achal Dave (Carnegie Mellon University), and Deva Ramanan (Carnegie Mellon University)

You Don't Only Look Once: Constructing Spatial-Temporal Memory for Integrated 3D Object Detection and Tracking 3165
Jiaming Sun (SenseTime), Yiming Xie (Zhejiang University), Siyu Zhang (SenseTime), Linghao Chen (Zhejiang University), Guofeng Zhang (Zhejiang University), Hujun Bao (Zhejiang University), and Xiaowei Zhou (Zhejiang University)

CrossDet: Crossline Representation for Object Detection 3175
Heqian Qiu (University of Electronic Science and Technology of China), Hongliang Li (University of Electronic Science and Technology of China), Qingbo Wu (University of Electronic Science and Technology of China), Jianhua Cui (University of Electronic Science and Technology of China), Zichen Song (University of Electronic Science and Technology of China), Lanxiao Wang (University of Electronic Science and Technology of China), and Minjian Zhang (University of Electronic Science and Technology of China)

Towards a Universal Model for Cross-Dataset Crowd Counting 3185
Zhiheng Ma (Xi'an Jiaotong University), Xiaopeng Hong (Xi'an Jiaotong University), Xing Wei (Xi'an Jiaotong University), Yunfeng Qiu (Xi'an Jiaotong University), and Yihong Gong (Xi'an Jiaotong University)

Exploiting Sample Correlation for Crowd Counting With Multi-Expert Network 3195
Xinyan Liu (University of Chinese Academy of Sciences), Guorong Li (University of Chinese Academy of Sciences), Zhenjun Han (University of Chinese Academy of Sciences), Weigang Zhang (Harbin Institute of Technology, Weihai), Yifan Yang (UCAS), Qingming Huang (University of Chinese Academy of Sciences), and Nicu Sebe (University of Trento)

Are We Missing Confidence in Pseudo-LiDAR Methods for Monocular 3D Object Detection? 3205
Andrea Simonelli (University of Trento), Samuel Rota Bulò (Facebook), Lorenzo Porzi (Facebook), Peter Kontschieder (Facebook), and Elisa Ricci (University of Trento)

Uniformity in Heterogeneity: Diving Deep Into Count Interval Partition for Crowd Counting 3214
Changan Wang (Tencent Youtu Lab), Qingyu Song (Zhejiang University), Boshen Zhang (Tencent), Yabiao Wang (Tencent), Ying Tai (Tencent YouTu), Xuyi Hu (University College London), Chengjie Wang (Tencent), Jilin Li (Tencent), Jiayi Ma (Wuhan University), and Yang Wu (Applied Research Center, (ARC)Tencent PCG)

Efficient Large Scale Inlier Voting for Geometric Vision Problems 3223
Dror Aiger (Google), Simon Lynen (Google), Jan Hosang (Google), and Bernhard Zeisl (Google)

Continual Learning for Image-Based Camera Localization 3232
Shuzhe Wang (Aalto University), Zakaria Laskar (Aalto University), Iaroslav Melekhov (Aalto University), Xiaotian Li (Aalto University), and Juho Kannala (Aalto University, Finland)

Query Adaptive Few-Shot Object Detection With Heterogeneous Graph Convolutional Networks 3243
Guangxing Han (Columbia University), Yicheng He (Columbia University), Shiyuan Huang (Columbia University), Jiawei Ma (Columbia University), and Shih-Fu Chang (Columbia University)

Multi-Source Domain Adaptation for Object Detection 3253
Xingxu Yao (Nankai University), Sicheng Zhao (University of California Berkeley), Pengfei Xu (Didi Chuxing), and Jufeng Yang (Nankai University)

RandomRooms: Unsupervised Pre-Training From Synthetic Shapes and Randomized Layouts for 3D Object Detection 3263
Yongming Rao (Tsinghua University), Benlin Liu (UCLA), Yi Wei (Tsinghua University), Jiwen Lu (Tsinghua University), Cho-Jui Hsieh (UCLA), and Jie Zhou (Tsinghua University)

Exploring Geometry-Aware Contrast and Clustering Harmonization for Self-Supervised 3D Object Detection 3273
Hanxue Liang (ETH, Zurich), Chenhan Jiang (Noah's Ark Lab, Huawei Inc.), Dapeng Feng (Sun Yat-sen University), Xin Chen (The University of Hong Kong), Hang Xu (Huawei Noah's Ark Lab), Xiaodan Liang (Sun Yat-sen University), Wei Zhang (Noah's Ark Lab, Huawei Technologies), Zhenguo Li (Huawei Noah's Ark Lab), and Luc Van Gool (ETH Zurich)

RePOSE: Fast 6D Object Pose Refinement via Deep Texture Rendering 3283
Shun Iwase (Carnegie Mellon University), Xingyu Liu (Carnegie Mellon University), Rawal Khirodkar (Carnegie Mellon University), Rio Yokota (Tokyo Institute of Technology, AIST- Tokyo Tech Real World Big-Data Computation Open Innovation Laboratory, (RWBC- OIL), National Institute of Advanced Industrial Science and Technology, (AIST)), and Kris M. Kitani (Carnegie Mellon University)

PICCOLO: Point Cloud-Centric Omnidirectional Localization 3293
Junho Kim (Seoul National University), Changwoon Choi (Seoul National University), Hojun Jang (Seoul National University), and Young Min Kim (Seoul National University)

GarmentNets: Category-Level Pose Estimation for Garments via Canonical Space Shape Completion 3304
Cheng Chi (Columbia University) and Shuran Song (Columbia University)

Personalized and Invertible Face De-Identification by Disentangled Identity Information Manipulation 3314
Jingyi Cao (Shanghai Jiao Tong University), Bo Liu (University of Technology Sydney, (UTS)), Yunqian Wen (Shanghai Jiao Tong University), Rong Xie (Shanghai Jiao Tong University), and Li Song (Shanghai Jiao Tong University)

Long-Term Temporally Consistent Unpaired Video Translation From Simulated Surgical 3D Data..... 3323
Dominik Rivoir (National Center for Tumor Diseases, (NCT) Dresden), Micha Pfeiffer (National Center for Tumor Diseases, (NCT), Dresden), Reuben Docea (National Center for Tumor Diseases, (NCT) Dresden), Fiona Kolbinger (University Hospital Carl Gustav Carus, TU Dresden), Carina Riediger (University Hospital of Dresden), Jürgen Weitz (University Hospital of Dresden), and Stefanie Speidel (National Center for Tumor Diseases, (NCT) Dresden)

Multi-Scale Matching Networks for Semantic Correspondence 3334
Dongyang Zhao (Fudan University), Ziyang Song (Fudan University), Zhenghao Ji (Fudan University), Gangming Zhao (The University of Hong Kong), Weifeng Ge (Fudan University), and Yizhou Yu (The University of Hong Kong)

Rethinking Counting and Localization in Crowds: A Purely Point-Based Framework 3345
Qingyu Song (Zhejiang University), Changan Wang (Tencent Youtu Lab), Zhengkai Jiang (Tencent Youtu Lab), Yabiao Wang (Tencent), Ying Tai (Tencent YouTu), Chengjie Wang (Tencent), Jilin Li (Tencent), Feiyue Huang (Tencent), and Yang Wu (Applied Research Center, (ARC)Tencent PCG)

Learning To Better Segment Objects From Unseen Classes With Unlabeled Videos 3355
Yuming Du (École des ponts ParisTech), Yang Xiao (École des ponts ParisTech), and Vincent Lepetit (Ecole des Ponts ParisTech)

Foreground Activation Maps for Weakly Supervised Object Localization 3365
Meng Meng (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China), Qi Tian (Huawei Cloud & AI), Yongdong Zhang (University of Science and Technology of China), and Feng Wu (University of Science and Technology of China)

ICON: Learning Regular Maps Through Inverse Consistency 3376
Hastings Greer (UNC Chapel Hill), Roland Kwitt (UNC), François-Xavier Vialard (University of Salzburg), and Marc Niethammer (University Paris-Est)

DiscoBox: Weakly Supervised Instance Segmentation and Semantic Correspondence From Box Supervision 3386
Shiyi Lan (University of Maryland), Zhiding Yu (NVIDIA), Christopher Choy (Nvidia), Subhashree Radhakrishnan (NVIDIA), Guilin Liu (NVIDIA), Yuke Zhu (Nvidia), Larry S. Davis (University of Maryland), and Anima Anandkumar (Caltech)

Exploring Classification Equilibrium in Long-Tailed Object Detection 3397
Chengjian Feng (Malong LLC), Yujie Zhong (University of Oxford), and Weilin Huang (Malong LLC)

Normalization Matters in Weakly Supervised Object Localization 3407
Jeesoo Kim (Seoul National University), Junsuk Choe (Sogang University), Sangdoo Yun (NAVER AI LAB), and Nojun Kwak (Seoul National University)

Training Multi-Object Detector by Estimating Bounding Box Distribution for Input Image 3417
Jaeyoung Yoo (Seoul National University), Hojun Lee (Seoul National University), Inseop Chung (Seoul National University), Geonseok Seo (Seoul National University), and Nojun Kwak (Seoul National University)

Semi-Supervised Active Learning With Temporal Output Discrepancy 3427
Siyu Huang (Nanyang Technological University), Tianyang Wang (Austin Peay State University), Haoyi Xiong (Baidu Research), Jun Huan (Styling AI), and Dejing Dou (Baidu)

FASA: Feature Augmentation and Sampling Adaptation for Long-Tailed Instance Segmentation . 3437
Yuhang Zang (Nanyang Technological University), Chen Huang (Apple), and Chen Change Loy (Nanyang Technological University)

Learning Hierarchical Graph Neural Networks for Image Clustering 3447
Yifan Xing (AMAZON CORPORATE LLC), Tong He (Amazon), Tianjun Xiao (Amazon), Yongxin Wang (Amazon), Yuanjun Xiong (Amazon), Wei Xia (Amazon), David Wipf (Amazon), Zheng Zhang (New York University Shanghai), and Stefano Soatto (AWS Amazon ML)

Big Self-Supervised Models Advance Medical Image Classification 3458
Shekoofeh Azizi (Google), Basil Mustafa (Google), Fiona Ryan (Georgia Institute of Technology), Zachary Beaver (Google), Jan Freyberg (Google Health), Jonathan Deaton (Google), Aaron Loh (Google), Alan Karthikesalingam (Google Health), Simon Kornblith (Google Brain), Ting Chen (Google), Vivek Natarajan (Google), and Mohammad Norouzi (Google Research, Brain Team)

Collaborative and Adversarial Learning of Focused and Dispersive Representations for Semi-Supervised Polyp Segmentation 3469
Huisi Wu (Shenzhen University), Guilian Chen (Shenzhen University), Zhenkun Wen (Shenzhen University), and Jing Qin (The Hong Kong Polytechnic University)

Preservational Learning Improves Self-Supervised Medical Image Models by Reconstructing Diverse Contexts 3479
Hong-Yu Zhou (The University of Hong Kong), Chixiang Lu (Huazhong University of Science and Technology), Sibei Yang (ShanghaiTech University), Xiaoguang Han (Shenzhen Research Institute of Big Data, the Chinese University of Hong Kong, (Shenzhen)), and Yizhou Yu (The University of Hong Kong)

TOOD: Task-Aligned One-Stage Object Detection 3490
Chengjian Feng (Malong LLC), Yujie Zhong (University of Oxford), Yu Gao (Malong LLC), Matthew R. Scott (Malong Technologies), and Weilin Huang (Malong LLC)

Oriented R-CNN for Object Detection 3500
Xingxing Xie (Northwestern Polytechnical University), Gong Cheng (Northwestern Polytechnical University), Jiabao Wang (Northwestern Polytechnical University), Xiwen Yao (Northwestern Polytechnical University), and Junwei Han (NWPU, China)

Towards Rotation Invariance in Object Detection 3510
Agastya Kalra (University of Waterloo), Guy Stoppi (Akasha Imaging), Bradley Brown (Akasha Imaging), Rishav Agarwal (University of Waterloo), and Achuta Kadambi (UCLA)

FMODetect: Robust Detection of Fast Moving Objects 3521
Denys Rozumnyi (ETH Zurich / CTU in Prague), Jiří Matas (CMP CTU FEE), Filip Šroubek (Institute of Information Theory and Automation, Czech Academy of Sciences), Marc Pollefeys (ETH Zurich / Microsoft), and Martin R. Oswald (ETH Zurich)

Visual Relationship Detection Using Part-and-Sum Transformers With Composite Queries 3530
Qi Dong (Amazon), Zhuowen Tu (UC San Diego), Haofu Liao (Amazon), Yuting Zhang (Amazon), Vijay Mahadevan (Amazon Inc.), and Stefano Soatto (AWS Amazon ML)

DualPoseNet: Category-Level 6D Object Pose and Size Estimation Using Dual Pose Network With Refined Learning of Pose Consistency 3540
Jiehong Lin (South China University of Technology), Zewei Wei (South China University of Technology), Zhihao Li (Noah's Ark Lab, Huawei Technologies Co., Ltd.), Songcen Xu (Noah's Ark Lab, Huawei Technologies Co., Ltd.), Kui Jia (South China University of Technology), and Yuanqing Li (South China University of Technology)

SimROD: A Simple Adaptation Method for Robust Object Detection 3550
Rindra Ramamonjison (Huawei Technologies Canada), Amin Banitalebi-Dehkordi (Huawei Technologies Canada Co., Ltd.), Xinyu Kang (University of British Columbia), Xiaolong Bai (Huawei), and Yong Zhang (Huawei Technologies Canada Co., Ltd.)

Disentangled High Quality Salient Object Detection 3560
Lv Tang (Nanjing University), Bo Li (Nanjing University), Yijie Zhong (Nanjing University), Shouhong Ding (Tencent), and Mofei Song (Southeast University)

G-DetKD: Towards General Distillation Framework for Object Detectors via Contrastive and Semantic-Guided Feature Imitation 3571
Lewei Yao (Huawei Noah's Ark Lab), Renjie Pi (Huawei Noah's Ark Lab), Hang Xu (Huawei Noah's Ark Lab), Wei Zhang (Noah's Ark Lab, Huawei Technologies), Zhenguo Li (Huawei Noah's Ark Lab), and Tong Zhang (Hong Kong University of Science and Technology)

TransFER: Learning Relation-Aware Facial Expression Representations With Transformers 3581
Fanglei Xue (University of Chinese Academy of Sciences), Qiangchang Wang (West Virginia University), and Guodong Guo (Baidu)

Rethinking Transformer-Based Set Prediction for Object Detection 3591
Zhiqing Sun (Carnegie Mellon University), Shengcao Cao (Carnegie Mellon University), Yiming Yang (Carnegie Mellon University), and Kris M. Kitani (Carnegie Mellon University)

Fast Convergence of DETR With Spatially Modulated Co-Attention 3601
Peng Gao (Chinese university of hong kong), Minghang Zheng (Peking University), Xiaogang Wang (Chinese University of Hong Kong, Hong Kong), Jifeng Dai (SenseTime), and Hongsheng Li (Chinese University of Hong Kong)

Reconcile Prediction Consistency for Balanced Object Detection 3611
Keyang Wang (Chongqing University, China) and Lei Zhang (Chongqing University, China)

Mutual Supervision for Dense Object Detection 3621
Ziteng Gao (Nanjing University), Limin Wang (Nanjing University), and Gangshan Wu (Nanjing University)

Conditional DETR for Fast Training Convergence 3631
Depu Meng (University of Science and Technology of China), Xiaokang Chen (Peking University), Zejia Fan (Peking University), Gang Zeng (Peking University), Houqiang Li (University of Science and Technology of China), Yuhui Yuan (Microsoft Research), Lei Sun (Microsoft Research Asia), and Jingdong Wang (Microsoft)

Meta Pairwise Relationship Distillation for Unsupervised Person Re-Identification 3641
Haoxuanye Ji (Xi'An JiaoTong University), Le Wang (Xi'an Jiaotong University), Sanping Zhou (Xi'an Jiaotong University), Wei Tang (University of Illinois at Chicago), Nanning Zheng (Xi'an Jiaotong University), and Gang Hua (Wormpex AI Research)

Teacher-Student Adversarial Depth Hallucination To Improve Face Recognition 3651
Hardik Uppal (Queen's University), Alireza Sepas-Moghaddam (Queen's University), Michael Greenspan (Queen's University), and Ali Etemad (Queen's University)

Fake It Till You Make It: Face Analysis in the Wild Using Synthetic Data Alone 3661
Erroll Wood (Microsoft), Tadas Baltrušaitis, Charlie Hewitt (Microsoft), Sebastian Dziadzio (Microsoft), Thomas J. Cashman (Microsoft), and Jamie Shotton (Microsoft)

Disentangled Representation for Age-Invariant Face Recognition: A Mutual Information Minimization Perspective 3672
Xuege Hou (Tsinghua University), Yali Li (Tsinghua University), and Shengjin Wang (Tsinghua University)

Cross-Encoder for Unsupervised Gaze Representation Learning 3682
Yunjia Sun (Institute of Computing Technology, Chinese Academy of Sciences), Jiabei Zeng (Institute of Computing Technology, Chinese Academy of Sciences), Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences)

VENet: Voting Enhancement Network for 3D Object Detection 3692
Qian Xie (Nanjing University of Aeronautics and Astronautics, China), Yu-Kun Lai (Cardiff University, UK), Jing Wu (Cardiff University, UK), Zhoutao Wang (Nanjing University of Aeronautics and Astronautics, China), Dening Lu (Nanjing University of Aeronautics and Astronautics, China), Mingqiang Wei (Nanjing University of Aeronautics and Astronautics, China), and Jun Wang (Nanjing University of Aeronautics and Astronautics, China)

Free-Form Description Guided 3D Visual Graph Network for Object Grounding in Point Cloud .. 3702
Mingtao Feng (Xidian Unversity), Zhen Li (Xidian University), Qi Li (Xidian University), Liang Zhang (School of Computer Science and Technology, Xidian University, China), XiangDong Zhang (School of Telecommunications Engineering, Xidian University), Guangming Zhu (School of Computer Science and Technology, Xidian University, China), Hui Zhang (Hunan University), Yaonan Wang (Hunan University), and Ajmal Mian (University of Western Australia)

Real-Time Vanishing Point Detector Integrating Under-Parameterized RANSAC and Hough Transform 3712
Jianping Wu (Suzhou Vocational University), Liang Zhang (Suzhou Vocational University), Ye Liu (Suzhou Vocational University), and Ke Chen (University of Kentucky)

Looking Here or There? Gaze Following in 360-Degree Images 3722
Yunhao Li (Shanghai Jiaotong University), Wei Shen (SJTU), Zhongpai Gao (Shanghai Jiao Tong University), Yucheng Zhu (Shanghai Jiao Tong University), Guangtao Zhai (Shanghai Jiao Tong University), and Guodong Guo (Baidu)

Towards Efficient Graph Convolutional Networks for Point Cloud Handling 3732
Yawei Li (ETH Zurich), He Chen (Johns Hopkins University), Zhaopeng Cui (Zhejiang University), Radu Timofte (ETH Zurich), Marc Pollefeys (ETH Zurich / Microsoft), Gregory S. Chirikjian (Johns Hopkins University), and Luc Van Gool (ETH Zurich)

Multi-Echo LiDAR for 3D Object Detection 3743
Yunze Man (Carnegie Mellon University), Xinshuo Weng (Carnegie Mellon University), Prasanna Kumar Sivakumar (DENSO International America, Inc.), Matthew O'Toole (Carnegie Mellon University), and Kris M. Kitani (Carnegie Mellon University)

CondLaneNet: A Top-To-Down Lane Detection Framework Based on Conditional Convolution .. 3753
Lizhe Liu (Alibaba Group), Xiaohao Chen (Alibaba Group), Siyu Zhu (Alibaba A.I. Labs), and Ping Tan (Simon Fraser University)

CrackFormer: Transformer Network for Fine-Grained Crack Detection 3763
Huajun Liu (Nanjing University of Science and Technology), Xiangyu Miao (Nanjing university of science and), Christoph Mertz (CMU), Chengzhong Xu (University of Macau), and Hui Kong (Nanjing University of Science and Technology)

DWKS: A Local Descriptor of Deformations Between Meshes and Point Clouds 3773
Robin Magnet (Ecole Polytechnique) and Maks Ovsjanikov (Ecole polytechnique)

Physics-Enhanced Machine Learning for Virtual Fluorescence Microscopy 3783
Colin L. Cooke (Duke University), Fanjie Kong (Duke University), Amey Chaware (Duke University), Kevin C. Zhou (Duke University), Kanghyun Kim (Duke University), Rong Xu (Duke University), D. Michael Ando (Google Research), Samuel J. Yang (Google), Pavan Chandra Konda (Duke University), and Roarke Horstmeyer (Duke University)

DAM: Discrepancy Alignment Metric for Face Recognition 3794
Jiaheng Liu (Beihang University), Yudong Wu (Sensetime Group Limited), Yichao Wu (Sensetime Group Limited), Chuming Li (sensetime), Xiaolin Hu (Tsinghua University), Ding Liang (Sensetime Group Limited), and Mengyu Wang (Peking University)

Topologically Consistent Multi-View Face Inference Using Volumetric Sampling 3804
Tianye Li (University of Southern California), Shichen Liu (University of Southern California), Timo Bolkart (Max Planck Institute for Intelligent Systems), Jiayi Liu (University of Southern California), Hao Li (Pinscreen / UC Berkeley), and Yajie Zhao (USC ICT)

Generalizing Gaze Estimation With Outlier-Guided Collaborative Adaptation 3815
Yunfei Liu (Beihang University), Ruicong Liu (Beihang University), Haofei Wang (Peng Cheng Laboratory), and Feng Lu (Beihang University)

Learn To Cluster Faces via Pairwise Classification 3825
Junfu Liu (Meituan), Di Qiu (Meituan), Pengfei Yan (Meituan), and Xiaolin Wei (Meituan)

End-to-End Robust Joint Unsupervised Image Alignment and Clustering 3834
Xiangrui Zeng (Carnegie Mellon University), Gregory Howe (Carnegie Mellon University), and Min Xu (Carnegie Mellon University)

FACIAL: Synthesizing Dynamic Talking Face With Implicit Attribute Learning 3847
Chenxu Zhang (The University of Texas at Dallas), Yifan Zhao (Beihang University), Yifei Huang (East China Normal University), Ming Zeng (School of Informatics, Xiamen University), Saifeng Ni (Samsung Research America), Madhukar Budagavi (Samsung Research America), and Xiaohu Guo (The University of Texas at Dallas)

Disentangled Lifespan Face Synthesis 3857
Sen He (University of Surrey), Wentong Liao (Leibniz University Hannover), Michael Ying Yang (University of Twente), Yi-Zhe Song (University of Surrey), Bodo Rosenhahn (Leibniz University Hannover), and Tao Xiang (University of Surrey)

Retrieve in Style: Unsupervised Facial Feature Transfer and Retrieval 3867
Min Jin Chong (Univeristy of Illinois at Urbana-Champaign), Wen-Sheng Chu (Google), Abhishek Kumar (Google Brain), and David Forsyth (Univeristy of Illinois at Urbana-Champaign)

Towards Face Encryption by Generating Adversarial Identity Masks 3877
Xiao Yang (Tsinghua University), Yinpeng Dong (Tsinghua University), Tianyu Pang (Tsinghua University), Hang Su (Tsinghua Univiersity), Jun Zhu (Tsinghua University), Yuefeng Chen (Alibaba Group), and Hui Xue (Alibaba)

Re-Aging GAN: Toward Personalized Face Age Transformation 3888
Farkhod Makhmudkhujaev (Inha University), Sungeun Hong (Inha University), and In Kyu Park (Inha University)

Recurrent Mask Refinement for Few-Shot Medical Image Segmentation 3898
Hao Tang (University of California Irvine), Xingwei Liu (University of California Irvine), Shanlin Sun (University of California, Irvine), Xiangyi Yan (University of California, Irvine), and Xiaohui Xie (University of California, Irvine)

Generative Adversarial Registration for Improved Conditional Deformable Templates 3909
Neel Dey (New York University), Mengwei Ren (New York University), Adrian V. Dalca (MIT), and Guido Gerig (NYU)

GLoRIA: A Multimodal Global-Local Representation Learning Framework for Label-Efficient Medical Image Recognition 3922
Shih-Cheng Huang (Stanford University, USA), Liyue Shen (Stanford University, USA), Matthew P. Lungren (Stanford University, USA), and Serena Yeung (Stanford University, USA)

Semantic Aware Data Augmentation for Cell Nuclei Microscopical Images With Artificial Neural Networks 3932
Alireza Naghizadeh (Rutgers University), Hongye Xu (Rutgers University), Mohab Mohamed (Rutgers University), Dimitris N. Metaxas (Rutgers University), and Dongfang Liu (Rutgers University)

T-AutoML: Automated Machine Learning for Lesion Segmentation Using Transformers in 3D Medical Imaging 3942
Dong Yang (NVIDIA Corporation), Andriy Myronenko (NVIDIA), Xiaosong Wang (NVIDIA Corporation), Ziyue Xu (NVIDIA), Holger R. Roth (NVIDIA), and Daguang Xu (NVIDIA Corporation)

RFNet: Region-Aware Fusion Network for Incomplete Multi-Modal Brain Tumor Segmentation 3955
Yuhang Ding (Baidu, China), Xin Yu (University of Technology Sydney), and Yi Yang (UTS)

Visual-Textual Attentive Semantic Consistency for Medical Report Generation 3965
Yi Zhou (Southeast University), Lei Huang (Beihang University), Tao Zhou (Nanjing University of Science and Technology), Huazhu Fu (IHPC, ASTAR), and Ling Shao (Inception Institute of Artificial Intelligence)

The Way to My Heart Is Through Contrastive Learning: Remote Photoplethysmography From Unlabelled Video 3975
John Gideon (Toyota Research Institute) and Simon Stent (Toyota Research Institute)

Multi-Class Cell Detection Using Spatial Context Representation 3985
Shahira Abousamra (Stony Brook University, USA), David Belinsky (Stony Brook University, USA), John Van Arnam (Stony Brook University, USA), Felicia Allard (Stony Brook University, USA), Eric Yee (Stony Brook University, USA), Rajarsi Gupta (Stony Brook University, USA), Tahsin Kurc (Stony Brook University, USA), Dimitris Samaras (Stony Brook University, USA), Joel Saltz (Stony Brook University, USA), and Chao Chen (Stony Brook University, USA)

Multimodal Co-Attention Transformer for Survival Prediction in Gigapixel Whole Slide Images 3995
Richard J. Chen (Harvard Medical School), Ming Y. Lu (Pathology, Brigham and Women's Hospital, Harvard Medical School), Wei-Hung Weng (MIT), Tiffany Y. Chen (Pathology, Brigham and Women's Hospital, Harvard Medical School), Drew F.K. Williamson (Pathology, Brigham and Women's Hospital, Harvard Medical School), Trevor Manz (Harvard Medical School), Maha Shady (Harvard Medical School), and Faisal Mahmood (Pathology, Brigham and Women's Hospital, Harvard Medical School)

CDNet: Centripetal Direction Network for Nuclear Instance Segmentation 4006
Hongliang He (School of Electronic and Computer Engineering, Peking University), Zhongyi Huang (School of Electronic and Computer Engineering, Peking University), Yao Ding (University of Chinese Academy of Sciences), Guoli Song (Peng Cheng Laboratory), Lin Wang (School of Electronic and Computer Engineering, Peking University), Qian Ren (School of Electronic and Computer Engineering, Peking University), Pengxu Wei (Sun Yat-sen University), Zhiqiang Gao (Peng Cheng Laboratory), and Jie Chen (Peking University)

Mutual-Complementing Framework for Nuclei Detection and Segmentation in Pathology Image 4016
Zunlei Feng (Zhejiang University), Zhonghua Wang (Zhejiang University), Xinchao Wang (National University of Singapore), Yining Mao (Zhejiang University), Thomas Li (The University of Hong Kong), Jie Lei (Zhejiang University of Technology), Yuexuan Wang (Zhejiang University/The University of Hong Kong), and Mingli Song (Zhejiang University)

Deep Survival Analysis With Longitudinal X-Rays for COVID-19 4026
Michelle Shu (Cornell University), Richard Strong Bowen (Cornell Tech), Charles Herrmann (Cornell), Gengmo Qi (Cornell University), Michele Santacatterina (TRIPODS Center for Data Science - Cornell University), and Ramin Zabih (Cornell Tech/Google Research)

Self-Supervised Cryo-Electron Tomography Volumetric Image Restoration From Single Noisy Volume With Sparsity Constraint 4036
Zhidong Yang (Institute of Computing Technology, Chinese Academy of Sciences), Fa Zhang (Institute of Computing Technology, Chinese Academy of Sciences), and Renmin Han (Research Center for Mathematics and Interdisciplinary Sciences, Shandong University)

CryoDRGN2: Ab Initio Neural Reconstruction of 3D Protein Structures From Real Cryo-EM Images 4046
Ellen D. Zhong (MIT), Adam Lerer (Facebook AI Research), Joseph H. Davis (MIT), and Bonnie Berger (MIT)

3D Pose Estimation and Recognition including Point Clouds

Hierarchical Conditional Flow: A Unified Framework for Image Super-Resolution and Image Rescaling 4056
Jingyun Liang (ETH Zurich), Andreas Lugmayr (ETH Zurich), Kai Zhang (ETH, Zurich), Martin Danelljan (ETH Zurich), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)

Learning Dual Priors for JPEG Compression Artifacts Removal 4066
Xueyang Fu (University of Science and Technology of China), Xi Wang (University of Science and Technology of China), Aiping Liu (University of Science and Technology of China), Junwei Han (Northwestern Polytechnical University), and Zheng-Jun Zha (University of Science and Technology of China)

Mutual Affine Network for Spatially Variant Kernel Estimation in Blind Image Super-Resolution 4076
Jingyun Liang (ETH Zurich), Guolei Sun (ETH Zurich), Kai Zhang (ETH, Zurich), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)

STAR: A Structure-Aware Lightweight Transformer for Real-Time Image Enhancement 4086
Zhaoyang Zhang (The Chinese University of Hong Kong), Yitong Jiang (the Chinese University of HK), Jun Jiang (SenseBrain Research, (USA)), Xiaogang Wang (Chinese University of Hong Kong, Hong Kong), Ping Luo (The University of Hong Kong), and Jinwei Gu (SenseBrain)

Perceptual Variousness Motion Deblurring With Light Global Context Refinement 4096
Jichun Li (Fudan University), Weimin Tan (Fudan University), and Bo Yan (Fudan University)

StarEnhancer: Learning Real-Time and Style-Aware Image Enhancement 4106
Yuda Song (Zhejiang University), Hui Qian (Zhejiang University), and Xin Du (Zhejiang University)

MFNet: Multi-Filter Directive Network for Weakly Supervised Salient Object Detection 4116
Yongri Piao (Dalian University of Technology), Jian Wang (Dalian University of Technology), Miao Zhang (Dalian University of Technology), and Huchuan Lu (Dalian University of Technology)

Uncertainty-Guided Transformer Reasoning for Camouflaged Object Detection 4126
Fan Yang (Group 42), Qiang Zhai (UESTC), Xin Li (Group 42), Rui Huang (UESTC), Ao Luo (Megvii), Hong Cheng (UESTC), and Deng-Ping Fan (Inception Institute of Artificial Intelligence)

Scene Context-Aware Salient Object Detection 4136
Avishek Siris (Swansea University), Jianbo Jiao (University of Oxford), Gary K.L. Tam (Swansea University), Xianghua Xie (Swansea University), and Rynson W.H. Lau (City University of Hong Kong)

Summarize and Search: Learning Consensus-Aware Dynamic Convolution for Co-Saliency Detection 4147
Ni Zhang (Northwestern Polytechnical University), Junwei Han (NWPU, China), Nian Liu (Inception Institute of Artificial Intelligence), and Ling Shao (Inception Institute of Artificial Intelligence)

Light Source Guided Single-Image Flare Removal From Unpaired Data 4157
Xiaotian Qiao (City University of Hong Kong), Gerhard P. Hancke (City University of Hong Kong), and Rynson W.H. Lau (City University of Hong Kong)

PlaneTR: Structure-Guided Transformers for 3D Plane Recovery 4166
Bin Tan (Wuhan University), Nan Xue (Wuhan University), Song Bai (University of Oxford), Tianfu Wu (NC State University), and Gui-Song Xia (Wuhan University)

ALL Snow Removed: Single Image Desnowing Algorithm Using Hierarchical Dual-Tree Complex Wavelet Representation and Contradict Channel Loss 4176
Wei-Ting Chen (National Taiwan University), Hao-Yu Fang (National Taiwan University), Cheng-Lin Hsieh (National Taiwan University), Cheng-Che Tsai (National Taiwan University), I-Hsiang Chen (National Taiwan University), Jian-Jiun Ding (National Taiwan University), and Sy-Yen Kuo (National Taiwan University)

Exploring Visual Engagement Signals for Representation Learning 4186
Menglin Jia (Cornell University), Zuxuan Wu (UMD), Austin Reiter (Facebook AI), Claire Cardie (Cornell University), Serge Belongie (University of Copenhagen), and Ser-Nam Lim (Facebook AI)

TransView: Inside, Outside, and Across the Cropping View Boundaries 4198
Zhiyu Pan (Huazhong Univ. of Sci.&Tech.), Zhiguo Cao (Huazhong Univ. of Sci.&Tech.), Kewei Wang (Huazhong Univ. of Sci.&Tech.), Hao Lu (Huazhong University of Science and Technology), and Weicai Zhong (Huawei CBG Consumer Cloud Service Big Data Platform Dept.)

Inverting a Rolling Shutter Camera: Bring Rolling Shutter Images to High Framerate Global Shutter Video 4208
Bin Fan (Northwestern Polytechnical University) and Yuchao Dai (Northwestern Polytechnical University)

Structure-Preserving Deraining With Residue Channel Prior Guidance 4218
Qiaosi Yi (East China Normal University), Juncheng Li (The Chinese University of Hong Kong), Qinyan Dai (East China Normal University), Faming Fang (East China Normal University), Guixu Zhang (East China Normal University), and Tieyong Zeng (The Chinese University of Hong Kong)

ReconfigISP: Reconfigurable Camera Image Processing Pipeline 4228
Ke Yu (The Chinese University of Hong Kong), Zexian Li (Beihang university), Yue Peng (JV), Chen Change Loy (Nanyang Technological University), and Jinwei Gu (SenseBrain)

Event-Intensity Stereo: Estimating Depth by the Best of Both Worlds 4238
Mohammad Mostafavi (GIST), Kuk-Jin Yoon (KAIST), and Jonghyun Choi (GIST)

End-to-End Piece-Wise Unwarping of Document Images 4248
Sagnik Das (Stony Brook University), Kunwar Yashraj Singh (Amazon), Jon Wu (Amazon), Erhan Bas (Amazon), Vijay Mahadevan (Amazon Inc.), Rahul Bhotika (Amazon), and Dimitris Samaras (Stony Brook University)

Context Reasoning Attention Network for Image Super-Resolution 4258
Yulun Zhang (Northeastern University), Donglai Wei (Harvard University), Can Qin (Northeastern University), Huan Wang (Northeastern University), Hanspeter Pfister (Harvard University), and Yun Fu (Northeastern University)

Dynamic High-Pass Filtering and Multi-Spectral Attention for Image Super-Resolution 4268
Salma Abdel Magid (Harvard University), Yulun Zhang (Northeastern University), Donglai Wei (Harvard University), Won-Dong Jang (Harvard University), Zudi Lin (Harvard University), Yun Fu (Northeastern University), and Hanspeter Pfister (Harvard University)

Pyramid Architecture Search for Real-Time Image Deblurring 4278
Xiaobin Hu (Technische Universität München), Wenqi Ren (Institute of Information Engineering, Chinese Academy of Sciences), Kaicheng Yu (EPFL), Kaihao Zhang (Australian National University), Xiaochun Cao (Chinese Academy of Sciences), Wei Liu (Tencent), and Bjoern Menze (TUM)

Learning Frequency-Aware Dynamic Network for Efficient Super-Resolution 4288
Wenbin Xie (Tsinghua University), Dehua Song (Huawei Noah's Ark Lab), Chang Xu (University of Sydney), Chunjing Xu (Huawei Noah's Ark Lab), Hui Zhang (Tsinghua University), and Yunhe Wang (Huawei Technologies)

Unsupervised Real-World Super-Resolution: A Domain Adaptation Perspective 4298
Wei Wang (Bytedance AI Lab), Haochen Zhang (University of California San Diego), Zehuan Yuan (Bytedance.Inc), and Changhu Wang (ByteDance.Inc)

Dynamic Attentive Graph Learning for Image Restoration 4308
Chong Mou (Peking University Shenzhen Graduate School), Jian Zhang (Peking University Shenzhen Graduate School), and Zhuoyuan Wu (PKU)

RGB-D Saliency Detection via Cascaded Mutual Information Minimization 4318
Jing Zhang (Australian National University), Deng-Ping Fan (Inception Institute of Artificial Intelligence), Yuchao Dai (Northwestern Polytechnical University), Xin Yu (University of Technology Sydney), Yiran Zhong (Australian National University), Nick Barnes (ANU), and Ling Shao (Inception Institute of Artificial Intelligence)

Learning RAW-to-sRGB Mappings With Inaccurately Aligned Supervision 4328
Zhilu Zhang (Harbin Institute of Technology), Haolin Wang (Harbin Institute of Technology), Ming Liu (Harbin Institute of Technology, China), Ruohao Wang (Harbin Institute of Technology), Jiawei Zhang (Sensetime Research), and Wangmeng Zuo (Harbin Institute of Technology, China)

Deep Structured Instance Graph for Distilling Object Detectors 4339
Yixin Chen (Chinese University of Hong Kong), Pengguang Chen (Chinese University of Hong Kong), Shu Liu (SmartMore), Liwei Wang (CUHK), and Jiaya Jia (Chinese University of Hong Kong)

Learning Unsupervised Metaformer for Anomaly Detection 4349
Jhih-Ciang Wu (Academia Sinica), Ding-Jie Chen (Academia Sinica), Chiou-Shann Fuh (National Taiwan University), and Tyng-Luh Liu (Academia Sinica)

Equivariant Imaging: Learning Beyond the Range Space 4359
Dongdong Chen (University of Edinburgh), Julián Tachella (University of Edinburgh), and Mike E. Davies (University of Edinburgh)

Multi-Level Curriculum for Training a Distortion-Aware Barrel Distortion Rectification Model 4369
Kang Liao (Beijing Jiaotong University), Chunyu Lin (Beijing Jiaotong University), Lixin Liao (Beijing Jiaotong University), Yao Zhao (Beijing Jiaotong University), and Weiyao Lin (Shanghai Jiao Tong university)

Zero-Shot Day-Night Domain Adaptation With a Physics Prior 4379
Attila Lengyel (Delft University of Technology), Sourav Garg (Queensland University of Technology), Michael Milford (ACRV and QUT, Australia), and Jan C. van Gemert (Delft University of Technology)

MixMix: All You Need for Data-Free Compression Are Feature and Data Mixing 4390
Yuhang Li (Yale University), Feng Zhu (University of Technology Sydney), Ruihao Gong (SenseTime), Mingzhu Shen (Sensetime Research), Xin Dong (H), Fengwei Yu (sensetime), Shaoqing Lu (Sensetime Research), and Shi Gu (UESTC)

Federated Learning for Non-IID Data via Unified Feature Learning and Optimization Objective Alignment 4400
Lin Zhang (Peking University), Yong Luo (Wuhan University), Yan Bai (Peking University), Bo Du (Wuhan University), and Ling-Yu Duan (Peking University)

Omniscient Video Super-Resolution 4409
Peng Yi (National Engineering Research Center for Multimedia Software, Wuhan University, China), Zhongyuan Wang (Wuhan University), Kui Jiang (Wuhan University), Junjun Jiang (Harbin Institute of Technology), Tao Lu (Wuhan Institute of Technology), Xin Tian (Wuhan University), and Jiayi Ma (Wuhan University)

Adaptive Unfolding Total Variation Network for Low-Light Image Enhancement 4419
Chuanjun Zheng (College of Computer Science and Software Engineering, Shenzhen University), Daming Shi (College of Computer Science and Software Engineering, Shenzhen University), and Wentian Shi (Shenzhen University)

Ultra-High-Definition Image HDR Reconstruction via Collaborative Bilateral Learning 4429
Zhuoran Zheng (Nanjing University of Science and Technology), Wenqi Ren (Institute of Information Engineering, Chinese Academy of Sciences), Xiaochun Cao (Chinese Academy of Sciences), Tao Wang (Huawei Noah's Ark Lab), and Xiuyi Jia (Nanjing University of Science and Technology)

Representative Color Transform for Image Enhancement 4439
Hanul Kim (Seoul National University of Science and Technology), Su-Min Choi (Chungnam National University), Chang-Su Kim (Korea university), and Yeong Jun Koh (Chungnam National University)

Super-Resolving Cross-Domain Face Miniatures by Peeking at One-Shot Exemplar 4449
Peike Li (AAII), Xin Yu (University of Technology Sydney), and Yi Yang (UTS)

Event Stream Super-Resolution via Spatiotemporal Constraint Learning 4460
Siqi Li (Tsinghua University), Yutong Feng (Tsinghua University), Yipeng Li (Tsinghua University), Yu Jiang (Jilin University, Changchun, China), Changqing Zou (Huawei Technologies Canada), and Yue Gao (Tsinghua University)

Self-Conditioned Probabilistic Learning of Video Rescaling 4470
Yuan Tian (Shanghai Jiao Tong University), Guo Lu (Beijing Institute of Technology), Xiongkuo Min (Shanghai Jiao Tong University), Zhaohui Che (Shanghai Jiao Tong University), Guangtao Zhai (Shanghai Jiao Tong University), Guodong Guo (Baidu), and Zhiyong Gao (Shanghai Jiao Tong University)

A New Journey From SDRTV to HDRTV 4480
Xiangyu Chen (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Zhengwen Zhang (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Jimmy S. Ren (SenseTime Research; Qing Yuan Research Institute, Shanghai Jiao Tong University), Lynhoo Tian (SenseTime), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), and Chao Dong (Shenzhen Institutes of Advanced Technology)

ResRep: Lossless CNN Pruning via Decoupling Remembering and Forgetting 4490
Xiaohan Ding (Tsinghua University), Tianxiang Hao (Tsinghua University), Jianchao Tan (Kwai Inc.), Ji Liu (Kwai Inc.), Jungong Han (Aberystwyth University), Yuchen Guo (Tsinghua University), and Guiguang Ding (Tsinghua University, China)

Efficient Video Compression via Content-Adaptive Super-Resolution 4501
Mehrdad Khani (MIT), Vibhaalakshmi Sivaraman (MIT CSAIL), and Mohammad Alizadeh (MIT CSAIL)

Bringing Events Into Video Deblurring With Non-Consecutively Blurry Frames 4511
Wei Shang (Harbin Institute of Technology), Dongwei Ren (Harbin Institute of Technology), Dongqing Zou (SenseTime ResearchQing Yuan Research Institute, Shanghai Jiao Tong University), Jimmy S. Ren (SenseTime Research; Qing Yuan Research Institute, Shanghai Jiao Tong University), Ping Luo (The University of Hong Kong), and Wangmeng Zuo (Harbin Institute of Technology, China)

SUNet: Symmetric Undistortion Network for Rolling Shutter Correction 4521
Bin Fan (Northwestern Polytechnical University), Yuchao Dai (Northwestern Polytechnical University), and Mingyi He (Northwestern Polytechnical University)

Robust Automatic Monocular Vehicle Speed Estimation for Traffic Surveillance 4531
Jerome Revaud (Naver Labs Europe) and Martin Humenberger (Naver Labs Europe)

Augmenting Depth Estimation With Geospatial Context 4542
Scott Workman (DZYNE Technologies) and Hunter Blanton (University of Kentucky)

Real-Time Video Inference on Edge Devices via Adaptive Model Streaming 4552
Mehrdad Khani (MIT), Pouya Hamadanian (MIT), Arash Nasr-Esfahany (MIT), and Mohammad Alizadeh (MIT CSAIL)

Score-Based Point Cloud Denoising 4563
Shitong Luo (Peking University) and Wei Hu (Peking University)

Rethinking Noise Synthesis and Modeling in Raw Denoising 4573
Yi Zhang (Chinese University of Hong Kong), Hongwei Qin (Sensetime), Xiaogang Wang (Chinese University of Hong Kong, Hong Kong), and Hongsheng Li (Chinese University of Hong Kong)

Extensions of Karger's Algorithm: Why They Fail in Theory and How They Are Useful in Practice 4582
Erik Jenner (Heidelberg University), Enrique Fita Sanmartín (Heidelberg University), and Fred A. Hamprecht (Heidelberg Collaboratory for Image Processing)

Low-Rank Tensor Completion by Approximating the Tensor Average Rank 4592
Zhanliang Wang (Ocean University of China), Junyu Dong (Ocean University of China), Xinguo Liu (Ocean University of China), and Xueying Zeng (Ocan University of China)

RDI-Net: Relational Dynamic Inference Networks 4601
Huanyu Wang (Zhejiang University), Songyuan Li (Zhejiang University), Shihao Su (Zhejiang University), Zequn Qin (Zhejiang University), and Xi Li (Zhejiang University)

Overfitting the Data: Compact Neural Video Delivery via Content-Aware Feature Modulation 4611
Jiaming Liu (Beijing University of Posts and Telecommunications), Ming Lu (Intel Labs China), Kaixin Chen (Beijing University of Posts and Telecommunications), Xiaoqi Li (Columbia university in the city of New york), Shizun Wang (Beijing University of Posts and Telecommunications), Zhaoqing Wang (University of Sydney), Enhua Wu (SKLCS, Institute of Software, Chinese Academy of Sciences, Beijing, ChinaFaculty of Science and Technology, University of Macau, Macao, China), Yurong Chen (Intel Labs China), Chuang Zhang (Beijing University of Posts and Telecommunications), and Ming Wu (Beijing University of Posts and Telecommunications)

Rethinking Coarse-To-Fine Approach in Single Image Deblurring 4621
Sung-Jin Cho (Korea University), Seo-Won Ji (Korea University), Jun-Pyo Hong (Korea University), Seung-Won Jung (Korea University), and Sung-Jea Ko (Korea University)

Cross-Patch Graph Convolutional Network for Image Denoising 4631
Yao Li (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), and Zheng-Jun Zha (University of Science and Technology of China)

PnP-DETR: Towards Efficient Visual Analysis With Transformers 4641
Tao Wang (National University of Singapore), Li Yuan (National University of Singapore), Yunpeng Chen (YITU Tech), Jiashi Feng (NUS), and Shuicheng Yan (National University of Singapore)

DCT-SNN: Using DCT To Distribute Spatial Information Over Time for Low-Latency Spiking Neural Networks 4651
Isha Garg (Purdue University), Sayeed Shafayet Chowdhury (Purdue University), and Kaushik Roy (Purdue Uniiversity)

Specificity-Preserving RGB-D Saliency Detection 4661
Tao Zhou (Nanjing University of Science and Technology), Huazhu Fu (IHPC, ASTAR), Geng Chen (Northwestern Polytechnical University), Yi Zhou (Southeast University), Deng-Ping Fan (Inception Institute of Artificial Intelligence), and Ling Shao (Inception Institute of Artificial Intelligence)

High-Fidelity Pluralistic Image Completion With Transformers 4672
Ziyu Wan (City University of Hong Kong), Jingbo Zhang (City University of Hong Kong), Dongdong Chen (Microsoft Cloud AI), and Jing Liao (City University of Hong Kong)

Mitigating Intensity Bias in Shadow Detection via Feature Decomposition and Reweighting 4682
Lei Zhu (City University of Hong Kong), Ke Xu (Shanghai Jiao Tong University), Zhanghan Ke (City University of Hong Kong), and Rynson W.H. Lau (City University of Hong Kong)

Light Field Saliency Detection With Dual Local Graph Learning and Reciprocative Guidance 4692
Nian Liu (Inception Institute of Artificial Intelligence), Wangbo Zhao (Northwestern Polytechnical University), Dingwen Zhang (NWPU), Junwei Han (NWPU, China), and Ling Shao (Inception Institute of Artificial Intelligence)

Visual Saliency Transformer 4702
Nian Liu (Inception Institute of Artificial Intelligence), Ni Zhang (Northwestern Polytechnical University), Kaiyuan Wan (Northwestern Polytechnical University), Ling Shao (Inception Institute of Artificial Intelligence), and Junwei Han (NWPU, China)

HiNet: Deep Image Hiding by Invertible Network 4713
Junpeng Jing (BUAA), Xin Deng (Beihang university), Mai Xu (BUAA), Jianyi Wang (BUAA), and Zhenyu Guan (BUAA)

CANet: A Context-Aware Network for Shadow Removal 4723
Zipei Chen (WuHan University), Chengjiang Long (JD Finance America Corporation), Ling Zhang (Wuhan University of Science and Technology), and Chunxia Xiao (Wuhan University)

Unpaired Learning for Deep Image Deraining With Rain Direction Regularizer 4733
Yang Liu (Dalian University of Technology), Ziyu Yue (Dalian University of Technology), Jinshan Pan (Nanjing University of Science and Technology), and Zhixun Su (Dalian University of Technology)

DivAug: Plug-In Automated Data Augmentation With Explicit Diversity Maximization 4742
Zirui Liu (Texas A&M University), Haifeng Jin (Texas A&M University), Ting-Hsiang Wang (Texas A&M University), Kaixiong Zhou (Texas A&M University), and Xia Hu (Texas A&M University)

Morphable Detector for Object Detection on Demand 4751
Xiangyun Zhao (Northwestern University), Xu Zou (Huazhong University of Science and Technology), and Ying Wu (Northwestern University)

