

2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2022)

**Waikoloa, Hawaii, USA
3 – 8 January 2022**

Pages 1-686



**IEEE Catalog Number: CFP22082-POD
ISBN: 978-1-6654-0916-2**

**Copyright © 2022 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:	CFP22082-POD
ISBN (Print-On-Demand):	978-1-6654-0916-2
ISBN (Online):	978-1-6654-0915-5
ISSN:	2472-6737

Additional Copies of This Publication Are Available From:

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

CURRAN ASSOCIATES INC.
proceedings
.com

2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) WACV 2022

Table of Contents

Message from the General and Program Chairs	xlv
Organizing Committee and Area Chairs	xlvi
Sponsors and Corporate Donors	xlvii

Computational Photography and Image Processing I

Evaluation of Correctness in Unsupervised Many-to-Many Image Translation	1
<i>Dina Bashkirova (Boston University), Ben Usman (Boston University), and Kate Saenko (Boston University)</i>	
Fast and Explicit Neural View Synthesis	11
<i>Pengsheng Guo (Apple), Miguel Angel Bautista (Apple), Alex Colburn (Apple Inc.), Liang Yang (Apple Inc), Daniel Ulbricht (Apple), Joshua M. Susskind (Apple), and Qi Shan (Apple Inc.)</i>	
Training a Task-Specific Image Reconstruction Loss	21
<i