Real-World Video Super-Resolution: A Benchmark Dataset and a Decomposition Based Learning Scheme 4761
Xi Yang (The Hong Kong Polytechnic University), Wangmeng Xiang (The Hong Kong Polytechnic University), Hui Zeng (The Hong Kong Polytechnic University), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)

Designing a Practical Degradation Model for Deep Blind Image Super-Resolution 4771
Kai Zhang (ETH Zurich), Jingyun Liang (ETH Zurich), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)

Learning a Single Network for Scale-Arbitrary Super-Resolution 4781
Longguang Wang (National University of Defense Technology), Yingqian Wang (National University of Defense Technology), Zaiping Lin (National University of Defense Technology), Jungang Yang (National University of Defense Technology), Wei An (National University of Defense Technology), and Yulan Guo (National University of Defense Technology)

Deep Blind Video Super-Resolution 4791
Jinshan Pan (Nanjing University of Science and Technology), Haoran Bai (Nanjing University of Science and Technology), Jiangxin Dong (Max Planck Institute for Informatics), Jiawei Zhang (Sensetime Research), and Jinhui Tang (Nanjing University of Science and Technology)

Achieving On-Mobile Real-Time Super-Resolution With Neural Architecture and Pruning Search 4801
Zheng Zhan (Northeastern University), Yifan Gong (Northeastern University), Pu Zhao (Northeastern University), Geng Yuan (Northeastern University), Wei Niu (William & Mary), Yushu Wu (Northeastern University), Tianyun Zhang (Cleveland State University), Malith Jayaweera (Northeastern University), David Kaeli (Northeastern University), Bin Ren (William & Mary), Xue Lin (Northeastern University), and Yanzhi Wang (Northeastern University)

SSH: A Self-Supervised Framework for Image Harmonization 4812
Yifan Jiang (University of Texas at Austin), He Zhang (Adobe), Jianming Zhang (Adobe Research), Yilin Wang (Adobe), Zhe Lin (Adobe Research), Kalyan Sunkavalli (Adobe Research), Simon Chen (Adobe Research), Sohrab Amirghodsi (Adobe Research), Sarah Kong (Adobe), and Zhangyang Wang (University of Texas at Austin)

Out-of-Boundary View Synthesis Towards Full-Frame Video Stabilization 4822
Yufei Xu (The University of Sydney, Australia), Jing Zhang (The University of Sydney, Australia), and Dacheng Tao (The University of Sydney, Australia)

R-SLAM: Optimizing Eye Tracking From Rolling Shutter Video of the Retina 4832
Jay Shenoy (UC Berkeley), James Fong (UC Berkeley), Jeffrey Tan (UC Berkeley), Austin Roorda (UC Berkeley), and Ren Ng (UC Berkeley)

Attentive and Contrastive Learning for Joint Depth and Motion Field Estimation 4842
Seokju Lee (KAIST), Francois Rameau (KAIST), Fei Pan (Korea Advanced Institute of Science and Technology), and In So Kweon (KAIST)

Panoptic Segmentation of Satellite Image Time Series With Convolutional Temporal Attention Networks 4852
Vivien Sainte Fare Garnot (IGN) and Loic Landrieu (IGN)

EvIntSR-Net: Event Guided Multiple Latent Frames Reconstruction and Super-Resolution 4862
Jin Han (Peking University), Yixin Yang (Peking University), Chu Zhou (Peking University), Chao Xu (Peking University), and Boxin Shi (Peking University)

Dense Deep Unfolding Network With 3D-CNN Prior for Snapshot Compressive Imaging 4872
Zhuoyuan Wu (PKU), Jian Zhang (Peking University Shenzhen Graduate School), and Chong Mou (Peking University Shenzhen Graduate School)

Video Matting via Consistency-Regularized Graph Neural Networks 4882
Tiantian Wang (University of California, Merced), Sifei Liu (NVIDIA), Yapeng Tian (University of Rochester), Kai Li (Northeastern University), and Ming-Hsuan Yang (University of California at Merced)

Collaborative Unsupervised Visual Representation Learning From Decentralized Data 4892
Weiming Zhuang (Nanyang Technological University), Xin Gan (Nanyang Technological University), Yonggang Wen (Nanyang Technological University), Shuai Zhang (Sensetime Ltd), and Shuai Yi (SenseTime Group Limited)

Full-Duplex Strategy for Video Object Segmentation 4902
Ge-Peng Ji (Wuhan University), Keren Fu (Sichuan University), Zhe Wu (Peng Cheng Laboratory), Deng-Ping Fan (Inception Institute of Artificial Intelligence), Jianbing Shen (Inception Institute of Artificial Intelligence), and Ling Shao (Inception Institute of Artificial Intelligence)

iNAS: Integral NAS for Device-Aware Salient Object Detection 4914
Yu-Chao Gu (Nankai University), Shang-Hua Gao (Nankai University), Xu-Sheng Cao (Nankai University), Peng Du (Huawei Technologies), Shao-Ping Lu (Nankai University), and Ming-Ming Cheng (Nankai University)

A Machine Teaching Framework for Scalable Recognition 4925
Pei Wang (UC San Diego) and Nuno Vasconcelos (UC San Diego)

The Benefit of Distraction: Denoising Camera-Based Physiological Measurements Using Inverse Attention 4935
Ewa M. Nowara (Rice University, USA), Daniel McDuff (Microsoft Research), and Ashok Veeraraghavan (Rice University, USA)

Adaptive Graph Convolution for Point Cloud Analysis 4945
Haoran Zhou (Nanjing University), Yidan Feng (Nanjing University of Aeronautics and Astronautics), Mingsheng Fang (Nanjing University), Mingqiang Wei (Nanjing University of Aeronautics and Astronautics), Jing Qin (The Hong Kong Polytechnic University), and Tong Lu (Nanjing University)

Weakly-Supervised Video Anomaly Detection With Robust Temporal Feature Magnitude Learning 4955
Yu Tian (Australian Institute for Machine Learning, University of Adelaide), Guansong Pang (University of Adelaide), Yuanhong Chen (Australian Institute for Machine Learning), Rajvinder Singh (University of Adelaide), Johan W. Verjans (SAHMRI), and Gustavo Carneiro (University of Adelaide)

Improving De-Raining Generalization via Neural Reorganization 4967
Jie Xiao (University of Science and Technology of China), Man Zhou (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), Aiping Liu (University of Science and Technology of China), and Zheng-Jun Zha (University of Science and Technology of China)

Towards Flexible Blind JPEG Artifacts Removal 4977
Jiaxi Jiang (ETH Zurich), Kai Zhang (ETH Zurich), and Radu Timofte (ETH Zurich)

Learning To Remove Refractive Distortions From Underwater Images 4987
Simron Thapa (Louisiana State University), Nianyi Li (Louisiana State University), and Jinwei Ye (Louisiana State University)

Location-Aware Single Image Reflection Removal 4997
Zheng Dong (Zhejiang university), Ke Xu (Shanghai Jiao Tong University), Yin Yang (Clemson University), Hujun Bao (Zhejiang University), Weiwei Xu (Zhejiang unviersity), and Rynson W.H. Lau (City University of Hong Kong)

DC-ShadowNet: Single-Image Hard and Soft Shadow Removal Using Unsupervised Domain-Classifier Guided Network 5007
Yeying Jin (nus), Aashish Sharma (National University of Singapore), and Robby T. Tan (National University of Singapore)

Polarimetric Helmholtz Stereopsis 5017
Yuqi Ding (Louisiana State University), Yu Ji (Plex-VR), Mingyuan Zhou (DGene), Sing Bing Kang (Zillow Group), and Jinwei Ye (Louisiana State University)

Self-Born Wiring for Neural Trees 5027
Ying Chen (Zhejiang University), Feng Mao (Alibaba Group), Jie Song (Zhejiang University), Xinchao Wang (National University of Singapore), Huiqiong Wang (Zhejiang University), and Mingli Song (Zhejiang University)

Student Customized Knowledge Distillation: Bridging the Gap Between Student and Teacher 5037
Yichen Zhu (Midea Group) and Yi Wang (Midea Group)

Adaptive Curriculum Learning 5047
Yajing Kong (The University of Sydney), Liu Liu (The University of Sydney), Jun Wang (City University of Hong Kong), and Dacheng Tao (The University of Sydney)

BlockPlanner: City Block Generation With Vectorized Graph Representation 5057
Linning Xu (CUHK), Yuanbo Xiangli (Chinese University of Hong Kong), Anyi Rao (The Chinese University of Hong Kong), Nanxuan Zhao (Chinese University of Hong Kong), Bo Dai (Nanyang Technological University), Ziwei Liu (Nanyang Technological University), and Dahua Lin (The Chinese University of Hong Kong)

Rethinking Deep Image Prior for Denoising 5067
Yeonsik Jo (GIST), Se Young Chun (Seoul National University), and Jonghyun Choi (GIST)

NASOA: Towards Faster Task-Oriented Online Fine-Tuning With a Zoo of Models 5077
Hang Xu (Huawei Noah's Ark Lab), Ning Kang (Huawei Noah's Ark Lab), Gengwei Zhang (Sun Yat-sen University), Chuanlong Xie (Huawei Noah's Ark Lab), Xiaodan Liang (Sun Yat-sen University), and Zhenguo Li (Huawei Noah's Ark Lab)

Learning Multiple Pixelwise Tasks Based on Loss Scale Balancing 5087
Jae-Han Lee (Korea University), Chul Lee (Dongguk University), and Chang-Su Kim (Korea university)

Pixel Difference Networks for Efficient Edge Detection 5097
Zhuo Su (Mr.), Wenzhe Liu (National University of Defense Technology), Zitong Yu (CMVS, University of Oulu), Dewen Hu (National University of Defense Technology), Qing Liao (Harbin Institute of Technology, (Shenzhen)), Qi Tian (Xidian University), Matti Pietikäinen (University of Oulu), and Li Liu (National University of Defense Technology)

Entropy Maximization and Meta Classification for Out-of-Distribution Detection in Semantic Segmentation 5108
Robin Chan (University of Wuppertal), Matthias Rottmann (University of Wuppertal), and Hanno Gottschalk (University of Wuppertal)

Spectral Leakage and Rethinking the Kernel Size in CNNs 5118
Nergis Tomen (Delft University of Technology) and Jan C. van Gemert (Delft University of Technology)

MUSIQ: Multi-Scale Image Quality Transformer 5128
Junjie Ke (Google), Qifei Wang (Google), Yilin Wang (Google Inc.), Peyman Milanfar (Google), and Feng Yang (Google Research)

BlockCopy: High-Resolution Video Processing With Block-Sparse Feature Propagation and Online Policies 5138
Thomas Verelst (KU Leuven) and Tinne Tuytelaars (KU Leuven)

SACoD: Sensor Algorithm Co-Design Towards Efficient CNN-Powered Intelligent PhlatCam 5148
Yonggan Fu (Rice University), Yang Zhang (IBM T. J. Watson Research), Yue Wang (Rice University), Zhihan Lu (Rice University), Vivek Boominathan (Rice University), Ashok Veeraraghavan (Rice University), and Yingyan Lin (Rice University)

Unsupervised Curriculum Domain Adaptation for No-Reference Video Quality Assessment 5158
Pengfei Chen (Xidian University / China University of Mining and Technology), Leida Li (Xidian University), Jinjian Wu (Xidian University), Weisheng Dong (Xidian University), and Guangming Shi (Xidian University)

Bit-Mixer: Mixed-Precision Networks With Runtime Bit-Width Selection 5168
Adrian Bulat (Samsung AI Center, Cambridge) and Georgios Tzimiropoulos (Queen Mary University of London)

ReCU: Reviving the Dead Weights in Binary Neural Networks 5178
Zihan Xu (Xiamen University, China), Mingbao Lin (Xiamen University, China), Jianzhuang Liu (Noah's Ark Lab, Huawei Technologies), Jie Chen (Peking University), Ling Shao (Inception Institute of Artificial Intelligence), Yue Gao (Tsinghua University), Yonghong Tian (Peking University), and Rongrong Ji (Xiamen University, China)

HIRE-SNN: Harnessing the Inherent Robustness of Energy-Efficient Deep Spiking Neural Networks by Training With Crafted Input Noise 5189
Souvik Kundu (University of Southern California), Massoud Pedram (University of Southern California), and Peter A. Beerel (University of Southern California)

FATNN: Fast and Accurate Ternary Neural Networks 5199
Peng Chen (The University of Adelaide), Bohan Zhuang (Monash University), and Chunhua Shen (University of Adelaide)

Towards Memory-Efficient Neural Networks via Multi-Level In Situ Generation 5209
Jiaqi Gu (University of Texas at Austin), Hanqing Zhu (University of Texas at Austin), Chenghao Feng (University of Texas at Austin), Mingjie Liu (The University of Texas at Austin), Zixuan Jiang (The University of Texas at Austin), Ray T. Chen (The University of Texas, Austin), and David Z. Pan (University of Texas at Austin)

GDP: Stabilized Neural Network Pruning via Gates With Differentiable Polarization 5219
Yi Guo (Kuaishou Technology), Huan Yuan (Kuaishou Technology), Jianchao Tan (Kwai Inc.), Zhangyang Wang (University of Texas at Austin), Sen Yang (Kwai Inc.), and Ji Liu (Kwai Inc.)

RMSMP: A Novel Deep Neural Network Quantization Framework With Row-Wise Mixed Schemes and Multiple Precisions 5231
Sung-En Chang (Northeastern University), Yanyu Li (Northeastern University), Mengshu Sun (Northeastern University), Weiwen Jiang (University of Notre Dame), Sijia Liu (Michigan State University), Yanzhi Wang (Northeastern University), and Xue Lin (Northeastern University)

Improving Low-Precision Network Quantization via Bin Regularization 5241
Tiantian Han (Xilinx Technology Beijing Limited), Dong Li (Xilinx), Ji Liu (Xilinx Beijing), Lu Tian (Xilinx, Inc.), and Yi Shan (Xilinx)

Distance-Aware Quantization 5251
Dohyung Kim (Yonsei University), Junghyup Lee (Yonsei University), and Bumsub Ham (Yonsei University)

Improving Neural Network Efficiency via Post-Training Quantization With Adaptive Floating-Point 5261
Fangxin Liu (Shanghai Jiao Tong University), Wenbo Zhao (Shanghai Jiao Tong University), Zhezhi He (Shanghai Jiao Tong University), Yanzhi Wang (Northeastern University), Zongwu Wang (Shanghai Jiao Tong University), Changzhi Dai (DeepBlue Technology, (Shanghai) Co., Ltd.), Xiaoyao Liang (SJTU), and Li Jiang (Shanghai Jiao Tong University)

Generalizable Mixed-Precision Quantization via Attribution Rank Preservation 5271
Ziwei Wang (Tsinghua University), Han Xiao (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)

Meta-Aggregator: Learning To Aggregate for 1-Bit Graph Neural Networks 5281
Yongcheng Jing (The University of Sydney), Yiding Yang (Stevens Institute of Technology), Xinchao Wang (National University of Singapore), Mingli Song (Zhejiang University), and Dacheng Tao (The University of Sydney)

Channel-Wise Knowledge Distillation for Dense Prediction 5291
Changyong Shu (Shanghai Em-Data Technology Co., Ltd.), Yifan Liu (University of Adelaide), Jianfei Gao (Shanghai Em-Data Technology Co., Ltd.), Zheng Yan (University of Adelaide), and Chunhua Shen (Shanghai Em-Data Technology Co., Ltd)

Improving Generalization of Batch Whitening by Convolutional Unit Optimization 5301
Yooshin Cho (KAIST), Hanbyel Cho (KAIST), Youngsoo Kim (KAIST), and Junmo Kim (KAIST)

Dynamic Dual Gating Neural Networks 5310
Fanrong Li (Institute of Automation, Chinese Academy of Sciences), Gang Li (Institute of Automation, Chinese Academy of Sciences), Xiangyu He (Institute of Automation, Chinese Academy of Sciences), and Jian Cheng (Chinese Academy of Sciences, China)

Once Quantization-Aware Training: High Performance Extremely Low-Bit Architecture Search .. 5320
Mingzhu Shen (Sensetime Research), Feng Liang (Sensetime Research), Ruihao Gong (SenseTime), Yuhang Li (Yale University), Chuming Li (sensetime), Chen Lin (University of Oxford), Fengwei Yu (Sensetime Research), Junjie Yan (Sensetime Group Limited), and Wanli Ouyang (The University of Sydney)

Towards Mixed-Precision Quantization of Neural Networks via Constrained Optimization 5330
Weihan Chen (CASIA), Peisong Wang (Institute of Automation, Chinese Academy of Sciences), and Jian Cheng (Chinese Academy of Sciences, China)

Sub-Bit Neural Networks: Learning To Compress and Accelerate Binary Neural Networks 5340
Yikai Wang (Tsinghua University), Yi Yang (Intel), Fuchun Sun (Tsinghua University), and Anbang Yao (Intel Labs China)

Cluster-Promoting Quantization With Bit-Drop for Minimizing Network Quantization Loss 5350
Jung Hyun Lee (KAIST), Jihun Yun (KAIST), Sung Ju Hwang (KAIST, AITRICS), and Eunho Yang (KAIST; AITRICS)

Fast and Efficient DNN Deployment via Deep Gaussian Transfer Learning 5360
Qi Sun (The Chinese University of Hong Kong), Chen Bai (The Chinese University of Hong Kong), Tinghuan Chen (The Chinese University of Hong Kong), Hao Geng (The Chinese University of Hong Kong), Xinyun Zhang (SmartMore), Yang Bai (CUHK), and Bei Yu (CUHK)

SmartShadow: Artistic Shadow Drawing Tool for Line Drawings 5371
Lvmin Zhang (Soochow University / Style2Paints Research), Jinyue Jiang (The University of Tokyo / Style2Paints Research), Yi Ji (Soochow University), and Chunping Liu (Soochow University)

BEV-Net: Assessing Social Distancing Compliance by Joint People Localization and Geometric Reasoning 5381
Zhirui Dai (UC San Diego), Yuepeng Jiang (UC San Diego), Yi Li (UC San Diego), Bo Liu (UCSD), Antoni B. Chan (City University of Hong Kong, Hong, Kong), and Nuno Vasconcelos (UC San Diego)

Towards Real-World Prohibited Item Detection: A Large-Scale X-Ray Benchmark 5392
Boying Wang (Institute of Software Chinese Academy of Sciences), Libo Zhang (Institute of Software Chinese Academy of Sciences), Longyin Wen (Bytedance, Inc.), Xianglong Liu (Beihang University), and Yanjun Wu (Institute of Software Chinese Academy of Sciences)

Explain Me the Painting: Multi-Topic Knowledgeable Art Description Generation 5402
Zechen Bai (University of Chinese Academy of Sciences), Yuta Nakashima (Osaka University), and Noa Garcia (Osaka University)

ZFlow: Gated Appearance Flow-Based Virtual Try-On With 3D Priors 5413
Ayush Chopra (MIT), Rishabh Jain (Bits Pilani, Pilani Campus), Mayur Hemani (Adobe Systems), and Balaji Krishnamurthy

3D Vision Applications and Systems

Structure-From-Sherds: Incremental 3D Reassembly of Axially Symmetric Pots From Unordered and Mixed Fragment Collections 5423
Je Hyeong Hong (Hanyang University), Seong Jong Yoo (KIST), Muhammad Arshad Zeeshan (Korea Institute of Science and Technology), Young Min Kim (Seoul National University), and Jinwook Kim (KIST)

Rotation Averaging in a Split Second: A Primal-Dual Method and a Closed-Form for Cycle Graphs 5432
Gabriel Moreira (Instituto Superior Técnico), Manuel Marques (Instituto Superior Tecnico, Portugal), and João Paulo Costeira (Instituto Superior Tecnico)

imGHUM: Implicit Generative Models of 3D Human Shape and Articulated Pose 5441
Thiemo Alldieck (Google), Hongyi Xu (Google), and Cristian Sminchisescu (Google)

DeePSD: Automatic Deep Skinning and Pose Space Deformation for 3D Garment Animation 5451
Hugo Bertiche (University of Barcelona), Meysam Madadi (CVC), Emilio Tylson (Univeristat de Barcelona), and Sergio Escalera (Computer Vision Center, (UAB) & University of Barcelona)

CanvasVAE: Learning To Generate Vector Graphic Documents 5461
Kota Yamaguchi (Cyberagent)

Distinctiveness Oriented Positional Equilibrium for Point Cloud Registration 5470
Taewon Min (KAIST), Chonghyuk Song (Carnegie Mellon University), Eunseok Kim (Agency for Defense Development), and Inwook Shim (Agency for Defense Development)

SnowflakeNet: Point Cloud Completion by Snowflake Point Deconvolution With Skip-Transformer 5479
Peng Xiang (Tsinghua University), Xin Wen (Tsinghua University and JD.com), Yu-Shen Liu (Tsinghua University), Yan-Pei Cao (Y-tech, Kuaishou Technology), Pengfei Wan (Kuaishou Technology), Wen Zheng (Kuaishou Technology), and Zhizhong Han (Wayne State University)

Superpoint Network for Point Cloud Oversegmentation 5490
Le Hui (Nanjing University of Science and Technology), Jia Yuan (Nanjing University of Science and Technology), Mingmei Cheng (Nanjing University of Science and Technology), Jin Xie (Nanjing University of Science and Technology), Xiaoya Zhang (NanJing University of Science and Technology), and Jian Yang (Nanjing University of Science and Technology)

4D Cloud Scattering Tomography 5500
Roi Ronen (Technion), Yoav Y. Schechner (Technion), and Eshkol Eytan (The Weizmann Institute of Science)

Feature Interactive Representation for Point Cloud Registration 5510
Bingli Wu (HuaZhong University of Science and Technology), Jie Ma (Huazhong University of Science and Technology), Gaojie Chen (Huazhong University of Science and Technology), and Pei An (HuaZhong University of Science and Technology)

Viewing Graph Solvability via Cycle Consistency 5520
Federica Arrigoni (University of Trento), Andrea Fusiello (UNIUD), Elisa Ricci (University of Trento), and Tomas Pajdla (Czech Technical University in Prague)

MBA-VO: Motion Blur Aware Visual Odometry 5530
Peidong Liu (ETH Zurich), Xingxing Zuo (Department of Computer Science, ETH Zürich), Viktor Larsson (ETH Zurich), and Marc Pollefeys (ETH Zurich / Microsoft)

Lightweight Multi-Person Total Motion Capture Using Sparse Multi-View Cameras 5540
Yuxiang Zhang (Tsinghua University), Zhe Li (Tsinghua University), Liang An (Tsinghua University), Mengcheng Li (Tsinghua University), Tao Yu (Tsinghua University), and Yebin Liu (Tsinghua University)

Orthographic-Perspective Epipolar Geometry 5550
Viktor Larsson (ETH Zurich), Marc Pollefeys (ETH Zurich / Microsoft), and Magnus Oskarsson (Lund University)

Minimal Solutions for Panoramic Stitching Given Gravity Prior 5559
Yaqing Ding (Nanjing University of Science and Technology), Daniel Barath (ETH Zürich), and Zuzana Kukelova (Czech Technical University in Prague)

UNISURF: Unifying Neural Implicit Surfaces and Radiance Fields for Multi-View Reconstruction 5569
Michael Oechsle (MPI-IS, University of Tuebingen and ETAS GmbH), Songyou Peng (ETH Zurich and MPI-IS), and Andreas Geiger (University of Tuebingen)

PR-RRN: Pairwise-Regularized Residual-Recursive Networks for Non-Rigid Structure-From-Motion 5580
Haitian Zeng (Baidu), Yuchao Dai (Northwestern Polytechnical University), Xin Yu (University of Technology Sydney), Xiaohan Wang (University of Technology, Sydney), and Yi Yang (UTS)

NerfingMVS: Guided Optimization of Neural Radiance Fields for Indoor Multi-View Stereo 5590
Yi Wei (Tsinghua University), Shaohui Liu (ETH Zurich), Yongming Rao (Tsinghua University), Wang Zhao (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)

H3D-Net: Few-Shot High-Fidelity 3D Head Reconstruction 5600
Eduard Ramon (Universitat Politècnica de Catalunya), Gil Triginer (Crisalix), Janna Escur (Crisalix), Albert Pumarola (IRI), Jaime Garcia (Crisalix SA), Xavier Giró-i-Nieto (Universitat Politecnica de Catalunya), and Francesc Moreno-Noguer (IRI)

Scene Synthesis via Uncertainty-Driven Attribute Synchronization 5610
Haitao Yang (University of Texas at Austin), Zaiwei Zhang (University of Texas at Austin), Siming Yan (The University of Texas at Austin), Haibin Huang (Kuaishou Technology), Chongyang Ma (Kuaishou Technology), Yi Zheng (Kuaishou Technology), Chandrajit Bajaj (University of Texas, Austin), and Qixing Huang (The University of Texas at Austin)

Out-of-Core Surface Reconstruction via Global TGV Minimization 5621
Nikolai Poliarnyi (Agisoft LLC)

Adaptive Surface Reconstruction With Multiscale Convolutional Kernels 5631
Benjamin Ummenhofer (Intel Labs) and Vladlen Koltun (Intel Labs)

Learning Icosahedral Spherical Probability Map Based on Bingham Mixture Model for Vanishing Point Estimation 5641
Haoang Li (The Chinese University of Hong Kong), Kai Chen (The Chinese University of Hong Kong), Pyojin Kim (Sookmyung Women's University), Kuk-Jin Yoon (KAIST), Zhe Liu (University of Cambridge), Kyungdon Joo (UNIST), and Yun-Hui Liu (CUHK)

3DeepCT: Learning Volumetric Scattering Tomography of Clouds 5651
Yael Sde-Chen (Technion), Yoav Y. Schechner (Technion), Vadim Holodovsky (Technion), and Eshkol Eytan (The Weizmann Institute of Science)

DeepPRO: Deep Partial Point Cloud Registration of Objects 5663
Donghoon Lee (Apple), Onur C. Hamsici (Apple), Steven Feng (Apple), Prachee Sharma (Apple Inc), and Thorsten Gernoth (Apple Inc.)

Pri3D: Can 3D Priors Help 2D Representation Learning? 5673
Ji Hou (Technical University of Munich), Saining Xie (Facebook AI Research), Benjamin Graham (Facebook Research), Angela Dai (Technical University of Munich), and Matthias Nießner (Technical University of Munich)

Extreme Structure From Motion for Indoor Panoramas Without Visual Overlaps 5683
Mohammad Amin Shabani (Simon Fraser University), Weilian Song (Simon Fraser University), Makoto Odamaki (Ricoh Company), Hirochika Fujiki (RICOH COMPANY, LTD.), and Yasutaka Furukawa (Simon Fraser University)

Dynamic View Synthesis From Dynamic Monocular Video 5692
Chen Gao (Virginia Tech), Ayush Saraf (Facebook), Johannes Kopf (Facebook), and Jia-Bin Huang (Virginia Tech)

Multi-View 3D Reconstruction With Transformers 5702
Dan Wang (University of British Columbia), Xinrui Cui (University of Science and Technology of China), Xun Chen (University of Science and Technology of China), Zhengxia Zou (University of Michigan), Tianyang Shi (Beihang University), Septimiu Salcudean (The University of British Columbia), Z. Jane Wang (University of British Columbia), and Rabab Ward (University of British Columbia)

EPP-MVSNet: Epipolar-Assembling Based Depth Prediction for Multi-View Stereo 5712
Xinjun Ma (Huawei), Yue Gong (Huawei Technologies Co., Ltd.), Qirui Wang (Huawei Technologies Co., Ltd.), Jingwei Huang (Huawei), Lei Chen (Hong Kong University of Science and Technology), and Fan Yu (Huawei Technologies Co. Ltd)

BARF: Bundle-Adjusting Neural Radiance Fields 5721
Chen-Hsuan Lin (Carnegie Mellon University), Wei-Chiu Ma (MIT), Antonio Torralba (MIT), and Simon Lucey (CMU)

PlenOctrees for Real-Time Rendering of Neural Radiance Fields 5732
Alex Yu (UC Berkeley), Ruilong Li (USC Institute for Creative Technologies, (ICT)), Matthew Tancik (UC Berkeley), Hao Li (Pinscreen / UC Berkeley), Ren Ng (UC Berkeley), and Angjoo Kanazawa (University of California Berkeley)

Neural Articulated Radiance Field 5742
Atsuhiro Noguchi (The University of Tokyo), Xiao Sun (Microsoft Research Asia), Stephen Lin (Microsoft Research), and Tatsuya Harada (The University of Tokyo / RIKEN)

Editing Conditional Radiance Fields 5753
Steven Liu (MIT), Xiuming Zhang (MIT), Zhoutong Zhang (MIT), Richard Zhang (Adobe), Jun-Yan Zhu (Carnegie Mellon University), and Bryan Russell (Adobe Research)

AD-NeRF: Audio Driven Neural Radiance Fields for Talking Head Synthesis 5764
Yudong Guo (University of Science and Technology of China), Keyu Chen (Netease Fuxi AI Lab), Sen Liang (Zhejiang University), Yong-Jin Liu (Tsinghua University), Hujun Bao (Zhejiang University), and Juyong Zhang (University of Science and Technology of China)

GTT-Net: Learned Generalized Trajectory Triangulation 5775
Xiangyu Xu (Stevens Institute of Technology) and Enrique Dunn (Stevens Institute of Technology)

Deep Hybrid Self-Prior for Full 3D Mesh Generation 5785
Xingkui Wei (Fudan University), Zhengqing Chen (Fudan University), Yanwei Fu (Fudan University), Zhaopeng Cui (Zhejiang University), and Yinda Zhang (Google)

ARAPReg: An As-Rigid-As Possible Regularization Loss for Learning Deformable Shape Generators 5795
Qixing Huang (The University of Texas at Austin), Xiangru Huang (University of Texas at Austin), Bo Sun (UT Austin), Zaiwei Zhang (University of Texas at Austin), Junfeng Jiang (Hohai University), and Chandrajit Bajaj (University of Texas, Austin)

3D Shape Generation and Completion Through Point-Voxel Diffusion 5806
Linqi Zhou (University of California, Los Angeles), Yilun Du (MIT), and Jiajun Wu (Stanford University)

LSD-StructureNet: Modeling Levels of Structural Detail in 3D Part Hierarchies 5816
Dominic Roberts (University of Illinois at Urbana-Champaign), Ara Danielyan (Autodesk AI Lab), Hang Chu (Autodesk Research), Mani Golparvar-Fard (University of Illinois at Urbana-Champaign), and David Forsyth (Univeristy of Illinois at Urbana-Champaign)

Self-Calibrating Neural Radiance Fields 5826
Yoonwoo Jeong (POSTECH), Seokjun Ahn (POSTECH), Christopher Choy (Nvidia), Anima Anandkumar (Caltech), Minsu Cho (POSTECH), and Jaesik Park (POSTECH)

Mip-NeRF: A Multiscale Representation for Anti-Aliasing Neural Radiance Fields 5835
Jonathan T. Barron (Google Research), Ben Mildenhall (Google Research), Matthew Tancik (UC Berkeley), Peter Hedman (Google Research), Ricardo Martin-Brualla (Google), and Pratul P. Srinivasan (Google Research)

Nerfies: Deformable Neural Radiance Fields 5845
Keunhong Park (University of Washington), Utkarsh Sinha (Google), Jonathan T. Barron (Google Research), Sofien Bouaziz (Facebook Reality Labs), Dan B Goldman (Google, Inc.), Steven M. Seitz (University of Washington), and Ricardo Martin-Brualla (Google)

Baking Neural Radiance Fields for Real-Time View Synthesis 5855
Peter Hedman (Google Research), Pratul P. Srinivasan (Google Research), Ben Mildenhall (Google Research), Jonathan T. Barron (Google Research), and Paul Debevec (Google Research)

Putting NeRF on a Diet: Semantically Consistent Few-Shot View Synthesis 5865
Ajay Jain (UC Berkeley), Matthew Tancik (UC Berkeley), and Pieter Abbeel (UC Berkeley)

PoGO-Net: Pose Graph Optimization With Graph Neural Networks 5875
Xinyi Li (Temple University) and Haibin Ling (Stony Brook University)

Radial Distortion Invariant Factorization for Structure From Motion 5886
José Pedro Iglesias (Chalmers) and Carl Olsson (Lund University, Sweden)

Gaussian Fusion: Accurate 3D Reconstruction via Geometry-Guided Displacement Interpolation. 5896
Duo Chen (Sichuan University), Zixin Tang (Sichuan University), Zhenyu Xu (School of Computer Science and Technology, Sichuan University), Yunan Zheng (Sichuan University), and Yiguang Liu (Sichuan University)

Dynamical Pose Estimation 5906
Heng Yang (MIT), Chris Doran (Cambridge University), and Jean-Jacques Slotine (MIT)

Calibrated and Partially Calibrated Semi-Generalized Homographies 5916
Snehal Bhayani (University of Oulu, Finland), Torsten Sattler (Czech Technical University in Prague), Daniel Barath (ETH Zürich), Patrik Beliansky (Faculty of Mathematics and Physics, Charles University, Prague), Janne Heikkilä (University of Oulu, Finland), and Zuzana Kukelova (Czech Technical University in Prague)

ELLIPSDF: Joint Object Pose and Shape Optimization With a Bi-Level Ellipsoid and Signed Distance Function Description 5926
Mo Shan (UCSD), Qiaojun Feng (University of California, San Diego), You-Yi Jau (UCSD), and Nikolay Atanasov (University of California, San Diego)

Learning Efficient Photometric Feature Transform for Multi-View Stereo 5936
Kaizhang Kang (Zhejiang University), Cihui Xie (Zhejiang University), Ruisheng Zhu (zhejiang university), Xiaohe Ma (zhejiang university), Ping Tan (Simon Fraser University), Hongzhi Wu (Zhejiang University), and Kun Zhou (Zhejiang University)

STR-GQN: Scene Representation and Rendering for Unknown Cameras Based on Spatial Transformation Routing 5946
Wen-Cheng Chen (National Cheng Kung University), Min-Chun Hu (National Tsing Hua University), and Chu-Song Chen (National Taiwan University)

Deep Permutation Equivariant Structure From Motion 5956
Dror Moran (Weizmann Institute of Science), Hodaya Koslowsky (Weizmann Institute of Science), Yoni Kasten (Weizmann Institute), Haggai Maron (NVIDIA Research), Meirav Galun (Weizmann Institute of Science), and Ronen Basri (Weizmann Institute of Science)

Pixel-Perfect Structure-From-Motion With Featuremetric Refinement 5967
Philipp Lindenberger (ETH Zurich), Paul-Edouard Sarlin (ETH Zurich), Viktor Larsson (ETH Zurich), and Marc Pollefeys (ETH Zurich / Microsoft)

ODAM: Object Detection, Association, and Mapping Using Posed RGB Video 5978
Kejie Li (University of Adelaide), Daniel DeTone (Facebook), Yu Fan Steven Chen (Facebook Reality Labs), Minh Vo (Facebook Reality Labs), Ian Reid (University of Adelaide, Australia), Hamid Rezatofighi (Monash University), Chris Sweeney (Facebook Reality Labs), Julian Straub (Facebook Reality Labs), and Richard Newcombe (Facebook)

SaccadeCam: Adaptive Visual Attention for Monocular Depth Sensing 5989
Brevin Tilmon (University of Florida) and Sanjeev J. Koppal (University of Florida)

Transfusion: A Novel SLAM Method Focused on Transparent Objects 5999
Yifan Zhu (Nankai University), Jiaxiong Qiu (Nankai University), and Bo Ren (Nankai University)

Matching in the Dark: A Dataset for Matching Image Pairs of Low-Light Scenes 6009
Wenzheng Song (Tohoku University, Japan), Masanori Suganuma (RIKEN AIP / Tohoku University), Xing Liu (Tohoku University, Japan), Noriyuki Shimobayashi (Tohoku University, Japan), Daisuke Maruta (Socionext Inc.), and Takayuki Okatani (Tohoku University/RIKEN AIP)

DepthInSpace: Exploitation and Fusion of Multiple Video Frames for Structured-Light Depth Estimation 6019
Mohammad Mahdi Johari (Idiap Research Institute, École polytechnique fédérale de Lausanne, (EPFL)), Camilla Carta (ams OSRAM), and François Fleuret (University of Geneva)

Stacked Homography Transformations for Multi-View Pedestrian Detection 6029
Liangchen Song (University at Buffalo), Jialian Wu (State University of New York at Buffalo), Ming Yang (Horizon Robotics), Qian Zhang (Horizon Robotics), Yuan Li (Google), and Junsong Yuan (State University of New York at Buffalo, USA)

Cross-Descriptor Visual Localization and Mapping 6038
Mihai Dusmanu (ETH Zurich), Ondrej Miksik (Microsoft), Johannes L. Schönberger (Microsoft), and Marc Pollefeys (ETH Zurich / Microsoft)

Minimal Cases for Computing the Generalized Relative Pose Using Affine Correspondences 6048
Banglei Guan (National University of Defense Technology), Ji Zhao (TuSimple), Daniel Barath (ETH Zürich), and Friedrich Fraundorfer (Graz University of Technology)

Digging Into Uncertainty in Self-Supervised Multi-View Stereo 6058
Hongbin Xu (South China University of Technology), Zhipeng Zhou (Alibaba Group), Yali Wang (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Wenxiong Kang (South China University of Technology), Baigui Sun (Alibaba Group), Hao Li (Alibaba Group), and Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences)

Differentiable Surface Rendering via Non-Differentiable Sampling 6068
Forrester Cole (Google Research), Kyle Genova (Google Research), Avneesh Sud (Google), Daniel Vlasic (Google), and Zhoutong Zhang (MIT)

Pyramid Point Cloud Transformer for Large-Scale Place Recognition 6078
Le Hui (Nanjing University of Science and Technology), Hang Yang (Nanjing University of Science and Technology), Mingmei Cheng (Nanjing University of Science and Technology), Jin Xie (Nanjing University of Science and Technology), and Jian Yang (Nanjing University of Science and Technology)

(Just) A Spoonful of Refinements Helps the Registration Error Go Down 6088
Sérgio Agostinho (Institute for Systems and Robotics, Instituto Superior Técnico, Universidade de Lisboa), Aljoša Ošep (TUM Munich), Alessio Del Bue (Istituto Italiano di Tecnologia, (IIT)), and Laura Leal-Taixé (TUM)

AdaFit: Rethinking Learning-Based Normal Estimation on Point Clouds 6098
Runsong Zhu (Wuhan University), Yuan Liu (The University of Hong Kong), Zhen Dong (Wuhan University), Yuan Wang (Wuhan University), Tengping Jiang (Wuhan University), Wenping Wang (The University of Hong Kong), and Bisheng Yang (Wuhan University)

Sampling Network Guided Cross-Entropy Method for Unsupervised Point Cloud Registration ... 6108
Haobo Jiang (Nanjing University of Science and Technology), Yaqi Shen (Nanjing University of Science and Technology), Jin Xie (Nanjing University of Science and Technology), Jun Li (Nanjing University of Science and Technology), Jianjun Qian (Nanjing University of Science and Technology), and Jian Yang (Nanjing University of Science and Technology)

A Robust Loss for Point Cloud Registration 6118
Zhi Deng (University of Science and Technology of China), Yuxin Yao (University of Science and Technology of China), Bailin Deng (Cardiff University), and Juyong Zhang (University of Science and Technology of China)

Rational Polynomial Camera Model Warping for Deep Learning Based Satellite Multi-View Stereo Matching 6128
Jian Gao (Wuhan University), Jin Liu (Wuhan University), and Shunping Ji (Wuhan University)

PatchMatch-RL: Deep MVS With Pixelwise Depth, Normal, and Visibility 6138
Jae Yong Lee (University of Illinois), Joseph DeGol (Microsoft), Chuhang Zou (Amazon), and Derek Hoiem (University of Illinois at Urbana-Champaign)

A Confidence-Based Iterative Solver of Depths and Surface Normals for Deep Multi-View Stereo 6148
Wang Zhao (Tsinghua University), Shaohui Liu (ETH Zurich), Yi Wei (Tsinghua University), Hengkai Guo (ByteDance AI Lab), and Yong-Jin Liu (Tsinghua University)

Just a Few Points Are All You Need for Multi-View Stereo: A Novel Semi-Supervised Learning Method for Multi-View Stereo 6158
Taekyung Kim (KAIST), Jaehoon Choi (University of Maryland), Seokeon Choi (KAIST), Dongki Jung (KAIST), and Changick Kim (KAIST)

AA-RMVSNet: Adaptive Aggregation Recurrent Multi-View Stereo Network 6167
Zizhuang Wei (Peking University), Qingtian Zhu (Peking University), Chen Min (Peking UniversityNational Innovation Institute of Defense Technology), Yisong Chen (Peking University), and Guoping Wang (Peking University)

Revisiting Stereo Depth Estimation From a Sequence-to-Sequence Perspective With Transformers 6177
Zhaoshuo Li (Johns Hopkins University), Xingtong Liu (Johns Hopkins University), Nathan Drenkow (The Johns Hopkins University), Andy Ding (Johns Hopkins University), Francis X. Creighton (Johns Hopkins University), Russell H. Taylor (Johns Hopkins University), and Mathias Unberath (Johns Hopkins University)

COTR: Correspondence Transformer for Matching Across Images 6187
Wei Jiang (University of British Columbia), Eduard Trulls (Google), Jan Hosang (Google), Andrea Tagliasacchi (Google Inc.), and Kwang Moo Yi (University of British Columbia)

On the Limits of Pseudo Ground Truth in Visual Camera Re-Localisation 6198
Eric Brachmann (Niantic), Martin Humenberger (Naver Labs Europe), Carsten Rother (University of Heidelberg), and Torsten Sattler (Czech Technical University in Prague)

iMAP: Implicit Mapping and Positioning in Real-Time 6209
Edgar Sucar (Imperial College London), Shikun Liu (Imperial College London), Joseph Ortiz (Imperial College London), and Andrew J. Davison (Imperial College London)

DeepMultiCap: Performance Capture of Multiple Characters Using Sparse Multiview Cameras ... 6219
Yang Zheng (Tsinghua University), Ruizhi Shao (Tsinghua University), Yuxiang Zhang (Tsinghua University), Tao Yu (Tsinghua University), Zerong Zheng (Tsinghua University), Qionghai Dai (Tsinghua University), and Yebin Liu (Tsinghua University)

Learning To Bundle-Adjust: A Graph Network Approach to Faster Optimization of Bundle Adjustment for Vehicular SLAM 6230
Tetsuya Tanaka (Socionext Inc.), Yukihiro Sasagawa (Socionext Inc.), and Takayuki Okatani (Tohoku University/RIKEN AIP)

Faster Multi-Object Segmentation Using Parallel Quadratic Pseudo-Boolean Optimization 6240
Niels Jeppesen (DTU Compute), Patrick M. Jensen (Technical University of Denmark), Anders N. Christensen (DTU Compute), Anders B. Dahl (Technical University of Denmark), and Vedrana A. Dahl (Technical University of Denmark)

Fusion Moves for Graph Matching 6250
Lisa Hutschenreiter (Heidelberg University), Stefan Haller (Heidelberg University), Lorenz Feineis (Heidelberg University), Carsten Rother (University of Heidelberg), Dagmar Kainmüller (BIH/MDC), and Bogdan Savchynskyy (Heidelberg University)

FFT-OT: A Fast Algorithm for Optimal Transportation 6260
Na Lei (Dalian University of Technology) and Xianfeng Gu (Stony Brook University)

Distilling Global and Local Logits With Densely Connected Relations 6270
Youmin Kim (Kyung Hee University), Jinbae Park (Kyung Hee University), YounHo Jang (Kyung Hee University), Muhammad Ali (Kyung Hee University), Tae-Hyun Oh (POSTECH), and Sung-Ho Bae (Kyung Hee University)

Learning To Match Features With Seeded Graph Matching Network 6281
Hongkai Chen (HKUST), Zixin Luo (HKUST), Jiahui Zhang (Alibaba A.I. Labs), Lei Zhou (HKUST), Xuyang Bai (HKUST), Zeyu Hu (Hong Kong University of Science and Technology), Chiew-Lan Tai (Hong Kong University of Science & Technology), and Long Quan (Hong Kong University of Science and Technology)

UASNet: Uncertainty Adaptive Sampling Network for Deep Stereo Matching 6291
Yamin Mao (Samsung Research Center of Beijing), Zhihua Liu (Samsung Research Center of Beijing), Weiming Li (Samsung Research Center of Beijing), Yuchao Dai (Northwestern Polytechnical University), Qiang Wang (Samsung Research Center of Beijing), Yun-Tae Kim (Samsung Advanced Institute of Technology), and Hong-Seok Lee (Samsung Advanced Institute of Technology)

Fast Light-Field Disparity Estimation With Multi-Disparity-Scale Cost Aggregation 6300
Zhicong Huang (Nanjing University), Xuemei Hu (Nanjing University), Zhou Xue (Bytedance), Weizhu Xu (Nanjing University), and Tao Yue (Nanjing University)

Making Higher Order MOT Scalable: An Efficient Approximate Solver for Lifted Disjoint Paths 6310
Andrea Hornakova (Max Planck Institute for Informatics), Timo Kaiser (Leibniz University of Hannover), Paul Swoboda (MPI fuer Informatik, Saarbruecken), Michal Rolinek (Max Planck Institute for Intelligent Systems), Bodo Rosenhahn (Leibniz University Hannover), and Roberto Henschel (Leibniz University of Hannover)

TMCOSS: Thresholded Multi-Criteria Online Subset Selection for Data-Efficient Autonomous Driving 6321
Soumi Das (Indian Institute of Technology, Kharagpur), Harikrishna Patibandla (IIT Kharagpur), Suparna Bhattacharya (Hewlett Packard Enterprise, Bangalore, India), Kshounis Bera (HPE), Niloy Ganguly (IIT Kharagpur), and Sourangshu Bhattacharya (IIT KHARAGPUR)

GNeRF: GAN-Based Neural Radiance Field Without Posed Camera 6331
Quan Meng (ShanghaiTech University), Anpei Chen (shanghaitech), Haimin Luo (Shanghaitech), Minye Wu (Shanghaitech University), Hao Su (UCSD), Lan Xu (ShanghaiTech), Xuming He (ShanghaiTech University), and Jingyi Yu (Shanghai Tech University)

Auto Graph Encoder-Decoder for Neural Network Pruning 6342
Sixing Yu (Iowa State University), Arya Mazaheri (TU Darmstadt), and Ali Jannesari (Iowa State University)

Gradient Normalization for Generative Adversarial Networks 6353
Yi-Lun Wu (National Chiao Tung University), Hong-Han Shuai (National Chiao Tung University), Zhi-Rui Tam (National Chiao Tung University), and Hong-Yu Chiu (National Chiao Tung University)

DRB-GAN: A Dynamic ResBlock Generative Adversarial Network for Artistic Style Transfer 6363
Wenju Xu (University of Kansas), Chengjiang Long (JD Finance America Corporation), Ruisheng Wang (University of Calgary, Canada), and Guanghui Wang (Ryerson University)

WarpedGANSpace: Finding Non-Linear RBF Paths in GAN Latent Space 6373
Christos Tzelepis (Queen Mary University of London), Georgios Tzimiropoulos (Queen Mary University of London), and Ioannis Patras (Queen Mary University of London)

Geometry-Aware Self-Training for Unsupervised Domain Adaptation on Object Point Clouds 6383
Longkun Zou (South China University of Technology), Hui Tang (South China University of Technology), Ke Chen (South China University of Technology), and Kui Jia (South China University of Technology)

Progressive Seed Generation Auto-Encoder for Unsupervised Point Cloud Learning 6393
Juyoung Yang (KAIST), Pyunghwan Ahn (KAIST), Doyeon Kim (KAIST), Haeil Lee (KAIST), and Junmo Kim (KAIST)

Guided Point Contrastive Learning for Semi-Supervised Point Cloud Semantic Segmentation 6403
Li Jiang (The Chinese University of Hong Kong), Shaoshuai Shi (The Chinese University of Hong Kong), Zhuotao Tian (The Chinese University of Hong Kong), Xin Lai (The Chinese University of Hong Kong), Shu Liu (SmartMore), Chi-Wing Fu (The Chinese University of Hong Kong), and Jiaya Jia (Chinese University of Hong Kong)

Bootstrap Your Own Correspondences 6413
Mohamed El Banani (University of Michigan) and Justin Johnson (University of Michigan)

Learning With Noisy Labels for Robust Point Cloud Segmentation 6423
Shuquan Ye (City University of Hong Kong), Dongdong Chen (Microsoft Cloud AI), Songfang Han (UCSD), and Jing Liao (City University of Hong Kong)

Synchronization of Group-Labelled Multi-Graphs 6433
Andrea Porfiri Dal Cin (Politecnico di Milano), Luca Magri (Politecnico di Milano), Federica Arrigoni (University of Trento), Andrea Fusiello (UNIUD), and Giacomo Boracchi (Politecnico di Milano)

Progressive Correspondence Pruning by Consensus Learning 6444
Chen Zhao (EPFL), Yixiao Ge (The Chinese University of Hong Kong), Feng Zhu (University of Science and Technology of China), Rui Zhao (The Chinese University of Hong Kong), Hongsheng Li (Chinese University of Hong Kong), and Mathieu Salzmann (EPFL)

Generalized Shuffled Linear Regression 6454
Feiran Li (Osaka University), Kent Fujiwara (LINE Corporation), Fumio Okura (Osaka University), and Yasuyuki Matsushita (Osaka University)

Collaborative Optimization and Aggregation for Decentralized Domain Generalization and Adaptation 6464
Guile Wu (Queen Mary University of London) and Shaogang Gong (Queen Mary University of London)

JEM++: Improved Techniques for Training JEM 6474
Xiulong Yang (Georgia State University) and Shihao Ji (Georgia State University)

SA-ConvONet: Sign-Agnostic Optimization of Convolutional Occupancy Networks 6484
Jiapeng Tang (South China University of Technology), Jiabao Lei (South China University of Technology), Dan Xu (The Hong Kong University of Science and Technology), Feiying Ma (alibaba), Kui Jia (South China University of Technology), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)

Vis2Mesh: Efficient Mesh Reconstruction From Unstructured Point Clouds of Large Scenes With Learned Virtual View Visibility 6494
Shuang Song (The Ohio State University), Zhaopeng Cui (Zhejiang University), and Rongjun Qin (The Ohio State University)

Learning Signed Distance Field for Multi-View Surface Reconstruction 6505
Jingyang Zhang (The Hong Kong University of Science and Technology), Yao Yao (The Hong Kong University of Science and Technology), and Long Quan (Hong Kong University of Science and Technology)

Spatio-Temporal Self-Supervised Representation Learning for 3D Point Clouds 6515
Siyuan Huang (UCLA), Yichen Xie (Shanghai Jiao Tong University), Song-Chun Zhu (UCLA), and Yixin Zhu (UCLA)

SPatchGAN: A Statistical Feature Based Discriminator for Unsupervised Image-to-Image Translation 6526
Xuning Shao (Netease Games AI Lab) and Weidong Zhang (Netease Games AI Lab)

Bias Loss for Mobile Neural Networks 6536
Lusine Abrahamyan (Vrije Universiteit Brussel), Valentin Ziatchin (PicsArt), Yiming Chen (Vrije Universiteit Brussel), and Nikos Deligiannis (Vrije Universiteit Brussel)

Self-Knowledge Distillation With Progressive Refinement of Targets 6547
Kyungyul Kim (LG CNS), ByeongMoon Ji (LG CNS), Doyoung Yoon (LG CNS), and Sangheum Hwang (Seoul National University of Science and Technology)

A Lazy Approach to Long-Horizon Gradient-Based Meta-Learning 6557
Muhammad Abdullah Jamal (University of Central Florida), Liqiang Wang (University of Central Florida), and Boqing Gong (Google)

Manifold Matching via Deep Metric Learning for Generative Modeling 6567
Mengyu Dai (Microsoft Corporation) and Haibin Hang (University of Delaware, Department of mathematical sciences)

Paint Transformer: Feed Forward Neural Painting With Stroke Prediction 6578
Songhua Liu (NJU), Tianwei Lin (Baidu Inc), Dongliang He (Baidu), Fu Li (Baidu), Ruifeng Deng (Baidu), Xin Li (Baidu), Errui Ding (Baidu Inc.), and Hao Wang (Rutgers University)

Robust Trust Region for Weakly Supervised Segmentation 6588
Dmitrii Marin (University of Waterloo) and Yuri Boykov (University of Waterloo)

Semantic Perturbations With Normalizing Flows for Improved Generalization 6599
Oguz Kaan Yüksel (EPFL), Sebastian U. Stich (EPFL), Martin Jaggi (EPFL), and Tatjana Chavdarova (UC Berkeley and EPFL)

Domain Generalization via Gradient Surgery 6610
Lucas Mansilla (CONICET / Universidad Nacional del Litoral), Rodrigo Echeveste (CONICET / Universidad Nacional del Litoral), Diego H. Milone (CONICET / Universidad Nacional del Litoral), and Enzo Ferrante (CONICET / Universidad Nacional del Litoral)

PixelPyramids: Exact Inference Models From Lossless Image Pyramids 6619
Shweta Mahajan (TU Darmstadt) and Stefan Roth (TU Darmstadt)

AdaAttN: Revisit Attention Mechanism in Arbitrary Neural Style Transfer 6629
Songhua Liu (NJU), Tianwei Lin (Baidu Inc), Dongliang He (Baidu), Fu Li (Baidu), Meiling Wang (Baidu), Xin Li (Baidu), Zhengxing Sun (Nanjing University), Qian Li (College of Meteorology and Oceanography, National University of Defense Technology), and Errui Ding (Baidu Inc.)

Evolving Search Space for Neural Architecture Search 6639
Yuanzheng Ci (The University of Sydney), Chen Lin (University of Oxford), Ming Sun (SenseTime Group Limited), Boyu Chen (The University of Sydney), Hongwen Zhang (Institute of Automation, Chinese Academy of Sciences), and Wanli Ouyang (The University of Sydney)

Searching for Robustness: Loss Learning for Noisy Classification Tasks 6650
Boyan Gao (University of Edinburgh), Henry Gouk (University of Edinburgh), and Timothy M. Hospedales (Edinburgh University)

Extending Neural P-Frame Codecs for B-Frame Coding 6660
Reza Pourreza (Qualcomm) and Taco Cohen (Qualcomm AI Research)

Generative Layout Modeling Using Constraint Graphs 6670
Wamiq Para (KAUST), Paul Guerrero (Adobe Research), Tom Kelly (university of leeds), Leonidas J. Guibas (Stanford University), and Peter Wonka (KAUST)

When Do GANs Replicate? On the Choice of Dataset Size 6681
Qianli Feng (The Ohio State University, USA), Chenqi Guo (The Ohio State University, USA), Fabian Benitez-Quiroz (Ohio State University, USA), and Aleix M. Martinez (The Ohio State University, USA)

ReStyle: A Residual-Based StyleGAN Encoder via Iterative Refinement 6691
Yuval Alaluf (Tel Aviv University), Or Patashnik (Tel Aviv University), and Daniel Cohen-Or (Tel Aviv University)

Orthogonal Jacobian Regularization for Unsupervised Disentanglement in Image Generation 6701
Yuxiang Wei (Harbin Institute of Technology), Yupeng Shi (Harbin Institute of Technology), Xiao Liu (Tomorrow Advancing Life), Zhilong Ji (Tomorrow Advancing Life), Yuan Gao (TAL), Zhongqin Wu (Tomorrow Advancing Life), and Wangmeng Zuo (Harbin Institute of Technology, China)

Dual Contrastive Loss and Attention for GANs 6711
Ning Yu (University of Maryland and Max Planck Institute for Informatics), Guilin Liu (NVIDIA), Aysegul Dundar (Bilkent University), Andrew Tao (NVIDIA), Bryan Catanzaro (NVIDIA), Larry S. Davis (University of Maryland), and Mario Fritz (CISPA Helmholtz Center for Information Security)

F-Drop&Match: GANs With a Dead Zone in the High-Frequency Domain 6723
Shin'ya Yamaguchi (NTT) and Sekitoshi Kanai (NTT Software Innovation Center)

Harnessing the Conditioning Sensorium for Improved Image Translation 6732
Cooper Nederhood (University of Chicago), Nicholas Kolkin (Toyota Technological Institute at Chicago), Deqing Fu (University of Chicago), and Jason Salavon (University of Chicago)

Multi-Class Multi-Instance Count Conditioned Adversarial Image Generation 6742
Amrutha Saseendran (Bosch Center for Artificial Intelligence), Kathrin Skubch (Bosch Center for Artificial Intelligence), and Margret Keuper (University of Mannheim)

DeepCAD: A Deep Generative Network for Computer-Aided Design Models 6752
Rundi Wu (Columbia University), Chang Xiao (Columbia University), and Changxi Zheng (Columbia University)

Scaling-Up Disentanglement for Image Translation 6763
Aviv Gabbay (Hebrew University) and Yedid Hoshen (The Hebrew University of Jerusalem)

Online Multi-Granularity Distillation for GAN Compression 6773
Yuxi Ren (ByteDance Inc), Jie Wu (ByteDance Inc), Xuefeng Xiao (ByteDance Inc), and Jianchao Yang (Bytedance Inc.)

Toward a Visual Concept Vocabulary for GAN Latent Space 6784
Sarah Schwettmann (MIT), Evan Hernandez (MIT), David Bau (MIT), Samuel Klein (Knowledge Futures Group), Jacob Andreas (MIT), and Antonio Torralba (MIT)

Image and Video Synthesis

Where2Act: From Pixels to Actions for Articulated 3D Objects 6793
Kaichun Mo (Stanford), Leonidas J. Guibas (Stanford University), Mustafa Mukadam (Facebook AI Research), Abhinav Gupta (CMU/FAIR), and Shubham Tulsiani (Facebook AI Research)

Multiscale Vision Transformers 6804
Haoqi Fan (Facebook AI Research), Bo Xiong (Facebook AI Research), Karttikeya Mangalam (UC Berkeley), Yanghao Li (Facebook AI Research), Zhicheng Yan (Facebook AI), Jitendra Malik (University of California at Berkeley), and Christoph Feichtenhofer (Facebook AI Research)

ViViT: A Video Vision Transformer 6816
Anurag Arnab (Google), Mostafa Dehghani (Google Brain), Georg Heigold, Chen Sun (Google), Mario Lučić (Google Brain), and Cordelia Schmid (Google)

End-to-End Dense Video Captioning With Parallel Decoding 6827
Teng Wang (The University of Hong Kong), Ruimao Zhang (The Chinese University of Hong Kong, Shenzhen), Zhichao Lu (Southern University of Science and Technology), Feng Zheng (SUSTech), Ran Cheng (Southern University of Science and Technology), and Ping Luo (The University of Hong Kong)

Learning To Cut by Watching Movies 6838
Alejandro Pardo (KAUST), Fabian Caba (Adobe Research), Juan Léon Alcázar (KAUST), Ali K. Thabet (Facebook), and Bernard Ghanem (KAUST)

Field of Junctions: Extracting Boundary Structure at Low SNR 6849
Dor Verbin (Harvard University) and Todd Zickler (Harvard University)

RINDNet: Edge Detection for Discontinuity in Reflectance, Illumination, Normal and Depth 6859
Mengyang Pu (Beijing Jiaotong University), Yaping Huang (Beijing Jiaotong University), Qingji Guan (Beijing Jiaotong University), and Haibin Ling (Stony Brook University)

Local Temperature Scaling for Probability Calibration 6869
Zhipeng Ding (Department of Computer Science, University of North Carolina at Chapel Hill), Xu Han (UNC Chapel Hill), Peirong Liu (UNC Chapel Hill), and Marc Niethammer (UNC)

An Elastica Geodesic Approach With Convexity Shape Prior 6880
Da Chen (Shandong Artificial Intelligence Institute, Qilu University of Technology, (Shandong Academy of Sciences)), Laurent D. Cohen (University Paris Dauphine), Jean-Marie Mirebeau (University Paris-Sud, University Paris-Saclay), and Xue-Cheng Tai (Hong Kong Baptist University)

Instances As Queries 6890
Yuxin Fang (Huazhong University of Science and Technology), Shusheng Yang (Huazhong University of Science and Technology), Xinggang Wang (Huazhong University of Science and Technology), Yu Li (Tencent), Chen Fang (Tencent), Ying Shan (Tencent), Bin Feng (Huazhong University of Science and Technology), and Wenyu Liu (Huazhong University of Science and Technology)

Seminar Learning for Click-Level Weakly Supervised Semantic Segmentation 6900
Hongjun Chen (Southern University of Science and Technology), Jinbao Wang (Southern University of Science and Technology), Hong Cai Chen (Southern University of Science and Technology), Xiantong Zhen (University of Amsterdam), Feng Zheng (SUSTech), Rongrong Ji (Xiamen University, China), and Ling Shao (Inception Institute of Artificial Intelligence)

Re-Distributing Biased Pseudo Labels for Semi-Supervised Semantic Segmentation: A Baseline Investigation 6910
Ruifei He (Zhejiang University), Jihan Yang (The University of Hong Kong), and Xiaojuan Qi (The University of Hong Kong)

Hypercorrelation Squeeze for Few-Shot Segmenation 6921
Juhong Min (POSTECH), Dahyun Kang (POSTECH), and Minsu Cho (POSTECH)

Self-Regulation for Semantic Segmentation 6933
Dong Zhang (Nanjing University of Science and Technology), Hanwang Zhang (Nanyang Technological University), Jinhui Tang (Nanjing University of Science and Technology), Xian-Sheng Hua (Damo Academy, Alibaba Group), and Qianru Sun (Singapore Management University)

Pseudo-Mask Matters in Weakly-Supervised Semantic Segmentation 6944
Yi Li (SenseTime Research), Zhanghui Kuang (Sensetime Ltd.), Liyang Liu (Tsinghua University), Yimin Chen (sensetime), and Wayne Zhang (SenseTime Research)

Prototypical Matching and Open Set Rejection for Zero-Shot Semantic Segmentation 6954
Hui Zhang (Nanyang Technological University) and Henghui Ding (Nanyang Technological University)

Leveraging Auxiliary Tasks With Affinity Learning for Weakly Supervised Semantic Segmentation 6964
Lian Xu (The University of Western Australia), Wanli Ouyang (The University of Sydney), Mohammed Bennamoun (University of Western Australia), Farid Boussaid (University of Western Australia), Ferdous Sohel (Murdoch University), and Dan Xu (The Hong Kong University of Science and Technology)

Unlocking the Potential of Ordinary Classifier: Class-Specific Adversarial Erasing Framework for Weakly Supervised Semantic Segmentation 6974
Hyeokjun Kweon (KAIST), Sung-Hoon Yoon (KAIST), Hyeonseong Kim (KAIST), Daehee Park (KAIST), and Kuk-Jin Yoon (KAIST)

Context Decoupling Augmentation for Weakly Supervised Semantic Segmentation 6984
Yukun Su (South China University of Technology), Ruizhou Sun (South China University of Technology), Guosheng Lin (Nanyang Technological University), and Qingyao Wu (South China University of Technology)

The Surprising Impact of Mask-Head Architecture on Novel Class Segmentation 6995
Vighnesh Birodkar (Google), Zhichao Lu (Google), Siyang Li (Google AI Perception), Vivek Rathod (Google), and Jonathan Huang (Google)

RECALL: Replay-Based Continual Learning in Semantic Segmentation 7006
Andrea Maracani (Istituto Italiano di Tecnologia), Umberto Michieli (University of Padova), Marco Toldo (University of Padova), and Pietro Zanuttigh (University of Padova)

C3-SemiSeg: Contrastive Semi-Supervised Segmentation via Cross-Set Learning and Dynamic Class-Balancing 7016
Yanning Zhou (The Chinese University of HongKong), Hang Xu (Huawei Noah's Ark Lab), Wei Zhang (Noah's Ark Lab, Huawei Technologies), Bin Gao (Noah's Ark Lab, Huawei Technologies), and Pheng-Ann Heng (The Chinese Univsersity of Hong Kong)

Generalize Then Adapt: Source-Free Domain Adaptive Semantic Segmentation 7026
Jogendra Nath Kundu (Indian Institute of Science), Akshay Kulkarni (Indian Institute of Science), Amit Singh (Indian Institute of Science), Varun Jampani (Google), and R. Venkatesh Babu (Indian Institute of Science)

Calibrated Adversarial Refinement for Stochastic Semantic Segmentation 7037
Elias Kassapis (University of Amsterdam; TomTom), Georgi Dikov (Qualcomm AI Research), Deepak K. Gupta (University of Amsterdam), and Cedric Nugteren (TomTom)

Self-Mutating Network for Domain Adaptive Segmentation in Aerial Images 7048
Kyungsu Lee (DGIST), Haeyun Lee (DGIST), and Jae Youn Hwang (DGIST)

AINet: Association Implantation for Superpixel Segmentation 7058
Yaxiong Wang (Xi'an Jiaotong University), Yunchao Wei (UTS), Xueming Qian (Xi'an Jiaotong University), Li Zhu (Xi'an Jiaotong University), and Yi Yang (UTS)

ShapeConv: Shape-Aware Convolutional Layer for Indoor RGB-D Semantic Segmentation 7068
Jinming Cao (Shandong University), Hanchao Leng (Shandong University), Dani Lischinski (The Hebrew University of Jerusalem), Daniel Cohen-Or (Tel Aviv University), Changhe Tu (Shandong University), and Yangyan Li (Alibaba Group)

Persistent Homology Based Graph Convolution Network for Fine-Grained 3D Shape Segmentation..... 7078
Chi-Chong Wong (University of Macau) and Chi-Man Vong (University of Macau)

Sparse-to-Dense Feature Matching: Intra and Inter Domain Cross-Modal Learning in Domain Adaptation for 3D Semantic Segmentation 7088
Duo Peng (Sichuan University), Yinjie Lei (Sichuan University), Wen Li (University of Electronic Science and Technology of China), Pingping Zhang (Dalian University of Technology), and Yulan Guo (National University of Defense Technology)

TempNet: Online Semantic Segmentation on Large-Scale Point Cloud Series 7098
Yunsong Zhou (Shanghai Jiao Tong University), Hongzi Zhu (Shanghai Jiao Tong University), Chunqin Li (Shanghai Jiao Tong University), Tiankai Cui (Shanghai Jiao Tong University), Shan Chang (Donghua University), and Minyi Guo (Shanghai Jiaotong University)

How Shift Equivariance Impacts Metric Learning for Instance Segmentation 7108
Josef Lorenz Rumberger (Max Delbrück Center for Molecular Medicine), Xiaoyan Yu (BIH/MDC), Peter Hirsch (BIH/MDC), Melanie Dohmen (BIH/MDC / Charite - Universitaetsmedizin Berlin), Vanessa Emanuela Guarino (BIH/MDC), Ashkan Mokarian (BIH/MDC), Lisa Mais (BIH/MDC), Jan Funke (HHMI Janelia Research Campus), and Dagmar Kainmüller (BIH/MDC)

Specialize and Fuse: Pyramidal Output Representation for Semantic Segmentation 7117
Chi-Wei Hsiao (National Tsing Hua University), Cheng Sun (National Tsing Hua University), Hwann-Tzong Chen (National Tsing Hua University), and Min Sun (NTHU)

Joint Topology-Preserving and Feature-Refinement Network for Curvilinear Structure Segmentation 7127
Mingfei Cheng (Beijing University of Posts and Telecommunications), Kaili Zhao (Beijing University of Posts and Telecommunications), Xuhong Guo (DOCOMO Beijing Labs), Yajing Xu (Beijing University of Posts and Telecommunications), and Jun Guo (Beijing University of Posts and Telecommunications)

SOTR: Segmenting Objects With Transformers 7137
Ruohao Guo (China Agricultural University), Dantong Niu (University of California, Berkeley), Liao Qu (China Agricultural University), and Zhenbo Li (China Agricultural University)

Cascade Image Matting With Deformable Graph Refinement 7147
Zijian Yu (Tsinghua University), Xuhui Li (Tsinghua University), Huijuan Huang (Kuaishou Technology), Wen Zheng (Kuaishou Technology), and Li Chen (Tsinghua University)

Self-Supervised Video Object Segmentation by Motion Grouping 7157
Charig Yang (University of Oxford), Hala Lamdouar (University of Oxford), Erika Lu (University of Oxford), Andrew Zisserman (University of Oxford), and Weidi Xie (University of Oxford)

ISNet: Integrate Image-Level and Semantic-Level Context for Semantic Segmentation 7169
Zhenchao Jin (University of Science and Technology of China), Bin Liu (University of Science and Technology of China), Qi Chu (University of Science and Technology of China), and Nenghai Yu (University of Science and Technology of China)

Cross-Sentence Temporal and Semantic Relations in Video Activity Localisation 7179
Jiabo Huang (Queen Mary University of London), Yang Liu (Peking University), Shaogang Gong (Queen Mary University of London), and Hailin Jin (Adobe Research)

Multiview Pseudo-Labeling for Semi-Supervised Learning From Video 7189
Bo Xiong (Facebook AI Research), Haoqi Fan (Facebook AI Research), Kristen Grauman (Facebook AI Research & UT Austin), and Christoph Feichtenhofer (Facebook AI Research)

Boundary-Sensitive Pre-Training for Temporal Localization in Videos 7200
Mengmeng Xu (KAUST), Juan-Manuel Pérez-Rúa (Samsung AI Centre), Victor Escorcia (Samsung AI Center), Brais Martínez (Samsung AI Research), Xiatian Zhu (Samsung), Li Zhang (Fudan University), Bernard Ghanem (KAUST), and Tao Xiang (University of Surrey)

Mining Contextual Information Beyond Image for Semantic Segmentation 7211
Zhenchao Jin (University of Science and Technology of China), Tao Gong (University of Science and Technology of China), Dongdong Yu (ByteDance), Qi Chu (University of Science and Technology of China), Jian Wang (ByteDance), Changhu Wang (ByteDance.Inc), and Jie Shao (Bytedance)

Complementary Patch for Weakly Supervised Semantic Segmentation 7222
Fei Zhang (Shanghai Jiao Tong University), Chaochen Gu (Shanghai Jiao Tong University), Chenyue Zhang (Shanghai Jiao Tong University), and Yuchao Dai (Northwestern Polytechnical University)

From Contexts to Locality: Ultra-High Resolution Image Segmentation via Locality-Aware Contextual Correlation 7232
Qi Li (Fuzhou University), Weixiang Yang (Fuzhou University), Wenxi Liu (Fuzhou University), Yuanlong Yu (Fuzhou University), and Shengfeng He (South China University of Technology)

Segmenter: Transformer for Semantic Segmentation 7242
Robin Strudel (INRIA Paris), Ricardo Garcia (INRIA Paris), Ivan Laptev (Inria/Google), and Cordelia Schmid (INRIA Paris)

Pixel Contrastive-Consistent Semi-Supervised Semantic Segmentation 7253
Yuanyi Zhong (University of Illinois at Urbana-Champaign), Bodi Yuan (X, The Moonshot Factory), Hong Wu (Google), Zhiqiang Yuan (Google), Jian Peng (University of Illinois at Urbana-Champaign), and Yu-Xiong Wang (University of Illinois at Urbana-Champaign)

ECS-Net: Improving Weakly Supervised Semantic Segmentation by Using Connections Between Class Activation Maps 7263
Kunyang Sun (Southeast University), Haoqing Shi (southeast university), Zhengming Zhang (Southeast University), and Yongming Huang (Southeast University)

Few-Shot Semantic Segmentation With Cyclic Memory Network 7273
Guo-Sen Xie (Inception Institute of Artificial Intelligence), Huan Xiong (Northeastern University), Jie Liu (Mohamed bin Zayed University of Artificial Intelligence, (MBZUAI)), Yazhou Yao (Nanjing University of Science and Technology), and Ling Shao (Inception Institute of Artificial Intelligence)

Exploring Cross-Image Pixel Contrast for Semantic Segmentation 7283
Wenguan Wang (Eidgenössische Technische Hochschule Zürich), Tianfei Zhou (ETH Zurich), Fisher Yu (ETH Zurich), Jifeng Dai (SenseTime), Ender Konukoglu (ETH Zurich), and Luc Van Gool (ETH Zurich)

Real-Time Instance Segmentation With Discriminative Orientation Maps 7294
Wentao Du (Zhejiang University), Zhiyu Xiang (Zhejiang University), Shuya Chen (Zhejiang University), Chengyu Qiao (Zhejiang University), Yiman Chen (Zhejiang University), and Tingming Bai (Zhejiang University)

Unsupervised Segmentation Incorporating Shape Prior via Generative Adversarial Networks 7304
Dahye Kim (Chung-Ang University) and Byung-Woo Hong (Chung-Ang University)

Unsupervised Point Cloud Object Co-Segmentation by Co-Contrastive Learning and Mutual Attention Sampling 7315
Cheng-Kun Yang (National Taiwan University), Yung-Yu Chuang (National Taiwan University), and Yen-Yu Lin (National Yang Ming Chiao Tung University)

Conditional Diffusion for Interactive Segmentation 7325
Xi Chen (Alibaba Group), Zhiyan Zhao (alibaba), Feiwu Yu (Alibaba Group), Yilei Zhang (Alibaba), and Manni Duan (alibaba-inc.com)

Weakly Supervised Temporal Anomaly Segmentation With Dynamic Time Warping 7335
Dongha Lee (POSTECH), Sehun Yu (POSTECH), Hyunjun Ju (POSTECH), and Hwanjo Yu (POSTECH)

Predictive Feature Learning for Future Segmentation Prediction 7345
Zihang Lin (Sun Yat-sen University), Jiangxin Sun (Sun Yat-sen University), Jian-Fang Hu (Sun Yat-sen University), Qizhi Yu (Zhejiang Laboratory), Jian-Huang Lai (Sun Yat-sen University), and Wei-Shi Zheng (Sun Yat-sen University, China)

Dynamic Network Quantization for Efficient Video Inference 7355
Ximeng Sun (Boston University), Rameswar Panda (MIT-IBM Watson AI Lab, IBM Research), Chun-Fu Richard Chen (MIT-IBM Watson AI Lab, IBM Research AI), Aude Oliva (MIT), Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research), and Kate Saenko (Boston University)

Graph-BAS3Net: Boundary-Aware Semi-Supervised Segmentation Network With Bilateral Graph Convolution 7366
Huimin Huang (Zhejiang University), Lanfen Lin (Zhejiang University), Yue Zhang (Zhejiang University), Yingying Xu (Zhejiang University), Jing Zheng (The first Affiliated hospital, Collegue of Medicine, Zhejiang university), XiongWei Mao (The Radiology Department, Zhejiang University Hospital), Xiaohan Qian (The First Affiliated Hospitol of Zhejiang University School of Medicine), Zhiyi Peng (Department of Radiology, The First Affiliated HospitalZhejiang University School of Medicine), Jianying Zhou (The First Affiliated Hospital of Zhejiang University School of Medicine), Yen-Wei Chen (Ritsumeikan University), and Ruofeng Tong (Zhejiang University)

A Weakly Supervised Amodal Segmenter With Boundary Uncertainty Estimation 7376
Khoi Nguyen (Oregon State University) and Sinisa Todorovic (Oregon State U)

Weakly Supervised Segmentation of Small Buildings With Point Labels 7386
Jae-Hun Lee (Korea University), ChanYoung Kim (Korea University), and Sanghoon Sull (Korea University)

Scribble-Supervised Semantic Segmentation by Uncertainty Reduction on Neural Representation and Self-Supervision on Neural Eigenspace 7396
Zhiyi Pan (Shandong University), Peng Jiang (Shandong University), Yunhai Wang (Shandong University), Changhe Tu (Shandong University), and Anthony G. Cohn (University of Leeds)

Scaling Semantic Segmentation Beyond 1K Classes on a Single GPU 7406
Shipra Jain (KTH Royal Institute of Technology), Danda Pani Paudel (ETH Zürich), Martin Danelljan (ETH Zurich), and Luc Van Gool (ETH Zurich)

Differentiable Convolution Search for Point Cloud Processing 7417
Xing Nie (NLPR, CASIA), Yongcheng Liu (Institute of Automation, Chinese Academy of Sciences), Shaohong Chen (Xidian University), Jianlong Chang (Huawei Cloud), Chunlei Huo (National Lab. Patter Recognition, Institute of Automation, Chinese Academy of Sciences), Gaofeng Meng (Chinese Academy of Sciences), Qi Tian (Huawei Cloud & AI), Weiming Hu (Institute of AutomationChinese Academy of Sciences), and Chunhong Pan (Institute of Automation, Chinese Academy of Sciences)

DRINet: A Dual-Representation Iterative Learning Network for Point Cloud Segmentation 7427
Maosheng Ye (HKUST), Shuangjie Xu (Deeproute.ai), Tongyi Cao (Deeproute.ai), and Qifeng Chen (HKUST)

CPFN: Cascaded Primitive Fitting Networks for High-Resolution Point Clouds 7437
Eric-Tuan Lê (University College London), Minhyuk Sung (KAIST), Duygu Ceylan (Adobe Research), Radomir Mech (Adobe Systems Incorporated), Tamy Boubekeur (Adobe Research), and Niloy J. Mitra (University College London)

Interpolation-Aware Padding for 3D Sparse Convolutional Neural Networks 7447
Yu-Qi Yang (Tsinghua University), Peng-Shuai Wang (Microsoft Research Asia), and Yang Liu (MSR, Beijing)

Spatio-Temporal Dynamic Inference Network for Group Activity Recognition 7456
Hangjie Yuan (Zhejiang University), Dong Ni (Zhejiang University), and Mang Wang (Alibaba Group)

Dynamic Divide-and-Conquer Adversarial Training for Robust Semantic Segmentation 7466
Xiaogang Xu (The Chinese University of Hong Kong), Hengshuang Zhao (University of Oxford), and Jiaya Jia (Chinese University of Hong Kong)

Towards Robustness of Deep Neural Networks via Regularization 7476
Yao Li (University of North Carolina at Chapel Hill), Martin Renqiang Min (NEC Labs America-Princeton), Thomas Lee (UC Davis), Wenchao Yu (UCLA), Erik Kruus (NEC Labs), Wei Wang (UCLA), and Cho-Jui Hsieh (UCLA)

AdvDrop: Adversarial Attack to DNNs by Dropping Information 7486
Ranjie Duan (Swinburne University of Technology), Yuefeng Chen (Alibaba Group), Dantong Niu (University of California, Berkeley), Yun Yang (Swinburne University of Technology), A. K. Qin (Swinburne University of Technology), and Yuan He (Alibaba Group)

S3VAADA: Submodular Subset Selection for Virtual Adversarial Active Domain Adaptation 7496
Harsh Rangwani (Indian Institute of Science), Arihant Jain (Indian Institute of Science), Sumukh K Aithal (PES University), and R. Venkatesh Babu (Indian Institute of Science)

Can Shape Structure Features Improve Model Robustness Under Diverse Adversarial Settings? .. 7506
Mingjie Sun (Carnegie Mellon University), Zichao Li (University of California, San Diego), Chaowei Xiao (University of Michigan, Ann Arbor), Haonan Qiu (Nanyang Technological University), Bhavya Kailkhura (Lawrence Livermore National Laboratory), Mingyan Liu (University of Michigan, Ann Arbor), and Bo Li (UIUC)

Self-Supervised Vessel Segmentation via Adversarial Learning 7516
Yuxin Ma (Shanghai Jiao Tong University), Yang Hua (Queen's University Belfast), Hanming Deng (Shanghai Jiao Tong University), Tao Song (Shanghai Jiao Tong University), Hao Wang (Louisiana State University), Zhengui Xue (Shanghai Jiao Tong University), Heng Cao (Shanghai General Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200030, China), Ruhui Ma (Shanghai Jiao Tong University), and Haibing Guan (Shanghai Jiao Tong University)

MultiSiam: Self-Supervised Multi-Instance Siamese Representation Learning for Autonomous Driving 7526
Kai Chen (HKUST), 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 (HKUST)

Tripartite Information Mining and Integration for Image Matting 7535
Yuhao Liu (Dalian University of Technology), Jiake Xie (LibAI lab), Xiao Shi (Agricultural Information Institute of CAAS), Yu Qiao (Dalian University of Technology), Yujie Huang (Fudan University), Yong Tang (PicUP.Ai), and Xin Yang (Dalian University of Technology)

OadTR: Online Action Detection With Transformers 7545
Xiang Wang (Huazhong University of Science and Technology), Shiwei Zhang (DAMO Academy, Alibaba Group), Zhiwu Qing (Huazhong University of Science and Technology), Yuanjie Shao (Huazhong University of Science and Technology), Zhengrong Zuo (Huazhong University of Science and Technology), Changxin Gao (Huazhong University of Science and Technology), and Nong Sang (Huazhong University of Science and Technology)

AdaMML: Adaptive Multi-Modal Learning for Efficient Video Recognition 7556
Rameswar Panda (MIT-IBM Watson AI Lab, IBM Research), Chun-Fu Richard Chen (MIT-IBM Watson AI Lab, IBM Research AI), Quanfu Fan (IBM Research), Ximeng Sun (Boston University), Kate Saenko (Boston University), Aude Oliva (MIT), and Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research)

Q-Match: Iterative Shape Matching via Quantum Annealing 7566
Marcel Seelbach Benkner (University of Siegen), Zorah Lähner (University of Siegen), Vladislav Golyanik (MPI for Informatics), Christof Wunderlich (University of Siegen), Christian Theobalt (MPI Informatik), and Michael Moeller (University of Siegen)

A Backdoor Attack Against 3D Point Cloud Classifiers 7577
Zhen Xiang (Pennsylvania State University), David J. Miller (Pennsylvania State University), Siheng Chen (Shanghai Jiao Tong University), Xi Li (The Pennsylvania State University), and George Kesidis (Penn State University)

Robustness Certification for Point Cloud Models 7588
Tobias Lorenz (CISPA Helmholtz Center for Information Security), Anian Ruoss (ETH Zurich), Mislav Balunović (ETH Zurich), Gagandeep Singh (VMware Research and UIUC), and Martin Vechev (ETH Zurich)

DRIVE: Deep Reinforced Accident Anticipation With Visual Explanation 7599
Wentao Bao (Rochester Institute of Technology), Qi Yu (Rochester Institute of Technology), and Yu Kong (Rochester Institute of Technology)

Where Are You Heading? Dynamic Trajectory Prediction With Expert Goal Examples 7609
He Zhao (York University) and Richard P. Wildes (York University)

Feature Importance-Aware Transferable Adversarial Attacks 7619
Zhibo Wang (Zhejiang University), Hengchang Guo (Wuhan University), Zhifei Zhang (Adobe Research), Wenxin Liu (Wuhan University), Zhan Qin (Zhejiang University), and Kui Ren (Zhejiang University)

TkML-AP: Adversarial Attacks to Top-k Multi-Label Learning 7629
Shu Hu (University at Buffalo), Lipeng Ke (University at Buffalo), Xin Wang (Keya Medical), and Siwei Lyu (University at Buffalo)

AGKD-BML: Defense Against Adversarial Attack by Attention Guided Knowledge Distillation and Bi-Directional Metric Learning 7638
Hong Wang (Stony Brook University), Yuefan Deng (Stony Brook University), Shinjae Yoo (Brookhaven National Laboratory), Haibin Ling (Stony Brook University), and Yuewei Lin (Brookhaven National Laboratory)

Just One Moment: Structural Vulnerability of Deep Action Recognition Against One Frame Attack 7648
Jaehui Hwang (Yonsei University), Jun-Hyuk Kim (Yonsei University), Jun-Ho Choi (Yonsei University), and Jong-Seok Lee (Yonsei University, Korea)

Sample Efficient Detection and Classification of Adversarial Attacks via Self-Supervised Embeddings 7657
Mazda Moayeri (University of Maryland) and Soheil Feizi (University of Maryland)

Adversarial Example Detection Using Latent Neighborhood Graph 7667
Ahmed Abusnaina (University of Central Florida), Yuhang Wu (Visa Inc.), Sunpreet Arora (Visa Research), Yizhen Wang (Visa Research), Fei Wang (Visa Research), Hao Yang (Visa Research), and David Mohaisen (University of Central Florida)

Parallel Rectangle Flip Attack: A Query-Based Black-Box Attack Against Object Detection 7677
Siyuan Liang (Chinese Academy of Sciences), Baoyuan Wu (The Chinese University of Hong Kong, Shenzhen; Shenzhen Research Institute of Big Data), Yanbo Fan (Tencent AI Lab), Xingxing Wei (Beihang University), and Xiaochun Cao (Chinese Academy of Sciences)

On Generating Transferable Targeted Perturbations 7688
Muzammal Naseer (Australian National University, (ANU)), Salman Khan (Australian National University, (ANU)), Munawar Hayat (Monash University), Fahad Shahbaz Khan (MBZUAI), and Fatih Porikli (Australian National University)

ProFlip: Targeted Trojan Attack With Progressive Bit Flips 7698
Huili Chen (UC San Diego), Cheng Fu (University of California, San Diego), Jishen Zhao (University of California, San Diego), and Farinaz Koushanfar (UC San Diego)

Towards Understanding the Generative Capability of Adversarially Robust Classifiers 7708
Yao Zhu (Zhejiang University), Jiacheng Ma (The University of Hong Kong), Jiacheng Sun (Huawei Noah's Ark Lab), Zewei Chen (Huawei Noah's Ark Lab), Rongxin Jiang (Zhejiang University), Yaowu Chen (Zhejiang University), and Zhenguo Li (Huawei Noah's Ark Lab)

Augmented Lagrangian Adversarial Attacks 7718
Jérôme Rony (ÉTS Montréal), Eric Granger (ETS Montreal), Marco Pedersoli (École de technologie supérieure), and Ismail Ben Ayed (ETS Montreal)

Meta Gradient Adversarial Attack 7728
Zheng Yuan (Institute of Computing Technology, Chinese Academy of Sciences), Jie Zhang (ICT, CAS), Yunpei Jia (ICT), Chuanqi Tan (tencent), Tao Xue (tencent), and Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences)

Reliably Fast Adversarial Training via Latent Adversarial Perturbation 7738
Geon Yeong Park (KAIST) and Sang Wan Lee (KAIST)

Adversarial Attacks on Multi-Agent Communication 7748
James Tu (University of Toronto), Tsunhsuan Wang (Massachusetts Institute of Technology), Jingkang Wang (University of Toronto), Sivabalan Manivasagam (University of Toronto), Mengye Ren (University of Toronto), and Raquel Urtasun (Uber ATG)

Consistency-Sensitivity Guided Ensemble Black-Box Adversarial Attacks in Low-Dimensional Spaces 7758
Jianhe Yuan (University of Missouri-Columbia) and Zhihai He (University of Missouri Columbia)

Meta-Attack: Class-Agnostic and Model-Agnostic Physical Adversarial Attack 7767
Weiwei Feng (University of Science and Technology of China), Baoyuan Wu (The Chinese University of Hong Kong, Shenzhen; Shenzhen Research Institute of Big Data), Tianzhu Zhang (University of Science and Technology of China), Yong Zhang (Tencent AI Lab), and Yongdong Zhang (University of Science and Technology of China)

Minimal Adversarial Examples for Deep Learning on 3D Point Clouds 7777
Jaeyeon Kim (HKUST), Binh-Son Hua (VinAI Research), Thanh Nguyen (Deakin University, Australia), and Sai-Kit Yeung (Hong Kong University of Science and Technology)

Relating Adversarially Robust Generalization to Flat Minima 7787
David Stutz (Max Planck Institute for Informatics), Matthias Hein (University of Tübingen), and Bernt Schiele (MPI Informatics)

Batch Normalization Increases Adversarial Vulnerability and Decreases Adversarial Transferability: A Non-Robust Feature Perspective 7798
Philipp Benz (KAIST), Chaoning Zhang (KAIST), and In So Kweon (KAIST)

Integer-Arithmetic-Only Certified Robustness for Quantized Neural Networks 7808
Haowen Lin (University of Southern California), Jian Lou (Emory University), Li Xiong (Emory University), and Cyrus Shahabi (University of Southern California)

On the Robustness of Vision Transformers to Adversarial Examples 7818
Kaleel Mahmood (University of Connecticut), Rigel Mahmood (University of Connecticut), and Marten van Dijk (CWI, Amsterdam)

Naturalistic Physical Adversarial Patch for Object Detectors 7828
Yu-Chih-Tuan Hu (National Taiwan University of Science and Technology), Bo-Han Kung (Academia Sinica), Daniel Stanley Tan (Academia Sinica), Jun-Cheng Chen (Academia Sinica), Kai-Lung Hua (NTUST), and Wen-Huang Cheng (National Chiao Tung University)

Exploiting Multi-Object Relationships for Detecting Adversarial Attacks in Complex Scenes 7838
Mingjun Yin (University of California, Riverside), Shasha Li (University of California, Riverside), Zikui Cai (University of California, Riverside), Chengyu Song (University of California, Riverside), M. Salman Asif (University of California, Riverside), Amit K. Roy-Chowdhury (University of California, Riverside), and Srikanth V. Krishnamurthy (University of California, Riverside)

Data-Free Universal Adversarial Perturbation and Black-Box Attack 7848
Chaoning Zhang (KAIST), Philipp Benz (KAIST), Adil Karjauv (KAIST), and In So Kweon (KAIST)

Removing Adversarial Noise in Class Activation Feature Space 7858
Dawei Zhou (Xidian University), Nannan Wang (Xidian University), Chunlei Peng (Xidian University), Xinbo Gao (Chongqing University of Posts and Telecommunications), Xiaoyu Wang (The Chinese University of Hong Kong, (Shenzhen)), Jun Yu (University of Science and Technology of China), and Tongliang Liu (The University of Sydney)

End-to-End Unsupervised Document Image Blind Denoising 7868
Mehrdad J. Gangeh (Ernst & Young), Marcin Plata (Ernst & Young), Hamid R. Motahari Nezhad (Macquarie University), and Nigel P Duffy (EY)

Fooling LiDAR Perception via Adversarial Trajectory Perturbation 7878
Yiming Li (New York University), Congcong Wen (University of Chinese Academy of Sciences), Felix Juefei-Xu (Alibaba Group, USA), and Chen Feng (New York University)

Sensor-Guided Optical Flow 7888
Matteo Poggi (University of Bologna), Filippo Aleotti (University of Bologna), and Stefano Mattoccia (University of Bologna)

Group-Aware Contrastive Regression for Action Quality Assessment 7899
Xumin Yu (Tsinghua University), Yongming Rao (Tsinghua University), Wenliang Zhao (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)

Online-Trained Upsampler for Deep Low Complexity Video Compression 7909
Jan P. Klopp (National Taiwan University), Keng-Chi Liu (Taiwan AI Labs), Shao-Yi Chien (National Taiwan University), and Liang-Gee Chen (National Taiwan University)

Contrast and Order Representations for Video Self-Supervised Learning 7919
Kai Hu (Carnegie Mellon University), Jie Shao (Fudan University), Yuan Liu (The University of Hong Kong), Bhiksha Raj (Carnegie Mellon University), Marios Savvides (Carnegie Mellon University), and Zhiqiang Shen (Carnegie Mellon University)

Temporal Cue Guided Video Highlight Detection With Low-Rank Audio-Visual Fusion 7930
Qinghao Ye (University of California, San Diego), Xiyue Shen (East China Normal University), Yuan Gao (University of Oxford), Zirui Wang (Hangzhou Dianzi University), Qi Bi (University of Amsterdam), Ping Li (Hangzhou Dianzi University), and Guang Yang (Imperial College London)

VideoLT: Large-Scale Long-Tailed Video Recognition 7940
Xing Zhang (Fudan University), Zuxuan Wu (UMD), Zejia Weng (Fudan University), Huazhu Fu (IHPC, ASTAR), Jingjing Chen (Fudan University), Yu-Gang Jiang (Fudan University), and Larry S. Davis (University of Maryland)

Cross-Category Video Highlight Detection via Set-Based Learning 7950
Minghao Xu (Shanghai Jiaotong University), Hang Wang (Shanghai Jiao Tong University), Bingbing Ni (Shanghai Jiao Tong University), Riheng Zhu (ByteDance.Inc), Zhenbang Sun (Bytedance), and Changhu Wang (ByteDance.Inc)

PR-Net: Preference Reasoning for Personalized Video Highlight Detection 7960
Runnan Chen (The University of Hong Kong), Penghao Zhou (Tencent Youtu Lab), Wenzhe Wang (Zhejiang University), Nenglun Chen (The University of Hong Kong), Pai Peng (Tencent Youtu Lab), Xing Sun (Tencent), and Wenping Wang (The University of Hong Kong)

Enhancing Self-Supervised Video Representation Learning via Multi-Level Feature Optimization 7970
Rui Qian (Shanghai Jiao Tong University), Yuxi Li (Tencent), Huabin Liu (Shanghai Jiao Tong University), John See (Heriot-Watt University Malaysia), Shuangrui Ding (Shanghai Jiao Tong University), Xian Liu (Zhejiang University), Dian Li (Tencent PCG), and Weiyao Lin (Shanghai Jiao Tong university)

Foreground-Action Consistency Network for Weakly Supervised Temporal Action Localization 7982
Linjiang Huang (The Chinese University of Hong Kong), Liang Wang (NLPR, China), and Hongsheng Li (Chinese University of Hong Kong)

Multimodal Clustering Networks for Self-Supervised Learning From Unlabeled Videos 7992
Brian Chen (Columbia University), Andrew Rouditchenko (MIT CSAIL), Kevin Duarte (University of Central Florida), Hilde Kuehne (University of Frankfurt), Samuel Thomas (IBM Research AI), Angie Boggust (MIT CSAIL), Rameswar Panda (MIT-IBM Watson AI Lab, IBM Research), Brian Kingsbury (IBM Research), Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research), David Harwath (MIT CSAIL), James Glass (Massachusetts Institute of Technology), Michael Picheny (NYU), and Shih-Fu Chang (Columbia University)

Temporal Action Detection With Multi-Level Supervision 8002
Baifeng Shi (UC Berkeley), Qi Dai (Microsoft Research), Judy Hoffman (Georgia Tech), Kate Saenko (Boston University), Trevor Darrell (UC Berkeley), and Huijuan Xu (Pennsylvania State University)

Searching for Two-Stream Models in Multivariate Space for Video Recognition 8013
Xinyu Gong (University of Texas at Austin), Heng Wang (Facebook AI), Mike Zheng Shou (Facebook AI), Matt Feiszli (Facebook Research), Zhangyang Wang (University of Texas at Austin), and Zhicheng Yan (Facebook AI)

Crossover Learning for Fast Online Video Instance Segmentation 8023
Shusheng Yang (Huazhong University of Science and Technology), Yuxin Fang (Huazhong University of Science and Technology), Xinggang Wang (Huazhong University of Science and Technology), Yu Li (Tencent), Chen Fang (Tencent), Ying Shan (Tencent), Bin Feng (Huazhong University of Science and Technology), and Wenyu Liu (Huazhong University of Science and Technology)

Domain Adaptive Video Segmentation via Temporal Consistency Regularization 8033
Dayan Guan (Nanyang Technological University), Jiaxing Huang (Nanyang Technological University), Aoran Xiao (Nanyang Technological University), and Shijian Lu (Nanyang Technological University)

Video Object Segmentation With Dynamic Memory Networks and Adaptive Object Alignment ... 8045
Shuxian Liang (Zhejiang University), Xu Shen (Alibaba Group), Jianqiang Huang (Alibaba Group), and Xian-Sheng Hua (Alibaba Group)

Generic Event Boundary Detection: A Benchmark for Event Segmentation 8055
Mike Zheng Shou (Facebook AI), Stan Weixian Lei (NUS), Weiyao Wang (Facebook Research), Deepti Ghadiyaram (Facebook), and Matt Feiszli (Facebook Research)

Weakly-Supervised Action Segmentation and Alignment via Transcript-Aware Union-of-Subspaces Learning 8065
Zijia Lu (Northeastern University) and Ehsan Elhamifar (Northeastern University)

ASCNet: Self-Supervised Video Representation Learning With Appearance-Speed Consistency ... 8076
Deng Huang (SCUT), Wenhao Wu (Baidu), Weiwen Hu (SCUT), Xu Liu (SCUT), Dongliang He (Baidu), Zhihua Wu (Baidu), Xiangmiao Wu (SCUT), Mingkui Tan (South China University of Technology), and Errui Ding (Baidu Inc.)

Detecting Human-Object Relationships in Videos 8086
Jingwei Ji (Stanford University), Rishi Desai (Stanford University), and Juan Carlos Niebles (Stanford University)

Unified Graph Structured Models for Video Understanding 8097
Anurag Arnab (Google), Chen Sun (Google), and Cordelia Schmid (Google)

Joint Visual and Audio Learning for Video Highlight Detection 8107
Taivanbat Badamdorj (University of Alberta), Mrigank Rochan (Huawei Technologies Canada), Yang Wang (University of Manitoba; Huawei Technologies Canada), and Li Cheng (ECE dept., University of Alberta)

TF-Blender: Temporal Feature Blender for Video Object Detection 8118
Yiming Cui (University of Florida), Liqi Yan (Fudan University), Zhiwen Cao (Purdue University), and Dongfang Liu (Rochester Institute of Technology)

Attention Is Not Enough: Mitigating the Distribution Discrepancy in Asynchronous Multimodal Sequence Fusion 8128
Tao Liang (Shenzhen Jinritoutiao Technology Co., Ltd), Guosheng Lin (Nanyang Technological University), Lei Feng (College of Computer Science, Chongqing University), Yan Zhang (University of Electronic Science and Technology of China), and Fengmao Lv (Southwest Jiaotong University)

HighlightMe: Detecting Highlights From Human-Centric Videos 8137
Uttaran Bhattacharya (University of Maryland, College Park), Gang Wu (Adobe Research), Stefano Petrangeli (Adobe), Viswanathan Swaminathan (Vishy) Swaminathan, (Adobe), and Dinesh Manocha (University of Maryland at College Park)

Interactive Prototype Learning for Egocentric Action Recognition 8148
Xiaohan Wang (University of Technology, Sydney), Linchao Zhu (University of Technology, Sydney), Heng Wang (Facebook AI), and Yi Yang (UTS)

Watch Only Once: An End-to-End Video Action Detection Framework 8158
Shoufa Chen (The University of Hong Kong), Peize Sun (The University of Hong Kong), Enze Xie (The University of Hong Kong), Chongjian Ge (The University of Hong Kong, (HKU)), Jiannan Wu (The University of Hong Kong), Lan Ma (TCL Corporate Research), Jiajun Shen (TCL Research), and Ping Luo (The University of Hong Kong)

Frequency-Aware Spatiotemporal Transformers for Video Inpainting Detection 8168
Bingyao Yu (Tsinghua University), Wanhua Li (Tsinghua University), Xiu Li (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)

3D Gestures and Detection

Parallel Detection-and-Segmentation Learning for Weakly Supervised Instance Segmentation 8178
Yunhang Shen (Xiamen University), Liujuan Cao (Xiamen University), Zhiwei Chen (Xiamen University), Baochang Zhang (Beihang University), Chi Su (Kingsoft Cloud), Yongjian Wu (Tencent Technology, (Shanghai) Co., Ltd), Feiyue Huang (Tencent), and Rongrong Ji (Xiamen University, China)

GistNet: A Geometric Structure Transfer Network for Long-Tailed Recognition 8189
Bo Liu (UCSD), Haoxiang Li (Wormpex AI Research), Hao Kang (Wormpex AI Research), Gang Hua (Wormpex AI Research), and Nuno Vasconcelos (UCSD, USA)

Semi-Supervised Semantic Segmentation With Pixel-Level Contrastive Learning From a Class-Wise Memory Bank 8199
Iñigo Alonso (Universidad de Zaragoza), Alberto Sabater (Universidad de Zaragoza), David Ferstl (Amazon PrimeAir), Luis Montesano (University of Zaragoza; Bitbrain), and Ana C. Murillo (Universidad de Zaragoza)

A Simple Baseline for Semi-Supervised Semantic Segmentation With Strong Data Augmentation . 8209
Jianlong Yuan (Alibaba Group), Yifan Liu (University of Adelaide), Chunhua Shen (University of Adelaide), Zhibin Wang (Alibaba Group), and Hao Li (Alibaba Group)

Self-Supervised Video Representation Learning With Meta-Contrastive Network 8219
Yuanze Lin (University of Washington), Xun Guo (Microsoft Research Asia), and Yan Lu (Microsoft Research Asia)

Continual Prototype Evolution: Learning Online From Non-Stationary Data Streams 8230
Matthias De Lange (KU Leuven) and Tinne Tuytelaars (KU Leuven)

CCT-Net: Category-Invariant Cross-Domain Transfer for Medical Single-to-Multiple Disease Diagnosis 8240
Yi Zhou (Southeast University), Lei Huang (Beihang University), Tao Zhou (Nanjing University of Science and Technology), and Ling Shao (Inception Institute of Artificial Intelligence)

Exploring Inter-Channel Correlation for Diversity-Preserved Knowledge Distillation 8251
Li Liu (Alibaba Group), Qingle Huang (Alibaba Group), Sihao Lin (Monash University), Hongwei Xie (Alibaba), Bing Wang (Alibaba group), Xiaojun Chang (RMIT University), and Xiaodan Liang (Sun Yat-sen University)

Online Continual Learning With Natural Distribution Shifts: An Empirical Study With Visual Data 8261
Zhipeng Cai (Intel Labs), Ozan Sener (Intel Labs), and Vladlen Koltun (Intel Labs)

Task Switching Network for Multi-Task Learning 8271
Guolei Sun (ETH Zurich), Thomas Probst (ETH Zurich), Danda Pani Paudel (ETH Zürich), Nikola Popović (ETH Zürich), Menelaos Kanakis (ETH Zurich), Jagruti Patel (ETH Zurich), Dengxin Dai (ETH Zurich), and Luc Van Gool (ETH Zurich)

Semantically Coherent Out-of-Distribution Detection 8281
Jingkang Yang (Nanyang Technology University), Haoqi Wang (SenseTime), Litong Feng (Sensetime Research), Xiaopeng Yan (SenseTime Research), Huabin Zheng (SenseTime Research), Wayne Zhang (SenseTime Research), and Ziwei Liu (Nanyang Technological University)

Trash To Treasure: Harvesting OOD Data With Cross-Modal Matching for Open-Set Semi-Supervised Learning 8290
Junkai Huang (Sun Yat-sen University), Chaowei Fang (Xidian University), Weikai Chen (Tencent America), Zhenhua Chai (Meituan), Xiaolin Wei (Meituan), Pengxu Wei (Sun Yat-sen University), Liang Lin (Sun Yat-sen University), and Guanbin Li (Sun Yat-sen University)

NAS-OoD: Neural Architecture Search for Out-of-Distribution Generalization 8300
Haoyue Bai (The Hong Kong University of Science and Technology), Fengwei Zhou (Huawei Noah's Ark Lab), Lanqing Hong (Huawei Noah's Ark Lab), Nanyang Ye (Shanghai Jiao Tong University), S.-H. Gary Chan (The Hong Kong University of Science and Technology), and Zhenguo Li (Huawei Noah's Ark Lab)

DRÆM – A Discriminatively Trained Reconstruction Embedding for Surface Anomaly Detection 8310
Vitjan Zavrtanik (University of Ljubljana), Matej Kristan (University of Ljubljana), and Danijel Skočaj (University of Ljubljana)

The Many Faces of Robustness: A Critical Analysis of Out-of-Distribution Generalization 8320
Dan Hendrycks (UC Berkeley), Steven Basart (University of Chicago), Norman Mu (University of California, Berkeley), Saurav Kadavath (UC Berkeley), Frank Wang (Google), Evan Dorundo (Google), Rahul Desai (UC Berkeley), Tyler Zhu (UC Berkeley), Samyak Parajuli (University of California, Berkeley), Mike Guo (UC Berkeley), Dawn Song (UC Berkeley), Jacob Steinhardt (UC Berkeley), and Justin Gilmer (Google Brain)

Keep CALM and Improve Visual Feature Attribution 8330
Jae Myung Kim (Seoul National University), Junsuk Choe (Sogang University), Zeynep Akata (University of Tübingen), and Seong Joon Oh (NAVER AI Lab)

Unsupervised Dense Deformation Embedding Network for Template-Free Shape Correspondence 8341
Ronghan Chen (Shenyang Institute of Automation, Chinese Academy of Sciences), Yang Cong (Chinese Academy of Sciences), and Jiahua Dong (Shenyang Institute of Automation, Chinese Academy of Sciences)

Online Pseudo Label Generation by Hierarchical Cluster Dynamics for Adaptive Person Re-Identification 8351
Yi Zheng (Zhejiang University), Shixiang Tang (The University of Sydney), Guolong Teng (Zhejiang University), Yixiao Ge (The Chinese University of Hong Kong), Kaijian Liu (Zhejiang University)), Jing Qin (The Hong Kong Polytechnic University), Donglian Qi (Zhejiang University), and Dapeng Chen (The Chinese University of Hong Kong)

Shape Self-Correction for Unsupervised Point Cloud Understanding 8362
Ye Chen (Shanghai Jiao Tong University), Jinxian Liu (Shanghai Jiao Tong University), Bingbing Ni (Shanghai Jiao Tong University), Hang Wang (Shanghai Jiao Tong University), Jiancheng Yang (Shanghai Jiao Tong University), Ning Liu (Shanghai Jiao Tong University), Teng Li (Anhui University), and Qi Tian (Huawei Cloud & AI)

DetCo: Unsupervised Contrastive Learning for Object Detection 8372
Enze Xie (The University of Hong Kong), Jian Ding (Wuhan University), Wenhai Wang (The Chinese University of Hong Kong), Xiaohang Zhan (Nanjing University), Hang Xu (Huawei Noah's Ark Lab), Peize Sun (The University of Hong Kong), Zhenguo Li (Huawei Noah's Ark Lab), and Ping Luo (The University of Hong Kong)

Binocular Mutual Learning for Improving Few-Shot Classification 8382
Ziqi Zhou (Megvii), Xi Qiu (Megvii Inc), Jiangtao Xie (Dalian University of Technology), Jianan Wu (Megvii Inc.), and Chi Zhang (Megvii Inc.)

Transductive Few-Shot Classification on the Oblique Manifold 8392
Guodong Qi (Zhejiang University, China; ZJU-League Research & Development Center), Huimin Yu (Zhejiang University, China; ZJU-League Research & Development Center), Zhaohui Lu (ZJU-League Research & Development Center), and Shuzhao Li (Zhejiang University, China; ZJU-League Research & Development Center)

Learning Rare Category Classifiers on a Tight Labeling Budget 8403
Ravi Teja Mullapudi (CMU), Fait Poms (Stanford), William R. Mark (Google), Deva Ramanan (Carnegie Mellon University), and Kayvon Fatahalian (Stanford)

Task-Aware Part Mining Network for Few-Shot Learning 8413
Jiamin Wu (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China), Yongdong Zhang (University of Science and Technology of China), and Feng Wu (University of Science and Technology of China)

Semi-Supervised Learning of Visual Features by Non-Parametrically Predicting View Assignments With Support Samples 8423
Mahmoud Assran (McGill University; Mila), Mathilde Caron (Facebook Artificial Intelligence Research), Ishan Misra (Facebook AI Research), Piotr Bojanowski (Facebook), Armand Joulin (Facebook AI Research), Nicolas Ballas (Facebook FAIR), and Michael Rabbat (Facebook FAIR)

A Multi-Mode Modulator for Multi-Domain Few-Shot Classification 8433
Yanbin Liu (University of Technology Sydney), Juho Lee (KAIST), Linchao Zhu (University of Technology, Sydney), Ling Chen (University of Technology, Sydney, Australia), Humphrey Shi (U of Oregon; UIUC), and Yi Yang (UTS)

LoFGAN: Fusing Local Representations for Few-Shot Image Generation 8443
Zheng Gu (Nanjing University), Wenbin Li (Nanjing University), Jing Huo (Nanjing University), Lei Wang (University of Wollongong, Australia), and Yang Gao (Nanjing University)

Interaction Compass: Multi-Label Zero-Shot Learning of Human-Object Interactions via Spatial Relations 8452
Dat Huynh (Northeastern University) and Ehsan Elhamifar (Northeastern University)

Unsupervised Few-Shot Action Recognition via Action-Appearance Aligned Meta-Adaptation ... 8464
Jay Patravali (Oregon State University), Gaurav Mittal (Microsoft), Ye Yu (Microsoft), Fuxin Li (Oregon State University), and Mei Chen (Microsoft)

A Hierarchical Transformation-Discriminating Generative Model for Few Shot Anomaly Detection 8475
Shelly Sheynin (Tel Aviv University), Sagie Benaim (Tel Aviv University), and Lior Wolf (Tel Aviv University, Israel)

Active Domain Adaptation via Clustering Uncertainty-Weighted Embeddings 8485
Viraj Prabhu (Georgia Tech), Arjun Chandrasekaran (Max Planck Institute for Intelligent Systems), Kate Saenko (Boston University), and Judy Hoffman (Georgia Tech)

Domain Adaptive Semantic Segmentation With Self-Supervised Depth Estimation 8495
Qin Wang (ETH Zurich), Dengxin Dai (ETH Zurich), Lukas Hoyer (ETH Zurich), Luc Van Gool (ETH Zurich), and Olga Fink (ETH Zurich)

Towards Discriminative Representation Learning for Unsupervised Person Re-Identification 8506
Takashi Isobe (Tsinghua University), Dong Li (Xilinx), Lu Tian (Xilinx, Inc.), Weihua Chen (alibaba group), Yi Shan (Xilinx), and Shengjin Wang (Tsinghua University)

Geometric Unsupervised Domain Adaptation for Semantic Segmentation 8517
Vitor Guizilini (Toyota Research Institute), Jie Li (Toyota Research Institute), Rareș Ambruș Rareș Ambruș (Toyota Research Institute), and Adrien Gaidon (Toyota Research Institute)

BiMaL: Bijective Maximum Likelihood Approach to Domain Adaptation in Semantic Scene Segmentation 8528
Thanh-Dat Truong (University of Arkansas), Chi Nhan Duong (Concordia University), Ngan Le (University of Arkansas), Son Lam Phung (University of Wollongong), Chase Rainwater (University of Arkansas), and Khoa Luu (University of Arkansas)

SENTRY: Selective Entropy Optimization via Committee Consistency for Unsupervised Domain Adaptation 8538
Viraj Prabhu (Georgia Tech), Shivam Khare (Georgia Institute of Technology), Deeksha Kartik (Georgia Institute of Technology), and Judy Hoffman (Georgia Tech)

Adversarial Robustness for Unsupervised Domain Adaptation 8548
Muhammad Awais (Kyung-Hee University), Fengwei Zhou (Huawei Noah's Ark Lab), Hang Xu (Huawei Noah's Ark Lab), Lanqing Hong (Huawei Noah's Ark Lab), Ping Luo (The University of Hong Kong), Sung-Ho Bae (Kyung Hee University), and Zhenguo Li (Huawei Noah's Ark Lab)

ECACL: A Holistic Framework for Semi-Supervised Domain Adaptation 8558
Kai Li (Northeastern University), Chang Liu (Northeastern University), Handong Zhao (Adobe Research), Yulun Zhang (Northeastern University), and Yun Fu (Northeastern University)

LabOR: Labeling Only if Required for Domain Adaptive Semantic Segmentation 8568
Inkyu Shin (KAIST), Dong-Jin Kim (KAIST), Jae Won Cho (KAIST), Sanghyun Woo (KAIST), Kwanyong Park (KAIST), and In So Kweon (KAIST)

Transporting Causal Mechanisms for Unsupervised Domain Adaptation 8579
Zhongqi Yue (Nanyang Technological University), Qianru Sun (Singapore Management University), Xian-Sheng Hua (Damo Academy, Alibaba Group), and Hanwang Zhang (Nanyang Technological University)

Self-Supervised Object Detection via Generative Image Synthesis 8589
Siva Karthik Mustikovela (Uni Heidelberg), Shalini De Mello (NVIDIA Research), Aayush Prakash (Nvidia), Umar Iqbal (NVIDIA Research), Sifei Liu (NVIDIA), Thu Nguyen-Phuoc (University of Bath), Carsten Rother (University of Heidelberg), and Jan Kautz (NVIDIA)

Skeleton2Mesh: Kinematics Prior Injected Unsupervised Human Mesh Recovery 8599
Zhenbo Yu (Shanghai Jiao Tong University), Junjie Wang (Shanghai Jiao Tong University), Jingwei Xu (Shanghai Jiao Tong University), Bingbing Ni (Shanghai Jiao Tong University), Chenglong Zhao (Shanghai Jiao Tong University), Minsi Wang (Shanghai Jiao Tong University), and Wenjun Zhang (Shanghai Jiao Tong University)

Intrinsic-Extrinsic Preserved GANs for Unsupervised 3D Pose Transfer 8610
Haoyu Chen (University of Oulu), Hao Tang (ETH Zurich), Henglin Shi (University of Oulu), Wei Peng (CMVS, University of Oulu), Nicu Sebe (University of Trento), and Guoying Zhao (University of Oulu)

Unsupervised Layered Image Decomposition Into Object Prototypes 8620
Tom Monnier (École des ponts Paristech), Elliot Vincent (École des ponts ParisTech / Inria), Jean Ponce (Inria), and Mathieu Aubry (École des ponts ParisTech)

Towards Alleviating the Modeling Ambiguity of Unsupervised Monocular 3D Human Pose Estimation 8631
Zhenbo Yu (Shanghai Jiao Tong University), Bingbing Ni (Shanghai Jiao Tong University), Jingwei Xu (Shanghai Jiao Tong University), Junjie Wang (Shanghai Jiao Tong University), Chenglong Zhao (Shanghai Jiao Tong University), and Wenjun Zhang (Shanghai Jiao Tong University)

Synthesized Feature Based Few-Shot Class-Incremental Learning on a Mixture of Subspaces 8641
Ali Cheraghian (Australian National University, (ANU)), Shafin Rahman (North South University), Sameera Ramasinghe (Australian National University), Pengfei Fang (The Australian National University), Christian Simon (Australian National University), Lars Petersson (Data61/CSIRO), and Mehrtash Harandi (Monash University)

Pseudo-Loss Confidence Metric for Semi-Supervised Few-Shot Learning 8651
Kai Huang (Northwestern Polytechnical University), Jie Geng (Northwestern Polytechnical University), Wen Jiang (Northwestern Polytechnical University), Xinyang Deng (Northwestern Polytechnical University), and Zhe Xu (Northwestern Polytechnical University)

DeFRCN: Decoupled Faster R-CNN for Few-Shot Object Detection 8661
Limeng Qiao (Megvii Inc.), Yuxuan Zhao (Megvii Inc.), Zhiyuan Li (Megvii Inc.), Xi Qiu (Megvii Inc), Jianan Wu (Megvii Inc.), and Chi Zhang (Megvii Inc.)

Curvature Generation in Curved Spaces for Few-Shot Learning 8671
Zhi Gao (Beijing Institute of Technology), Yuwei Wu (Beijing Institute of Technology, (BIT), China), Yunde Jia (Beijing Institute of Technology), and Mehrtash Harandi (Monash University)

Learning To Hallucinate Examples From Extrinsic and Intrinsic Supervision 8681
Liangke Gui (Carnegie Mellon University), Adrien Bardes (Inria), Ruslan Salakhutdinov (Carnegie Mellon University), Alexander Hauptmann (Carnegie Mellon University), Martial Hebert (Carnegie Mellon School of Computer Science), and Yu-Xiong Wang (University of Illinois at Urbana-Champaign)

Semantics Disentangling for Generalized Zero-Shot Learning 8692
Zhi Chen (The University of Queensland), Yadan Luo (University of Queensland), Ruihong Qiu (The University of Queensland), Sen Wang (The University of Queensland), Zi Huang (University of Queensland), Jingjing Li (University of Electronic Science and Technology of China), and Zheng Zhang (Harbin Institute of Technology, Shenzhen)

Mining Latent Classes for Few-Shot Segmentation 8701
Lihe Yang (Nanjing University), Wei Zhuo (Tencent), Lei Qi (Southeast University), Yinghuan Shi (Nanjing University), and Yang Gao (Nanjing University)

Discriminative Region-Based Multi-Label Zero-Shot Learning 8711
Sanath Narayan (Inception Institute of Artificial Intelligence), Akshita Gupta (IIAI), Salman Khan (Australian National University, (ANU)), Fahad Shahbaz Khan (MBZUAI), Ling Shao (Inception Institute of Artificial Intelligence), and Mubarak Shah (University of Central Florida)

Simpler Is Better: Few-Shot Semantic Segmentation With Classifier Weight Transformer 8721
Zhihe Lu (University of Surrey), Sen He (University of Surrey), Xiatian Zhu (University of Surrey), Li Zhang (Fudan University), Yi-Zhe Song (University of Surrey), and Tao Xiang (University of Surrey)

Iterative Label Cleaning for Transductive and Semi-Supervised Few-Shot Learning 8731
Michalis Lazarou (Imperial College London), Tania Stathaki (Imperial College London), and Yannis Avrithis (Inria)

PIT: Position-Invariant Transform for Cross-FoV Domain Adaptation 8741
Qiqi Gu (Shanghai Jiao Tong University), Qianyu Zhou (Shanghai Jiao Tong University), Minghao Xu (Shanghai Jiaotong University), Zhengyang Feng (Shanghai Jiao Tong University), Guangliang Cheng (Sensetime Group Limited), Xuequan Lu (Deakin University), Jianping Shi (Sensetime Group Limited), and Lizhuang Ma (Shanghai Jiao Tong University)

Domain-Invariant Disentangled Network for Generalizable Object Detection 8751
Chuang Lin (Monash University), Zehuan Yuan (Bytedance.Inc), Sicheng Zhao (University of California Berkeley), Peize Sun (The University of Hong Kong), Changhu Wang (bytedance), and Jianfei Cai (Monash University)

Deep Transport Network for Unsupervised Video Object Segmentation 8761
Kaihua Zhang (NUIST), Zicheng Zhao (nuist), Dong Liu (Netflix Inc), Qingshan Liu (Nanjing University of Information Science & Technology), and Bo Liu (JD.com)

Divide-and-Assemble: Learning Block-Wise Memory for Unsupervised Anomaly Detection 8771
Jinlei Hou (Zhejiang University), Yingying Zhang (Hikvision Research Institute), Qiaoyong Zhong (Hikvision Research Institute), Di Xie (Hikvision Research Institute), Shiliang Pu (Hikvision Research Institute), and Hong Zhou (Zhejiang University)

BAPA-Net: Boundary Adaptation and Prototype Alignment for Cross-Domain Semantic Segmentation 8781
Yahao Liu (University of Electronic Science and Technology of China), Jinhong Deng (University of Electronic Science and Technology of China), Xinchen Gao (University of Electronic Science and Technology of China), Wen Li (University of Electronic Science and Technology of China), and Lixin Duan (University of Electronic Science and Technology of China)

Variational Feature Disentangling for Fine-Grained Few-Shot Classification 8792
Jingyi Xu (Stony brook university), Hieu Le (Stony Brook University), Mingzhen Huang (Stony Brook University), ShahRukh Athar (Stony Brook University), and Dimitris Samaras (Stony Brook University)

Relational Embedding for Few-Shot Classification 8802
Dahyun Kang (POSTECH), Heeseung Kwon (POSTECH), Juhong Min (POSTECH), and Minsu Cho (POSTECH)

Composable Augmentation Encoding for Video Representation Learning 8814
Chen Sun (Google), Arsha Nagrani (Oxford), Yonglong Tian (MIT), and Cordelia Schmid (Google)

A Broad Study on the Transferability of Visual Representations With Contrastive Learning 8825
Ashraful Islam (Rensselaer Polytechnic Institute), Chun-Fu Richard Chen (MIT-IBM Watson AI Lab, IBM Research AI), Rameswar Panda (MIT-IBM Watson AI Lab, IBM Research), Leonid Karlinsky (IBM-Research), Richard Radke (Rensselaer Polytechnic Institute), and Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research)

Multi-Task Self-Training for Learning General Representations 8836
Golnaz Ghiasi (Google Brain), Barret Zoph (Google), Ekin D. Cubuk (Google Brain), Quoc V. Le (Google Brain), and Tsung-Yi Lin (Google Brain)

Unsupervised Domain Adaptive 3D Detection With Multi-Level Consistency 8846
Zhipeng Luo (Nanyang Technological University), Zhongang Cai (SenseTime International Pte Ltd), Changqing Zhou (Nanyang Technological University), Gongjie Zhang (Nanyang Technological University), Haiyu Zhao (SenseTime International Pte Ltd), Shuai Yi (SenseTime Group Limited), Shijian Lu (Nanyang Technological University), Hongsheng Li (Chinese University of Hong Kong), Shanghang Zhang (UC Berkeley), and Ziwei Liu (Nanyang Technological University)

mDALU: Multi-Source Domain Adaptation and Label Unification With Partial Datasets 8856
Rui Gong (ETH Zurich), Dengxin Dai (ETH Zurich), Yuhua Chen (ETH Zurich), Wen Li (University of Electronic Science and Technology of China), and Luc Van Gool (ETH Zurich)

A Simple Feature Augmentation for Domain Generalization 8866
Pan Li (Queen Mary University of London), Da Li (Samsung), Wei Li (Queen Mary University of London), Shaogang Gong (Queen Mary University of London), Yanwei Fu (Fudan University), and Timothy M. Hospedales (Edinburgh University)

Collaborative Learning With Disentangled Features for Zero-Shot Domain Adaptation 8876
Won Young Jhoo (Sungkyunkwan University) and Jae-Pil Heo (Sungkyunkwan University)

Deep Co-Training With Task Decomposition for Semi-Supervised Domain Adaptation 8886
Luyu Yang (University of Maryland, College Park), Yan Wang (Cornell University), Mingfei Gao (Salesforce Research), Abhinav Shrivastava (University of Maryland), Kilian Q. Weinberger (Cornell University), Wei-Lun Chao (The Ohio State University), and Ser-Nam Lim (Facebook AI)

Weak Adaptation Learning: Addressing Cross-Domain Data Insufficiency With Weak Annotator 8897
Shichao Xu (Northwestern University), Lixu Wang (Northwestern University), Yixuan Wang (Northwestern University), and Qi Zhu (Northwestern University)

Contrastive Coding for Active Learning Under Class Distribution Mismatch 8907
Pan Du (Renmin University of China), Suyun Zhao (Renmin University of China), Hui Chen (Renmin University of China), Shuwen Chai (Renmin University of China), Hong Chen (Renmin University, China), and Cuiping Li (Renmin University of China)

Gradient Distribution Alignment Certificates Better Adversarial Domain Adaptation 8917
Zhiqiang Gao (Xi'an Jiaotong-Liverpool Uiversity), Shufei Zhang (University of Liverpool), Kaizhu Huang (Xi'an Jiaotong-Liverpool Univ.), Qiufeng Wang (Xi'an Jiaotong-Liverpool University), and Chaoliang Zhong (Fujitsu R&D Center, Co., LTD, Beijing, China)

Meta Learning on a Sequence of Imbalanced Domains With Difficulty Awareness 8927
Zhenyi Wang (University at Buffalo), Tiehang Duan (SUNY Buffalo), Le Fang (University at Buffalo), Qiuling Suo (State University of New York at Buffalo), and Mingchen Gao (University at Buffalo, SUNY)

Confidence Calibration for Domain Generalization Under Covariate Shift 8938
Yunye Gong (SRI International), Xiao Lin (SRI International), Yi Yao (SRI International), Thomas G. Dietterich (Oregon State University), Ajay Divakaran (SRI, USA), and Melinda Gervasio (SRI International)

Active Universal Domain Adaptation 8948
Xinhong Ma (National Laboratory of Pattern Recognition Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Junyu Gao (CASIA), and Changsheng Xu (CASIA)

Generalized Source-Free Domain Adaptation 8958
Shiqi Yang (Computer Vision Center), Yaxing Wang (Computer Vision Center), Joost van de Weijer (Computer Vision Center), Luis Herranz (Computer Vision Center), and Shangling Jui (Huawei Kirin Solution)

RDA: Robust Domain Adaptation via Fourier Adversarial Attacking 8968
Jiaxing Huang (Nanyang Technological University), Dayan Guan (Nanyang Technological University), Aoran Xiao (Nanyang Technological University), and Shijian Lu (Nanyang Technological University)

OVANet: One-vs-All Network for Universal Domain Adaptation 8980
Kuniaki Saito (Boston University) and Kate Saenko (Boston University)

Adaptive Adversarial Network for Source-Free Domain Adaptation 8990
Haifeng Xia (Tulane University), Handong Zhao (Adobe Research), and Zhengming Ding (Tulane University)

Generalized and Incremental Few-Shot Learning by Explicit Learning and Calibration Without Forgetting 9000
Anna Kukleva (MPII), Hilde Kuehne (University of Frankfurt), and Bernt Schiele (MPI Informatics)

On the Importance of Distractors for Few-Shot Classification 9010
Rajshekhar Das (Carnegie Mellon University), Yu-Xiong Wang (University of Illinois at Urbana-Champaign), and José M. F. Moura (Carnegie Mellon University)

Mixture-Based Feature Space Learning for Few-Shot Image Classification 9021
Arman Afrasiyabi (Université Laval), Jean-François Lalonde (Université Laval), and Christian Gagné (Université Laval)

Coarsely-Labeled Data for Better Few-Shot Transfer 9032
Cheng Perng Phoo (Cornell University) and Bharath Hariharan (Cornell University)

Meta-Baseline: Exploring Simple Meta-Learning for Few-Shot Learning 9042
Yinbo Chen (UC San Diego), Zhuang Liu (UC Berkeley), Huijuan Xu (Pennsylvania State University), Trevor Darrell (UC Berkeley), and Xiaolong Wang (UCSD)

Multi-Target Adversarial Frameworks for Domain Adaptation in Semantic Segmentation 9052
Antoine Saporta (Sorbonne University), Tuan-Hung Vu (Valeo.ai), Matthieu Cord (Valeo.ai), and Patrick Pérez (Sorbonne University)

Dual Path Learning for Domain Adaptation of Semantic Segmentation 9062
Yiting Cheng (Fudan University), Fangyun Wei (Microsoft Research Asia), Jianmin Bao (Microsoft Research Asia), Dong Chen (Microsoft Research Asia), Fang Wen (Microsoft Research Asia), and Wenqiang Zhang (Fudan University)

Uncertainty-Aware Pseudo Label Refinery for Domain Adaptive Semantic Segmentation 9072
Yuxi Wang (CASIA), Junran Peng (Chinese Academy of Sciences), and ZhaoXiang Zhang (Chinese Academy of Sciences, China)

Semantic Concentration for Domain Adaptation 9082
Shuang Li (Beijing Institute of Technology), Mixue Xie (Beijing Institute of Technology), Fangrui Lv (Beijing Institute of Technology), Chi Harold Liu (Beijing Institute of Technology), Jian Liang (Alibaba Group), Chen Qin (Institute for Digital Communications, School of Engineering, University of Edinburgh, Edinburgh, UK), and Wei Li (Peking University)

Multi-Anchor Active Domain Adaptation for Semantic Segmentation 9092
Munan Ning (Tencent), Donghuan Lu (Tencent), Dong Wei (Tencent Jarvis Lab), Cheng Bian (Tencent), Chenglang Yuan (Tencent), Shuang Yu (Tencent), Kai Ma (Tencent), and Yefeng Zheng (Tencent)

CDS: Cross-Domain Self-Supervised Pre-Training 9103
Donghyun Kim (Boston University), Kuniaki Saito (Boston University), Tae-Hyun Oh (POSTECH), Bryan A. Plummer (Boston University), Stan Sclaroff (Boston University), and Kate Saenko (Boston University)

Knowledge Mining and Transferring for Domain Adaptive Object Detection 9113
Kun Tian (Institute of Automation, Chinese Academy of Sciences), Chenghao Zhang (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China), Ying Wang (Institute of Automation, Chinese Academy of Sciences), Shiming Xiang (Chinese Academy of Sciences, China), and Chunhong Pan (Institute of Automation, Chinese Academy of Sciences)

Robust Object Detection via Instance-Level Temporal Cycle Confusion 9123
Xin Wang (UC Berkeley), Thomas E. Huang (ETH Zürich), Benlin Liu (UCLA), Fisher Yu (ETH Zurich), Xiaolong Wang (UCSD), Joseph E. Gonzalez (UC Berkeley), and Trevor Darrell (UC Berkeley)

The Pursuit of Knowledge: Discovering and Localizing Novel Categories Using Dual Memory 9133
Sai Saketh Rambhatla (University of Maryland), Rama Chellappa (Johns Hopkins University), and Abhinav Shrivastava (University of Maryland)

A Style and Semantic Memory Mechanism for Domain Generalization 9144
Yang Chen (University of Science and Technology of China), Yu Wang (JD AI Research), Yingwei Pan (JD AI Research), Ting Yao (JD AI Research), Xinmei Tian (USTC), and Tao Mei (AI Research of JD.com)

Re-Energizing Domain Discriminator With Sample Relabeling for Adversarial Domain Adaptation 9154
Xin Jin (University of Science and Technology of China), Cuiling Lan (Microsoft Research), Wenjun Zeng (Microsoft Research), and Zhibo Chen (University of Science and Technology of China)

Tune It the Right Way: Unsupervised Validation of Domain Adaptation via Soft Neighborhood Density 9164
Kuniaki Saito (Boston University), Donghyun Kim (Boston University), Piotr Teterwak (Google), Stan Sclaroff (Boston University), Trevor Darrell (UC Berkeley), and Kate Saenko (Boston University)

Exploring Robustness of Unsupervised Domain Adaptation in Semantic Segmentation 9174
Jinyu Yang (University of Texas at Arlington), Chunyuan Li (University of Texas at Arlington), Weizhi An (UTA), Hehuan Ma (University of Texas at Arlington), Yuzhi Guo (University of Texas at Arlington), Yu Rong (Tencent AI Lab), Peilin Zhao (Tencent AI Lab), and Junzhou Huang (University of Texas at Arlington)

Seeking Similarities Over Differences: Similarity-Based Domain Alignment for Adaptive Object Detection 9184
Farzaneh Rezaeianaran (MPI), Rakshith Shetty (Max Planck Institute of Informatics), Rahaf Aljundi (Toyota Motor Europe), Daniel Olmeda Reino (Toyota Motor Europe), Shanshan Zhang (Max Planck Institute for Informatics), and Bernt Schiele (MPI Informatics)

Information-Theoretic Regularization for Multi-Source Domain Adaptation 9194
Geon Yeong Park (KAIST) and Sang Wan Lee (KAIST)

Graph Contrastive Clustering 9204
Huasong Zhong (Damo Academy, Alibaba Group), Jianlong Wu (Shandong University), Chong Chen (Alibaba), Jianqiang Huang (Peking University), Minghua Deng (Alibaba Group), Liqiang Nie (Shandong University), Zhouchen Lin (Peking University), and Xian-Sheng Hua (Alibaba Group)

Multi-VAE: Learning Disentangled View-Common and View-Peculiar Visual Representations for Multi-View Clustering 9214
Jie Xu (University of Electronic Science and Technology of China), Yazhou Ren (University of Electronic Science and Technology of China), Huayi Tang (University of Electronic Science and Technology of China), Xiaorong Pu (University of Electronic Science and Technology of China), Xiaofeng Zhu (Guangxi Normal University), Ming Zeng (Carnegie Mellon University), and Lifang He (Lehigh University)

Long Short View Feature Decomposition via Contrastive Video Representation Learning 9224
Nadine Behrmann (Bosch Center for Artificial Intelligence), Mohsen Fayyaz (Microsoft), Juergen Gall (University of Bonn), and Mehdi Noroozi (Bosch Gmb)

Video Pose Distillation for Few-Shot, Fine-Grained Sports Action Recognition 9234
James Hong (Stanford University), Matthew Fisher (Adobe Research), Michaël Gharbi (Adobe Research), and Kayvon Fatahalian (Stanford)

Semi-Supervised Single-Stage Controllable GANs for Conditional Fine-Grained Image Generation 9244
Tianyi Chen (South China University of Technology), Yi Liu (South China University of Technology), Yunfei Zhang (South China University of Technology), Si Wu (South China University of Technology), Yong Xu (South China University of Technology), Feng Liangbing (Cosmos Vision Technology Co., Ltd), and Hau San Wong (City University of Hong Kong)

Influence Selection for Active Learning 9254
Zhuoming Liu (SenseTime Group Limited), Hao Ding (Johns Hopkins University), Huaping Zhong (Sensetime Group Limited), Weijia Li (The Chinese University of Hong Kong), Jifeng Dai (SenseTime), and Conghui He (Sensetime Inc.)

A Unified Objective for Novel Class Discovery 9264
Enrico Fini (University of Trento), Enver Sangineto (University of Trento), Stéphane Lathuilière (Telecom-Paris), Zhun Zhong (University of Trento), Moin Nabi (SAP), and Elisa Ricci (University of Trento)

Localized Simple Multiple Kernel K-Means 9273
Xinwang Liu (National University of Defense Technology), Sihang Zhou (NUDT), Li Liu (National University of Defense Technology), Chang Tang (China University of Geosciences), Siwei Wang (National University of Defense Technology), Jiyuan Liu (National University of Defense Technology), and Yi Zhang (National University of Defense Technology)

Energy-Based Open-World Uncertainty Modeling for Confidence Calibration 9282
Yezhen Wang (Microsoft Research Lab), Bo Li (Microsoft Research Asia), Tong Che (MILA), Kaiyang Zhou (Nanyang Technological University), Ziwei Liu (Nanyang Technological University), and Dongsheng Li (Microsoft Research Asia)

Me-Momentum: Extracting Hard Confident Examples From Noisily Labeled Data 9292
Yingbin Bai (The University of Sydney) and Tongliang Liu (The University of Sydney)

Towards Novel Target Discovery Through Open-Set Domain Adaptation 9302
Taotao Jing (Tulane University), Hongfu Liu (Brandeis University), and Zhengming Ding (Tulane University)

Partial Video Domain Adaptation With Partial Adversarial Temporal Attentive Network 9312
*Yuecong Xu (Institute for Infocomm Research, A*STAR, Singapore), Jianfei Yang (Nanyang Technological University), Haozhi Cao (Nanyang Technological University), Zhenghua Chen (Institute for Infocomm Research, A*STAR, Singapore), Qi Li (Nanyang Technological University), and Kezhi Mao (Nanyang Technological University)*

Vector-Decomposed Disentanglement for Domain-Invariant Object Detection 9322
Aming Wu (Tianjin University), Rui Liu (Tianjin University), Yahong Han (Tianjin University), Linchao Zhu (University of Technology, Sydney), and Yi Yang (UTS)

STEM: An Approach to Multi-Source Domain Adaptation With Guarantees 9332
Van-Anh Nguyen (Vietnam National University-University of Science), Tuan Nguyen (Monash University), Trung Le (Monash University), Quan Hung Tran (Adobe Research), and Dinh Phung (Monash University)

MT-ORL: Multi-Task Occlusion Relationship Learning 9344
Panhe Feng (Beijing University of Posts and Telecommunications, ByteDance), Qi She (Bytedance AI Lab), Lei Zhu (Beijing University of Posts and Telecommunications), Jiaxin Li (Bytedance, TikTok), Lin Zhang (Carnegie Mellon University), Zijian Feng (ByteDance AI Lab), Changhu Wang (ByteDance.Inc), Chunpeng Li (Beijing University of Posts and Telecommunications), Xuejing Kang (Beijing University of Posts and Telecommunications), and Anlong Ming (Beijing University of Posts and Telecommunications)

Always Be Dreaming: A New Approach for Data-Free Class-Incremental Learning 9354
James Smith (Georgia Institute of Technology), Yen-Chang Hsu (Samsung Research America), Jonathan Balloch (Georgia Institute of Technology), Yilin Shen (Samsung Research America), Hongxia Jin (Samsung Research America), and Zsolt Kira (Georgia Institute of Technology)

Rehearsal Revealed: The Limits and Merits of Revisiting Samples in Continual Learning 9365
Eli Verwimp (KU Leuven), Matthias De Lange (KU Leuven), and Tinne Tuytelaars (KU Leuven)

Densely Guided Knowledge Distillation Using Multiple Teacher Assistants 9375
Wonchul Son (Ajou University), Jaemin Na (Ajou University), Junyong Choi (Ajou Univ), and Wonjun Hwang (Ajou University)

Testing Using Privileged Information by Adapting Features With Statistical Dependence 9385
Kwang In Kim (UNIST) and James Tompkin (Brown University)

Seasonal Contrast: Unsupervised Pre-Training From Uncurated Remote Sensing Data 9394
Oscar Mañas (ElementAI / ServiceNow), Alexandre Lacoste (ServiceNow), Xavier Giró-i-Nieto (Universitat Politecnica de Catalunya), David Vazquez (ServiceNow), and Pau Rodríguez (ElementAI / ServiceNow)

Boosting the Generalization Capability in Cross-Domain Few-Shot Learning via Noise-Enhanced Supervised Autoencoder 9404
Hanwen Liang (Huawei Noah's Ark Laboratory), Qiong Zhang (University of British Columbia), Peng Dai (Huawei Technologies Canada), and Juwei Lu (Huawei Noah's Ark Laboratory)

Meta Navigator: Search for a Good Adaptation Policy for Few-Shot Learning 9415
Chi Zhang (Nanyang Technological University), Henghui Ding (Nanyang Technological University), Guosheng Lin (Nanyang Technological University), Ruibo Li (Nanyang Technological University), Changhu Wang (ByteDance.Inc), and Chunhua Shen (University of Adelaide)

Few-Shot Image Classification: Just Use a Library of Pre-Trained Feature Extractors and a Simple Classifier 9425
Arkabandhu Chowdhury (Amazon.com), Mingchao Jiang (Rice University), Swarat Chaudhuri (The University of Texas at Austin), and Chris Jermaine (Rice University)

Few-Shot and Continual Learning With Attentive Independent Mechanisms 9435
Eugene Lee (National Chiao Tung University), Cheng-Han Huang (National Chiao Tung University), and Chen-Yi Lee (National Chiao Tung University)

Meta-Learning With Task-Adaptive Loss Function for Few-Shot Learning 9445
Sungyong Baik (Seoul National University), Janghoon Choi (Seoul National University), Heewon Kim (Seoul National University), Dohee Cho (Seoul Nationoal University), Jaesik Min (Hyundai Motor Group), and Kyoung Mu Lee (Seoul National University)

CoMatch: Semi-Supervised Learning With Contrastive Graph Regularization 9455
Junnan Li (Salesforce), Caiming Xiong (Salesforce Research), and Steven C.H. Hoi (Salesforce)

Learning From Noisy Data With Robust Representation Learning 9465
Junnan Li (Salesforce), Caiming Xiong (Salesforce Research), and Steven C.H. Hoi (Salesforce)

Distributional Robustness Loss for Long-Tail Learning 9475
Dvir Samuel (Bar-Ilan University) and Gal Chechik (Bar Ilan University)

Solving Inefficiency of Self-Supervised Representation Learning 9485
Guangrun Wang (University of Oxford), Keze Wang (Sun Yat-sen University), Guangcong Wang (Nanyang Technological University), Philip H.S. Torr (University of Oxford), and Liang Lin (Sun Yat-sen University)

Co2L: Contrastive Continual Learning 9496
Hyuntak Cha (KAIST), Jaeho Lee (KAIST), and Jinwoo Shin (KAIST)

Universal Representation Learning From Multiple Domains for Few-Shot Classification 9506
Wei-Hong Li (University of Edinburgh), Xialei Liu (Computer Vision Center Barcelona), and Hakan Bilen (University of Edinburgh)

Exploiting a Joint Embedding Space for Generalized Zero-Shot Semantic Segmentation 9516
Donghyeon Baek (Yonsei University), Youngmin Oh (Yonsei University), and Bumsub Ham (Yonsei University)

Field-Guide-Inspired Zero-Shot Learning 9526
Utkarsh Mall (Cornell University), Bharath Hariharan (Cornell University), and Kavita Bala (Cornell University)

SIGN: Spatial-Information Incorporated Generative Network for Generalized Zero-Shot Semantic Segmentation 9536
Jiaxin Cheng (USC Information Sciences Institute), Soumyaroop Nandi (Information Sciences Institute, University of Southern California), Prem Natarajan (Amazon.com Inc.), and Wael Abd-Almageed (Information Sciences Institute)

Universal-Prototype Enhancing for Few-Shot Object Detection 9547
Aming Wu (Tianjin University), Yahong Han (Tianjin University), Linchao Zhu (University of Technology, Sydney), and Yi Yang (UTS)

Representation Learning and Classification

On Compositions of Transformations in Contrastive Self-Supervised Learning 9557
Mandela Patrick (University of Oxford), Yuki M. Asano (University of Oxford), Polina Kuznetsova (University of Oxford), Ruth Fong (Facebook), João F. Henriques (University of Oxford), Geoffrey Zweig (Facebook), and Andrea Vedaldi (University of Oxford / Facebook AI Research)

With a Little Help From My Friends: Nearest-Neighbor Contrastive Learning of Visual Representations 9568
Debidatta Dwibedi (Google), Yusuf Aytar (DeepMind), Jonathan Tompson (Google), Pierre Sermanet (Google), and Andrew Zisserman (University of Oxford)

On Feature Decorrelation in Self-Supervised Learning 9578
Tianyu Hua (Tsinghua University), Wenxiao Wang (Tsinghua University), Zihui Xue (University of Texas at Austin), Sucheng Ren (South China University of Technology), Yue Wang (MIT), and Hang Zhao (Tsinghua University)

ISD: Self-Supervised Learning by Iterative Similarity Distillation 9589
Ajinkya Tejankar (UMBC), Soroush Abbasi Koohpayegani (University of Maryland Baltimore County), Vipin Pillai (UMBC), Paolo Favaro (University of Bern), and Hamed Pirsiavash (UMBC)

SelfReg: Self-Supervised Contrastive Regularization for Domain Generalization 9599
Daehee Kim (Kookmin University), Youngjun Yoo (Kookmin University), Seunghyun Park (Clova AI Research, NAVER Corp.), Jinkyu Kim (Korea University), and Jaekoo Lee (Kookmin University)

Concept Generalization in Visual Representation Learning 9609
Mert Bulent Sariyildiz (NAVER LABS Europe), Yannis Kalantidis (NAVER Labs Europe), Diane Larlus (Naver Labs Europe), and Karteek Alahari (Inria)

An Empirical Study of Training Self-Supervised Vision Transformers 9620
Xinlei Chen (FAIR), Saining Xie (Facebook AI Research), and Kaiming He (Facebook AI Research)

Emerging Properties in Self-Supervised Vision Transformers 9630
Mathilde Caron (Facebook Artificial Intelligence Research), Hugo Touvron (Facebook AI Research), Ishan Misra (Facebook AI Research), Hervé Jégou (Facebook AI Research), Julien Mairal (INRIA), Piotr Bojanowski (Facebook), and Armand Joulin (Facebook AI Research)

Do Image Classifiers Generalize Across Time? 9641
Vaishaal Shankar (UC Berkeley), Achal Dave (Carnegie Mellon University), Rebecca Roelofs (UC Berkeley), Deva Ramanan (CMU), Benjamin Recht (UC Berkeley), and Ludwig Schmidt (Toyota Research Institute)

Joint Inductive and Transductive Learning for Video Object Segmentation 9650
Yunyao Mao (University of Science and Technology of China), Ning Wang (University of Science and Technology of China), Wengang Zhou (University of Science and Technology of China), and Houqiang Li (University of Science and Technology of China)

Learning Facial Representations From the Cycle-Consistency of Face 9660
Jia-Ren Chang (National Chiao Tung University; aetherAI), Yong-Sheng Chen (National Chiao Tung University), and Wei-Chen Chiu (National Chiao Tung University)

UVStyle-Net: Unsupervised Few-Shot Learning of 3D Style Similarity Measure for B-Reps 9670
Peter Meltzer (University College London), Hooman Shayani (Autodesk AI Lab), Amir Khasahmadi (Autodesk AI Lab), Pradeep Kumar Jayaraman (Autodesk Research), Aditya Sanghi (Autodesk; Autodesk AI Lab), and Joseph Lambourne (Autodesk AI Lab)

Seeing Dynamic Scene in the Dark: A High-Quality Video Dataset With Mechatronic Alignment . 9680
Ruixing Wang (The Chinese University of Hong Kong), Xiaogang Xu (The Chinese University of Hong Kong), Chi-Wing Fu (The Chinese University of Hong Kong), Jiangbo Lu (SmartMore Corporation), Bei Yu (CUHK), and Jiaya Jia (Chinese University of Hong Kong)

Large Scale Interactive Motion Forecasting for Autonomous Driving: The Waymo Open Motion Dataset 9690
Scott Ettinger (Waymo), Shuyang Cheng (Waymo), Benjamin Caine (Google), Chenxi Liu (Waymo), Hang Zhao (Waymo), Sabeek Pradhan (Waymo), Yuning Chai (Waymo), Ben Sapp (Waymo), Charles R. Qi (Waymo), Yin Zhou (Waymo), Zoey Yang (Waymo, LLC.), Aurélien Chouard (Waymo), Pei Sun (Waymo), Jiquan Ngiam (Google Brain), Vijay Vasudevan (Google Brain), Alexander McCauley, Jonathon Shlens (Google), and Dragomir Anguelov (Waymo)

Contact-Aware Retargeting of Skinned Motion 9700
Ruben Villegas (Adobe Research), Duygu Ceylan (Adobe Research), Aaron Hertzmann (Adobe Research), Jimei Yang (Adobe), and Jun Saito (Adobe)

Video Autoencoder: Self-Supervised Disentanglement of Static 3D Structure and Motion 9710
Zihang Lai (CMU), Sifei Liu (NVIDIA), Alexei A. Efros (UC Berkeley), and Xiaolong Wang (UCSD)

Dissecting Image Crops 9721
Basile Van Hoorick (Columbia University) and Carl Vondrick (Columbia University)

A Hierarchical Variational Neural Uncertainty Model for Stochastic Video Prediction 9731
Moitreya Chatterjee (University of Illinois at Urbana Champaign), Narendra Ahuja (University of Illinois at Urbana-Champaign, USA), and Anoop Cherian (MERL)

X-World: Accessibility, Vision, and Autonomy Meet 9742
Jimuyang Zhang (Boston University), Minglan Zheng (Boston University), Matthew Boyd (Boston University), and Eshed Ohn-Bar (Boston University)

Learning To Estimate Hidden Motions With Global Motion Aggregation 9752
Shihao Jiang (Australian National University), Dylan Campbell (University of Oxford), Yao Lu (Australian National University), Hongdong Li (Australian National University, Australia), and Richard Hartley (Australian National University, Australia)

Unsupervised Point Cloud Pre-Training via Occlusion Completion 9762
Hanchen Wang (University of Cambridge), Qi Liu (University of Oxford), Xiangyu Yue (University of California, Berkeley), Joan Lasenby (University of Cambridge), and Matt J. Kusner (University College London)

Learn-To-Race: A Multimodal Control Environment for Autonomous Racing 9773
James Herman (Carnegie Mellon University), Jonathan Francis (Carnegie Mellon University; Bosch Research Pittsburgh), Siddha Ganju (Nvidia Corporation), Bingqing Chen (Carnegie Mellon University), Anirudh Koul (Pinterest), Abhinav Gupta (Carnegie Mellon University), Alexey Skabelkin (Arrival), Ivan Zhukov (Arrival), Max Kumskoy (Arrival), and Eric Nyberg (CMU)

LOKI: Long Term and Key Intentions for Trajectory Prediction 9783
Harshayu Girase (University of California, Berkeley), Haiming Gang (Honda Research Institute), Srikanth Malla (Honda Research Institute), Jiachen Li (University of California, Berkeley), Akira Kanehara (Honda R&D Co., Ltd.), Karttikeya Mangalam (UC Berkeley), and Chiho Choi (Honda Research Institute US)

AgentFormer: Agent-Aware Transformers for Socio-Temporal Multi-Agent Forecasting 9793
Ye Yuan (Carnegie Mellon University), Xinshuo Weng (Carnegie Mellon University), Yanglan Ou (Pennsylvania State University), and Kris M. Kitani (Carnegie Mellon University)

Human Trajectory Prediction via Counterfactual Analysis 9804
Guangyi Chen (Tsinghua University), Junlong Li (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)

Visio-Temporal Attention for Multi-Camera Multi-Target Association 9814
Yu-Jhe Li (Carnegie Mellon University), Xinshuo Weng (Carnegie Mellon University), Yan Xu (Carnegie Mellon University), and Kris M. Kitani (Carnegie Mellon University)

CrowdDriven: A New Challenging Dataset for Outdoor Visual Localization 9825
Ara Jafarzadeh (Chalmers), Manuel López Antequera (Facebook), Pau Gargallo (Facebook), Yubin Kuang (Facebook), Carl Toft (Chalmers), Fredrik Kahl (Chalmers), and Torsten Sattler (Czech Technical University in Prague)

High-Performance Discriminative Tracking With Transformers 9836
Bin Yu (Institute of Automation, Chinese Academy of Sciences), Ming Tang (Institute of Automation, Chinese Academy of Sciences), Linyu Zheng (Institute of Automation, Chinese Academy of Sciences), Guibo Zhu (Institute of Automation, Chinese Academy of Sciences), Jinqiao Wang (Institute of Automation, Chinese Academy of Sciences), Hao Feng (Alibaba Group), Xuetao Feng (Alibaba Group), and Hanqing Lu (NLPR, Institute of Automation, CAS)

Saliency-Associated Object Tracking 9846
Zikun Zhou (Harbin Institute of Technology, Shenzhen), Wenjie Pei (Harbin Institute of Technology, Shenzhen), Xin Li (Peng Cheng Laboratory), Hongpeng Wang (Harbin institute of technology, shenzhen), Feng Zheng (SUSTech), and Zhenyu He (Harbin Institute of Technology, (Shenzhen))

Track Without Appearance: Learn Box and Tracklet Embedding With Local and Global Motion Patterns for Vehicle Tracking 9856
Gaoang Wang (Zhejiang University), Renshu Gu (Hangzhou Dianzi University), Zuozhu Liu (Zhejiang-UIUC Institute), Weijie Hu (Guangdong University of Petrochemical Technology), Mingli Song (Zhejiang University), and Jenq-Neng Hwang (University of Washington)

Towards Interpretable Deep Metric Learning With Structural Matching 9867
Wenliang Zhao (Tsinghua University), Yongming Rao (Tsinghua University), Ziyi Wang (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)

On Equivariant and Invariant Learning of Object Landmark Representations 9877
Zezhou Cheng (University of Massachusetts, Amherst), Jong-Chyi Su (University of Massachusetts, Amherst), and Subhransu Maji (University of Massachusetts, Amherst)

Deep Matching Prior: Test-Time Optimization for Dense Correspondence 9887
Sunghwan Hong (Korea University) and Seungryong Kim (Korea University)

Learning Better Visual Data Similarities via New Grouplet Non-Euclidean Embedding 9898
Yanfu Zhang (University of Pittsburgh), Lei Luo (UTA), Wenhan Xian (University of Pittsburgh), and Heng Huang (University of Pittsburgh & JD Explore Academy)

Clustering by Maximizing Mutual Information Across Views 9908
Kien Do (Deakin Unviersity), Truyen Tran (Deakin University), and Svetha Venkatesh (Deakin University)

Learning Compatible Embeddings 9919
Qiang Meng (AiBee), Chixiang Zhang (Aibee Inc.), Xiaoqiang Xu (Aibee), and Feng Zhou (Aibee)

Contrasting Contrastive Self-Supervised Representation Learning Pipelines 9929
Klemen Kotar (Allen Institute for AI), Gabriel Ilharco (University of Washington), Ludwig Schmidt (Toyota Research Institute), Kiana Ehsani (Allen Institute for Artificial Intelligence), and Roozbeh Mottaghi (Allen Institute for AI)

Modelling Neighbor Relation in Joint Space-Time Graph for Video Correspondence Learning 9940
Zixu Zhao (The Chinese University of Hong Kong), Yueming Jin (The Chinese University of Hong Kong), and Pheng-Ann Heng (The Chinese Univsersity of Hong Kong)

Time-Equivariant Contrastive Video Representation Learning 9950
Simon Jenni (Adobe Research) and Hailin Jin (Adobe Research)

Co-Scale Conv-Attentional Image Transformers 9961
Weijian Xu (University of California, San Diego), Yifan Xu (UC San Diego), Tyler Chang (UC San Diego), and Zhuowen Tu (UC San Diego)

T-SVDNet: Exploring High-Order Prototypical Correlations for Multi-Source Domain Adaptation 9971
Ruihuang Li (Tianjin university), Xu Jia (Dalian University of Technology), Jianzhong He (Peking University), Shuaijun Chen (Huawei Noah's Ark Lab), and Qinghua Hu (Tianjin University)

Field Convolutions for Surface CNNs 9981
Thomas W. Mitchel (Johns Hopkins University), Vladimir G. Kim (Adobe), and Michael Kazhdan (Johns Hopkins University)

Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows 9992
Ze Liu (USTC), Yutong Lin (Xi'an Jiaotong University), Yue Cao (Microsoft Research), Han Hu (Microsoft Research Asia), Yixuan Wei (Tsinghua University), Zheng Zhang (MSRA, Huazhong University of Science and Technolog), Stephen Lin (Microsoft Research), and Baining Guo (MSR Asia)

InSeGAN: A Generative Approach to Segmenting Identical Instances in Depth Images 10003
Anoop Cherian (MERL), Gonçalo Dias Pais (Instituto Superior Técnico, Lisboa), Siddarth Jain (MERL), Tim K. Marks (Mitsubishi Electric Research Laboratories, USA), and Alan Sullivan (MERL)

Rethinking and Improving Relative Position Encoding for Vision Transformer 10013
Kan Wu (Sun Yat-sen University), Houwen Peng (Microsoft Research), Minghao Chen (Stony Brook University), Jianlong Fu (Microsoft Research), and Hongyang Chao (Sun Yat-sen University)

Weakly Supervised Contrastive Learning 10022
Mingkai Zheng (SenseTime), Fei Wang (University of Science and Technology of China), Shan You (SenseTime), Chen Qian (SenseTime), Changshui Zhang (Tsinghua University), Xiaogang Wang (Chinese University of Hong Kong, Hong Kong), and Chang Xu (University of Sydney)

Unsupervised Semantic Segmentation by Contrasting Object Mask Proposals 10032
Wouter Van Gansbeke (KU Leuven), Simon Vandenhende (KU Leuven), Stamatios Georgoulis (ETH Zurich), and Luc Van Gool (ETH Zurich)

Divide and Contrast: Self-Supervised Learning From Uncurated Data 10043
Yonglong Tian (MIT), Olivier J. Hénaff (DeepMind), and Aäron van den Oord (DeepMind)

Rethinking Self-Supervised Correspondence Learning: A Video Frame-Level Similarity Perspective 10055
Jiarui Xu (University of California San Diego) and Xiaolong Wang (UCSD)

Efficient Visual Pretraining With Contrastive Detection 10066
Olivier J. Hénaff (DeepMind), Skanda Koppula (DeepMind), Jean-Baptiste Alayrac (DeepMind), Aaron van den Oord (DeepMind), Oriol Vinyals (DeepMind), and João Carreira (DeepMind)

SketchAA: Abstract Representation for Abstract Sketches 10077
Lan Yang (Beijing University of Posts and Telecommunications), Kaiyue Pang (University of Surrey), Honggang Zhang (Beijing University of Posts and Telecommunications), and Yi-Zhe Song (University of Surrey)

LaLaLoc: Latent Layout Localisation in Dynamic, Unvisited Environments 10087
Henry Howard-Jenkins (University of Oxford), Jose-Raul Ruiz-Sarmiento (Ingenieria de Sistemas y Automatica), and Victor Adrian Prisacariu (University of Oxford)

OpenForensics: Large-Scale Challenging Dataset for Multi-Face Forgery Detection and Segmentation In-the-Wild 10097
Trung-Nghia Le (National Institute of Informatics), Huy H. Nguyen (SOKENDAI), Junichi Yamagishi (National Institute of Informatics), and Isao Echizen (National Institute of Informatics)

FloorPlanCAD: A Large-Scale CAD Drawing Dataset for Panoptic Symbol Spotting 10108
Zhiwen Fan (Alibaba A.I. Labs), Lingjie Zhu (Alibaba A.I. Labs), Honghua Li (Alibaba A.I. Labs), Xiaohao Chen (Alibaba Group), Siyu Zhu (Alibaba A.I. Labs), and Ping Tan (Simon Fraser University)

H2O: Two Hands Manipulating Objects for First Person Interaction Recognition 10118
Taein Kwon (ETH Zurich), Bugra Tekin (Microsoft), Jan Stühmer (Samsung AI Centre, Cambridge), Federica Bogo (Microsoft), and Marc Pollefeys (ETH Zurich / Microsoft)

Contrastive Learning of Image Representations With Cross-Video Cycle-Consistency 10129
Haiping Wu (McGill University) and Xiaolong Wang (UCSD)

Self-Supervised Visual Representations Learning by Contrastive Mask Prediction 10140
Yucheng Zhao (University of Science and Technology of China), Guangting Wang (University of Science and Technology of China), Chong Luo (MSRA), Wenjun Zeng (Microsoft Research), and Zheng-Jun Zha (University of Science and Technology of China)

Temporal Knowledge Consistency for Unsupervised Visual Representation Learning 10150
Weixin Feng (Beijing University of Posts and Telecommunications), Yuanjiang Wang (Megvii), Lihua Ma (megvii), Ye Yuan (Megvii), and Chi Zhang (Megvii Inc.)

Geography-Aware Self-Supervised Learning 10161
Kumar Ayush (Stanford University), Burak Uzkent (Stanford University), Chenlin Meng (Stanford University), Kumar Tanmay (IIT Kharagpur), Marshall Burke (Stanford University), David Lobell (Stanford University), and Stefano Ermon (Stanford University)

Self-Supervised Representation Learning From Flow Equivariance 10171
Yuwen Xiong (University of Toronto), Mengye Ren (University of Toronto), Wenyuan Zeng (Uber ATG, University of Toronto), and Raquel Urtasun (Uber ATG)

Improve Unsupervised Pretraining for Few-Label Transfer 10181
Suichan Li (University of Science and Technology of China), Dongdong Chen (Microsoft Cloud AI), Yinpeng Chen (Microsoft), Lu Yuan (Microsoft), Lei Zhang (International Digital Economy Academy, (IDEA)), Qi Chu (University of Science and Technology of China), Bin Liu (University of Science and Technology of China), and Nenghai Yu (University of Science and Technology of China)

Improving Robustness Against Common Corruptions With Frequency Biased Models 10191
Tonmoy Saikia (University of Freiburg), Cordelia Schmid (Inria/Google), and Thomas Brox (University of Freiburg)

Temporal-Wise Attention Spiking Neural Networks for Event Streams Classification 10201
Man Yao (Xi'an Jiaotong University), Huanhuan Gao (Xi'an Jiaotong University), Guangshe Zhao (Xi'an Jiaotong University), Dingheng Wang (Xi'an Jiaotong University), Yihan Lin (Tsinghua University), Zhaoxu Yang (Xi'an Jiaotong University), and Guoqi Li (Tsinghua University)

Understanding Robustness of Transformers for Image Classification 10211
Srinadh Bhojanapalli (Google AI), Ayan Chakrabarti (Google Research), Daniel Glasner (Google), Daliang Li (Google Research), Thomas Unterthiner (Google Research, Brain Team), and Andreas Veit (Google)

Learning Conditional Knowledge Distillation for Degraded-Reference Image Quality Assessment 10222
Heliang Zheng (USTC), Huan Yang (Microsoft Research), Jianlong Fu (Microsoft Research), Zheng-Jun Zha (University of Science and Technology of China), and Jiebo Luo (U. Rochester)

Self-Supervised Pretraining of 3D Features on Any Point-Cloud 10232
Zaiwei Zhang (University of Texas at Austin), Rohit Girdhar (Facebook AI Research), Armand Joulin (Facebook AI Research), and Ishan Misra (Facebook AI Research)

Active Learning for Deep Object Detection via Probabilistic Modeling 10244
Jiwoong Choi (Seoul National University), Ismail Elezi (Technical University of Munich, (TUM)), Hyuk-Jae Lee (SNU), Clement Farabet (NVIDIA), and Jose M. Alvarez (NVIDIA)

ACAV100M: Automatic Curation of Large-Scale Datasets for Audio-Visual Video Representation Learning 10254
Sangho Lee (Seoul National University), Jiwan Chung (Seoul National University), Youngjae Yu (Seoul National University Vision and Learning Lab), Gunhee Kim (Seoul National University), Thomas Breuel (NVIDIA), Gal Chechik (Nvidia), and Yale Song (Microsoft Research)

Benchmark Platform for Ultra-Fine-Grained Visual Categorization Beyond Human Performance 10265
Xiaohan Yu (Griffith University, Australia), Yang Zhao (The University of Adelaide), Yongsheng Gao (Griffith University), Xiaohui Yuan (Wuhan University of Technology, China), and Shengwu Xiong (Wuhan University of Technology)

Video Annotation for Visual Tracking via Selection and Refinement 10276
Kenan Dai (Dalian University of Technology), Jie Zhao (Dalian University of Technology), Lijun Wang (Dalian University of Technology), Dong Wang (Dalian University of Technology), Jianhua Li (Dalian University of Technology), Huchuan Lu (Dalian University of Technology), Xuesheng Qian (CSA Intellicloud Ltd), and Xiaoyun Yang (Remark Holdings)

Improving Contrastive Learning by Visualizing Feature Transformation 10286
Rui Zhu (The Chinese University of HongKong, Shenzhen), Bingchen Zhao (Tongji University), Jingen Liu (JD AI Research), Zhenglong Sun (The Chinese University of HongKong, Shenzhen), and Chang Wen Chen (SUNY at Buffalo, USA)

Rethinking Preventing Class-Collapsing in Metric Learning With Margin-Based Losses 10296
Elad Levi (Nexar), Tete Xiao (UC Berkeley), Xiaolong Wang (UCSD), and Trevor Darrell (UC Berkeley)

Mean Shift for Self-Supervised Learning 10306
Soroush Abbasi Koohpayegani (University of Maryland Baltimore County), Ajinkya Tejankar (UMBC), and Hamed Pirsiavash (UMBC)

Instance Similarity Learning for Unsupervised Feature Representation 10316
Ziwei Wang (Tsinghua University), Yunsong Wang (Tsinghua Univiersity), Ziyi Wu (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)

Warp Consistency for Unsupervised Learning of Dense Correspondences 10326
Prune Truong (ETH Zurich), Martin Danelljan (ETH Zurich), Fisher Yu (ETH Zurich), and Luc Van Gool (ETH Zurich)

Refining Activation Downsampling With SoftPool 10337
Alexandros Stergiou (Utrecht University), Ronald Poppe (Utrecht University), and Grigorios Kalliatakis (University of Warwick)

Adversarial Unsupervised Domain Adaptation With Conditional and Label Shift: Infer, Align and Iterate 10347
Xiaofeng Liu (CMU), Zhenhua Guo (Graduate School at Shenzhen, Tsinghua University, China), Site Li (CMU), Fangxu Xing (Massachusetts General Hospital / Harvard Medical School), Jane You (HK), C.-C. Jay Kuo (USC), Georges El Fakhri (Massachusetts General Hospital / Harvard Medical School), and Jonghye Woo (Massachusetts General Hospital / Harvard Medical School)

RANK-NOSH: Efficient Predictor-Based Architecture Search via Non-Uniform Successive Halving 10357
Ruochen Wang (University of California, Los Angeles), Xiangning Chen (University of California, Los Angeles), Minhao Cheng (UCLA), Xiaocheng Tang (DiDi AI Labs), and Cho-Jui Hsieh (UCLA)

Distilling Holistic Knowledge With Graph Neural Networks 10367
Sheng Zhou (Zhejiang University), Yucheng Wang (Zhejiang University), Defang Chen (Zhejiang University), Jiawei Chen (University of Science and Technology of China), Xin Wang (Tsinghua University), Can Wang (Zhejiang University), and Jiajun Bu (Zhejiang University)

BuildingNet: Learning To Label 3D Buildings 10377
Pratheba Selvaraju (University of Massachusetts, Amherst), Mohamed Nabail (University of Massachusetts Amherst), Marios Loizou (University of Cyprus), Maria Maslioukova (University of Cyprus), Melinos Averkiou (University of Cyprus), Andreas Andreou (University of Cyprus), Siddhartha Chaudhuri (Adobe Research), and Evangelos Kalogerakis (UMass Amherst)

Curious Representation Learning for Embodied Intelligence 10388
Yilun Du (MIT), Chuang Gan (MIT-IBM Watson AI Lab), and Phillip Isola (MIT)

ViewNet: Unsupervised Viewpoint Estimation From Conditional Generation 10398
Octave Mariotti (University of Edinburgh), Oisin Mac Aodha (University of Edinburgh), and Hakan Bilen (University of Edinburgh)

Motion-Augmented Self-Training for Video Recognition at Smaller Scale 10409
Kirill Gavrilyuk (University of Amsterdam), Mihir Jain (Qualcomm AI Research), Ilia Karmanov (Qualcomm Research), and Cees G. M. Snoek (University of Amsterdam)

PARTS: Unsupervised Segmentation With Slots, Attention and Independence Maximization 10419
Daniel Zoran (DeepMind), Rishabh Kabra (DeepMind), Alexander Lerchner (DeepMind), and Danilo J. Rezende (Google DeepMind)

Learning Spatio-Temporal Transformer for Visual Tracking 10428
Bin Yan (Dalian University of Technology), Houwen Peng (Microsoft Research), Jianlong Fu (Microsoft Research), Dong Wang (Dalian University of Technology), and Huchuan Lu (Dalian University of Technology)

Temporally-Coherent Surface Reconstruction via Metric-Consistent Atlases 10438
Jan Bednarik (EPFL), Vladimir G. Kim (Adobe), Siddhartha Chaudhuri (Adobe Research), Shaifali Parashar (EPFL), Mathieu Salzmann (EPFL), Pascal Fua (EPFL, Switzerland), and Noam Aigerman (Adobe)

SGMNet: Learning Rotation-Invariant Point Cloud Representations via Sorted Gram Matrix 10448
Jianyun Xu (HIKVISION), Xin Tang (Hikvision), Yushi Zhu (Hikvision Research Institute), Jie Sun (Hikvision), and Shiliang Pu (Hikvision Research Institute)

Point-Set Distances for Learning Representations of 3D Point Clouds 10458
Trung Nguyen (VinAI Research), Quang-Hieu Pham (Woven Planet North America - Level 5), Tam Le (RIKEN AIP), Tung Pham (VinAI Research), Nhat Ho (University of Texas at Austin), and Binh-Son Hua (VinAI Research)

Exploring Simple 3D Multi-Object Tracking for Autonomous Driving 10468
Chenxu Luo (Johns Hopkins University), Xiaodong Yang (NVIDIA Research), and Alan Yuille (Johns Hopkins University)

High-Resolution Optical Flow From 1D Attention and Correlation 10478
Haofei Xu (University of Science and Technology of China), Jiaolong Yang (Microsoft Research), Jianfei Cai (Monash University), Juyong Zhang (University of Science and Technology of China), and Xin Tong (Microsoft)

Not All Operations Contribute Equally: Hierarchical Operation-Adaptive Predictor for Neural Architecture Search 10488
Ziye Chen (School of Compuer Science, Wuhan University), Yibing Zhan (Jingdong), Baosheng Yu (The University of Sydney), Mingming Gong (University of Melbourne), and Bo Du (Wuhan University)

Poly-NL: Linear Complexity Non-Local Layers With 3rd Order Polynomials 10498
Francesca Babiloni (Huawei Noah's Ark Lab), Ioannis Marras (Huawei Noah's Ark Lab), Filippos Kokkinos (University College London), Jiankang Deng (Huawei), Grigorios Chrysos (École Polytechnique Fédérale de Lausanne), and Stefanos Zafeiriou (Huawei Technologies Co. Ltd)

Impact of Aliasing on Generalization in Deep Convolutional Networks 10509
Cristina Vasconcelos (Google), Hugo Larochelle (Google), Vincent Dumoulin (Google), Rob Romijnders (Google AI), Nicolas Le Roux (Google), and Ross Goroshin (Google)

Region Similarity Representation Learning 10519
Tete Xiao (UC Berkeley), Colorado J Reed (University of California, Berkeley), Xiaolong Wang (UCSD), Kurt Keutzer (EECS, UC Berkeley), and Trevor Darrell (UC Berkeley)

Personalized Image Semantic Segmentation 10529
Yu Zhang (Nankai University), Chang-Bin Zhang (Nankai University), Peng-Tao Jiang (Nankai University), Ming-Ming Cheng (Nankai University), and Feng Mao (Alibaba Group)

Space-Time Crop & Attend: Improving Cross-Modal Video Representation Learning 10540
Mandela Patrick (University of Oxford), Po-Yao Huang (University of Oxford), Ishan Misra (Carnegie Mellon University), Florian Metze (Facebook AI Research), Andrea Vedaldi (Facebook), Yuki M. Asano (University of Oxford), and João F. Henriques (University of Oxford / Facebook AI Research)

Partner-Assisted Learning for Few-Shot Image Classification 10553
Jiawei Ma (Columbia University), Hanchen Xie (USC Information Sciences Institute), Guangxing Han (Columbia University), Shih-Fu Chang (Columbia University), Aram Galstyan (USC Information Sciences Institute), and Wael Abd-Almageed (Information Sciences Institute)

Focus on the Positives: Self-Supervised Learning for Biodiversity Monitoring 10563
Omiros Pantazis (University College London), Gabriel J. Brostow (University College London), Kate E. Jones (University College London), and Oisin Mac Aodha (University of Edinburgh)

Weakly Supervised Representation Learning With Coarse Labels 10573
Yuanhong Xu (Alibaba Group), Qi Qian (Alibaba Group), Hao Li (Alibaba Group), Rong Jin (alibaba group), and Juhua Hu (University of Washington Tacoma)

Webly Supervised Fine-Grained Recognition: Benchmark Datasets and an Approach 10582
Zeren Sun (Nanjing University of Science and Technology), Yazhou Yao (Nanjing University of Science and Technology), Xiu-Shen Wei (Nanjing University of Science and Technology), Yongshun Zhang (Nanjing University), Fumin Shen (UESTC), Jianxin Wu (Nanjing University), Jian Zhang (UTS), and Heng Tao Shen (University of Electronic Science and Technology of China, (UESTC))

von Mises-Fisher Loss: An Exploration of Embedding Geometries for Supervised Learning 10592
Tyler R. Scott (University of Colorado Boulder), Andrew C. Gallagher (Google), and Michael C. Mozer (Google Research / University of Colorado)

Contrastive Learning for Label Efficient Semantic Segmentation 10603
Xiangyun Zhao (Northwestern University), Raviteja Vemulapalli (Google), Philip Andrew Mansfield (Google), Boqing Gong (Google), Bradley Green (Google Inc.), Lior Shapira (Google), and Ying Wu (Northwestern University)

LoOp: Looking for Optimal Hard Negative Embeddings for Deep Metric Learning 10614
Bhavya Vasudeva (ISI Kolkata), Puneesh Deora (ISI Kolkata), Saumik Bhattacharya (Indian Institute of Technology Kharagpur), Umapada Pal (Indian Statistical Institute, Kolkata), and Sukalpa Chanda (Østfold University College)

Do Different Deep Metric Learning Losses Lead to Similar Learned Features? 10624
Konstantin Kobs (Universität Würzburg), Michael Steininger (University of Würzburg), Andrzej Dulny (University of Würzburg), and Andreas Hotho (Universitat Wurzburg)

DiagViB-6: A Diagnostic Benchmark Suite for Vision Models in the Presence of Shortcut and Generalization Opportunities 10635
Elias Eulig (Bosch Center for Artificial Intelligence), Piyapat Saranrittichai (Bosch Center for Artificial Intelligence), Chaithanya Kumar Mummadi (Bosch Center for Artificial Intelligence), Kilian Rambach (Bosch Center for Artificial Intelligence), William Beluch (Bosch Center for Artificial Intelligence), Xiahan Shi (Bosch Center for Artificial Intelligence), and Volker Fischer (Bosch Center for Artificial Intelligence)

Kernel Methods in Hyperbolic Spaces 10645
Pengfei Fang (The Australian National University), Mehrtash Harandi (Monash University), and Lars Petersson (Data61/CSIRO)

Lipschitz Continuity Guided Knowledge Distillation 10655
Yuzhang Shang (Illinois Institute of Technology), Bin Duan (Illinois Institute of Technology), Ziliang Zong (Texas State University), Liqiang Nie (Shandong University), and Yan Yan (Illinois Institute of Technology)

Learning With Privileged Tasks 10665
Yuru Song (UCSD), Zan Lou (Sensetime), Shan You (SenseTime), Erkun Yang (University of North Carolina at Chapel Hill), Fei Wang (University of Science and Technology of China), Chen Qian (SenseTime), Changshui Zhang (Tsinghua University), and Xiaogang Wang (Chinese University of Hong Kong, Hong Kong)

Lifelong Infinite Mixture Model Based on Knowledge-Driven Dirichlet Process 10675
Fei Ye (University of york) and Adrian G. Bors (University of York)

Low-Shot Validation: Active Importance Sampling for Estimating Classifier Performance on Rare Categories 10685
Fait Poms (Stanford), Vishnu Sarukkai (Stanford), Ravi Teja Mullapudi (CMU), Nimit S. Sohoni (Stanford University), William R. Mark (Google), Deva Ramanan (Carnegie Mellon University), and Kayvon Fatahalian (Stanford)

Cloud Transformers: A Universal Approach to Point Cloud Processing Tasks 10695
Kirill Mazur (Samsung) and Victor Lempitsky (Samsung)

DepthTrack: Unveiling the Power of RGBD Tracking 10705
Song Yan (Tampere University), Jinyu Yang (Southern University of Science and Technology; University of Birmingham), Jani Käpylä (Tampere University), Feng Zheng (SUSTech), Aleš Leonardis (University of Birmingham), and Joni-Kristian Kämäräinen (Tampere University)

Transparent Object Tracking Benchmark 10714
Heng Fan (Stony Brook University), Halady Akhilesha Miththanthaya (Stony Brook University), Harshit Harshit (Stony Brook University), Siranjiv Ramana Rajan (Stony Brook University), Xiaoqiong Liu (StonyBrook University), Zhilin Zou (Stony Brook University), Yuewei Lin (Brookhaven National Laboratory), and Haibin Ling (Stony Brook University)

KoDF: A Large-Scale Korean DeepFake Detection Dataset 10724
Patrick Kwon (MoneyBrain Inc.), Jaeseong You (MoneyBrain Inc.), Gyuhyeon Nam (MoneyBrain Inc.), Sungwoo Park (MoneyBrain Inc.), and Gyeongsu Chae (MoneyBrain Inc.)

Dynamic Surface Function Networks for Clothed Human Bodies 10734
Andrei Burov (Technical University of Munich), Matthias Nießner (Technical University of Munich), and Justus Thies (Max Planck Institute for Intelligent Systems)

ACDC: The Adverse Conditions Dataset With Correspondences for Semantic Driving Scene Understanding 10745
Christos Sakaridis (ETH Zurich), Dengxin Dai (ETH Zurich), and Luc Van Gool (ETH Zurich)

Unidentified Video Objects: A Benchmark for Dense, Open-World Segmentation 10756
Weiyao Wang (Facebook Research), Matt Feiszli (Facebook Research), Heng Wang (Facebook Research), and Du Tran (Facebook AI)

Omnidata: A Scalable Pipeline for Making Multi-Task Mid-Level Vision Datasets From 3D Scans 10766
Ainaz Eftekhar (Sharif University of Technology), Alexander Sax (University of California, Berkeley), Jitendra Malik (University of California at Berkeley), and Amir Zamir (Swiss Federal Institute of Technology, (EPFL))

End-to-End Video Instance Segmentation via Spatial-Temporal Graph Neural Networks 10777
Tao Wang (Shanghai Jiao Tong University), Ning Xu (Adobe Research), Kean Chen (Shanghai Jiao Tong University), and Weiyao Lin (Shanghai Jiao Tong university)

Separable Flow: Learning Motion Cost Volumes for Optical Flow Estimation 10787
Feihu Zhang (University of Oxford), Oliver J. Woodford (Snap Inc), Victor Adrian Prisacariu (University of Oxford), and Philip H.S. Torr (University of Oxford)

ORBIT: A Real-World Few-Shot Dataset for Teachable Object Recognition 10798
Daniela Massiceti (Microsoft Research), Luisa Zintgraf (University of Oxford), John Bronskill (University of Cambridge), Lida Theodorou (City University of London), Matthew Tobias Harris (City University of London), Edward Cutrell (Microsoft), Cecily Morrison, Katja Hofmann (Microsoft Research), and Simone Stumpf (City University of London)

Wanderlust: Online Continual Object Detection in the Real World 10809
Jianren Wang (Carnegie Mellon University), Xin Wang (UC Berkeley), Yue Shang-Guan (Zhejiang University), and Abhinav Gupta (CMU/FAIR)

Learning To Adversarially Blur Visual Object Tracking 10819
Qing Guo (College of Intelligence and Computing, Tianjin University), Ziyi Cheng (Kyushu University), Felix Juefei-Xu (Alibaba Group, USA), Lei Ma (University of Alberta), Xiaofei Xie (Nanyang Technological University), Yang Liu (Nanyang Technology University, Singapore), and Jianjun Zhao (Kyushu University)

MOTSynth: How Can Synthetic Data Help Pedestrian Detection and Tracking? 10829
Matteo Fabbri (University of Modena and Reggio Emilia), Guillem Brasó (TUM), Gianluca Maugeri (University of Modena and Reggio Emilia), Orcun Cetintas (Technical University of Munich), Riccardo Gasparini (University of Modena and Reggio Emilia), Aljoša Ošep (TUM Munich), Simone Calderara (University of Modena and Reggio Emilia, Italy), Laura Leal-Taixé (TUM), and Rita Cucchiara (Università di Modena e Reggio Emilia)

Learning To Track With Object Permanence 10840
Pavel Tokmakov (Toyota Research Institute), Jie Li (Toyota Research Institute), Wolfram Burgard (Toyota Research Institute), and Adrien Gaidon (Toyota Research Institute)

StereOBJ-1M: Large-Scale Stereo Image Dataset for 6D Object Pose Estimation 10850
Xingyu Liu (Carnegie Mellon University), Shun Iwase (Carnegie Mellon University), and Kris M. Kitani (Carnegie Mellon University)

SynFace: Face Recognition With Synthetic Data 10860
Haibo Qiu (The University of Sydney), Baosheng Yu (The University of Sydney), Dihong Gong (Tencent AI Lab), Zhifeng Li (Tencent AI Lab), Wei Liu (Tencent), and Dacheng Tao (The University of Sydney)

UltraPose: Synthesizing Dense Pose With 1 Billion Points by Human-Body Decoupling 3D Model..... 10871
Haonan Yan (Beijing Momo Technology Co., Ltd.), Jiaqi Chen (Sun Yat-sen University), Xujie Zhang (Sun Yat-sen University), Shengkai Zhang (Beijing Momo Technology Co., Ltd.), Nianhong Jiao (Beijing Momo Technology Co., Ltd.), Xiaodan Liang (Sun Yat-sen University), and Tianxiang Zheng (Beijing Momo Technology Co., Ltd.)

Common Objects in 3D: Large-Scale Learning and Evaluation of Real-Life 3D Category Reconstruction 10881
Jeremy Reizenstein (Facebook AI Research), Roman Shapovalov (University College London), Philipp Henzler (Facebook AI Research), Luca Sbordone (Facebook AI), Patrick Labatut (Facebook AI Research), and David Novotny (Facebook AI Research)

Hypersim: A Photorealistic Synthetic Dataset for Holistic Indoor Scene Understanding 10892
Mike Roberts (Apple), Jason Ramapuram (Apple Inc), Anurag Ranjan (Apple), Atulit Kumar (Apple), Miguel Angel Bautista (Apple), Nathan Paczan (Apple), Russ Webb (Apple), and Joshua M. Susskind (Apple)

Towards Real-World X-Ray Security Inspection: A High-Quality Benchmark and Lateral Inhibition Module for Prohibited Items Detection 10903
Renshuai Tao (Beihang University), Yanlu Wei (Beihang University), Xiangjian Jiang (Beihang University), Hainan Li (Beihang University), Haotong Qin (Beihang University), Jiakai Wang (Beihang University), Yuqing Ma (BUAA), Libo Zhang (Institute of Software Chinese Academy of Sciences), and Xianglong Liu (Beihang University)

3D-FRONT: 3D Furnished Rooms With layOuts and semaNTics 10913
Huan Fu (Alibaba Group), Bowen Cai (Alibaba Group), Lin Gao (Institute of Computing Technology, Chinese Academy of Sciences), Ling-Xiao Zhang (Institute of Computing Technology, Chinese Academy of Sciences), Jiaming Wang (Alibaba Group), Cao Li (Alibaba Group), Qixun Zeng (Alibaba Group), Chengyue Sun (Alibaba Group), Rongfei Jia (Alibaba Group), Binqiang Zhao (Alibaba), and Hao Zhang (Simon Fraser University)

BV-Person: A Large-Scale Dataset for Bird-View Person Re-Identification 10923
Cheng Yan (Tianjin University), Guansong Pang (University of Adelaide), Lei Wang (University of Wollongong, Australia), Jile Jiao (Alibaba Group), Xuetao Feng (Alibaba Group), Chunhua Shen (University of Adelaide), and Jingjing Li (University of Alberta)

FloW: A Dataset and Benchmark for Floating Waste Detection in Inland Waters 10933
Yuwei Cheng (Tsinghua University), Jiannan Zhu (Northwestern Polytechnical University), Mengxin Jiang (Tsinghua University), Jie Fu (University of Montreal, Mila), Changsong Pang (Northwestern Polytechnical University), Peidong Wang (Tsinghua University), Kris Sankaran (University of Wisconsin-Madison), Olawale Onabola (MIla), Yimin Liu (Tsinghua University), Dianbo Liu (MIT), and Yoshua Bengio (Mila)

BioFors: A Large Biomedical Image Forensics Dataset 10943
Ekraam Sabir (USC ISI), Soumyaroop Nandi (Information Sciences Institute, University of Southern California), Wael Abd-Almageed (Information Sciences Institute), and Prem Natarajan (Amazon)

3D Point Clouds Stereo, and Multiview

The Power of Points for Modeling Humans in Clothing 10954
Qianli Ma (Max Planck Institute for Intelligent Systems), Jinlong Yang (Max Planck Institute for Intelligent Systems), Siyu Tang (ETH Zurich), and Michael J. Black (Max Planck Institute for Intelligent Systems)

Action-Conditioned 3D Human Motion Synthesis With Transformer VAE 10965
Mathis Petrovich (Ecole des Ponts), Michael J. Black (Max Planck Institute for Intelligent Systems), and Gül Varol (Ecole des Ponts ParisTech)

EventHPE: Event-Based 3D Human Pose and Shape Estimation 10976
Shihao Zou (University of Alberta), Chuan Guo (University of Alberta), Xinxin Zuo (University of Alberta), Sen Wang (University of Alberta), Pengyu Wang (Shandong University), Xiaoqin Hu (Celepixel Technology), Shoushun Chen (Celepixel Technology), Minglun Gong (University of Guelph), and Li Cheng (ECE dept., University of Alberta)

Egocentric Pose Estimation From Human Vision Span 10986
Hao Jiang (Facebook) and Vamsi Krishna Ithapu (Facebook Reality Labs)

Unsupervised 3D Pose Estimation for Hierarchical Dance Video Recognition 10995
Xiaodan Hu (University of Illinois at Urbana-Champaign) and Narendra Ahuja (University of Illinois at Urbana-Champaign, USA)

Human Pose Regression With Residual Log-Likelihood Estimation 11005
Jiefeng Li (Shanghai Jiao Tong University), Siyuan Bian (Shanghai Jiao Tong University), Ailing Zeng (The Chinese University of Hong Kong), Can Wang (SenseTime), Bo Pang (Shanghai Jiao Tong University), Wentao Liu (Sensetime), and Cewu Lu (Shanghai Jiao Tong University)

SPEC: Seeing People in the Wild With an Estimated Camera 11015
Muhammed Kocabas (Max Planck Institute for Intelligent Systems / ETH Zurich), Chun-Hao P. Huang (Max Planck Institute for Intelligent Systems), Joachim Tesch (Max Planck Institute for Intelligent Systems), Lea Müller (Max Planck Institute for Intelligent Systems), Otmar Hilliges (ETH Zurich), and Michael J. Black (Max Planck Institute for Intelligent Systems)

ARCH++: Animation-Ready Clothed Human Reconstruction Revisited 11026
Tong He (UCLA), Yuanlu Xu (Facebook Reality Labs), Shunsuke Saito (Facebook), Stefano Soatto (UCLA), and Tony Tung (Facebook Reality Labs)

Keypoint Communities 11037
Duncan Zauss (EPFL), Sven Kreiss (EPFL), and Alexandre Alahi (EPFL)

Removing the Bias of Integral Pose Regression 11047
Kerui Gu (National University of Singapore), Linlin Yang (University of Bonn), and Angela Yao (National University of Singapore)

Speech Drives Templates: Co-Speech Gesture Synthesis With Learned Templates 11057
Shenhan Qian (ShanghaiTech University), Zhi Tu (Shanghaitech University), Yihao Zhi (Shanghaitech University), Wen Liu (ShanghaiTech University), and Shenghua Gao (Shanghaitech University)

SignBERT: Pre-Training of Hand-Model-Aware Representation for Sign Language Recognition . 11067
Hezhen Hu (University of Science and Technology of China), Weichao Zhao (University of Science and Technology of China), Wengang Zhou (University of Science and Technology of China), Yuechen Wang (University of Science and Technology of China), and Houqiang Li (University of Science and Technology of China)

CPF: Learning a Contact Potential Field To Model the Hand-Object Interaction 11077
Lixin Yang (Shanghai Jiao Tong University), Xinyu Zhan (Shanghai Jiao Tong University), Kailin Li (Shanghai Jiao Tong University), Wenqiang Xu (Shanghai Jiao Tong University), Jiefeng Li (Shanghai Jiao Tong University), and Cewu Lu (Shanghai Jiao Tong University)

Hand-Object Contact Consistency Reasoning for Human Grasps Generation 11087
Hanwen Jiang (UCSD), Shaowei Liu (UCSD), Jiashun Wang (UCSD), and Xiaolong Wang (UCSD)

SOMA: Solving Optical Marker-Based MoCap Automatically 11097
Nima Ghorbani (Max Planck Institute for Intelligent Systems Tübingen) and Michael J. Black (Max Planck Institute for Intelligent Systems)

PARE: Part Attention Regressor for 3D Human Body Estimation 11107
Muhammed Kocabas (Max Planck Institute for Intelligent Systems / ETH Zurich), Chun-Hao P. Huang (Max Planck Institute for Intelligent Systems), Otmar Hilliges (ETH Zurich), and Michael J. Black (Max Planck Institute for Intelligent Systems)

Learning Realistic Human Reposing Using Cyclic Self-Supervision With 3D Shape, Pose, and Appearance Consistency 11118
Soubhik Sanyal (Max Planck Institute for Intelligent Systems), Alex Vorobiov (Amazon), Timo Bolkart (Max Planck Institute for Intelligent Systems), Matthew Loper (Amazon), Betty Mohler (Amazon Research Tuebingen), Larry S. Davis (AMAZON COM SERVICES INC), Javier Romero (Facebook), and Michael J. Black (Max Planck Institute for Intelligent Systems)

Graph-Based 3D Multi-Person Pose Estimation Using Multi-View Images 11128
Size Wu (University of Science and Technology of China), Sheng Jin (SenseTime Research), Wentao Liu (Sensetime), Lei Bai (University of Sydney), Chen Qian (SenseTime), Dong Liu (University of Science and Technology of China), and Wanli Ouyang (The University of Sydney)

Shape-Aware Multi-Person Pose Estimation From Multi-View Images 11138
Zijian Dong (ETH Zurich), Jie Song (ETH Zurich), Xu Chen (ETH Zürich), Chen Guo (ETH Zurich), and Otmar Hilliges (ETH Zurich)

Camera Distortion-Aware 3D Human Pose Estimation in Video With Optimization-Based Meta-Learning 11149
Hanbyel Cho (KAIST), Yooshin Cho (KAIST), Jaemyung Yu (KAIST), and Junmo Kim (KAIST)

Monocular, One-Stage, Regression of Multiple 3D People 11159
Yu Sun (Harbin Institute of technology), Qian Bao (AI Research of JD.com), Wu Liu (AI Research of JD.com), Yili Fu (HIT), Michael J. Black (Max Planck Institute for Intelligent Systems), and Tao Mei (AI Research of JD.com)

End-to-End Detection and Pose Estimation of Two Interacting Hands 11169
Dong Uk Kim (UNIST), Kwang In Kim (UNIST), and Seungryul Baek (UNIST)

Probabilistic Monocular 3D Human Pose Estimation With Normalizing Flows 11179
Tom Wehrbein (Leibniz University Hannover), Marco Rudolph (Leibniz University Hannover), Bodo Rosenhahn (Leibniz University Hannover), and Bastian Wandt (University of British Columbia)

Space-Time-Separable Graph Convolutional Network for Pose Forecasting 11189
Theodoros Sofianos (Sapienza University), Alessio Sampieri (Sapienza University), Luca Franco (Sapienza University), and Fabio Galasso (Sapienza University)

Hierarchical Kinematic Probability Distributions for 3D Human Shape and Pose Estimation From Images in the Wild 11199
Akash Sengupta (University of Cambridge), Ignas Budvytis (Department of Engineering, University of Cambridge), and Roberto Cipolla (University of Cambridge)

Self-Supervised 3D Hand Pose Estimation From Monocular RGB via Contrastive Learning 11210
Adrian Spurr (ETH Zurich), Aneesh Dahiya (ETH Zurich), Xi Wang (ETH Zurich), Xucong Zhang (ETH Zurich), and Otmar Hilliges (ETH Zurich)

An Empirical Study of the Collapsing Problem in Semi-Supervised 2D Human Pose Estimation . 11220
Rongchang Xie (Peking University), Chunyu Wang (Microsoft Research asia), Wenjun Zeng (Microsoft Research), and Yizhou Wang (PKU)

Learning To Regress Bodies From Images Using Differentiable Semantic Rendering 11230
Sai Kumar Dwivedi (Max Planck Institute for Intelligent Systems), Nikos Athanasiou (Max Planck Institute for Intelligent Systems), Muhammed Kocabas (Max Planck Institute for Intelligent Systems / ETH Zurich), and Michael J. Black (Max Planck Institute for Intelligent Systems)

HandFoldingNet: A 3D Hand Pose Estimation Network Using Multiscale-Feature Guided Folding of a 2D Hand Skeleton 11240
Wencan Cheng (Sungkyunkwan University), Jae Hyun Park (Sungkyunkwan University), and Jong Hwan Ko (Sungkyunkwan University)

Learning Causal Representation for Training Cross-Domain Pose Estimator via Generative Interventions 11250
Xiheng Zhang (Zhejiang University), Yongkang Wong (National University of Singapore), Xiaofei Wu (Noah's Ark Lab, Huawei Technologies Company, Ltd., China), Juwei Lu (Huawei Noah's Ark Laboratory), Mohan Kankanhalli (National University of Singapore), Xiangdong Li (Zhejiang University), and Weidong Geng (Zhejiang University)

Hand Image Understanding via Deep Multi-Task Learning 11261
Xiong Zhang (JOYY Inc.), Hongsheng Huang (Joyy), Jianchao Tan (Kwai Inc.), Hongmin Xu (Y-Lab, Kwai), Cheng Yang (JOYY Inc.), Guozhu Peng (JOYY, Inc), Lei Wang (JOYY Inc), and Ji Liu (Kwai Inc.)

Audio2Gestures: Generating Diverse Gestures From Speech Audio With Conditional Variational Autoencoders 11273
Jing Li (Harbin Institute of Technology, Shenzhen), Di Kang (Tencent), Wenjie Pei (Harbin Institute of Technology, Shenzhen), Xuefei Zhe (Tencent AI lab), Ying Zhang (Tencent), Zhenyu He (Harbin Institute of Technology, (Shenzhen)), and Linchao Bao (Tencent AI Lab)

Self-Mutual Distillation Learning for Continuous Sign Language Recognition 11283
Aiming Hao (Institute of Computing Technology, Chinese Academy of Sciences), Yuecong Min (Institute of Computing Technology, Chinese Academy of Sciences), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences)

TokenPose: Learning Keypoint Tokens for Human Pose Estimation 11293
Yanjie Li (Tsinghua University), Shoukui Zhang (Megvii), Zhicheng Wang (Megvii), Sen Yang (Southeast University), Wankou Yang (Southeast University), Shu-Tao Xia (Tsinghua University), and Erjin Zhou (Megvii Research)

The Animation Transformer: Visual Correspondence via Segment Matching 11303
Evan Casey (Cadmium), Víctor Pérez (Cadmium), and Zhuoru Li (Project HAT)

Contextually Plausible and Diverse 3D Human Motion Prediction 11313
Sadegh Aliakbarian (Microsoft), Fatemeh Saleh (Australian National University, (ANU)), Lars Petersson (Data61/CSIRO), Stephen Gould (Australian National University, Australia), and Mathieu Salzmann (EPFL)

Learning Motion Priors for 4D Human Body Capture in 3D Scenes 11323
Siwei Zhang (ETH Zurich), Yan Zhang (ETH Zurich), Federica Bogo (Microsoft), Marc Pollefeys (ETH Zurich / Microsoft), and Siyu Tang (ETH Zurich)

Interacting Two-Hand 3D Pose and Shape Reconstruction From Single Color Image 11334
Baowen Zhang (Institute of Software Chinese Academy of Sciences), Yangang Wang (Southeast University), Xiaoming Deng (Institute of Software, Chinese Academy of Sciences), Yinda Zhang (Google), Ping Tan (Simon Fraser University), Cuixia Ma (Institute of Software Chinese Academy of Sciences; University of Chinese Academy of Sciences), and Hongan Wang (Institute of Software, Chinese Academy of Sciences)

SemiHand: Semi-Supervised Hand Pose Estimation With Consistency 11344
Linlin Yang (University of Bonn), Shicheng Chen (National University of Singapore), and Angela Yao (National University of Singapore)

Stochastic Scene-Aware Motion Prediction 11354
Mohamed Hassan (Max Planck Institute for Intelligent Systems), Duygu Ceylan (Adobe Research), Ruben Villegas (Adobe Research), Jun Saito (Adobe), Jimei Yang (Adobe), Yi Zhou (University of Southern California), and Michael J. Black (Max Planck Institute for Intelligent Systems)

Neural Architecture Search for Joint Human Parsing and Pose Estimation 11365
Dan Zeng (Shanghai University), Yuhang Huang (Shanghai University), Qian Bao (AI Research of JD.com), Junjie Zhang (Shanghai University), Chi Su (Kingsoft Cloud), and Wu Liu (AI Research of JD.com)

Weakly Supervised Text-Based Person Re-Identification 11375
Shizhen Zhao (Huazhong University of Science and Technology), Changxin Gao (Huazhong University of Science and Technology), Yuanjie Shao (Huazhong University of Science and Technology), Wei-Shi Zheng (Sun Yat-sen University, China), and Nong Sang (Huazhong University of Science and Technology)

Single Image 3D Shape Retrieval via Cross-Modal Instance and Category Contrastive Learning.. 11385
Ming-Xian Lin (Institute of Computing Technology, Chinese Academy of Sciences), Jie Yang (Institute of Computing Technology, Chinese Academy of Sciences), He Wang (Stanford University), Yu-Kun Lai (Cardiff University), Rongfei Jia (Alibaba Group), Binqiang Zhao (Alibaba), and Lin Gao (Institute of Computing Technology, Chinese Academy of Sciences)

Learning Deep Local Features With Multiple Dynamic Attentions for Large-Scale Image Retrieval 11396
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), and Houqiang Li (University of Science and Technology of China)

Cherry-Picking Gradients: Learning Low-Rank Embeddings of Visual Data via Differentiable Cross-Approximation 11406
Mikhail Usvyatsov (ETH Zürich), Anastasia Makarova (ETH Zurich), Rafael Ballester-Ripoll (IE University), Maxim Rakhuba (HSE University), Andreas Krause (ETH Zürich), and Konrad Schindler (ETH)

Learning Skeletal Graph Neural Networks for Hard 3D Pose Estimation 11416
Ailing Zeng (The Chinese University of Hong Kong), Xiao Sun (Microsoft Research Asia), Lei Yang (Sensetime Group Limited), Nanxuan Zhao (Chinese University of Hong Kong), Minhao Liu (The Chinese University of Hong Kong), and Qiang Xu (The Chinese University of Hong Kong)

PyMAF: 3D Human Pose and Shape Regression With Pyramidal Mesh Alignment Feedback Loop 11426
Hongwen Zhang (Institute of Automation, Chinese Academy of Sciences), Yating Tian (Nanjing University), Xinchi Zhou (The University of Sydney), Wanli Ouyang (The University of Sydney), Yebin Liu (Tsinghua University), Limin Wang (Nanjing University), and Zhenan Sun (Chinese of Academy of Sciences)

Revitalizing Optimization for 3D Human Pose and Shape Estimation: A Sparse Constrained Formulation 11437
Taosha Fan (Northwestern University), Kalyan Vasudev Alwala (Facebook), Donglai Xiang (Carnegie Mellon University), Weipeng Xu (Facebook Reality Labs Research), Todd Murphey (Northwestern Univ.), and Mustafa Mukadam (Facebook AI Research)

MSR-GCN: Multi-Scale Residual Graph Convolution Networks for Human Motion Prediction .. 11447
Lingwei Dang (South China University of Technology), Yongwei Nie (South China University of Technology), Chengjiang Long (JD Finance America Corporation), Qing Zhang (Sun Yat-sen University), and Guiqing Li (South China University of Technology)

Modulated Graph Convolutional Network for 3D Human Pose Estimation 11457
Zhiming Zou (University of Illinois at Chicago) and Wei Tang (University of Illinois at Chicago)

HuMoR: 3D Human Motion Model for Robust Pose Estimation 11468
Davis Rempe (Stanford University), Tolga Birdal (TU Munich), Aaron Hertzmann (Adobe Research), Jimei Yang (Adobe), Srinath Sridhar (Brown University), and Leonidas J. Guibas (Stanford University)

Estimating Egocentric 3D Human Pose in Global Space 11480
Jian Wang (Max Planck Institute for Informatics), Lingjie Liu (Max Planck Institute for Informatics), Weipeng Xu (Facebook Reality Labs Research), Kripasindhu Sarkar (Max Plank Institute for Informatics), and Christian Theobalt (MPI Informatik)

EM-POSE: 3D Human Pose Estimation From Sparse Electromagnetic Trackers 11490
Manuel Kaufmann (ETH Zurich), Yi Zhao (Facebook Reality Labs), Chengcheng Tang (Facebook Reality Labs), Lingling Tao (Facebook Reality Labs), Christopher Twigg (Facebook Reality Labs), Jie Song (ETH Zurich), Robert Wang (Facebook Reality Labs), and Otmar Hilliges (ETH Zurich)

Normalized Human Pose Features for Human Action Video Alignment 11501
Jingyuan Liu (Hong Kong Univerisity of Science and Technology), Mingyi Shi (Hong Kong University), Qifeng Chen (HKUST), Hongbo Fu (City University of Hong Kong), and Chiew-Lan Tai (Hong Kong University of Science & Technology)

Physics-Based Human Motion Estimation and Synthesis From Videos 11512
Kevin Xie (University of Toronto), Tingwu Wang (University of Toronto; Vector Institute), Umar Iqbal (NVIDIA Research), Yunrong Guo (NVIDIA), Sanja Fidler (University of Toronto, NVIDIA), and Florian Shkurti (University of Toronto)

Visual Alignment Constraint for Continuous Sign Language Recognition 11522
Yuecong Min (Institute of Computing Technology, Chinese Academy of Sciences), Aiming Hao (Institute of Computing Technology, Chinese Academy of Sciences), Xiujuan Chai (Agricultural Information Institute, Chinese), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences)

Aligning Subtitles in Sign Language Videos 11532
Hannah Bull (LIMSI, (CNRS)), Triantafyllos Afouras (University of Oxford), Gül Varol (Ecole des Ponts ParisTech), Samuel Albanie (University of Oxford), Liliane Momeni (University of Oxford), and Andrew Zisserman (University of Oxford)

TACo: Token-Aware Cascade Contrastive Learning for Video-Text Alignment 11542
Jianwei Yang (Microsoft Research), Yonatan Bisk (Carnegie Mellon University), and Jianfeng Gao (Microsoft Research)

Support-Set Based Cross-Supervision for Video Grounding 11553
Xinpeng Ding (The Hong Kong University of Science and Technology), Nannan Wang (Xidian University), Shiwei Zhang (DAMO Academy, Alibaba Group), De Cheng (Xidian University), Xiaomeng Li (HKUST), Ziyuan Huang (National University of Singapore), Mingqian Tang (Alibaba Group), and Xinbo Gao (Xidian University)

TeachText: CrossModal Generalized Distillation for Text-Video Retrieval 11563
Ioana Croitoru (Institute of Mathematics of the Romanian Academy), Simion-Vlad Bogolin (Institute of Mathematics of the Romanian Academy), Marius Leordeanu (University "Politehnica" of Bucharest), Hailin Jin (Adobe Research), Andrew Zisserman (University of Oxford), Samuel Albanie (University of Oxford), and Yang Liu (Peking University)

SNARF: Differentiable Forward Skinning for Animating Non-Rigid Neural Implicit Shapes 11574
Xu Chen (ETH Zürich), Yufeng Zheng (ETH Zurich), Michael J. Black (Max Planck Institute for Intelligent Systems), Otmar Hilliges (ETH Zurich), and Andreas Geiger (University of Tuebingen)

Probabilistic Modeling for Human Mesh Recovery 11585
Nikos Kolotouros (University of Pennsylvania), Georgios Pavlakos (UC Berkeley), Dinesh Jayaraman (University of Pennsylvania), and Kostas Daniilidis (University of Pennsylvania)

Deep Virtual Markers for Articulated 3D Shapes 11595
Hyomin Kim (POSTECH), Jungeon Kim (POSTECH), Jaewon Kam (POSTECH), Jaesik Park (POSTECH), and Seungyong Lee (POSTECH)

Self-Supervised Transfer Learning for Hand Mesh Recovery From Binocular Images 11606
Zheng Chen (Dalian University of Technology), Sihan Wang (Dalian University of Technology), Yi Sun (Dalian University of Technology), and Xiaohong Ma (Dalian University of Technology)

Neural TMDlayer: Modeling Instantaneous Flow of Features via SDE Generators 11615
Zihang Meng (University of Wisconsin Madison), Vikas Singh (University of Wisconsin Madison), and Sathya N. Ravi (University of Illinois at Chicago)

A Unified 3D Human Motion Synthesis Model via Conditional Variational Auto-Encoder 11625
Yujun Cai (Nanyang Technological University), Yiwei Wang (National University of Singapore), Yiheng Zhu (ByteDance AI Lab), Tat-Jen Cham (Nanyang Technological University), Jianfei Cai (Monash University), Junsong Yuan (State University of New York at Buffalo, USA), Jun Liu (Singapore University of Technology and Design), Chuanxia Zheng (Nanyang Technological University), Sijie Yan (Chinese University of Hong Kong), Henghui Ding (Nanyang Technological University), Xiaohui Shen (ByteDance), Ding Liu (Bytedance), and Nadia Magnenat Thalmann (Nanyang Technological University)

3D Human Pose Estimation With Spatial and Temporal Transformers 11636
Ce Zheng (University of Central Florida), Sijie Zhu (University of Central Florida), Matias Mendieta (University of Central Florida), Taojiannan Yang (University of Central Florida), Chen Chen (University of Central Florida), and Zhengming Ding (Tulane University)

TravelNet: Self-Supervised Physically Plausible Hand Motion Learning From Monocular Color Images 11646
Zimeng Zhao (Southeast University), Xi Zhao (Southeast University), and Yangang Wang (Southeast University)

DECA: Deep Viewpoint-Equivariant Human Pose Estimation Using Capsule Autoencoders 11657
Nicola Garau (UNITN), Niccolò Bisagno (Università di Trento), Piotr Bródka (UNITN), and Nicola Conci (UNITN)

Full-Body Motion From a Single Head-Mounted Device: Generating SMPL Poses From Partial Observations 11667
Andrea Dittadi (Technical University of Denmark), Sebastian Dziadzio (Microsoft), Darren Cosker (University of Bath), Ben Lundell (Microsoft), Thomas J. Cashman (Microsoft), and Jamie Shotton (Microsoft)

Towards Accurate Alignment in Real-Time 3D Hand-Mesh Reconstruction 11678
Xiao Tang (The Chinese University of Hong Kong), Tianyu Wang (The Chinese University of Hong Kong), and Chi-Wing Fu (The Chinese University of Hong Kong)

Neural-GIF: Neural Generalized Implicit Functions for Animating People in Clothing 11688
Garvita Tiwari (MPI-INF), Nikolaos Sarafianos (Facebook Reality Labs), Tony Tung (Facebook Reality Labs), and Gerard Pons-Moll (University of Tübingen)

Motion Adaptive Pose Estimation From Compressed Videos 11699
Zhipeng Fan (NYU TANDON SCHOOL of ENGINEERING), Jun Liu (Singapore University of Technology and Design), and Yao Wang (New York University)

DensePose 3D: Lifting Canonical Surface Maps of Articulated Objects to the Third Dimension.... 11709
Roman Shapovalov (Facebook AI Research), David Novotny (Facebook AI Research), Benjamin Graham (Facebook Research), Patrick Labatut (Facebook AI Research), and Andrea Vedaldi (University of Oxford / Facebook AI Research)

Online Knowledge Distillation for Efficient Pose Estimation 11720
Zheng Li (Hangzhou Normal University), Jingwen Ye (Zhejiang University), Mingli Song (Zhejiang University), Ying Huang (Hangzhou Normal University), and Zhigeng Pan (Nanjing University of Information Science and Technology)

Improving Robustness of Facial Landmark Detection by Defending Against Adversarial Attacks 11731
Congcong Zhu (Shanghai University), Xiaoqiang Li (Shanghai University), Jide Li (Shanghai University), and Songmin Dai (Shanghai University)

Ranking Models in Unlabeled New Environments 11741
Xiaoxiao Sun (Australian National University), Yunzhong Hou (Australian National University), Weijian Deng (Australian National University), Hongdong Li (Australian National University, Australia), and Liang Zheng (Australian National University)

DOLG: Single-Stage Image Retrieval With Deep Orthogonal Fusion of Local and Global Features 11752
Min Yang (Baidu), Dongliang He (Baidu), Miao Fan (Baidu), Baorong Shi (Institute of Computing Technology, Chinese Academy of Science; University of Chinese Academy of Sciences), Xuetong Xue (Baidu), Fu Li (Baidu), Errui Ding (Baidu Inc.), and Jizhou Huang (Baidu)

Product1M: Towards Weakly Supervised Instance-Level Product Retrieval via Cross-Modal Pretraining 11762
Xunlin Zhan (Sun Yat-Sen University), Yangxin Wu (Sun Yat-sen University), Xiao Dong (Sun Yat-sen University), Yunchao Wei (UTS), Minlong Lu (Alibaba Group), Yichi Zhang (Alibaba Group), Hang Xu (Huawei Noah's Ark Lab), and Xiaodan Liang (Sun Yat-sen University)

Learning With Memory-Based Virtual Classes for Deep Metric Learning 11772
Byungsoo Ko (NAVER/LINE Corp.), Geonmo Gu (NAVER corp), and Han-Gyu Kim (NAVER)

TransPose: Keypoint Localization via Transformer 11782
Sen Yang (Southeast University), Zhibin Quan (Southeast University), Mu Nie (Southeast University), and Wankou Yang (Southeast University)

Explainable Person Re-Identification With Attribute-Guided Metric Distillation 11793
Xiaodong Chen (University of science and technology of China), Xinchen Liu (AI Research of JD.com), Wu Liu (AI Research of JD.com), Xiao-Ping Zhang (University of science and technology of China), Yongdong Zhang (Ryerson University), and Tao Mei (AI Research of JD.com)

CM-NAS: Cross-Modality Neural Architecture Search for Visible-Infrared Person Re-Identification 11803
Chaoyou Fu (Institute of Automation, Chinese Academy of Sciences), Yibo Hu (JD AI Research), Xiang Wu (Reconova), Hailin Shi (JD AI Research), Tao Mei (JD.com), and Ran He (Institute of Automation, Chinese Academy of Sciences)

Occlude Them All: Occlusion-Aware Attention Network for Occluded Person Re-ID 11813
Peixian Chen (Xiamen University), Wenfeng Liu (xmu), Pingyang Dai (Xiamen University), Jianzhuang Liu (Noah's Ark Lab, Huawei Technologies), Qixiang Ye (University of Chinese Academy of Sciences, China), Mingliang Xu (Zhengzhou University), Qi'an Chen (Xiamen University), and Rongrong Ji (Xiamen University, China)

Memory-Augmented Dynamic Neural Relational Inference 11823
Dong Gong (The University of Adelaide), Zhen Zhang (University of Adelaide), Javen Qinfeng Shi (University of Adelaide), and Anton van den Hengel (University of Adelaide)

The Center of Attention: Center-Keypoint Grouping via Attention for Multi-Person Pose Estimation 11833
Guillem Brasó (Technical University Munich), Nikita Kister (Technical University of Munich), and Laura Leal-Taixé (TUM)

IDM: An Intermediate Domain Module for Domain Adaptive Person Re-ID 11844
Yongxing Dai (Peking University), Jun Liu (Singapore University of Technology and Design), Yifan Sun (Megvii Technology), Zekun Tong (National University of Singapore), Chi Zhang (Megvii Inc.), and Ling-Yu Duan (Peking University)

Occluded Person Re-Identification With Single-Scale Global Representations 11855
Cheng Yan (Beihang University), Guansong Pang (University of Adelaide), Jile Jiao (Alibaba Group), Xiao Bai (Beihang University), Xuetao Feng (Alibaba Group), and Chunhua Shen (University of Adelaide)

Learning To Know Where To See: A Visibility-Aware Approach for Occluded Person Re-Identification 11865
Jinrui Yang (SUN YAT-SEN UNIVERSITY), Jiawei Zhang (Sun Yat-sen University), Fufu Yu (Tencent Youtu), Xinyang Jiang (Tencent), Mengdan Zhang (Youtu, Tencent), Xing Sun (Tencent), Ying-Cong Chen (Massachusetts Institute of Technology), and Wei-Shi Zheng (Sun Yat-sen University, China)

Clothing Status Awareness for Long-Term Person Re-Identification 11875
Yan Huang (University of Technology Sydney), Qiang Wu (University of Technology Sydney), JingSong Xu (University of Technology Sydney), Yi Zhong (Beijing Institution of Technology), and ZhaoXiang Zhang (Chinese Academy of Sciences, China)

Beyond Road Extraction: A Dataset for Map Update Using Aerial Images 11885
Favyen Bastani (MIT CSAIL) and Samuel Madden (MIT)

HiT: Hierarchical Transformer With Momentum Contrast for Video-Text Retrieval 11895
Song Liu (Peking University), Haoqi Fan (Facebook AI Research), Shengsheng Qian (Institute of Automation, Chinese Academy of Sciences), Yiru Chen (Kuaishou), Wenkui Ding (Kuaishou), and Zhongyuan Wang (Kuaishou)

ALADIN: All Layer Adaptive Instance Normalization for Fine-Grained Style Similarity 11906
Dan Ruta (University of Surrey), Saeid Motiian (Adobe), Baldo Faieta (Adobe), Zhe Lin (Adobe Research), Hailin Jin (Adobe Research), Alex Filipkowski (Adobe), Andrew Gilbert (University of Surrey), and John Collomosse (Adobe Research)

Rethinking Spatial Dimensions of Vision Transformers 11916
Byeongho Heo (NAVER AI LAB), Sangdoo Yun (NAVER AI LAB), Dongyoon Han (NAVER AI Lab), Sanghyuk Chun (NAVER AI LAB), Junsuk Choe (Sogang University), and Seong Joon Oh (NAVER AI Lab)

Stochastic Transformer Networks With Linear Competing Units: Application To End-to-End SL Translation 11926
Andreas Voskou (Cyprus University of Technology), Konstantinos P. Panousis (Cyprus University of Technology), Dimitrios Kosmopoulos (University of Patras), Dimitris N. Metaxas (Rutgers), and Sotirios Chatzis (Cyprus University of Technology)

Building-GAN: Graph-Conditioned Architectural Volumetric Design Generation 11936
Kai-Hung Chang (Autodesk Research), Chin-Yi Cheng (Autodesk Research), Jieliang Luo (Autodesk Research), Shingo Murata (Obayashi), Mehdi Nourbakhsh (Autodesk Research), and Yoshito Tsuji (Obayashi AI Design Lab)

LIRA: Learnable, Imperceptible and Robust Backdoor Attacks 11946
Khoa Doan (Virginia Tech), Yingjie Lao (Clemson University), Weijie Zhao (Baidu Research), and Ping Li (Baidu)

Inferring High-Resolution Traffic Accident Risk Maps Based on Satellite Imagery and GPS Trajectories 11957
Songtao He (Massachusetts Institute of Technology), Mohammad Amin Sadeghi (HBKU QCRI), Sanjay Chawla (QCRI), Mohammad Alizadeh (MIT CSAIL), Hari Balakrishnan (Massachusetts Institute of Technology), and Samuel Madden (MIT)

EC-DARTS: Inducing Equalized and Consistent Optimization Into DARTS 11966
Qinqin Zhou (Xiamen University), Xiawu Zheng (Xiamen University), Liujuan Cao (Xiamen University), Bineng Zhong (Guangxi Normal University), Teng Xi (Baidu Inc.), Gang Zhang (Baidu Inc.), Errui Ding (Baidu Inc.), Mingliang Xu (Zhengzhou University), and Rongrong Ji (Xiamen University, China)

PT-CapsNet: A Novel Prediction-Tuning Capsule Network Suitable for Deeper Architectures 11976
Chenbin Pan (Syracuse University) and Senem Velipasalar (Syracuse University)

Weakly Supervised Person Search With Region Siamese Networks 11986
Chuchu Han (HUST), Kai Su (Bytedance), Dongdong Yu (ByteDance), Zehuan Yuan (Bytedance.Inc), Changxin Gao (Huazhong University of Science and Technology), Nong Sang (Huazhong University of Science and Technology), Yi Yang (UTS), and Changhu Wang (ByteDance.Inc)

ASMR: Learning Attribute-Based Person Search With Adaptive Semantic Margin Regularizer ... 11996
Boseung Jeong (POSTECH), Jicheol Park (POSTECH), and Suha Kwak (POSTECH)

Pyramid Spatial-Temporal Aggregation for Video-Based Person Re-Identification 12006
Yingquan Wang (JiangSu University), Pingping Zhang (Dalian University of Technology), Shang Gao (Dalian University of Technology), Xia Geng (Jiangsu University), Hu Lu (jiangsu university), and Dong Wang (Dalian University of Technology)

Video-Based Person Re-Identification With Spatial and Temporal Memory Networks 12016
Chanho Eom (Yonsei University), Geon Lee (Yonsei University), Junghyup Lee (Yonsei University), and Bumsub Ham (Yonsei University)

Learning by Aligning: Visible-Infrared Person Re-Identification Using Cross-Modal Correspondences 12026
Hyunjong Park (Yonsei University), Sanghoon Lee (Yonsei University), Junghyup Lee (Yonsei University), and Bumsub Ham (Yonsei University)

Universal Cross-Domain Retrieval: Generalizing Across Classes and Domains 12036
Soumava Paul (IIT Kharagpur), Titir Dutta (Indian Institute of Science, Bangalore), and Soma Biswas (Indian Institute of Science, Bangalore)

Deep Relational Metric Learning 12045
Wenzhao Zheng (Tsinghua University), Borui Zhang (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)

Deep Symmetric Network for Underexposed Image Enhancement With Recurrent Attentional Learning 12055
Lin Zhao (Nankai University), Shao-Ping Lu (Nankai University), Tao Chen (Elephant Technologies), Zhenglu Yang (Nankai University), and Ariel Shamir (The Interdisciplinary Center)

Self-Supervised Product Quantization for Deep Unsupervised Image Retrieval 12065
Young Kyun Jang (Seoul National University) and Nam Ik Cho (Seoul National University)

Learning Specialized Activation Functions With the Piecewise Linear Unit 12075
Yucong Zhou (Huawei), Zezhou Zhu (Beijing University of Posts and Telecommunications), and Zhao Zhong (HUAWEI)

Instance-Level Image Retrieval Using Reranking Transformers 12085
Fuwen Tan (University of Virginia), Jiangbo Yuan (eBay Inc.), and Vicente Ordonez (Rice University)

Face Image Retrieval With Attribute Manipulation 12096
Alireza Zaeemzadeh (University of Central Florida), Shabnam Ghadar (Adobe), Baldo Faieta (Adobe), Zhe Lin (Adobe Research), Nazanin Rahnavard (University of Central Florida), Mubarak Shah (University of Central Florida), and Ratheesh Kalarot (Adobe)

Video Geo-Localization Employing Geo-Temporal Feature Learning and GPS Trajectory Smoothing 12106
Krishna Regmi (University of Central Florida) and Mubarak Shah (University of Central Florida)

Telling the What While Pointing to the Where: Multimodal Queries for Image Retrieval 12116
Soravit Changpinyo (Google Research), Jordi Pont-Tuset (Google), Vittorio Ferrari (Google Research), and Radu Soricut (Google)

Learning Attribute-Driven Disentangled Representations for Interactive Fashion Retrieval 12127
Yuxin Hou (Aalto University), Eleonora Vig (Amazon), Michael Donoser (Amazon), and Loris Bazzani (Amazon)

Bayesian Triplet Loss: Uncertainty Quantification in Image Retrieval 12138
Frederik Warburg (The Technical University of Denmark), Martin Jørgensen (Technical University of Denmark), Javier Civera (Universidad de Zaragoza), and Søren Hauberg (Technical University of Denmark, Denmark)

Viewpoint Invariant Dense Matching for Visual Geolocalization 12149
Gabriele Berton (Istituto Italiano di Tecnologia), Carlo Masone (Istituto Italiano di Tecnologia), Valerio Paolicelli (Istituto Italiano di Tecnologia), and Barbara Caputo (Politecnico di Torino)

Vision Transformers for Dense Prediction 12159
René Ranftl (Intel Labs), Alexey Bochkovskiy (Intel ISL), and Vladlen Koltun (Intel Labs)

Robustness via Cross-Domain Ensembles 12169
Teresa Yeo (EPFL), Oğuzhan Fatih Kar (EPFL), and Amir Zamir (Swiss Federal Institute of Technology, (EPFL))

Vector Neurons: A General Framework for SO(3)-Equivariant Networks 12180
Congyue Deng (Stanford University), Or Litany (NVIDIA), Yueqi Duan (Stanford University), Adrien Poulenard (Stanford), Andrea Tagliasacchi (Google Inc.), and Leonidas J. Guibas (Stanford University)

Product Quantizer Aware Inverted Index for Scalable Nearest Neighbor Search 12190
Haechan Noh (Sungkyunkwan University), Taeho Kim (SungKyunkwan University), and Jae-Pil Heo (Sungkyunkwan University)

Direct Differentiable Augmentation Search 12199
Aoming Liu (ETH Zurich), Zehao Huang (TuSimple), Zhiwu Huang (Singapore Management University), and Naiyan Wang (TuSimple)

Distilling Optimal Neural Networks: Rapid Search in Diverse Spaces 12209
Bert Moons (Qualcomm), Parham Noorzad (Qualcomm AI Research), Andrii Skliar (Qualcomm AI Research), Giovanni Mariani (Qualcomm AI Research), Dushyant Mehta (Qualcomm), Chris Lott (Qualcomm), and Tijmen Blankevoort (Qualcomm)

FairNAS: Rethinking Evaluation Fairness of Weight Sharing Neural Architecture Search 12219
Xiangxiang Chu (Meituan), Bo Zhang (Meituan), and Ruijun Xu (NA)

Homogeneous Architecture Augmentation for Neural Predictor 12229
Yuqiao Liu (Sichuan University), Yehui Tang (Peking University), and Yanan Sun (Sichuan university)

LeViT: A Vision Transformer in ConvNet's Clothing for Faster Inference 12239
Benjamin Graham (Facebook Research), Alaaeldin El-Nouby (Facebook AI Research), Hugo Touvron (Facebook AI Research), Pierre Stock (Facebook AI Research), Armand Joulin (Facebook AI Research), Hervé Jégou (Facebook AI Research), and Matthijs Douze (Facebook AI Research)

AutoFormer: Searching Transformers for Visual Recognition 12250
Minghao Chen (Stony Brook University), Houwen Peng (Microsoft Research), Jianlong Fu (Microsoft Research), and Haibin Ling (Stony Brook University)

BossNAS: Exploring Hybrid CNN-Transformers With Block-Wisely Self-Supervised Neural Architecture Search 12261
Changlin Li (Monash University), Tao Tang (Sun Yat-sen Universtiy), Guangrun Wang (University of Oxford), Jiefeng Peng (Sun Yat-sen University), Bing Wang (Alibaba group), Xiaodan Liang (Sun Yat-sen University), and Xiaojun Chang (RMIT University)

Unifying Nonlocal Blocks for Neural Networks 12272
Lei Zhu (Beijing University of Posts and Telecommunications), Qi She (Bytedance AI Lab), Duo Li (HKUST), Yanye Lu (Peking University), Xuejing Kang (Beijing University of Posts and Telecommunications), Jie Hu (Institute of Software Chinese Academy of Sciences), and Changhu Wang (ByteDance.Inc)

Adaptive Convolutions With Per-Pixel Dynamic Filter Atom 12282
Ze Wang (Purdue University), Zichen Miao (Purdue University), Jun Hu (Facebook), and Qiang Qiu (Purdue University)

Learning Latent Architectural Distribution in Differentiable Neural Architecture Search via Variational Information Maximization 12292
Yaoming Wang (Shanghai Jiao Tong University), Yuchen Liu (Shanghai Jiao Tong university), Wenrui Dai (Shanghai Jiao Tong University), Chenglin Li (Shanghai Jiao Tong University), Junni Zou (Shanghai Jiao Tong University), and Hongkai Xiong (Shanghai Jiao Tong University)

AdvRush: Searching for Adversarially Robust Neural Architectures 12302
Jisoo Mok (Seoul National University), Byunggook Na (Seoul National University), Hyeokjun Choe (Seoul National University), and Sungroh Yoon (Seoul National University)

Orthogonal Projection Loss 12313
Kanchana Ranasinghe (Stony Brook University), Muzammal Naseer (Australian National University, (ANU)), Munawar Hayat (Monash University), Salman Khan (Australian National University, (ANU)), and Fahad Shahbaz Khan (MBZUAI)

One-Pass Multi-View Clustering for Large-Scale Data 12324
Jiyuan Liu (National University of Defense Technology), Xinwang Liu (National University of Defense Technology), Yuexiang Yang (National University of Defense Technology), Li Liu (National University of Defense Technology), Siqi Wang (National University of Defense Technology), Weixuan Liang (National University of Defense Technology), and Jiangyong Shi (National University of Defense Technology)

Pi-NAS: Improving Neural Architecture Search by Reducing Supernet Training Consistency Shift 12334
Jiefeng Peng (Sun Yat-sen University), Jiqi Zhang (Sun Yat-sen University), Changlin Li (Monash University), Guangrun Wang (University of Oxford), Xiaodan Liang (Sun Yat-sen University), and Liang Lin (Sun Yat-sen University)

Low-level Vision and Computational Photography

Gravity-Aware Monocular 3D Human-Object Reconstruction 12345
Rishabh Dabral (IIT Bombay), Soshi Shimada (MPI for Informatics), Arjun Jain (Indian Institute of Technology Bombay), Christian Theobalt (MPI Informatik), and Vladislav Golyanik (MPI for Informatics)

Uncertainty-Aware Human Mesh Recovery From Video by Learning Part-Based 3D Dynamics .. 12355
Gun-Hee Lee (Korea University) and Seong-Whan Lee (Korea University)

EventHands: Real-Time Neural 3D Hand Pose Estimation From an Event Stream 12365
Viktor Rudnev (Max Planck Institute for Informatics), Vladislav Golyanik (MPI for Informatics), Jiayi Wang (Max Planck Institut Informatik), Hans-Peter Seidel (Max Planck Institute for Informatics), Franziska Mueller (Google Inc.), Mohamed Elgharib (Max Planck Institute for Informatics), and Christian Theobalt (MPI Informatik)

SO-Pose: Exploiting Self-Occlusion for Direct 6D Pose Estimation 12376
Yan Di (Technical University of Munich), Fabian Manhardt (Google), Gu Wang (Tsinghua University), Xiangyang Ji (Tsinghua University), Nassir Navab (Technische Universität München, Germany), and Federico Tombari (Google, TU Munich)

Single View Physical Distance Estimation Using Human Pose 12386
Xiaohan Fei (Amazon), Henry Wang (Amazon), Lin Lee Cheong (Amazon), Xiangyu Zeng (Columbia University), Meng Wang (Amazon), and Joseph Tighe (Amazon)

Reconstructing Hand-Object Interactions in the Wild 12397
Zhe Cao (UC Berkeley), Ilija Radosavovic (UC Berkeley), Angjoo Kanazawa (University of California Berkeley), and Jitendra Malik (University of California at Berkeley)

Structured Outdoor Architecture Reconstruction by Exploration and Classification 12407
Fuyang Zhang (Simon Fraser University), Xiang Xu (Simon Fraser University), Nelson Nauata (Simon Fraser University), and Yasutaka Furukawa (Simon Fraser University)

Sat2Vid: Street-View Panoramic Video Synthesis From a Single Satellite Image 12416
Zuoyue Li (ETH Zurich), Zhenqiang Li (The University of Tokyo), Zhaopeng Cui (Zhejiang University), Rongjun Qin (The Ohio State University), Marc Pollefeys (ETH Zurich / Microsoft), and Martin R. Oswald (ETH Zurich)

3DIAS: 3D Shape Reconstruction With Implicit Algebraic Surfaces 12426
Mohsen Yavartanoo (Seoul National University), Jaeyoung Chung (Seoul National University), Reyhaneh Neshatavar (Seoul National University), and Kyoung Mu Lee (Seoul National University)

3DStyleNet: Creating 3D Shapes With Geometric and Texture Style Variations 12436
Kangxue Yin (NVIDIA), Jun Gao (University of Toronto, Nvidia), Maria Shugrina (University of Toronto), Sameh Khamis (Nvidia), and Sanja Fidler (University of Toronto, NVIDIA)

Unsupervised Learning of Fine Structure Generation for 3D Point Clouds by 2D Projections Matching 12446
Chao Chen (Tsinghua University), Zhizhong Han (Wayne State University), Yu-Shen Liu (Tsinghua University), and Matthias Zwicker (University of Maryland)

CSG-Stump: A Learning Friendly CSG-Like Representation for Interpretable Shape Parsing 12458
Daxuan Ren (Nanyang Technological University), Jianmin Zheng (Nanyang Technological University), Jianfei Cai (Monash University), Jiatong Li (Sensetime), Haiyong Jiang (NTU), Zhongang Cai (SenseTime International Pte Ltd), Junzhe Zhang (Nanyang Technological University), Liang Pan (Nanyang Technological University), Mingyuan Zhang (Beijing SenseTime Technology Development Limited), Haiyu Zhao (SenseTime International Pte Ltd), and Shuai Yi (SenseTime Group Limited)

ME-PCN: Point Completion Conditioned on Mask Emptiness 12468
Bingchen Gong (The University of Hong Kong), Yinyu Nie (Technical University of Munich), Yiqun Lin (The Chinese University of Hong Kong, Shenzhen), Xiaoguang Han (Shenzhen Research Institute of Big Data, the Chinese University of Hong Kong, (Shenzhen)), and Yizhou Yu (The University of Hong Kong)

PoinTr: Diverse Point Cloud Completion With Geometry-Aware Transformers 12478
Xumin Yu (Tsinghua University), Yongming Rao (Tsinghua University), Ziyi Wang (Tsinghua University), Zuyan Liu (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)

RFNet: Recurrent Forward Network for Dense Point Cloud Completion 12488
Tianxin Huang (Zhejiang University), Hao Zou (Zhejiang University), Jinhao Cui (Zhejiang Unversity), Xuemeng Yang (Zhejiang University), Mengmeng Wang (Zhejiang University), Xiangrui Zhao (Zhejiang University), Jiangning Zhang (Zhejiang University), Yi Yuan (NetEase Fuxi AI Lab), Yifan Xu (NetEase Fuxi AI Lab), and Yong Liu (Zhejiang University)

SLIDE: Single Image 3D Photography With Soft Layering and Depth-Aware Inpainting 12498
Varun Jampani (Google), Huiwen Chang (Google), Kyle Sargent (Google), Abhishek Kar (Google), Richard Tucker (Google), Michael Krainin (Google), Dominik Kaeser (Google), William T. Freeman (Google), David Salesin (Google Research), Brian Curless (University of Washington), and Ce Liu (Google)

Worldsheet: Wrapping the World in a 3D Sheet for View Synthesis From a Single Image 12508
Ronghang Hu (Facebook), Nikhila Ravi (Facebook AI Research), Alexander C. Berg (University of North Carolina, USA), and Deepak Pathak (Carnegie Mellon University)

Learning Indoor Inverse Rendering With 3D Spatially-Varying Lighting 12518
Zian Wang (University of Toronto), Jonah Philion (University of Toronto, NVIDIA), Sanja Fidler (University of Toronto, NVIDIA), and Jan Kautz (NVIDIA)

3D Building Reconstruction From Monocular Remote Sensing Images 12528
Weijia Li (The Chinese University of Hong Kong), Lingxuan Meng (University of Electronic Science and Technology of China), Jinwang Wang (Wuhan University), Conghui He (Sensetime Inc.), Gui-Song Xia (Wuhan University), and Dahua Lin (The Chinese University of Hong Kong)

In-the-Wild Single Camera 3D Reconstruction Through Moving Water Surfaces 12538
Jinhui Xiong (KAUST) and Wolfgang Heidrich (KAUST)

RetrievalFuse: Neural 3D Scene Reconstruction With a Database 12548
Yawar Siddiqui (Technical University of Munich), Justus Thies (Max Planck Institute for Intelligent Systems), Fangchang Ma (Apple Inc.), Qi Shan (Apple Inc.), Matthias Nießner (Technical University of Munich), and Angela Dai (Technical University of Munich)

MINE: Towards Continuous Depth MPI With NeRF for Novel View Synthesis 12558
Jiaxin Li (Bytedance, TikTok), Zijian Feng (ByteDance AI Lab), Qi She (Bytedance AI Lab), Henghui Ding (Nanyang Technological University), Changhu Wang (ByteDance.Inc), and Gim Hee Lee (National University of Singapore)

Patch2CAD: Patchwise Embedding Learning for In-the-Wild Shape Retrieval From a Single Image 12569
Weicheng Kuo (Google), Anelia Angelova (Google), Tsung-Yi Lin (Google Brain), and Angela Dai (Technical University of Munich)

Toward Realistic Single-View 3D Object Reconstruction With Unsupervised Learning From Multiple Images 12580
Long-Nhat Ho (VinAI Research), Anh Tuan Tran (VinAI Research), Quynh Phung (Vinai Research), and Minh Hoai (Stony Brook University)

Geometric Deep Neural Network Using Rigid and Non-Rigid Transformations for Human Action Recognition 12591
Rasha Friji (ENSI Tunisia\ Talan Tunisia), Hassen Drira (cristal-IMT Lille Douai), Faten Chaieb (INSAT), Hamza Kchok (INSAT), and Sebastian Kurtek (The Ohio State University)

Bridging Unsupervised and Supervised Depth From Focus via All-in-Focus Supervision 12601
Ning-Hsu Wang (National Tsing Hua University; MediaTek), Ren Wang (MediaTek Inc.), Yu-Lun Liu (National Taiwan University; MediaTek), Yu-Hao Huang (Mediatek Inc.), Yu-Lin Chang (MediaTek Inc.), Chia-Ping Chen (MediaTek Inc.), and Kevin Jou (MediaTek Inc.)

DeepPanoContext: Panoramic 3D Scene Understanding With Holistic Scene Context Graph and Relation-Based Optimization 12612
Cheng Zhang (UESTC), Zhaopeng Cui (Zhejiang University), Cai Chen (University of Electronic Science and Technology of China), Shuaicheng Liu (UESTC; Megvii), Bing Zeng (University of Electronic Science and Technology of China), Hujun Bao (Zhejiang University), and Yinda Zhang (Google)

Fine-Grained Semantics-Aware Representation Enhancement for Self-Supervised Monocular Depth Estimation 12622
Hyunyoung Jung (Seoul National University), Eunhyeok Park (POSTECH), and Sungjoo Yoo (Seoul National University)

Deep Implicit Surface Point Prediction Networks 12633
Rahul Venkatesh (Carnegie Mellon University), Tejan Karmali (Indian Institute of Science, Bengaluru), Sarthak Sharma (Verisk Analytics), Aurobrata Ghosh (Verisk AI, Verisk Analytics), R. Venkatesh Babu (Indian Institute of Science), László A. Jeni (Carnegie Mellon University), and Maneesh Singh (Verisk Analytics)

StructDepth: Leveraging the Structural Regularities for Self-Supervised Indoor Depth Estimation 12643
Boying Li (Shanghai Jiao Tong university), Yuan Huang (Shang Hai Jiao Tong University), Zeyu Liu (Shanghai Jiao Tong University), Danping Zou (Shang Jiao Tong University), and Wenxian Yu (Shanghai Jiao Tong university)

Learning Anchored Unsigned Distance Functions With Gradient Direction Alignment for Single-View Garment Reconstruction 12654
Fang Zhao (Inception Institute of Artificial Intelligence), Wenhao Wang (Beihang University), Shengcai Liao (Inception Institute of Artificial Intelligence), and Ling Shao (Inception Institute of Artificial Intelligence)

NeRD: Neural Reflectance Decomposition From Image Collections 12664
Mark Boss (University of Tübingen), Raphael Braun (University of Tübingen), Varun Jampani (Google), Jonathan T. Barron (Google Research), Ce Liu (Google), and Hendrik P.A. Lensch (University of Tübingen)

NPMs: Neural Parametric Models for 3D Deformable Shapes 12675
Pablo Palafox (Technical University Munich), Aljaž Božič (Technical University Munich), Justus Thies (Max Planck Institute for Intelligent Systems), Matthias Nießner (Technical University of Munich), and Angela Dai (Technical University of Munich)

4DComplete: Non-Rigid Motion Estimation Beyond the Observable Surface 12686
Yang Li (The University of Tokyo), Hikari Takehara (Nara Institue of Science and Technology), Takafumi Taketomi (Cyberagent), Bo Zheng (Huawei), and Matthias Nießner (Technical University of Munich)

Holistic Pose Graph: Modeling Geometric Structure Among Objects in a Scene Using Graph Inference for 3D Object Prediction 12697
Jiwei Xiao (Institute of Computing Technology, Chinese Academy of Sciences), Ruiping Wang (Institute of Computing Technology, Chinese Academy of Sciences), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences)

Can Scale-Consistent Monocular Depth Be Learned in a Self-Supervised Scale-Invariant Manner? 12707
Lijun Wang (Dalian University of Technology), Yifan Wang (Dalian University of Technology), Linzhao Wang (Huawei Technologies Co., Ltd.), Yunlong Zhan (Huawei Technologies Co., Ltd.), Ying Wang (Huawei Technologies Co., Ltd.), and Huchuan Lu (Dalian University of Technology)

Self-Supervised Monocular Depth Estimation for All Day Images Using Domain Separation 12717
Lina Liu (Zhejiang University), Xibin Song (Baidu), Mengmeng Wang (Zhejiang University), Yong Liu (Zhejiang University), and Liangjun Zhang (baidu)

Unsupervised Depth Completion With Calibrated Backprojection Layers 12727
Alex Wong (University of California, Los Angeles) and Stefano Soatto (UCLA)

PX-NET: Simple and Efficient Pixel-Wise Training of Photometric Stereo Networks 12737
Fotios Logothetis (Toshiba research), Ignas Budvytis (Department of Engineering, University of Cambridge), Roberto Mecca (cambridge university), and Roberto Cipolla (University of Cambridge)

Boosting Monocular Depth Estimation With Lightweight 3D Point Fusion 12747
Lam Huynh (University of Oulu), Phong Nguyen (University of Oulu), Jiří Matas (CMP CTU FEE), Esa Rahtu (Tampere University), and Janne Heikkilä (University of Oulu, Finland)

R-MSFM: Recurrent Multi-Scale Feature Modulation for Monocular Depth Estimating 12757
Zhongkai Zhou (Hohai University), Xinnan Fan (Hohai University), Pengfei Shi (Hohai University), and Yuanxue Xin (Hohai University)

MonoIndoor: Towards Good Practice of Self-Supervised Monocular Depth Estimation for Indoor Environments 12767
Pan Ji (OPPO US Research Center), Runze Li (University of California Riverside), Bir Bhanu (University of California Riverside), and Yi Xu (OPPO US Research Center)

DnD: Dense Depth Estimation in Crowded Dynamic Indoor Scenes 12777
Dongki Jung (KAIST), Jaehoon Choi (University of Maryland), Yonghan Lee (NAVER LABS), Deokhwa Kim (Naverlabs), Changick Kim (KAIST), Dinesh Manocha (UMD), and Donghwan Lee (NAVER LABS)

Adaptive Confidence Thresholding for Monocular Depth Estimation 12788
Hyesong Choi (Ewha w. university), Hunsang Lee (Yonsei Univ.), Sunkyung Kim (Ewha Womans University), Sunok Kim (Korea Aerospace University), Seungryong Kim (Korea University), Kwanghoon Sohn (Yonsei Univ.), and Dongbo Min (Ewha Womans University)

Towards High Fidelity Monocular Face Reconstruction With Rich Reflectance Using Self-Supervised Learning and Ray Tracing 12799
Abdallah Dib (InterDigital), Cédric Thébault (InterDigital), Junghyun Ahn (InterDigital), Philippe-Henri Gosselin (InterDigital), Christian Theobalt (MPI Informatik), and Louis Chevallier (InterDigital)

Sparse Needlets for Lighting Estimation With Spherical Transport Loss 12810
Fangneng Zhan (Nanyang Technological University), Changgong Zhang (Alibaba Group), Wenbo Hu (The Chinese University of Hong Kong), Shijian Lu (Nanyang Technological University), Feiying Ma (alibaba), Xuansong Xie (Alibaba), and Ling Shao (Inception Institute of Artificial Intelligence)

SurfaceNet: Adversarial SVBRDF Estimation From a Single Image 12820
Giuseppe Vecchio (University of Catania), Simone Palazzo (University of Catania), and Concetto Spampinato (University of Catania)

Adaptive Surface Normal Constraint for Depth Estimation 12829
Xiaoxiao Long (The University of Hong Kong), Cheng Lin (The University of Hong Kong), Lingjie Liu (Max Planck Institute for Informatics), Wei Li (Peking University), Christian Theobalt (MPI Informatik), Ruigang Yang (University of Kentucky, USA), and Wenping Wang (The University of Hong Kong)

VaPiD: A Rapid Vanishing Point Detector via Learned Optimizers 12839
Shichen Liu (University of Southern California), Yichao Zhou (UC Berkeley), and Yajie Zhao (USC ICT)

GyroFlow: Gyroscope-Guided Unsupervised Optical Flow Learning 12849
Haipeng Li (megvii.com), Kunming Luo (Megvii), and Shuaicheng Liu (UESTC; Megvii)

Towards Interpretable Deep Networks for Monocular Depth Estimation 12859
Zunzhi You (Sun Yat-sen University), Yi-Hsuan Tsai (NEC Labs America), Wei-Chen Chiu (National Chiao Tung University), and Guanbin Li (Sun Yat-sen University)

Hierarchical Memory Matching Network for Video Object Segmentation 12869
Hongje Seong (Yonsei University), Seoung Wug Oh (Adobe Research), Joon-Young Lee (Adobe Research), Seongwon Lee (Yonsei university), Suhyeon Lee (Yonsei University), and Euntai Kim (Yonsei University)

PIAP-DF: Pixel-Interested and Anti Person-Specific Facial Action Unit Detection Net With Discrete Feedback Learning 12879
Yang Tang (Beijing University of Posts and Telecommunications), Wangding Zeng (Beijing University of Posts and Telecommunications), Dafei Zhao (University of Ottawa), and Honggang Zhang (Beijing University of Posts and Telecommunications)

Context-Sensitive Temporal Feature Learning for Gait Recognition 12889
Xiaohu Huang (Huazhong University of Science and Technology), Duowang Zhu (Huazhong University of Science and Technology), Hao Wang (Huazhong University of Science and Technology), Xinggang Wang (Huazhong University of Science and Technology), Bo Yang (Wuhan FiberHome Digital Technology Co., Ltd), Botao He (Wuhan FiberHome Digital Technology Co., Ltd), Wenyu Liu (Huazhong University of Science and Technology), and Bin Feng (Huazhong University of Science and Technology)

DeepGaze IIE: Calibrated Prediction in and Out-of-Domain for State-of-the-Art Saliency Modeling 12899
Akis Linardos (University of Barcelona), Matthias Kümmerer (University of Tübingen), Ori Press (University of Tübingen), and Matthias Bethge (University of Tübingen)

I2UV-HandNet: Image-to-UV Prediction Network for Accurate and High-Fidelity 3D Hand Mesh Modeling 12909
Ping Chen (iQIYI, Inc.), Yujin Chen (Wuhan University), Dong Yang (iQIYI, Inc.), Fangyin Wu (iQIYI, Inc.), Qin Li (iQIYI, Inc), Qingpei Xia (IQIYI, Inc.), and Yong Tan (IQIYI, Inc.)

Mesh Graphormer 12919
Kevin Lin (Microsoft), Lijuan Wang (Microsoft), and Zicheng Liu (Microsoft)

CodeNeRF: Disentangled Neural Radiance Fields for Object Categories 12929
Wonbong Jang (University College London) and Lourdes Agapito (University College London)

Non-Rigid Neural Radiance Fields: Reconstruction and Novel View Synthesis of a Dynamic Scene From Monocular Video 12939
Edgar Tretschk (Max-Planck-Institut für Informatik), Ayush Tewari (Max Planck Institute for Informatics), Vladislav Golyanik (MPI for Informatics), Michael Zollhöfer (Facebook Reality Labs), Christoph Lassner (Facebook), and Christian Theobalt (MPI Informatik)

THUNDR: Transformer-Based 3D Human Reconstruction With Markers 12951
Mihai Zanfir (IMAR), Andrei Zanfir (Google), Eduard Gabriel Bazavan (Google), William T. Freeman (Google), Rahul Sukthankar (Google), and Cristian Sminchisescu (Lund University)

Discovering 3D Parts From Image Collections 12961
Chun-Han Yao (University of California at Merced), Wei-Chih Hung (Waymo), Varun Jampani (Google), and Ming-Hsuan Yang (University of California at Merced)

Planar Surface Reconstruction From Sparse Views 12971
Linyi Jin (University of Michigan), Shengyi Qian (University of Michigan), Andrew Owens (University of Michigan), and David F. Fouhey (University of Michigan)

A-SDF: Learning Disentangled Signed Distance Functions for Articulated Shape Representation 12981
Jiteng Mu (University of California, San Diego), Weichao Qiu (Johns Hopkins University), Adam Kortylewski (Johns Hopkins University), Alan Yuille (Johns Hopkins University), Nuno Vasconcelos (UCSD, USA), and Xiaolong Wang (UCSD)

SIMstack: A Generative Shape and Instance Model for Unordered Object Stacks 12992
Zoe Landgraf (Imperial College London), Raluca Scona (Imperial College London), Tristan Laidlow (Imperial College London), Stephen James (Imperial College London), Stefan Leutenegger (Imperial College London), and Andrew J. Davison (Imperial College London)

Sketch2Mesh: Reconstructing and Editing 3D Shapes From Sketches 13003
Benoit Guillard (EPFL), Edoardo Remelli (epfl), Pierre Yvernay (EPFL), and Pascal Fua (EPFL, Switzerland)

Encoder-Decoder With Multi-Level Attention for 3D Human Shape and Pose Estimation 13013
Ziniu Wan (Carnegie Mellon University), Zhengjia Li (tongji university), Maoqing Tian (SenseTime Group Limited), Jianbo Liu (Chinese university of Hong Kong), Shuai Yi (SenseTime Group Limited), and Hongsheng Li (Chinese University of Hong Kong)

Object Tracking by Jointly Exploiting Frame and Event Domain 13023
Jiqing Zhang (Dalian University of Technology), Xin Yang (Dalian University of Technology), Yingkai Fu (Dalian University of Technology), Xiaopeng Wei (Dalian University of Technology), Baocai Yin (Dalian University of Technology), and Bo Dong (SRI International)

Learning an Augmented RGB Representation With Cross-Modal Knowledge Distillation for Action Detection 13033
Rui Dai (INRIA), Srijan Das (Stony Brook University), and François Bremond (Inria Sophia Antipolis, France)

Learning Self-Similarity in Space and Time As Generalized Motion for Video Action Recognition 13045
Heeseung Kwon (POSTECH), Manjin Kim (POSTECH), Suha Kwak (POSTECH), and Minsu Cho (POSTECH)

Motion Guided Attention Fusion To Recognize Interactions From Videos 13056
Tae Soo Kim (Johns Hopkins University), Jonathan Jones (Johns Hopkins University), and Gregory D. Hager (The Johns Hopkins University)

Multiresolution Deep Implicit Functions for 3D Shape Representation 13067
Zhang Chen (ShanghaiTech University), Yinda Zhang (Google), Kyle Genova (Google Research), Sean Fanello (Google), Sofien Bouaziz (Facebook Reality Labs), Christian Häne (Google), Ruofei Du (Google), Cem Keskin (Facebook), Thomas Funkhouser (Google Research), and Danhang Tang (Google)

Geometric Granularity Aware Pixel-To-Mesh 13077
Yue Shi (Shanghai Jiaotong University), Bingbing Ni (Shanghai Jiao Tong University), Jinxian Liu (Shanghai Jiao Tong University), Dingyi Rong (Shanghai Jiao Tong University), Ye Qian (Shanghai Jiao Tong University), and Wenjun Zhang (Shanghai Jiao Tong University)

Efficient and Differentiable Shadow Computation for Inverse Problems 13087
Linjie Lyu (MPII), Marc Habermann (Max Planck Institute for Informatics), Lingjie Liu (Max Planck Institute for Informatics), Mallikarjun B R (Max Planck Institute for Informatics), Ayush Tewari (Max Planck Institute for Informatics), and Christian Theobalt (MPI Informatik)

Motion Basis Learning for Unsupervised Deep Homography Estimation With Subspace Projection..... 13097
Nianjin Ye (Megvii), Chuan Wang (Face++, (Megvii)), Haoqiang Fan (Megvii Incface++), and Shuaicheng Liu (UESTC; Megvii)

SLIM: Self-Supervised LiDAR Scene Flow and Motion Segmentation 13106
Stefan Andreas Baur (Daimler AG), David Josef Emmerichs (Mercedes-Benz AG), Frank Moosmann (Daimler AG), Peter Pinggera (Tensoreye), Björn Ommer (Heidelberg University), and Andreas Geiger (University of Tuebingen)

Estimating and Exploiting the Aleatoric Uncertainty in Surface Normal Estimation 13117
Gwangbin Bae (University of Cambridge), Ignas Budvytis (Department of Engineering, University of Cambridge), and Roberto Cipolla (University of Cambridge)

On Exposing the Challenging Long Tail in Future Prediction of Traffic Actors 13127
Osama Makansi (University of Freiburg), Özgün Çiçek (University of Freiburg), Yassine Marrakchi (University of Freiburg), and Thomas Brox (University of Freiburg)

MG-GAN: A Multi-Generator Model Preventing Out-of-Distribution Samples in Pedestrian Trajectory Prediction 13138
Patrick Dendorfer (TUM), Sven Elflein (TUM), and Laura Leal-Taixé (TUM)

Unlimited Neighborhood Interaction for Heterogeneous Trajectory Prediction 13148
Fang Zheng (Xi'an Jiaotong University), Le Wang (Xi'an Jiaotong University), Sanping Zhou (Xi'an Jiaotong University), Wei Tang (University of Illinois at Chicago), Zhenxing Niu (Alibaba Group), Nanning Zheng (Xi'an Jiaotong University), and Gang Hua (Wormpex AI Research)

MEDIRL: Predicting the Visual Attention of Drivers via Maximum Entropy Deep Inverse Reinforcement Learning 13158
Sonia Baee (University of Virginia), Erfan Pakdamanian (University of Virginia), Inki Kim (University of Illinois at Urbana Champaign), Lu Feng (University of Virginia), Vicente Ordonez (Rice University), and Laura Barnes (University of Virginia)

Voxel-Based Network for Shape Completion by Leveraging Edge Generation 13169
Xiaogang Wang (National University of Singapore), Marcelo H Ang (National University of Singapore), and Gim Hee Lee (National University of Singapore)

Box-Aware Feature Enhancement for Single Object Tracking on Point Clouds 13179
Chaoda Zheng (The Chinese University of Hong Kong, Shen Zhen), Xu Yan (The Chinese University of Hong Kong, Shenzhen), Jiantao Gao (Shanghai University), Weibing Zhao (The Chinese University of Hong Kong, Shenzhen), Wei Zhang (Baidu Inc), Zhen Li (Chinese University of Hong Kong, Shenzhen), and Shuguang Cui (The Chinese University of Hong Kong, Shenzhen)

CAPTRA: CAtegory-Level Pose Tracking for Rigid and Articulated Objects From Point Clouds . 13189
Yijia Weng (Peking University), He Wang (Stanford University), Qiang Zhou (ShanDong University), Yuzhe Qin (University of California San Diego), Yueqi Duan (Stanford University), Qingnan Fan (Tencent AI Lab), Baoquan Chen (Peking University), Hao Su (UC, San Diego), and Leonidas J. Guibas (Stanford University)

A General Recurrent Tracking Framework Without Real Data 13199
Shuai Wang (Beihang University), Hao Sheng (Beihang University), Yang Zhang (Beijing University of Chemical Technology), Yubin Wu (Beihang University), and Zhang Xiong (Beihang University)

PCAM: Product of Cross-Attention Matrices for Rigid Registration of Point Clouds 13209
Anh-Quan Cao (Inria), Gilles Puy (Valeo), Alexandre Boulch (valeo.ai), and Renaud Marlet (Ecole des Ponts ParisTech)

M3D-VTON: A Monocular-to-3D Virtual Try-On Network 13219
Fuwei Zhao (Sun Yat-Sen University), Zhenyu Xie (Sun Yat-sen University), Michael Kampffmeyer (UiT The Arctic University of Norway), Haoye Dong (Sun Yat-sen University), Songfang Han (UCSD), Tianxiang Zheng (Beijing Momo Technology Co., Ltd.), Tao Zhang (Beijing Momo Technology Co., Ltd.), and Xiaodan Liang (Sun Yat-sen University)

Three Steps to Multimodal Trajectory Prediction: Modality Clustering, Classification and Synthesis 13230
Jianhua Sun (Shanghai Jiaotong Univ), Yuxuan Li (Shanghai Jiaotong University), Hao-Shu Fang (SJTU), and Cewu Lu (Shanghai Jiao Tong University)

Square Root Marginalization for Sliding-Window Bundle Adjustment 13240
Nikolaus Demmel (TU Munich), David Schubert (Technical University of Munich), Christiane Sommer (Technical University of Munich), Daniel Cremers (TU Munich), and Vladyslav Usenko (TU Munich)

Provably Approximated Point Cloud Registration 13249
Ibrahim Jubran (The University of Haifa), Alaa Maalouf (The University of Haifa), Ron Kimmel (Technion), and Dan Feldman (The University of Haifa)

Likelihood-Based Diverse Sampling for Trajectory Forecasting 13259
Yecheng Jason Ma (University of Pennsylvania), Jeevana Priya Inala (MIT), Dinesh Jayaraman (University of Pennsylvania), and Osbert Bastani (University of Pennsylvania)

Self-Supervised 3D Face Reconstruction via Conditional Estimation 13269
Yandong Wen (Carnegie Mellon University), Weiyang Liu (University of Cambridge), Bhiksha Raj (Carnegie Mellon University), and Rita Singh (Carnegie Mellon University)

Motion Prediction Using Trajectory Cues 13279
Zhenguang Liu (Zhejiang University), Pengxiang Su (Jilin University, Changchun, China), Shuang Wu (Nanyang Technological University), Xuanjing Shen (College of Computer Science and Technology, Jilin University, Changchun, China), Haipeng Chen (Jilin University), Yanbin Hao (University of Science and Technology of China), and Meng Wang (Hefei University of Technology)

Generating Smooth Pose Sequences for Diverse Human Motion Prediction 13289
Wei Mao (the Australian National University), Miaomiao Liu (The Australian National University), and Mathieu Salzmann (EPFL)

Spatially Conditioned Graphs for Detecting Human-Object Interactions 13299
Frederic Z. Zhang (Australian National University), Dylan Campbell (University of Oxford), and Stephen Gould (Australian National University, Australia)

Self-Supervised 3D Skeleton Action Representation Learning With Motion Consistency and Continuity 13308
Yukun Su (South China University of Technology), Guosheng Lin (Nanyang Technological University), and Qingyao Wu (South China University of Technology)

Learn To Match: Automatic Matching Network Design for Visual Tracking 13319
Zhipeng Zhang (Chinese Academy of Sciences), Yihao Liu (University of Chinese Academy of Sciences), Xiao Wang (Peng Cheng Laboratory), Bing Li (National Laboratory of Pattern Recognition, (NLPR), Institute of Automation, Chinese Academy of Sciences), and Weiming Hu (Institute of AutomationChinese Academy of Sciences)

Evidential Deep Learning for Open Set Action Recognition 13329
Wentao Bao (Rochester Institute of Technology), Qi Yu (Rochester Institute of Technology), and Yu Kong (Rochester Institute of Technology)

Channel-Wise Topology Refinement Graph Convolution for Skeleton-Based Action Recognition 13339
Yuxin Chen (Institute of Automation, Chinese Academy of Sciences), Ziqi Zhang (CASIA), Chunfeng Yuan (NLPR), Bing Li (National Laboratory of Pattern Recognition, (NLPR), Institute of Automation, Chinese Academy of Sciences), Ying Deng (School of Aeronautical Manufacturing Engineering, Nanchang Hangkong University), and Weiming Hu (Institute of AutomationChinese Academy of Sciences)

Consistency-Aware Graph Network for Human Interaction Understanding 13349
Zhenhua Wang (Zhejiang University of Technology), Jiajun Meng (Zhejiang University of Technology), Dongyan Guo (Zhejiang University of Technology), Jianhua Zhang (Tianjin University of Technology), Javen Qinfeng Shi (University of Adelaide), and Shengyong Chen (Tianjin University of Technology)

GeomNet: A Neural Network Based on Riemannian Geometries of SPD Matrix Space and Cholesky Space for 3D Skeleton-Based Interaction Recognition 13359
Xuan Son Nguyen (Ensea)

TRiPOD: Human Trajectory and Pose Dynamics Forecasting in the Wild 13370
Vida Adeli (Ferdowsi University of Mashhad), Mahsa Ehsanpour (The University of Adelaide), Ian Reid (University of Adelaide, Australia), Juan Carlos Niebles (Stanford University), Silvio Savarese (Stanford University), Ehsan Adeli (Stanford University), and Hamid Rezatofighi (Monash University)

AI Choreographer: Music Conditioned 3D Dance Generation With AIST++ 13381
Ruilong Li (USC Institute for Creative Technologies, (ICT)), Shan Yang (Google), David A. Ross (Google), and Angjoo Kanazawa (University of California Berkeley)

AdaSGN: Adapting Joint Number and Model Size for Efficient Skeleton-Based Action Recognition 13393
Lei Shi (Institute of Automation, Chinese Academy of Sciences), Yifan Zhang (Institute of Automation, Chinese Academy of Sciences), Jian Cheng (Chinese Academy of Sciences, China), and Hanqing Lu (NLPR, Institute of Automation, CAS)

Skeleton Cloud Colorization for Unsupervised 3D Action Representation Learning 13403
Siyuan Yang (Nanyang Technological University), Jun Liu (Singapore University of Technology and Design), Shijian Lu (Nanyang Technological University), Meng Hwa Er (Nanyang Technological University), and Alex C. Kot (Nanyang Technological University)

Else-Net: Elastic Semantic Network for Continual Action Recognition From Skeleton Data 13414
Tianjiao Li (Singapore University of Technology and Design), Qiuhong Ke (The University of Melbourne), Hossein Rahmani (Lancaster University), Rui En Ho (Singapore University of Technology and Design), Henghui Ding (Nanyang Technological University), and Jun Liu (Singapore University of Technology and Design)

Learning Target Candidate Association To Keep Track of What Not To Track 13424
Christoph Mayer (ETH Zurich), Martin Danelljan (ETH Zurich), Danda Pani Paudel (ETH Zürich), and Luc Van Gool (ETH Zurich)

Divide and Conquer for Single-Frame Temporal Action Localization 13435
Chen Ju (Cooperative Medianet Innovation Center, Shang hai Jiao Tong University), Peisen Zhao (Cooperative Medianet Innovation Center, Shang hai Jiao Tong University), Siheng Chen (Shanghai Jiao Tong University), Ya Zhang (Cooperative Medianet Innovation Center, Shang hai Jiao Tong University), Yanfeng Wang (Cooperative medianet innovation center of Shanghai Jiao Tong University), and Qi Tian (Huawei Cloud & AI)

HAA500: Human-Centric Atomic Action Dataset With Curated Videos 13445
Jihoon Chung (HKUST), Cheng-hsin Wuu (Carnegie Mellon University), Hsuan-ru Yang (HKUST), Yu-Wing Tai (Kuaishou Technology / HKUST), and Chi-Keung Tang (Hong Kong University of Science and Technology)

Discovering Human Interactions With Large-Vocabulary Objects via Query and Multi-Scale Detection 13455
Suchen Wang (Nanyang Technological University), Kim-Hui Yap (Nanyang Technological University), Henghui Ding (Nanyang Technological University), Jiyan Wu (Nanyang Technological University), Junsong Yuan (State University of New York at Buffalo, USA), and Yap-Peng Tan (Nanyang Technological University, Singapore)

Social Fabric: Tubelet Compositions for Video Relation Detection 13465
Shuo Chen (University of Amsterdam), Zenglin Shi (University of Amsterdam), Pascal Mettes (University of Amsterdam), and Cees G. M. Snoek (University of Amsterdam)

The Spatio-Temporal Poisson Point Process: A Simple Model for the Alignment of Event Camera Data 13475
Cheng Gu (TU-Berlin), Erik Learned-Miller (University of Massachusetts, Amherst), Daniel Sheldon (University of Massachusetts, Amherst), Guillermo Gallego (TU Berlin), and Pia Bideau (TU Berlin)

Anticipative Video Transformer 13485
Rohit Girdhar (Facebook AI Research) and Kristen Grauman (Facebook AI Research & UT Austin)

Enriching Local and Global Contexts for Temporal Action Localization 13496
Zixin Zhu (Xi'an jiaotong University), Wei Tang (University of Illinois at Chicago), Le Wang (Xi'an Jiaotong University), Nanning Zheng (Xi'an Jiaotong University), and Gang Hua (Wormpex AI Research)

Relaxed Transformer Decoders for Direct Action Proposal Generation 13506
Jing Tan (Nanjing University), Jiaqi Tang (Nanjing University), Limin Wang (Nanjing University), and Gangshan Wu (Nanjing University)

MultiSports: A Multi-Person Video Dataset of Spatio-Temporally Localized Sports Actions 13516
Yixuan Li (Nanjing University), Lei Chen (Nanjing University), Runyu He (Nanjing University), Zhenzhi Wang (Nanjing University), Gangshan Wu (Nanjing University), and Limin Wang (Nanjing University)

Learning To Track Objects From Unlabeled Videos 13526
Jilai Zheng (Shanghai Jiao Tong University), Chao Ma (Shanghai Jiao Tong University), Houwen Peng (Microsoft Research), and Xiaokang Yang (Shanghai Jiao Tong University of China)

Generating Masks From Boxes by Mining Spatio-Temporal Consistencies in Videos 13536
Bin Zhao (ETH Zurich), Goutam Bhat (ETH Zurich), Martin Danelljan (ETH Zurich), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)

Channel Augmented Joint Learning for Visible-Infrared Recognition 13547
Mang Ye (Wuhan University), Weijian Ruan (Shenzhen Institute of Advanced Technology Chniese Academy of Sciences), Bo Du (Wuhan University), and Mike Zheng Shou (Facebook AI)

VidTr: Video Transformer Without Convolutions 13557
Yanyi Zhang (Amazon), Xinyu Li (Rutgers University), Chunhui Liu (Amazon), Bing Shuai (Amazon), Yi Zhu (Amazon), Biagio Brattoli (Heidelberg University), Hao Chen (Amazon), Ivan Marsic (Rutgers University), and Joseph Tighe (Amazon)

A Hybrid Video Anomaly Detection Framework via Memory-Augmented Flow Reconstruction and Flow-Guided Frame Prediction 13568
Zhian Liu (South China University of Technology), Yongwei Nie (South China University of Technology), Chengjiang Long (JD Finance America Corporation), Qing Zhang (Sun Yat-sen University), and Guiqing Li (South China University of Technology)

Assignment-Space-Based Multi-Object Tracking and Segmentation 13578
Anwesa Choudhuri (UIUC), Girish Chowdhary (University of Illinois at Urbana Champaign), and Alexander G. Schwing (UIUC)

D2-Net: Weakly-Supervised Action Localization via Discriminative Embeddings and Denoised Activations 13588
Sanath Narayan (Inception Institute of Artificial Intelligence), Hisham Cholakkal (MBZUAI), Munawar Hayat (Monash University), Fahad Shahbaz Khan (MBZUAI), Ming-Hsuan Yang (University of California at Merced), and Ling Shao (Inception Institute of Artificial Intelligence)

Learning Cross-Modal Contrastive Features for Video Domain Adaptation 13598
Donghyun Kim (Boston University), Yi-Hsuan Tsai (NEC Labs America), Bingbing Zhuang (NEC Labs), Xiang Yu (NEC Labs), Stan Sclaroff (Boston University), Kate Saenko (Boston University), and Manmohan Chandraker (UC San Diego)

Selective Feature Compression for Efficient Activity Recognition Inference 13608
Chunhui Liu (Amazon), Xinyu Li (Amazon), Hao Chen (Amazon), Davide Modolo (Amazon), and Joseph Tighe (Amazon)

Elaborative Rehearsal for Zero-Shot Action Recognition 13618
Shizhe Chen (INRIA) and Dong Huang (Carnegie Mellon University)

Learning Action Completeness From Points for Weakly-Supervised Temporal Action Localization 13628
Pilhyeon Lee (Yonsei University) and Hyeran Byun (Yonsei University)

Video Self-Stitching Graph Network for Temporal Action Localization 13638
Chen Zhao (KAUST), Ali K. Thabet (Facebook), and Bernard Ghanem (KAUST)

GroupFormer: Group Activity Recognition With Clustered Spatial-Temporal Transformer 13648
Shuaicheng Li (Sensetime), Qianggang Cao (Sensetime), Lingbo Liu (Sun Yat-sen University), Kunlin Yang (Sensetime Group Limited), Shinan Liu (sensetime), Jun Hou (SenseTime Group Limited), and Shuai Yi (SenseTime Group Limited)

Multi-Modal Multi-Action Video Recognition 13658
Zhensheng Shi (Ocean University of China), Ju Liang (Ocean University of China), Qianqian Li (Ocean University of China), Haiyong Zheng (Ocean University of China), Zhaorui Gu (Ocean University of China), Junyu Dong (Ocean University of China), and Bing Zheng (Ocean University of China)

Target Adaptive Context Aggregation for Video Scene Graph Generation 13668
Yao Teng (Nanjing University), Limin Wang (Nanjing University), Zhifeng Li (Tencent AI Lab), and Gangshan Wu (Nanjing University)

Class-Incremental Learning for Action Recognition in Videos 13678
Jaeyoo Park (Seoul National University), Minsoo Kang (Seoul National University), and Bohyung Han (Seoul National University)

TAM: Temporal Adaptive Module for Video Recognition 13688
Zhaoyang Liu (National Key Lab for Novel Software Technology, Nanjing University, China), Limin Wang (Nanjing University), Wayne Wu (SenseTime Research), Chen Qian (SenseTime), and Tong Lu (Nanjing University)

Efficient Action Recognition via Dynamic Knowledge Propagation 13699
Hanul Kim (Seoul National University of Science and Technology), Mihir Jain (Qualcomm AI Research), Jun-Tae Lee (Qualcomm AI Research), Sungrack Yun (Qualcomm AI Research), and Fatih Porikli (Qualcomm AI Research)

CAG-QIL: Context-Aware Actionness Grouping via Q Imitation Learning for Online Temporal Action Localization 13709
Hyolim Kang (Yonsei University), Kyungmin Kim (Yonsei University), Yumin Ko (Yonsei University), and Seon Joo Kim (Yonsei University)

Class Semantics-Based Attention for Action Detection 13719
Deepak Sridhar (Huawei Technologies Canada, Noah's Ark Lab), Niamul Quader (Huawei Noah's Ark Laboratory), Srikanth Muralidharan (Huawei Noah's Ark Laboratory), Yaoxin Li (University), Peng Dai (Huawei Technologies Canada), and Juwei Lu (Huawei Noah's Ark Laboratory)

Transfer/Low-shot/Semi/Unsupervised Learning

Image Synthesis via Semantic Composition 13729
Yi Wang (Chinese University of Hong Kong), Lu Qi (The Chinese University of Hong Kong), Ying-Cong Chen (Massachusetts Institute of Technology), Xiangyu Zhang (Megvii Technology), and Jiaya Jia (Chinese University of Hong Kong)

PIRenderer: Controllable Portrait Image Generation via Semantic Neural Rendering 13739
Yurui Ren (Shenzhen Graduate School, Peking University), Ge Li (SECE, Shenzhen Graduate School, Peking University), Yuanqi Chen (SECE, Peking University), Thomas H. Li (Advanced Institute of Information Technology, Peking University), and Shan Liu (Tencent America)

Image Shape Manipulation From a Single Augmented Training Sample 13749
Yael Vinker (The Hebrew University of Jerusalem), Eliahu Horwitz (The Hebrew University of Jerusalem), Nir Zabari (The Hebrew University), and Yedid Hoshen (The Hebrew University of Jerusalem)

Learning Object-Compositional Neural Radiance Field for Editable Scene Rendering 13759
Bangbang Yang (Zhejiang University), Yinda Zhang (Google), Yinghao Xu (Chinese University of Hong Kong), Yijin Li (Zhejiang University), Han Zhou (Zhejiang University), Hujun Bao (Zhejiang University), Guofeng Zhang (Zhejiang University), and Zhaopeng Cui (Zhejiang University)

A Latent Transformer for Disentangled Face Editing in Images and Videos 13769
Xu Yao (Telecom Paris), Alasdair Newson (Telecom Paris), Yann Gousseau (Telecom Paris), and Pierre Hellier (Interdigital, (Technicolor))

Talk-To-Edit: Fine-Grained Facial Editing via Dialog 13779
Yuming Jiang (Nanyang Technological University), Ziqi Huang (Nanyang Technological University), Xingang Pan (The Chinese University of Hong Kong), Chen Change Loy (Nanyang Technological University), and Ziwei Liu (Nanyang Technological University)

FashionMirror: Co-Attention Feature-Remapping Virtual Try-On With Sequential Template Poses 13789
Chieh-Yun Chen (National Chiao Tung University), Ling Lo (National Chiao Tung University), Pin-Jui Huang (National Chiao Tung University), Hong-Han Shuai (National Chiao Tung University), and Wen-Huang Cheng (National Chiao Tung University)

Image Synthesis From Layout With Locality-Aware Mask Adaption 13799
Zejian Li (Zhejiang University), Jingyu Wu (Zhejiang University), Immanuel Koh (Singapore University of Technology and Design), Yongchuan Tang (Zhejiang University, Zhejiang Lab), and Lingyun Sun (Zhejiang University)

Learned Spatial Representations for Few-Shot Talking-Head Synthesis 13809
Moustafa Meshry (University of Maryland), Saksham Suri (University of Maryland, College Park), Larry S. Davis (University of Maryland), and Abhinav Shrivastava (University of Maryland)

Motion-Aware Dynamic Architecture for Efficient Frame Interpolation 13819
Myungsub Choi (Seoul National University), Suyoung Lee (Seoul National University), Heewon Kim (Seoul National University), and Kyoung Mu Lee (Seoul National University)

3D Human Texture Estimation From a Single Image With Transformers 13829
Xiangyu Xu (Nanyang Technological University) and Chen Change Loy (Nanyang Technological University)

Structure-Transformed Texture-Enhanced Network for Person Image Synthesis 13839
Munan Xu (Shenzhen Graduate School, Peking University), Yuanqi Chen (SECE, Peking University), Shan Liu (Tencent America), Thomas H. Li (Advanced Institute of Information Technology, Peking University), and Ge Li (SECE, Shenzhen Graduate School, Peking University)

Learning To Stylize Novel Views 13849
Hsin-Ping Huang (University of California, Merced), Hung-Yu Tseng (Facebook), Saurabh Saini (International Institute of Information Technology Hyderabad), Maneesh Singh (Verisk Analytics), and Ming-Hsuan Yang (University of California at Merced)

Learning Generative Models of Textured 3D Meshes From Real-World Images 13859
Dario Pavllo (ETH Zurich), Jonas Kohler (ETHZ), Thomas Hofmann (ETH Zurich), and Aurelien Lucchi (ETH Zurich)

VariTex: Variational Neural Face Textures 13870
Marcel C. Bühler (ETH Zürich), Abhimitra Meka (Google), Gengyan Li (ETH Zürich), Thabo Beeler (Disney Research, Studios), and Otmar Hilliges (ETH Zurich)

Multiple Heads Are Better Than One: Few-Shot Font Generation With Multiple Localized Experts 13880
Song Park (Yonsei University), Sanghyuk Chun (NAVER AI LAB), Junbum Cha (Clova AI Research, NAVER Corp.), Bado Lee (Clova AI Research, NAVER Corp.), and Hyunjung Shim (Yonsei University)

From Continuity to Editability: Inverting GANs With Consecutive Images 13890
Yangyang Xu (South China University of Technology), Yong Du (Ocean University of China), Wenpeng Xiao (South China University of Technology), Xuemiao Xu (South China University of Technology), and Shengfeng He (South China University of Technology)

Focal Frequency Loss for Image Reconstruction and Synthesis 13899
Liming Jiang (Nanyang Technological University), Bo Dai (Nanyang Technological University), Wayne Wu (SenseTime Research), and Chen Change Loy (Nanyang Technological University)

Frequency Domain Image Translation: More Photo-Realistic, Better Identity-Preserving 13910
Mu Cai (University of Wisconsin-Madison), Hong Zhang (SenseTime Group Ltd), Huijuan Huang (Kuaishou Technology), Qichuan Geng (State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University), Yixuan Li (University of Wisconsin-Madison), and Gao Huang (Tsinghua)

GAN Inversion for Out-of-Range Images With Geometric Transformations 13921
Kyoungkook Kang (POSTECH), Seongtae Kim (POSTECH), and Sunghyun Cho (POSTECH)

Detail Me More: Improving GAN's Photo-Realism of Complex Scenes 13930
Raghudeep Gadde (Amazon), Qianli Feng (The Ohio State University), and Aleix M. Martinez (Amazon)

DAE-GAN: Dynamic Aspect-Aware GAN for Text-to-Image Synthesis 13940
Shulan Ruan (University of Science and Technology of China), Yong Zhang (Tencent AI Lab), Kun Zhang (Hefei University of Technology), Yanbo Fan (Tencent AI Lab), Fan Tang (Jilin University), Qi Liu (University of Science and Technology of China, China), and Enhong Chen (University of Science and Technology of China)

Labels4Free: Unsupervised Segmentation Using StyleGAN 13950
Rameen Abdal (KAUST), Peihao Zhu (KAUST), Niloy J. Mitra (University College London), and Peter Wonka (KAUST)

Diagonal Attention and Style-Based GAN for Content-Style Disentanglement in Image Generation and Translation 13960
Gihyun Kwon (KAIST) and Jong Chul Ye (Department of Bio and Brain Engineering, KAIST, Korea)

Learning High-Fidelity Face Texture Completion Without Complete Face Texture 13970
Jongyoo Kim (Microsoft Research Asia), Jiaolong Yang (Microsoft Research), and Xin Tong (Microsoft)

Deep Halftoning With Reversible Binary Pattern 13980
Menghan Xia (The Chinese University of Hong Kong), Wenbo Hu (The Chinese University of Hong Kong), Xueting Liu (Caritas Institute of Higher Education), and Tien-Tsin Wong (The Chinese University of Hong Kong)

TransferI2I: Transfer Learning for Image-to-Image Translation From Small Datasets 13990
Yaxing Wang (Computer Vision Center), Héctor Laria (Computer Vision Center), Joost van de Weijer (Computer Vision Center), Laura Lopez-Fuentes (Universitat de les Illes Balears), and Bogdan Raducanu (Computer Version Center)

Instance-Wise Hard Negative Example Generation for Contrastive Learning in Unpaired Image-to-Image Translation 14000
Weilun Wang (University of Science and Technology of China), Wengang Zhou (University of Science and Technology of China), Jianmin Bao (Microsoft Research Asia), Dong Chen (Microsoft Research Asia), and Houqiang Li (University of Science and Technology of China)

Multi-Scale Separable Network for Ultra-High-Definition Video Deblurring 14010
Senyou Deng (Institute of Information Engineering, Chinese Academy of Sciences), Wenqi Ren (Institute of Information Engineering, Chinese Academy of Sciences), Yanyang Yan (Institute of Information Engineering, Chinese Academy of Sciences), Tao Wang (Huawei Noah's Ark Lab), Fenglong Song (Huawei Noah's Ark Lab), and Xiaochun Cao (Chinese Academy of Sciences)

FuseFormer: Fusing Fine-Grained Information in Transformers for Video Inpainting 14020
Rui Liu (Chinese University of Hong Kong), Hanming Deng (Sensetime), Yangyi Huang (Zhejiang University), Xiaoyu Shi (CUHK), Lewei Lu (Sensetime Research), Wenxiu Sun (SenseTime Research and Tetras.AI), Xiaogang Wang (Chinese University of Hong Kong, Hong Kong), Jifeng Dai (SenseTime), and Hongsheng Li (Chinese University of Hong Kong)

Sketch Your Own GAN 14030
Sheng-Yu Wang (Carnegie Mellon University), David Bau (MIT), and Jun-Yan Zhu (Carnegie Mellon University)

Omni-GAN: On the Secrets of cGANs and Beyond 14041
Peng Zhou (Shanghai Jiao Tong University), Lingxi Xie (Huawei Inc.), Bingbing Ni (Shanghai Jiao Tong University), Cong Geng (Shanghai Jiao Tong University), and Qi Tian (Huawei Cloud & AI)

GANcraft: Unsupervised 3D Neural Rendering of Minecraft Worlds 14052
Zekun Hao (Cornell University), Arun Mallya (NVIDIA), Serge Belongie (University of Copenhagen), and Ming-Yu Liu (NVIDIA)

GAN-Control: Explicitly Controllable GANs 14063
Alon Shoshan (Technion), Nadav Bhonker (Amazon), Igor Kviatkovsky (Amazon), and Gérard Medioni (USC)

Towards Discovery and Attribution of Open-World GAN Generated Images 14074
Sharath Girish (University of Maryland, College Park), Saksham Suri (University of Maryland, College Park), Sai Saketh Rambhatla (University of Maryland), and Abhinav Shrivastava (University of Maryland)

PixelSynth: Generating a 3D-Consistent Experience From a Single Image 14084
Chris Rockwell (University of Michigan), David F. Fouhey (University of Michigan), and Justin Johnson (University of Michigan)

WaveFill: A Wavelet-Based Generation Network for Image Inpainting 14094
Yingchen Yu (Nanyang Technological University), Fangneng Zhan (Nanyang Technological University), Shijian Lu (Nanyang Technological University), Jianxiong Pan (Alibaba Group), Feiying Ma (alibaba), Xuansong Xie (Alibaba), and Chunyan Miao (NTU)

MVSNeRF: Fast Generalizable Radiance Field Reconstruction From Multi-View Stereo 14104
Anpei Chen (ShanghaiTech University), Zexiang Xu (Adobe Research), Fuqiang Zhao (ShanghaiTech University), Xiaoshuai Zhang (University of California, San Diego), Fanbo Xiang (University of California, San Diego), Jingyi Yu (Shanghai Tech University), and Hao Su (University of California, San Diego)

Image Inpainting via Conditional Texture and Structure Dual Generation 14114
Xiefan Guo (Beihang University), Hongyu Yang (Beihang University), and Di Huang (Beihang University, China)

Aligning Latent and Image Spaces To Connect the Unconnectable 14124
Ivan Skorokhodov (KAUST), Grigorii Sotnikov (National Research University Higher School of Economics), and Mohamed Elhoseiny (KAUST)

Rethinking the Truly Unsupervised Image-to-Image Translation 14134
Kyungjune Baek (Yonsei University), Yunjey Choi (NAVER AI Lab), Youngjung Uh (Yonsei University), Jaejun Yoo (EPFL), and Hyunjung Shim (Yonsei University)

CR-Fill: Generative Image Inpainting With Auxiliary Contextual Reconstruction 14144
Yu Zeng (Johns Hopkins University), Zhe Lin (Adobe Research), Huchuan Lu (Dalian University of Technology), and Vishal M. Patel (Johns Hopkins University)

Unaligned Image-to-Image Translation by Learning to Reweight 14154
Shaoan Xie (Carnegie Mellon University), Mingming Gong (University of Melbourne), Yanwu Xu (University of Pittsburgh), and Kun Zhang (Carnegie Mellon University)

Image Manipulation Detection by Multi-View Multi-Scale Supervision 14165
Xinru Chen (Renmin University), Chengbo Dong (Renmin University of China), Jiaqi Ji (Renmin University of China), Juan Cao (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), and Xirong Li (Renmin University of China)

From Two to One: A New Scene Text Recognizer With Visual Language Modeling Network 14174
Yuxin Wang (University of Science and Technology of China), Hongtao Xie (University of Science and Technology of China), Shancheng Fang (University of Science and Technology of China), Jing Wang (Huawei Cloud & AI), Shenggao Zhu (Huawei), and Yongdong Zhang (University of Science and Technology of China)

Neural Strokes: Stylized Line Drawing of 3D Shapes 14184
Difan Liu (UMass-Amherst), Matthew Fisher (Adobe Research), Aaron Hertzmann (Adobe Research), and Evangelos Kalogerakis (UMass Amherst)

Modulated Periodic Activations for Generalizable Local Functional Representations 14194
Ishit Mehta (University of California San Diego), Michaël Gharbi (Adobe Research), Connelly Barnes (Adobe), Eli Shechtman (Adobe Research, US), Ravi Ramamoorthi (University of California San Diego), and Manmohan Chandraker (UC San Diego)

SIGNET: Efficient Neural Representation for Light Fields 14204
Brandon Yushan Feng (University of Maryland) and Amitabh Varshney (University of Maryland)

Searching for Controllable Image Restoration Networks 14214
Heewon Kim (Seoul National University), Sungyong Baik (Seoul National University), Myungsub Choi (Seoul National University), Janghoon Choi (Seoul National University), and Kyoung Mu Lee (Seoul National University)

Cortical Surface Shape Analysis Based on Alexandrov Polyhedra 14224
Min Zhang (Brigham and Women's Hospital, Harvard Medical School), Yang Guo (Stony Brook University), Na Lei (Dalian University of Technology), Zhou Zhao (Stony Brook University), Jianfeng Wu (Arizona State University), Xiaoyin Xu (Stony Brook University), Yalin Wang (Arizona State University), and Xianfeng Gu (Brigham and Women's Hospital, Harvard Medical School)

Toward Spatially Unbiased Generative Models 14233
Jooyoung Choi (Seoul National University), Jungbeom Lee (Seoul National University), Yonghyun Jeong (Samsung SDS), and Sungroh Yoon (Seoul National University)

LatentCLR: A Contrastive Learning Approach for Unsupervised Discovery of Interpretable Directions 14243
Oğuz Kaan Yüksel (EPFL), Enis Simsar (Technical University of Munich), Ezgi Gülperi Er (Bogazici University), and Pinar Yanardag (Bogazici University)

Semantically Robust Unpaired Image Translation for Data With Unmatched Semantics Statistics 14253
Zhiwei Jia (University of California, San Diego), Bodi Yuan (X, The Moonshot Factory), Kangkang Wang (X), Hong Wu (Google), David Clifford (Google), Zhiqiang Yuan (Google), and Hao Su (UCSD)

Unsupervised Image Generation With Infinite Generative Adversarial Networks 14264
Hui Ying (Zhejiang University), He Wang (Leeds University), Tianjia Shao (Zhejiang University), Yin Yang (Clemson University), and Kun Zhou (Zhejiang University)

Reality Transform Adversarial Generators for Image Splicing Forgery Detection and Localization 14274
Xiuli Bi (Chongqing University of Posts and Telecommunications), Zhipeng Zhang (Chongqing university of post and telecommunications), and Bin Xiao (Chongqing University of Posts and Telecommunications)

Unconstrained Scene Generation With Locally Conditioned Radiance Fields 14284
Terrance DeVries (Apple), Miguel Angel Bautista (Apple), Nitish Srivastava (Apple), Graham W. Taylor (University of Guelph), and Joshua M. Susskind (Apple)

Animatable Neural Radiance Fields for Modeling Dynamic Human Bodies 14294
Sida Peng (Zhejiang University), Junting Dong (Zhejiang University), Qianqian Wang (Cornell), Shangzhan Zhang (Zhejiang University), Qing Shuai (Zhejiang University), Xiaowei Zhou (Zhejiang University), and Hujun Bao (Zhejiang University)

Neural Radiance Flow for 4D View Synthesis and Video Processing 14304
Yilun Du (MIT), Yinan Zhang (Stanford University), Hong-Xing Yu (Stanford University), Joshua B. Tenenbaum (MIT), and Jiajun Wu (Stanford University)

KiloNeRF: Speeding Up Neural Radiance Fields With Thousands of Tiny MLPs 14315
Christian Reiser (MPI-IS and University of Tübingen), Songyou Peng (ETH Zurich and MPI-IS), Yiyi Liao (MPI-IS and University of Tübingen), and Andreas Geiger (University of Tuebingen)

FastNeRF: High-Fidelity Neural Rendering at 200FPS 14326
Stephan J. Garbin (University College London), Marek Kowalski (Microsoft), Matthew Johnson (Microsoft), Jamie Shotton (Microsoft), and Julien Valentin (Microsoft)

Geometry-Free View Synthesis: Transformers and No 3D Priors 14336
Robin Rombach (Heidelberg University), Patrick Esser (Heidelberg University), and Björn Ommer (Heidelberg University)

ILVR: Conditioning Method for Denoising Diffusion Probabilistic Models 14347
Jooyoung Choi (Seoul National University), Sungwon Kim (Seoul National University), Yonghyun Jeong (Samsung SDS), Youngjune Gwon (Samsung SDS), and Sungroh Yoon (Seoul National University)

Towards Vivid and Diverse Image Colorization With Generative Color Prior 14357
Yanze Wu (Tencent), Xintao Wang (Tencent), Yu Li (Tencent), Honglun Zhang (Applied Research Center, Tencent PCG), Xun Zhao (Tencent Company), and Ying Shan (Tencent)

Physics-Based Differentiable Depth Sensor Simulation 14367
Benjamin Planche (Siemens Technology) and Rajat Vikram Singh (NVIDIA)

HeadGAN: One-Shot Neural Head Synthesis and Editing 14378
Michail Christos Doukas (Imperial College London), Stefanos Zafeiriou (Imperial College London), and Viktoriia Sharmanska (Imperial College London)

EigenGAN: Layer-Wise Eigen-Learning for GANs 14388
Zhenliang He (Key Lab of Intelligent Information Processing of Chinese Academy of Sciences, (CAS), Institute of Computing Technology, CAS), Meina Kan (Institute of Computing Technology, Chinese Academy of Sciences), and Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences)

Collaging Class-Specific GANs for Semantic Image Synthesis 14398
Yuheng Li (UC Davis), Yijun Li (Adobe Research), Jingwan Lu (Adobe Research), Eli Shechtman (Adobe Research, US), Yong Jae Lee (University of California, Davis), and Krishna Kumar Singh (Adobe Research)

Latent Transformations via NeuralODEs for GAN-Based Image Editing 14408
Valentin Khrulkov (Yandex), Leyla Mirvakhabova (Skolkovo Institute of Science and Technology), Ivan Oseledets (Skolkovo Institute of Science and Technology), and Artem Babenko (Yandex)

Dual Projection Generative Adversarial Networks for Conditional Image Generation 14418
Ligong Han (Rutgers University), Martin Renqiang Min (NEC Labs America-Princeton), Anastasis Stathopoulos (Rutgers University), Yu Tian (Rutgers), Ruijiang Gao (University of Texas at Austin), Asim Kadav (NEC Labs), and Dimitris N. Metaxas (Rutgers)

Artificial Fingerprinting for Generative Models: Rooting Deepfake Attribution in Training Data 14428
Ning Yu (University of Maryland and Max Planck Institute for Informatics), Vladislav Skripniuk (CISPA Helmholtz Center for Information Security), Sahar Abdelnabi (CISPA Helmholtz Center for Information Security), and Mario Fritz (CISPA Helmholtz Center for Information Security)

Infinite Nature: Perpetual View Generation of Natural Scenes From a Single Image 14438
Andrew Liu (Google), Richard Tucker (Google), Varun Jampani (Google), Ameesh Makadia (Google Research), Noah Snavely (Cornell University and Google AI), and Angjoo Kanazawa (University of California Berkeley)

Occlusion-Aware Video Object Inpainting 14448
Lei Ke (HKUST), Yu-Wing Tai (Kuaishou Technology / HKUST), and Chi-Keung Tang (Hong Kong University of Science and Technology)

ELF-VC: Efficient Learned Flexible-Rate Video Coding 14459
Oren Rippel (WaveOne, Inc.), Alexander G. Anderson (WaveOne, Inc.), Kedar Tatwawadi (WaveOne Inc), Sanjay Nair (WaveOne, Inc.), Craig Lytle (WaveOne), and Lubomir Bourdev (WaveOne, Inc.)

XVFI: eXtreme Video Frame Interpolation 14469
Hyeonjun Sim (KAIST), Jihyong Oh (KAIST), and Munchurl Kim (Korea Advanced Institute of Science and Technology)

OSCAR-Net: Object-Centric Scene Graph Attention for Image Attribution 14479
Eric Nguyen (Stanford University), Tu Bui (University of Surrey), Viswanathan Swaminathan (Vishy) Swaminathan, (Adobe), and John Collomosse (Adobe Research)

Learning a Sketch Tensor Space for Image Inpainting of Man-Made Scenes 14489
Chenjie Cao (Fudan University) and Yanwei Fu (Fudan University)

Embedding Novel Views in a Single JPEG Image 14499
Yue Wu (HKUST), Guotao Meng (HKUST), and Qifeng Chen (HKUST)

EgoRenderer: Rendering Human Avatars From Egocentric Camera Images 14508
Tao Hu (university of maryland), Kripasindhu Sarkar (Max Plank Institute for Informatics), Lingjie Liu (Max Planck Institute for Informatics), Matthias Zwicker (University of Maryland), and Christian Theobalt (MPI Informatik)

Asymmetric Bilateral Motion Estimation for Video Frame Interpolation 14519
Junheum Park (Korea University), Chul Lee (Dongguk University), and Chang-Su Kim (Korea university)

STRIVE: Scene Text Replacement in Videos 14529
Vijay Kumar B G (PARC), Jeyasri Subramanian (Palo Alto Research Center), Varnith Chordia (Palo Alto Research Center), Eugene Bart (Xerox PARC), Shaobo Fang (Apple), Kelly Guan (Stanford University), and Raja Bala (PARC)

Parallel Multi-Resolution Fusion Network for Image Inpainting 14539
Wentao Wang (Shanghai Jiao Tong University), Jianfu Zhang (RIKEN AIP; Shanghai Jiao Tong University), Li Niu (Shanghai Jiao Tong University), Haoyu Ling (Shanghai Jiao Tong University), Xue Yang (Shanghai Jiao Tong University), and Liqing Zhang (Shanghai Jiao Tong University)

Towards Complete Scene and Regular Shape for Distortion Rectification by Curve-Aware Extrapolation 14549
Kang Liao (Beijing Jiaotong University), Chunyu Lin (Beijing Jiaotong University), Yunchao Wei (UTS), Feng Li (Beijing Jiaotong University), Shangrong Yang (Beijing Jiaotong University), and Yao Zhao (Beijing Jiaotong University)

Internal Video Inpainting by Implicit Long-Range Propagation 14559
Hao Ouyang (HKUST), Tengfei Wang (HKUST), and Qifeng Chen (HKUST)

Training Weakly Supervised Video Frame Interpolation With Events 14569
Zhiyang Yu (Harbin Institute of Technology), Yu Zhang (Beihang University), Deyuan Liu (Peking University; SenseTime Research and Tetras.AI), Dongqing Zou (SenseTime ResearchQing Yuan Research Institute, Shanghai Jiao Tong University), Xijun Chen (Harbin Institute of Technology), Yebin Liu (Tsinghua University), and Jimmy S. Ren (SenseTime Research; Qing Yuan Research Institute, Shanghai Jiao Tong University)

Flow-Guided Video Inpainting With Scene Templates 14579
Dong Lao (KAUST), Peihao Zhu (KAUST), Peter Wonka (KAUST), and Ganesh Sundaramoorthi (Kaust)

Domain-Aware Universal Style Transfer 14589
Kibeom Hong (Yonsei university), Seogkyu Jeon (Yonsei university), Huan Yang (Microsoft Research), Jianlong Fu (Microsoft Research), and Hyeran Byun (Yonsei University)

StyleFormer: Real-Time Arbitrary Style Transfer via Parametric Style Composition 14598
Xiaolei Wu (Beihang University), Zhihao Hu (Beihang University), Lu Sheng (Beihang University), and Dong Xu (University of Sydney)

Bridging the Gap Between Label- and Reference-Based Synthesis in Multi-Attribute Image-to-Image Translation 14608
Qiusheng Huang (East China Normal University), Zhilin Zheng (PingAn Technology), Xueqi Hu (East China Normal University), Li Sun (East China Normal University), and Qingli Li (East China Normal University)

Dressing in Order: Recurrent Person Image Generation for Pose Transfer, Virtual Try-On and Outfit Editing 14618
Aiyu Cui (University of Illinois at Urbana-Champaign), Daniel McKee (UIUC), and Svetlana Lazebnik (UIUC)

Gait Recognition via Effective Global-Local Feature Representation and Local Temporal Aggregation 14628
Beibei Lin (Beijing Jiaotong University), Shunli Zhang (Beijing Jiaotong University), and Xin Yu (University of Technology Sydney)

Unpaired Learning for High Dynamic Range Image Tone Mapping 14637
Yael Vinker (The Hebrew University of Jerusalem), Inbar Huberman-Spiegelglas (The Hebrew University of Jerusalem), and Raanan Fattal (Hebrew University of Jerusalem, Israel)

Deep Edge-Aware Interactive Colorization Against Color-Bleeding Effects 14647
Eungyeup Kim (Korea Advanced Institute of Science and Technology), Sanghyeon Lee (Korea Advanced Institute of Science and Technology), Jeonghoon Park (Korea Advanced Institute of Science and Technology), Somi Choi (Korea Advanced Institute of Science and Technology), Choonghyun Seo (NAVER WEBTOON Corp.), and Jaegul Choo (Korea Advanced Institute of Science and Technology)

Neural Image Compression via Attentional Multi-Scale Back Projection and Frequency Decomposition 14657
Ge Gao (Samsung R&D Institute China Xi'an), Pei You (Samsung R&D Institute China Xi'an), Rong Pan (Samsung R&D Institute China Xi'an), Shunyuan Han (Samsung R&D Institute China Xi'an), Yuanyuan Zhang (Samsung R&D Institute China Xi'an), Yuchao Dai (Northwestern Polytechnical University), and Hojae Lee (Samsung Electronics)

Dynamic Cross Feature Fusion for Remote Sensing Pansharpening 14667
Xiao Wu (University of Electronic Science and Technology of China), Ting-Zhu Huang (School of Mathematical Sciences, University of Electronic Science and Technology of China), Liang-Jian Deng (School of Mathematical Sciences, University of Electronic Science and Technology of China), and Tian-Jing Zhang (Yingcai Honors College, University of Electronic Science and Technology of China)

Attention-Based Multi-Reference Learning for Image Super-Resolution 14677
Marco Pesavento (University of Surrey), Marco Volino (University of Surrey), and Adrian Hilton (University of Surrey)

iPOKE: Poking a Still Image for Controlled Stochastic Video Synthesis 14687
Andreas Blattmann (Heidelberg University), Timo Milbich (Heidelberg University), Michael Dorkenwald (Heidelberg University), and Björn Ommer (Heidelberg University)

Point-Based Modeling of Human Clothing 14698
Ilya Zakharkin (Samsung), Kirill Mazur (Samsung), Artur Grigorev (Samsung), and Victor Lempitsky (Samsung)

SLAMP: Stochastic Latent Appearance and Motion Prediction 14708
Adil Kaan Akan (Koc University), Erkut Erdem (Hacettepe University), Aykut Erdem (Koc University), and Fatma Güney (Koc University)

Pathdreamer: A World Model for Indoor Navigation 14718
Jing Yu Koh (Google Research), Honglak Lee (LG AI Research / University of Michigan), Yinfei Yang (Google Research), Jason Baldridge (Google), and Peter Anderson (Google)

Click To Move: Controlling Video Generation With Sparse Motion 14729
Pierfrancesco Ardino (University of Trento), Marco De Nadai (Fondazione Bruno Kessler), Bruno Lepri (FBK, Trento, Italy), Elisa Ricci (University of Trento), and Stéphane Lathuilière (Telecom-Paris)

Accelerating Atmospheric Turbulence Simulation via Learned Phase-to-Space Transform 14739
Zhiyuan Mao (Purdue University), Nicholas Chimitt (Purdue University), and Stanley H. Chan (Purdue University, USA)

Benchmarking Ultra-High-Definition Image Super-Resolution 14749
Kaihao Zhang (Australian National University), Dongxu Li (THE AUSTRALIAN NATIONAL UNIVERSITY), Wenhan Luo (Tencent), Wenqi Ren (Institute of Information Engineering, Chinese Academy of Sciences), Björn Stenger (Rakuten Institute of Technology), Wei Liu (Tencent), Hongdong Li (Australian National University, Australia), and Ming-Hsuan Yang (University of California at Merced)

Painting From Part 14759
Dongsheng Guo (Ocean University of China), Haoru Zhao (Ocean University of China), Yunhao Cheng (Ocean University of China), Haiyong Zheng (Ocean University of China), Zhaorui Gu (Ocean University of China), and Bing Zheng (Ocean University of China)

Gait Recognition in the Wild: A Benchmark 14769
Zheng Zhu (Tsinghua University), Xianda Guo (xforward.ai), Tian Yang (XForwardAI), Junjie Huang (Xforwardai), Jiankang Deng (Imperial College London), Guan Huang (Institute of Automation, Chinese Academy of Sciences), Dalong Du (XForwardAI Technology Co., Ltd), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)

Joint Audio-Visual Deepfake Detection 14780
Yipin Zhou (Facebook AI) and Ser-Nam Lim (Facebook AI)

DisUnknown: Distilling Unknown Factors for Disentanglement Learning 14790
Sitao Xiang (University of Southern California), Yuming Gu (University of Southern California), Pengda Xiang (University of Southern California), Menglei Chai (Snap Inc.), Hao Li (Pinscreen / UC Berkeley), Yajie Zhao (USC ICT), and Mingming He (USC ICT)

Membership Inference Attacks Are Easier on Difficult Problems 14800
Avital Shafran (The Hebrew University of Jerusalem), Shmuel Peleg (The Hebrew University), and Yedid Hoshen (The Hebrew University of Jerusalem)

Understanding and Evaluating Racial Biases in Image Captioning 14810
Dora Zhao (Princeton University), Angelina Wang (Princeton University), and Olga Russakovsky (Princeton University)

Robust Watermarking for Deep Neural Networks via Bi-Level Optimization 14821
Peng Yang (Baidu), Yingjie Lao (Clemson University), and Ping Li (Baidu Research)

Detection and Continual Learning of Novel Face Presentation Attacks 14831
Mohammad Rostami (University of Southern California), Leonidas Spinoulas (USC/ISI), Mohamed Hussein (USC/ISI), Joe Mathai (Information Sciences Institute), and Wael Abd-Almageed (Information Sciences Institute)

Manifold Alignment for Semantically Aligned Style Transfer 14841
Jing Huo (Nanjing University), Shiyin Jin (Nanjing University), Wenbin Li (Nanjing University), Jing Wu (Cardiff), Yu-Kun Lai (Cardiff University), Yinghuan Shi (Nanjing University), and Yang Gao (Nanjing University)

Image Harmonization With Transformer 14850
Zonghui Guo (Ocean University of China), Dongsheng Guo (Ocean University of China), Haiyong Zheng (Ocean University of China), Zhaorui Gu (Ocean University of China), Bing Zheng (Ocean University of China), and Junyu Dong (Ocean University of China)

Diverse Image Style Transfer via Invertible Cross-Space Mapping 14860
Haibo Chen (Zhejiang University), Lei Zhao (Zhejiang University), Huiming Zhang (Zhejiang University), Zhizhong Wang (Zhejiang University), Zhiwen Zuo (Zhejiang University), Ailin Li (College of Computer Science and Technology, Zhejiang University), Wei Xing (Zhejiang University), and Dongming Lu (Zhejiang University)

LocalTrans: A Multiscale Local Transformer Network for Cross-Resolution Homography Estimation 14870
Ruizhi Shao (Tsinghua University), Gaochang Wu (Northeastern University), Yuemei Zhou (Tsinghua University), Ying Fu (Beijing Institute of Technology), Lu Fang (Tsinghua University), and Yebin Liu (Tsinghua University)

SemIE: Semantically-Aware Image Extrapolation 14880
Bholeshwar Khurana (Indian Institute of Technology, Kanpur), Soumya Ranjan Dash (Indian Institute of Technology Kanpur), Abhishek Bhatia (Indian Institute of Technology, Kanpur), Aniruddha Mahapatra (Adobe), Hrituraj Singh (Adobe Research), and Kuldeep Kulkarni (Adobe Research)

Calibrating Concepts and Operations: Towards Symbolic Reasoning on Real Images 14890
Zhuowan Li (Johns Hopkins University), Elias Stengel-Eskin (Johns Hopkins University), Yixiao Zhang (Johns Hopkins University), Cihang Xie (University of California, Santa Cruz), Quan Hung Tran (Adobe Research), Benjamin Van Durme (Johns Hopkins University), and Alan Yuille (Johns Hopkins University)

3D Local Convolutional Neural Networks for Gait Recognition 14900
Zhen Huang (University of Science and Technology of China), Dixiu Xue (Alibaba Group), Xu Shen (Alibaba Group), Xinmei Tian (USTC), Houqiang Li (University of Science and Technology of China), Jianqiang Huang (Alibaba Group), and Xian-Sheng Hua (Damo Academy, Alibaba Group)

Learning Instance-Level Spatial-Temporal Patterns for Person Re-Identification 14910
Min Ren (Center for Research on Intelligent Perception and Computing, (CRIPAC), Institute of Automation, Chinese Academy of Sciences, (CASIA), University of Chinese Academy of Sciences(UCAS)), Lingxiao He (AI Research of JD.com), Xingyu Liao (AI Research of JD.com), Wu Liu (AI Research of JD.com), Yunlong Wang (Center for Research on Intelligent Perception and Computing, (CRIPAC) National Laboratory of Pattern Recognition, (NLPR) Institute of Automation, Chinese Academy of Sciences, (CASIA)), and Tieniu Tan (NLPR, China)

Joint Visual Semantic Reasoning: Multi-Stage Decoder for Text Recognition 14920
Ayan Kumar Bhunia (University of Surrey), Aneeshan Sain (University of Surrey), Amandeep Kumar (West Bengal University of Technology), Shuvozit Ghose (Institute of Engineering and Management), Pinaki Nath Chowdhury (University of Surrey), and Yi-Zhe Song (University of Surrey)

Towards the Unseen: Iterative Text Recognition by Distilling From Errors 14930
Ayan Kumar Bhunia (University of Surrey), Pinaki Nath Chowdhury (University of Surrey), Aneeshan Sain (University of Surrey), and Yi-Zhe Song (University of Surrey)

ICE: Inter-Instance Contrastive Encoding for Unsupervised Person Re-Identification 14940
Hao Chen (INRIA), Benoit Lagadec (European System Integration), and François Bremond (Inria Sophia Antipolis, France)

Discover the Unknown Biased Attribute of an Image Classifier 14950
Zhiheng Li (University of Rochester) and Chenliang Xu (University of Rochester)

Understanding and Mitigating Annotation Bias in Facial Expression Recognition 14960
Yunliang Chen (University of California, Berkeley) and Jungseock Joo (University of California Los Angeles)

BiaSwap: Removing Dataset Bias With Bias-Tailored Swapping Augmentation 14972
Eungyeup Kim (Korea Advanced Institute of Science and Technology), Jihyeon Lee (Korea Advanced Institute of Science and Technology), and Jaegul Choo (Korea Advanced Institute of Science and Technology)

Learning Bias-Invariant Representation by Cross-Sample Mutual Information Minimization 14982
Wei Zhu (University of Rochester), Haitian Zheng (University of Rochester), Haofu Liao (Amazon), Weijian Li (University of Rochester), and Jiebo Luo (U. Rochester)

TransReID: Transformer-Based Object Re-Identification 14993
Shuting He (Zhejiang University), Hao Luo (Alibaba Group), Pichao Wang (Alibaba Group), Fan Wang (Alibaba Group), Hao Li (Alibaba Group), and Wei Jiang (Department of Control Science and Engineering, Zhejiang University)

Learning Self-Consistency for Deepfake Detection 15003
Tianchen Zhao (University of Michigan), Xiang Xu (Amazon), Mingze Xu (Amazon), Hui Ding (Amazon AWS Rekognition), Yuanjun Xiong (Amazon), and Wei Xia (Amazon)

Self-Supervised Domain Adaptation for Forgery Localization of JPEG Compressed Images 15014
Yuan Rao (Sun Yat-sen University, China) and Jiangqun Ni (Sun Yat-sen University, China)

Exploring Temporal Coherence for More General Video Face Forgery Detection 15024
Yinglin Zheng (School of Informatics, Xiamen University), Jianmin Bao (Microsoft Research Asia), Dong Chen (Microsoft Research Asia), Ming Zeng (School of Informatics, Xiamen University), and Fang Wen (Microsoft Research Asia)

TransForensics: Image Forgery Localization With Dense Self-Attention 15035
Jing Hao (Hikvision Research Institute), Zhixin Zhang (Hikvision Research Institute), Shicai Yang (Hikvision Research Institute), Di Xie (Hikvision Research Institute), and Shiliang Pu (Hikvision Research Institute)

Adaptive Label Noise Cleaning With Meta-Supervision for Deep Face Recognition 15045
Yaobin Zhang (Beijing University of Posts and Telecommunications), Weihong Deng (Beijing University of Posts and Telecommunications), Yaoyao Zhong (Beijing University of Posts and Telecommunications), Jiani Hu (Beijing University of Posts and Telecommunications), Xian Li (Canon Innovative Solution, (Beijing) Co., Ltd.), Dongyue Zhao (Canon Innovative Solution, (Beijing) Co., Ltd.), and Dongchao Wen (Canon Innovative Solution, (Beijing) Co., Ltd.)

Ensemble Attention Distillation for Privacy-Preserving Federated Learning 15056
Xuan Gong (University at Buffalo), Abhishek Sharma (United Imaging Intelligence), Srikrishna Karanam (United Imaging Intelligence), Ziyan Wu (United Imaging Intelligence), Terrence Chen (United Imaging Intelligence), David Doermann (University at Buffalo), and Arun Innanje (United Imaging Intelligence)

PASS: Protected Attribute Suppression System for Mitigating Bias in Face Recognition 15067
Prithviraj Dhar (Johns Hopkins University), Joshua Gleason (Univ of Maryland), Aniket Roy (Johns Hopkins University), Carlos D. Castillo (Johns Hopkins University), and Rama Chellappa (Johns Hopkins University)

Multi-Expert Adversarial Attack Detection in Person Re-Identification Using Context Inconsistency 15077
Xueping Wang (Hunan University), Shasha Li (University of California, Riverside), Min Liu (Hunan University), Yaonan Wang (Hunan University), and Amit K. Roy-Chowdhury (University of California, Riverside)

ID-Reveal: Identity-Aware DeepFake Video Detection 15088
Davide Cozzolino (University Federico II of Naples), Andreas Rössler (Technical University of Munich), Justus Thies (Max Planck Institute for Intelligent Systems), Matthias Nießner (Technical University of Munich), and Luisa Verdoliva (University Federico II of Naples)

Adversarial Learning

Social NCE: Contrastive Learning of Socially-Aware Motion Representations 15098
Yuejiang Liu (EPFL), Qi Yan (EPFL), and Alexandre Alahi (EPFL)

Hierarchical Object-to-Zone Graph for Object Navigation 15110
Sixian Zhang (ICT, China Academy of Science), Xinhang Song (ICT), Yubing Bai (ICT, China Academy of Science), Weijie Li (ICT, China Academy of Science), Yakui Chu (Huawei Application Innovate Laboratory), and Shuqiang Jiang (ICT, China Academy of Science)

GridToPix: Training Embodied Agents With Minimal Supervision 15121
Unnat Jain (UIUC), Iou-Jen Liu (University of Illinois at Urbana-Champaign), Svetlana Lazebnik (UIUC), Aniruddha Kembhavi (Allen Institute for Artificial Intelligence), Luca Weihs (Allen Institute for Artificial Intelligence), and Alexander G. Schwing (UIUC)

Active Learning for Lane Detection: A Knowledge Distillation Approach 15132
Fengchao Peng (Huawei Noah's Ark Lab), Chao Wang (Huawei Noah's Ark Lab), Jianzhuang Liu (Noah's Ark Lab, Huawei Technologies), and Zhen Yang (Huawei Noah's Ark Lab)

Waypoint Models for Instruction-Guided Navigation in Continuous Environments 15142
Jacob Krantz (Oregon State University), Aaron Gokaslan (Cornell), Dhruv Batra (Georgia Tech & Facebook AI Research), Stefan Lee (Oregon State University), and Oleksandr Maksymets (Facebook AI Research)

Geometry-Based Distance Decomposition for Monocular 3D Object Detection 15152
Xuepeng Shi (Imperial College London), Qi Ye (Zhejiang Univeristy), Xiaozhi Chen (DJI), Chuangrong Chen (DJI), Zhixiang Chen (Imperial College London), and Tae-Kyun Kim (Imperial College London)

GRF: Learning a General Radiance Field for 3D Representation and Rendering 15162
Alex Trevithick (Williams College) and Bo Yang (The Hong Kong Polytechnic University)

Change Is Everywhere: Single-Temporal Supervised Object Change Detection in Remote Sensing Imagery 15173
Zhuo Zheng (Wuhan University), Ailong Ma (Wuhan University), Liangpei Zhang (Wuhan University), and Yanfei Zhong (Wuhan University)

Indoor Scene Generation From a Collection of Semantic-Segmented Depth Images 15183
Ming-Jia Yang (Beihang University), Yu-Xiao Guo (MSRA), Bin Zhou (State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University), and Xin Tong (Microsoft)

Globally Optimal and Efficient Manhattan Frame Estimation by Delimiting Rotation Search Space 15193
Wuwei Ge (ZongMu Technology, (Shanghai) Limited Company), Yu Song (Beijing Jiaotong University), Baichao Zhang (ZongMu Technology, (Shanghai) Limited Company), and Zehua Dong (ZongMu Technology, (Shanghai) Limited Company)

End-to-End Urban Driving by Imitating a Reinforcement Learning Coach 15202
Zhejun Zhang (ETH Zurich), Alexander Liniger (ETH Zurich), Dengxin Dai (ETH Zurich), Fisher Yu (ETH Zurich), and Luc Van Gool (ETH Zurich)

From Goals, Waypoints & Paths to Long Term Human Trajectory Forecasting 15213
Karttikeya Mangalam (UC Berkeley), Yang An (Technical University of Munich), Harshayu Girase (University of California, Berkeley), and Jitendra Malik (University of California at Berkeley)

VSAC: Efficient and Accurate Estimator for H and F 15223
Maksym Ivashechkin (CMP CTU FEE), Daniel Barath (CMP CTU FEE), and Jiří Matas (ETH Zürich)

BabelCalib: A Universal Approach to Calibrating Central Cameras 15233
Yaroslava Lochman (Chalmers University of Technology), Kostiantyn Liepieshov (Ukrainian Catholic University), Jianhui Chen (Facebook), Michal Perdoch (Facebook), Christopher Zach (Chalmers University), and James Pritts (Chalmers University)

Robust Small Object Detection on the Water Surface Through Fusion of Camera and Millimeter Wave Radar 15243
Yuwei Cheng (Tsinghua University), Hu Xu (Northwestern Polytechnical University), and Yimin Liu (Tsinghua University)

FIERY: Future Instance Prediction in Bird's-Eye View From Surround Monocular Cameras 15253
Anthony Hu (University of Cambridge), Zak Murez, Nikhil Mohan (Wayve), Sofía Dudas (Wayve), Jeffrey Hawke (Wayve), Vijay Badrinarayanan (Wayve), Roberto Cipolla (University of Cambridge), and Alex Kendall (Wayve)

Fog Simulation on Real LiDAR Point Clouds for 3D Object Detection in Adverse Weather 15263
Martin Hahner (ETH Zurich), Christos Sakaridis (ETH Zurich), Dengxin Dai (ETH Zurich), and Luc Van Gool (ETH Zurich)

LSG-CPD: Coherent Point Drift With Local Surface Geometry for Point Cloud Registration 15273
Weixiao Liu (Johns Hopkins University), Hongtao Wu (Johns Hopkins University), and Gregory S. Chirikjian (Johns Hopkins University, USA)

DenseTNT: End-to-End Trajectory Prediction From Dense Goal Sets 15283
Junru Gu (Tsinghua University), Chen Sun (Google), and Hang Zhao (Tsinghua University)

Hierarchical Disentangled Representation Learning for Outdoor Illumination Estimation and Editing 15293
Piaopiao Yu (National Key Lab for Novel Software Technology, Nanjing University), Jie Guo (Nanjing University), Fan Huang (Nanjing University), Cheng Zhou (Nanjing University), Hongwei Che (Guangdong OPPO Mobile Telecommunications Corp Ltd), Xiao Ling (Guangdong OPPO Mobile Telecommunications Corp Ltd), and Yanwen Guo (Nanjing University)

Continuous Copy-Paste for One-Stage Multi-Object Tracking and Segmentation 15303
Zhenbo Xu (University of Science and Technology of China), Ajin Meng (University of Science and Technology of China), Zhenbo Shi (University of Science and Technology of China), Wei Yang (University of Science and Technology of China), Zhi Chen (University of Science and Technology of China), and Liusheng Huang (University of Science and Technology of China)

Deep Metric Learning for Open World Semantic Segmentation 15313
Jun Cen (The Hong Kong University of Science and Technology), Peng Yun (Hong Kong University of Science and Technology), Junhao Cai (HKUST), Michael Yu Wang (HKUST), and Ming Liu (HKUST)

PrimitiveNet: Primitive Instance Segmentation With Local Primitive Embedding Under Adversarial Metric 15323
Jingwei Huang (Huawei), Yanfeng Zhang (Huawei Technologies Co., Ltd.), and Mingwei Sun (Wuhan University)

Scribble-Supervised Semantic Segmentation Inference 15334
Jingshan Xu (Nanjing University of Science and Technology), Chuanwei Zhou (Nanjing University of Science and Technology), Zhen Cui (Nanjing University of Science and Technology), Chunyan Xu (Nanjing University of Science and Technology), Yuge Huang (Tencent YouTu), Pengcheng Shen (Tencent), Shaoxin Li (Tencent), and Jian Yang (Nanjing University of Science and Technology)

Scaling Up Instance Annotation via Label Propagation 15344
Dim P. Papadopoulos (MIT), Ethan Weber (MIT), and Antonio Torralba (MIT)

THDA: Treasure Hunt Data Augmentation for Semantic Navigation 15354
Oleksandr Maksymets (Facebook AI Research), Vincent Cartillier (Georgia Tech), Aaron Gokaslan (Cornell), Erik Wijmans (Georgia Tech), Wojciech Galuba (Oregon State University), Stefan Lee (Facebook), and Dhruv Batra (Georgia Tech & Facebook AI Research)

Pose Invariant Topological Memory for Visual Navigation 15364
Asuto Taniguchi (Ricoh Company, LTD.), Fumihiro Sasaki (Ricoh Company, LTD.), and Ryota Yamashina (Ricoh Company, LTD.)

FLAR: A Unified Prototype Framework for Few-Sample Lifelong Active Recognition 15374
Lei Fan (Northwestern University), Peixi Xiong (Northwestern University), Wei Wei (Northwestern University), and Ying Wu (Northwestern University)

Learning of Visual Relations: The Devil Is in the Tails 15384
Alakh Desai (UCSD), Tz-Ying Wu (UCSD), Subarna Tripathi (Intel Labs), and Nuno Vasconcelos (UCSD, USA)

Rethinking 360° Image Visual Attention Modelling With Unsupervised Learning. 15394
Yasser Abdelaziz Dahou Djilali (Dublin City University (DCU)), Tarun Krishna (Dublin City University (DCU)), Kevin McGuinness (Dublin City University (DCU)), and Noel E. O'Connor (Dublin City University (DCU))

Standardized Max Logits: A Simple yet Effective Approach for Identifying Unexpected Road Obstacles in Urban-Scene Segmentation 15405
Sanghun Jung (KAIST), Jungsoo Lee (Graduate School of Artificial Intelligence, KAIST), Daehoon Gwak (KAIST), Sungha Choi (LG AI Research), and Jaegul Choo (Korea Advanced Institute of Science and Technology)

4D-Net for Learned Multi-Modal Alignment 15415
AJ Piergiovanni (Google), Vincent Casser (Waymo), Michael S. Ryoo (Google; Stony Brook University), and Anelia Angelova (Google)

SPG: Unsupervised Domain Adaptation for 3D Object Detection via Semantic Point Generation . 15426
Qiangeng Xu (USC), Yin Zhou (Waymo), Weiyue Wang (Waymo), Charles R. Qi (Waymo), and Dragomir Anguelov (Waymo)

HiFT: Hierarchical Feature Transformer for Aerial Tracking 15437
Ziang Cao (Tongji University), Changhong Fu (Tongji University), Junjie Ye (Tongji University), Bowen Li (Tongji University), and Yiming Li (New York University)

Hierarchical Aggregation for 3D Instance Segmentation 15447
Shaoyu Chen (Huazhong University of Science and Technology), Jiemin Fang (Huazhong University of Science and Technology), Qian Zhang (Horizon Robotics), Wenyu Liu (Huazhong University of Science and Technology), and Xinggang Wang (Huazhong University of Science and Technology)

Learning Inner-Group Relations on Point Clouds 15457
Haoxi Ran (Tencent), Wei Zhuo (Tencent), Jun Liu (tencent), and Li Lu (Sichuan University)

VMNet: Voxel-Mesh Network for Geodesic-Aware 3D Semantic Segmentation 15468
Zeyu Hu (Hong Kong University of Science and Technology), Xuyang Bai (HKUST), Jiaxiang Shang (HKUST), Runze Zhang (Tencent), Jiayu Dong (Tencent Interactive Entertainment Group), Xin Wang (Tencent), Guangyuan Sun (TENCENT), Hongbo Fu (City University of Hong Kong), and Chiew-Lan Tai (Hong Kong University of Science & Technology)

Warp-Refine Propagation: Semi-Supervised Auto-Labeling via Cycle-Consistency 15479
Aditya Ganeshan (Preferred Networks Inc), Alexis Vallet (Preferred Networks, Inc.), Yasunori Kudo (Preferred Networks, Inc.), Shin-ichi Maeda (Preferred Networks, inc.), Tommi Kerola (Preferred Networks, Inc.), Rares Ambrus (Toyota Research Institute), Dennis Park (Toyota Research Institute), and Adrien Gaidon (Toyota Research Institute)

ReDAL: Region-Based and Diversity-Aware Active Learning for Point Cloud Semantic Segmentation 15490
Tsung-Han Wu (National Taiwan University), Yueh-Cheng Liu (National Taiwan University), Yu-Kai Huang (National Taiwan University), Hsin-Ying Lee (National Taiwan University), Hung-Ting Su (National Taiwan University), Ping-Chia Huang (National Taiwan University), and Winston H. Hsu (National Taiwan University)

Perturbed Self-Distillation: Weakly Supervised Large-Scale Point Cloud Semantic Segmentation 15500
Yachao Zhang (Xiamen University), Yanyun Qu (XMU), Yuan Xie (East China Normal University), Zonghao Li (School of Informatics Xiamen University), Shanshan Zheng (Xiamen University), and Cuihua Li (Xiamen university)

Revealing the Reciprocal Relations Between Self-Supervised Stereo and Monocular Depth Estimation 15509
Zhi Chen (University of Science and Technology of China), Xiaoqing Ye (baidu), Wei Yang (University of Science and Technology of China), Zhenbo Xu (University of Science and Technology of China), Xiao Tan (Baidu Inc.), Zhikang Zou (Huazhong university of science and technology), Errui Ding (Baidu Inc.), Xinming Zhang (University of Science and Technology of China), and Liusheng Huang (University of Science and Technology of China)

FOVEA: Foveated Image Magnification for Autonomous Navigation 15519
Chittesh Thavamani (Carnegie Mellon University), Mengtian Li (Carnegie Mellon University), Nicolas Cebron (Argo AI), and Deva Ramanan (Carnegie Mellon University)

Spatial Uncertainty-Aware Semi-Supervised Crowd Counting 15529
Yanda Meng (University of Liverpool), Hongrun Zhang (University of Liverpool), Yitian Zhao (Cixi Institute of Biomedical Engineering, Ningbo Institute of Industrial Technology, Chinese Academy of Sciences), Xiaoyun Yang (Remark Holdings), Xuesheng Qian (CSA Intellicloud Ltd), Xiaowei Huang (Liverpool University), and Yalin Zheng (University of Liverpool)

Excavating the Potential Capacity of Self-Supervised Monocular Depth Estimation 15540
Rui Peng (Peking University), Ronggang Wang (Peking University), Yawen Lai (Peking University), Luyang Tang (Peking University), and Yangang Cai (Peking University)

Crowd Counting With Partial Annotations in an Image 15550
*Yanyu Xu (Institute of High Performance Computing, A*Star), Ziming Zhong (ShanghaiTech University), Dongze Lian (Shanghaitech University), Jing Li (ShanghaiTech University), Zhengxin Li (ShanghaiTech University), Xinxing Xu (Institute of High Performance Computing, A*Star), and Shenghua Gao (Shanghaitech University)*

Personalized Trajectory Prediction via Distribution Discrimination 15560
Guangyi Chen (Tsinghua University), Junlong Li (Tsinghua University), Nuoxing Zhou (Tsinghua University), Liangliang Ren (Tsinghua University), and Jiwen Lu (Tsinghua University)

Learning To Drive From a World on Rails 15570
Dian Chen (UT Austin), Vladlen Koltun (Intel Labs), and Philipp Krähenbühl (UT Austin)

Bifold and Semantic Reasoning for Pedestrian Behavior Prediction 15580
Amir Rasouli (Huawei Technologies Canada), Mohsen Rohani (Huawei Noah's Ark), and Jun Luo (Huawei Technologies Canada Co. Ltd.)

Procedure Planning in Instructional Videos via Contextual Modeling and Model-Based Policy Learning 15591
Jing Bi (University of Rochester), Jiebo Luo (U. Rochester), and Chenliang Xu (University of Rochester)

Prediction by Anticipation: An Action-Conditional Prediction Method Based on Interaction Learning 15601
Ershad Banijamali (UNIVERSITY of WATERLOO), Mohsen Rohani (Huawei Noah's Ark), Elmira Amirloo (Huawei Technologies Canada Co. Ltd.), Jun Luo (Huawei Technologies Canada Co. Ltd.), and Pascal Poupart (University of Waterloo)

Robust 2D/3D Vehicle Parsing in Arbitrary Camera Views for CVIS 15611
Hui Miao (Beihang University), Feixiang Lu (Baidu Research), Zongdai Liu (Baidu Research), Liangjun Zhang (Baidu Research Institute), Dinesh Manocha (UMD), and Bin Zhou (State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University)

AutoShape: Real-Time Shape-Aware Monocular 3D Object Detection 15621
Zongdai Liu (Baidu Research), Dingfu Zhou (Baidu), Feixiang Lu (Baidu Research), Jin Fang (Baidu), and Liangjun Zhang (Baidu Research Institute)

Road Anomaly Detection by Partial Image Reconstruction With Segmentation Coupling 15631
Tomas Vojir (Czech Technical University, Prague), Tomáš Šipka (CTU FEE), Rahaf Aljundi (Toyota Motor Europe), Nikolay Chumerin (Toyota Motor Europe), Daniel Olmeda Reino (Toyota Motor Europe), and Jiri Matas (CMP CTU FEE)

Structured Bird's-Eye-View Traffic Scene Understanding From Onboard Images 15641
Yigit Baran Can (ETH Zurich), Alexander Liniger (ETH Zurich), Danda Pani Paudel (ETH Zürich), and Luc Van Gool (ETH Zurich)

Multi-View Radar Semantic Segmentation 15651
Arthur Ouaknine (Telecom Paris), Alasdair Newson (Telecom Paris), Patrick Pérez (Valeo.ai), Florence Tupin (Telecom Paris), and Julien Rebut (Valeo AI)

VIL-100: A New Dataset and a Baseline Model for Video Instance Lane Detection 15661
Yujun Zhang (Tianjin University), Lei Zhu (The Chinese University of Hong Kong), Wei Feng (College of Intelligence and Computing, Tianjin University, China), Huazhu Fu (IHPC, ASTAR), Mingqian Wang (Tianjin University), Qingxia Li (Automotive Data of China, (Tianjin) Co., Ltd), Cheng Li (Tianjin University), and Song Wang (University of South Carolina)

RobustNav: Towards Benchmarking Robustness in Embodied Navigation 15671
Prithvijit Chattopadhyay (Georgia Institute of Technology), Judy Hoffman (Georgia Tech), Roozbeh Mottaghi (Allen Institute for AI), and Aniruddha Kembhavi (Allen Institute for Artificial Intelligence)

Triggering Failures: Out-of-Distribution Detection by Learning From Local Adversarial Attacks in Semantic Segmentation 15681
Victor Besnier (valeo), Andrei Bursuc (valeo.ai), David Picard (ENPC), and Alexandre Briot (Valeo)

Robustness and Generalization via Generative Adversarial Training 15691
Omid Poursaeed (Cornell University), Tianxing Jiang (Cornell University), Harry Yang (Facebook), Serge Belongie (University of Copenhagen), and Ser-Nam Lim (Facebook AI)

Learnable Boundary Guided Adversarial Training 15701
Jiequan Cui (Chinese University of Hong Kong), Shu Liu (SmartMore), Liwei Wang (Tencent), and Jiaya Jia (Chinese University of Hong Kong)

Safety-Aware Motion Prediction With Unseen Vehicles for Autonomous Driving 15711
Xuanchi Ren (HKUST), Tao Yang (Xi'an JiaoTong University), Li Erran Li (Columbia University), Alexandre Alahi (EPFL), and Qifeng Chen (HKUST)

Toward Human-Like Grasp: Dexterous Grasping via Semantic Representation of Object-Hand .. 15721
Tianqiang Zhu (Dalian University of Technology), Rina Wu (Dalian University of Technology), Xiangbo Lin (Dalian University of Technology), and Yi Sun (Dalian University of Technology)

Act the Part: Learning Interaction Strategies for Articulated Object Part Discovery 15732
Samir Yitzhak Gadre (Columbia University), Kiana Ehsani (Allen Institute for Artificial Intelligence), and Shuran Song (Columbia University)

H2O: A Benchmark for Visual Human-Human Object Handover Analysis 15742
Ruolin Ye (Shanghai Jiao Tong University), Wenqiang Xu (Shanghai Jiao Tong University), Zhendong Xue (Shanghai Jiao Tong University), Tutian Tang (Shanghai Jiao Tong University), Yanfeng Wang (Cooperative medianet innovation center of Shanghai Jiao Tong University), and Cewu Lu (Shanghai Jiao Tong University)

The Functional Correspondence Problem 15752
Zihang Lai (CMU), Senthil Purushwalkam (Carnegie Mellon University), and Abhinav Gupta (CMU/FAIR)

Continual Neural Mapping: Learning an Implicit Scene Representation From Sequential Observations 15762
Zike Yan (Peking University), Yuxin Tian (Intel Labs China), Xuesong Shi (Intel Labs China), Ping Guo (Intel Labs China), Peng Wang (Intel Labs China), and Hongbin Zha (Peking University, China)

NEAT: Neural Attention Fields for End-to-End Autonomous Driving 15773
Kashyap Chitta (MPI-IS and University of Tuebingen), Aditya Prakash (Max Planck Institute for Intelligent Systems, Tuebingen), and Andreas Geiger (University of Tuebingen)

MGNet: Monocular Geometric Scene Understanding for Autonomous Driving 15784
Markus Schön (Ulm University), Michael Buchholz (Ulm University), and Klaus Dietmayer (Ulm University)

Visual Distant Supervision for Scene Graph Generation 15796
Yuan Yao (Tsinghua University), Ao Zhang (Tsinghua University), Xu Han (Tsinghua University), Mengdi Li (University of Hamburg), Cornelius Weber (Hamburg University), Zhiyuan Liu (Tsinghua University), Stefan Wermter (University of Hamburg), and Maosong Sun (Tsinghua University)

Generative Compositional Augmentations for Scene Graph Prediction 15807
Boris Knyazev (University of Guelph), Harm de Vries (Element AI), Cătălina Cangea (University of Cambridge), Graham W. Taylor (University of Guelph), Aaron Courville (Universite de Montreal), and Eugene Belilovsky (Concordia University)

In-Place Scene Labelling and Understanding With Implicit Scene Representation 15818
Shuaifeng Zhi (Imperial College London), Tristan Laidlow (Imperial College London), Stefan Leutenegger (Imperial College London), and Andrew J. Davison (Imperial College London)

Interaction via Bi-Directional Graph of Semantic Region Affinity for Scene Parsing 15828
Henghui Ding (Nanyang Technological University), Hui Zhang (Nanyang Technological University), Jun Liu (Singapore University of Technology and Design), Jiaxin Li (Bytedance, TikTok), Zijian Feng (ByteDance AI Lab), and Xudong Jiang (Nanyang Technological University)

Enhanced Boundary Learning for Glass-Like Object Segmentation 15839
Hao He (NLPR CASIA), Xiangtai Li (Peking University), Guangliang Cheng (Sensetime Group Limited), Jianping Shi (Sensetime Group Limited), Yunhai Tong (Peking University), Gaofeng Meng (Chinese Academy of Sciences), Véronique Prinet (Institute of Automation, Chinese Academy of Sciences), and LuBin Weng (CASIC)

Exploring Relational Context for Multi-Task Dense Prediction 15849
David Brüggemann (ETH Zurich), Menelaos Kanakis (ETH Zurich), Anton Obukhov (ETH Zurich), Stamatios Georgoulis (ETH Zurich), and Luc Van Gool (ETH Zurich)

Segmentation-Grounded Scene Graph Generation 15859
Siddhesh Khandelwal (University of British Columbia), Mohammed Suhail (University of British Columbia), and Leonid Sigal (University of British Columbia)

Visual Graph Memory With Unsupervised Representation for Visual Navigation 15870
Obin Kwon (Seoul National University), Nuri Kim (Seoul National University), Yunho Choi (Seoul National University), Hwiyeon Yoo (Seoul National University), Jeongho Park (Seoul National University), and Songhwai Oh (Seoul National University)

Topic Scene Graph Generation by Attention Distillation From Caption 15880
Wenbin Wang (Institute of Computing Technology, Chinese Academy of Sciences), Ruiping Wang (Institute of Computing Technology, Chinese Academy of Sciences), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences)

Grounding Consistency: Distilling Spatial Common Sense for Precise Visual Relationship Detection 15891
Markos Diomataris (DeepLab), Nikolaos Gkanatsios (Carnegie Mellon University), Vassilis Pitsikalis (DeepLab), and Petros Maragos (National Technical University of Athens)

Exploring Long Tail Visual Relationship Recognition With Large Vocabulary 15901
Sherif Abdelkarim (KAUST), Aniket Agarwal (IIT Roorkee), Panos Achlioptas (Stanford University), Jun Chen (King Abdullah University of Science and Technology), Jiaji Huang (Baidu Research), Boyang Li (Nanyang Technological University), Kenneth Church (Baidu), and Mohamed Elhoseiny (KAUST)

Context-Aware Scene Graph Generation With Seq2Seq Transformers 15911
Yichao Lu (Layer 6 AI), Himanshu Rai (Layer6 AI), Jason Chang (Layer6 AI), Boris Knyazev (University of Guelph), Guangwei Yu (Layer6 AI), Shashank Shekhar (University of Guelph), Graham W. Taylor (University of Guelph), and Maksims Volkovs (Layer6 AI)

Episodic Transformer for Vision-and-Language Navigation 15922
Alexander Pashevich (INRIA), Cordelia Schmid (Google), and Chen Sun (Google)

Interpretation of Emergent Communication in Heterogeneous Collaborative Embodied Agents . 15933
Shivansh Patel (Indian Institute of Technology Kanpur), Saim Wani (Indian Institute of Technology Kanpur), Unnat Jain (UIUC), Alexander G. Schwing (UIUC), Svetlana Lazebnik (UIUC), Manolis Savva (Simon Fraser University), and Angel X. Chang (Simon Fraser University)

Graspness Discovery in Clutters for Fast and Accurate Grasp Detection 15944
Chenxi Wang (Shanghai Jiao Tong University), Hao-Shu Fang (SJTU), Minghao Gou (Shanghai Jiao Tong University), Hongjie Fang (Shanghai Jiao Tong University), Jin Gao (Shanghai Jiaotong University), and Cewu Lu (Shanghai Jiao Tong University)

Pose Correction for Highly Accurate Visual Localization in Large-Scale Indoor Spaces 15954
Janghun Hyeon (Korea University), Joohyung Kim (Korea University), and Nakju Doh (Korea University)

Exploiting Scene Graphs for Human-Object Interaction Detection 15964
Tao He (Monash University), Lianli Gao (The University of Electronic Science and Technology of China), Jingkuan Song (UESTC), and Yuan-Fang Li (Monash University)

Deep Hough Voting for Robust Global Registration 15974
Junha Lee (Postech), Seungwook Kim (POSTECH), Minsu Cho (POSTECH), and Jaesik Park (POSTECH)

P2-Net: Joint Description and Detection of Local Features for Pixel and Point Matching 15984
Bing Wang (University of Oxford), Changhao Chen (National University of Defense Technology), Zhaopeng Cui (Zhejiang University), Jie Qin (Inception Institute of Artificial Intelligence), Chris Xiaoxuan Lu (University of Edinburgh), Zhengdi Yu (Durham University), Peijun Zhao (University of Oxford), Zhen Dong (Wuhan University), Fan Zhu (Inception Institute of Artificial Intelligence), Niki Trigoni (University of Oxford), and Andrew Markham (University of Oxford)

HRegNet: A Hierarchical Network for Large-Scale Outdoor LiDAR Point Cloud Registration 15994
Fan Lu (Tongji University), Guang Chen (Tongji University), Yinlong Liu (Technische Universität München), Lijun Zhang (Tongji University), Sanqing Qu (Tongji University), Shu Liu (ETHZ), and Rongqi Gu (westwell-lab)

RPVNet: A Deep and Efficient Range-Point-Voxel Fusion Network for LiDAR Point Cloud Segmentation 16004
Jianyun Xu (HIKVISION), Ruixiang Zhang (Zhejiang University), Jian Dou (hikvision), Yushi Zhu (Hikvision Research Institute), Jie Sun (Hikvision), and Shiliang Pu (Hikvision Research Institute)

MonteFloor: Extending MCTS for Reconstructing Accurate Large-Scale Floor Plans 16014
Sinisa Stekovic (Graz University of Technology), Mahdi Rad (TU Graz), Friedrich Fraundorfer (Graz University of Technology), and Vincent Lepetit (Ecole des Ponts ParisTech)

Self-Supervised Real-to-Sim Scene Generation 16024
Aayush Prakash (Nvidia), Shoubhik Debnath (Nvidia), Jean-Francois Lafleche (Nvidia), Eric Cameracci (Nvidia), Gavriel State (NVIDIA), Stan Birchfield (NVIDIA), and Marc T. Law (NVIDIA)

Regularizing Nighttime Weirdness: Efficient Self-Supervised Monocular Depth Estimation in the Dark 16035
Kun Wang (Nanjing University of Science and Technology), Zhenyu Zhang (Tencent), Zhiqiang Yan (Nanjing University of Science and Tenchnology), Xiang Li (Nanjing University of Science and Technology), Baobei Xu (HikVision), Jun Li (Nanjing University of Science and Technology), and Jian Yang (Nanjing University of Science and Technology)

Variational Attention: Propagating Domain-Specific Knowledge for Multi-Domain Learning in Crowd Counting 16045
Binghui Chen (Alibaba), Zhaoyi Yan (Harbin Institute of Technology), Ke Li (Artificial Intelligence Center, DAMO Academy, Alibaba Group, Hangzhou, China), Pengyu Li (Alibaba Group), Biao Wang (Alibaba), Wangmeng Zuo (Harbin Institute of Technology, China), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)

GP-S3Net: Graph-Based Panoptic Sparse Semantic Segmentation Network 16056
Ryan Razani (Huawei Technology Canada), Ran Cheng (mcgill university), Enxu Li (Huawei), Ehsan Taghavi (Huawei Technologies Canada), Yuan Ren (Huawei), and Liu Bingbing (Huawei Noah's Ark Lab, Canada)

VolumeFusion: Deep Depth Fusion for 3D Scene Reconstruction 16066
Jaesung Choe (KAIST), Sunghoon Im (DGIST), Francois Rameau (KAIST), Minjun Kang (KAIST), and In So Kweon (KAIST)

RAIN: Reinforced Hybrid Attention Inference Network for Motion Forecasting 16076
Jiachen Li (University of California, Berkeley), Fan Yang (Carnegie Mellon University), Hengbo Ma (University of California, Berkeley), Srikanth Malla (Honda Research Institute), Masayoshi Tomizuka (University of California, Berkeley), and Chiho Choi (Honda Research Institute US)

LookOut: Diverse Multi-Future Prediction and Planning for Self-Driving 16087
Alexander Cui (Uber ATG, California Institute of Technology), Sergio Casas (Uber ATG / University of Toronto), Abbas Sadat (Uber ATG), Renjie Liao (University of Toronto), and Raquel Urtasun (Uber ATG)

Auxiliary Tasks and Exploration Enable ObjectGoal Navigation 16097
Joel Ye (Georgia Institute of Technology), Dhruv Batra (Georgia Tech & Facebook AI Research), Abhishek Das (Facebook AI Research), and Erik Wijmans (Georgia Tech)

The Surprising Effectiveness of Visual Odometry Techniques for Embodied PointGoal Navigation 16107
Xiaoming Zhao (University of Illinois at Urbana-Champaign), Harsh Agrawal (Georgia Institute of Technology), Dhruv Batra (Georgia Tech & Facebook AI Research), and Alexander G. Schwing (UIUC)

Glimpse-Attend-and-Explore: Self-Attention for Active Visual Exploration 16117
Soroush Seifi (KU Leuven), Abhishek Jha (K.U. Leuven), and Tinne Tuytelaars (KU Leuven)

Bayesian Deep Basis Fitting for Depth Completion With Uncertainty 16127
Chao Qu (University of Pennsylvania), Wenxin Liu (University of Pennsylvania), and Camillo J. Taylor (University of Pennsylvania)

Admix: Enhancing the Transferability of Adversarial Attacks 16138
Xiaosen Wang (Huazhong University of Science and Technology), Xuanran He (Nanyang Technological University), Jingdong Wang (Microsoft), and Kun He (Huazhong University of Science and Technology)

Aha! Adaptive History-Driven Attack for Decision-Based Black-Box Models 16148
Jie Li (Xiamen University), Rongrong Ji (Xiamen University, China), Peixian Chen (Xiamen University), Baochang Zhang (Beihang University), Xiaopeng Hong (Xi'an Jiaotong University), Ruixin Zhang (Tencent), Shaoxin Li (Tencent), Jilin Li (Tencent), Feiyue Huang (Tencent), and Yongjian Wu (Tencent Technology, (Shanghai) Co., Ltd)

Knowledge-Enriched Distributional Model Inversion Attacks 16158
Si Chen (Virginia Tech), Mostafa Kahla (Virginia Tech), Ruoxi Jia (Virginia Tech), and Guo-Jun Qi (Futurewei Technologies)

Attack As the Best Defense: Nullifying Image-to-Image Translation GANs via Limit-Aware Adversarial Attack 16168
Chin-Yuan Yeh (National Taiwan University), Hsi-Wen Chen (National Taiwan University), Hong-Han Shuai (National Chiao Tung University), De-Nian Yang (Academia Sinica), and Ming-Syan Chen (National Taiwan University)

Full-Velocity Radar Returns by Radar-Camera Fusion 16178
Yunfei Long (Michigan State University), Daniel Morris (Michigan State University), Xiaoming Liu (Michigan State University), Marcos Castro (Ford Motor Company), Punarjay Chakravarty (Ford Motor Company), and Praveen Narayanan (Ford Motor Company)

PU-EVA: An Edge-Vector Based Approximation Solution for Flexible-Scale Point Cloud Upsampling 16188
Luqing Luo (University of Macau), Lulu Tang (University of Macau), Wanyi Zhou (South China University of Technology), Shizheng Wang (Chinese Academy of Sciences), and Zhi-Xin Yang (University of Macau)

A Closer Look at Rotation-Invariant Deep Point Cloud Analysis 16198
Feiran Li (Osaka University), Kent Fujiwara (LINE Corporation), Fumio Okura (Osaka University), and Yasuyuki Matsushita (Osaka University)

CTRL-C: Camera Calibration TRansformer With Line-Classification 16208
Jinwoo Lee (Kakao Brain), Hyunsung Go (Kookmin University), Hyunjoon Lee (Kakao Brain), Sunghyun Cho (POSTECH), Minhyuk Sung (KAIST), and Junho Kim (Kookmin University)

SurfGen: Adversarial 3D Shape Synthesis With Explicit Surface Discriminators 16218
Andrew Luo (Carnegie Mellon University), Tianqin Li (Carnegie Mellon University), Wen-Hao Zhang (University of Chicago), and Tai Sing Lee (Carnegie Mellon University)

Adaptive Focus for Efficient Video Recognition 16229
Yulin Wang (Tsinghua University), Zhaoxi Chen (Tsinghua University), Haojun Jiang (Tsinghua University), Shiji Song (Tsinghua), Yizeng Han (Tsinghua University), and Gao Huang (Tsinghua)

Point Transformer 16239
Hengshuang Zhao (University of Oxford), Li Jiang (The Chinese University of Hong Kong), Jiaya Jia (Chinese University of Hong Kong), Philip H.S. Torr (University of Oxford), and Vladlen Koltun (Intel Labs)

Transformer-Based Attention Networks for Continuous Pixel-Wise Prediction 16249
Guanglei Yang (Harbin Institute of Technology), Hao Tang (ETH Zurich), Mingli Ding (Harbin institute of Technology), Nicu Sebe (University of Trento), and Elisa Ricci (University of Trento)

Perception-Aware Multi-Sensor Fusion for 3D LiDAR Semantic Segmentation 16260
Zhuangwei Zhuang (South China University of Technology), Rong Li (South China University of Technology), Kui Jia (South China University of Technology), Qicheng Wang (Shenzhen Youjia Innov Tech Co., Ltd), Yuanqing Li (South China University of Technology), and Mingkui Tan (South China University of Technology)

Region-Aware Contrastive Learning for Semantic Segmentation 16271
Hanzhe Hu (Peking University), Jinshi Cui (Peking University, China), and Liwei Wang (Peking University)

Refining Action Segmentation With Hierarchical Video Representations 16282
Hyemin Ahn (Technical University of Munich) and Dongheui Lee (Technical University of Munich)

Condensing a Sequence to One Informative Frame for Video Recognition 16291
Zhaofan Qiu (JD.com), Ting Yao (JD AI Research), Yan Shu (University of Science and Techology of China), Chong-Wah Ngo (Singapore Management University), and Tao Mei (AI Research of JD.com)

Vision-Language Transformer and Query Generation for Referring Segmentation 16301
Henghui Ding (Nanyang Technological University), Chang Liu (Nanyang Technological University), Suchen Wang (Nanyang Technological University), and Xudong Jiang (Nanyang Technological University)

Salient Object Ranking With Position-Preserved Attention 16311
Hao Fang (ZJU), Daoxin Zhang (Alibaba Group), Yi Zhang (Zhejiang University), Minghao Chen (Zhejiang University), Jiawei Li (Alibaba Group), Yao Hu (Alibaba Youku Cognitive and Intelligent Lab), Deng Cai (ZJU), and Xiaofei He (Zhejiang University)

Few-Shot Visual Relationship Co-Localization 16322
Revant Teotia (Indian Institute of Technology, Jodhpur), Vaibhav Mishra (Indian Institute of Technology, Jodhpur), Mayank Maheshwari (Indian Institute of Technology, Jodhpur), and Anand Mishra (Indian Institute of Technology, Jodhpur)

Graph-to-3D: End-to-End Generation and Manipulation of 3D Scenes Using Scene Graphs 16332
Helisa Dhamo (Technical University of Munich), Fabian Manhardt (Google), Nassir Navab (TU Munich, Germany), and Federico Tombari (Google, TU Munich)

Unconditional Scene Graph Generation 16342
Sarthak Garg (Technische Universität München), Helisa Dhamo (Technical University of Munich), Azade Farshad (Technical University of Munich), Sabrina Musatian (Technical University of Munich), Nassir Navab (TU Munich, Germany), and Federico Tombari (Google, TU Munich)

Spatial-Temporal Transformer for Dynamic Scene Graph Generation 16352
Yuren Cong (Leibniz University Hannover), Wentong Liao (Leibniz University Hannover), Hanno Ackermann (Leibniz University Hannover), Bodo Rosenhahn (Leibniz University Hannover), and Michael Ying Yang (University of Twente)

From General to Specific: Informative Scene Graph Generation via Balance Adjustment 16363
Yuyu Guo (UESTC), Lianli Gao (The University of Electronic Science and Technology of China), Xuanhan Wang (University of Electronic Science and Technology of China), Yuxuan Hu (Southwest University), Xing Xu (University of Electronic Science and Technology of China), Xu Lu (Kwai), Heng Tao Shen (University of Electronic Science and Technology of China, (UESTC)), and Jingkuan Song (UESTC)

A Simple Baseline for Weakly-Supervised Scene Graph Generation 16373
Jing Shi (University of Rochester), Yiwu Zhong (University of Wisconsin-Madison), Ning Xu (Adobe Research), Yin Li (University of Wisconsin-Madison), and Chenliang Xu (University of Rochester)

Cross-Modality Person Re-Identification via Modality Confusion and Center Aggregation 16383
Xin Hao (Beijing Institute of Technology), Sanyuan Zhao (Beijing Institute of Technology), Mang Ye (Wuhan University), and Jianbing Shen (Inception Institute of Artificial Intelligence)

Practical Relative Order Attack in Deep Ranking 16393
Mo Zhou (Xi'an Jiaotong University), Le Wang (Xi'an Jiaotong University), Zhenxing Niu (Alibaba Group), Qilin Zhang (HERE Technologies), Yinghui Xu (Alibaba DAMO Academy), Nanning Zheng (Xi'an Jiaotong University), and Gang Hua (Wormpex AI Research)

Low Curvature Activations Reduce Overfitting in Adversarial Training 16403
Vasu Singla (University of Maryland), Sahil Singla (University of Maryland), Soheil Feizi (University of Maryland), and David Jacobs (University of Maryland, USA)

Defending Against Universal Adversarial Patches by Clipping Feature Norms 16414
Cheng Yu (Tsinghua University), Jiansheng Chen (University of Science and Technology Beijing), Youze Xue (Tsinghua University), Yuyang Liu (Tsinghua University), Weitao Wan (Tsinghua University), Jiayu Bao (Tsinghua University), and Huimin Ma (University of Science and Technology Beijing)

Revisiting Adversarial Robustness Distillation: Robust Soft Labels Make Student Better 16423
Bojia Zi (Fudan University), Shihao Zhao (Fudan University), Xingjun Ma (Deakin University), and Yu-Gang Jiang (Fudan University)

CLEAR: Clean-Up Sample-Targeted Backdoor in Neural Networks 16433
Liuwan Zhu (Old Dominion University), Rui Ning (Old Dominion University), Chunsheng Xin (Old Dominion University), Chonggang Wang (InterDigital), and Hongyi Wu (Old Dominion University)

Invisible Backdoor Attack With Sample-Specific Triggers 16443
Yuezun Li (Ocean University of China), Yiming Li (Tsinghua University), Baoyuan Wu (The Chinese University of Hong Kong, Shenzhen; Shenzhen Research Institute of Big Data), Longkang Li (CUHK), Ran He (Institute of Automation, Chinese Academy of Sciences), and Siwei Lyu (University at Buffalo)

Rethinking the Backdoor Attacks' Triggers: A Frequency Perspective 16453
Yi Zeng (Virginia Tech), Won Park (University of Michigan), Z. Morley Mao (University of Michigan), and Ruoxi Jia (Virginia Tech)

Black-Box Detection of Backdoor Attacks With Limited Information and Data 16462
Yinpeng Dong (Tsinghua University), Xiao Yang (Tsinghua University), Zhijie Deng (Tsinghua University), Tianyu Pang (Tsinghua University), Zihao Xiao (RealAI, Inc.), Hang Su (Tsinghua Univiersity), and Jun Zhu (Tsinghua University)

PointBA: Towards Backdoor Attacks in 3D Point Cloud 16472
*Xinke Li (National University of Singapore), Zhirui Chen (National University of Singapore), Yue Zhao (National University of Singapore), Zekun Tong (National University of Singapore), Yabang Zhao (National University of Singapore), Andrew Lim (National University of Singapore), and Joey Tianyi Zhou (Institute of High Performance Computing, A*STAR)*

Author Index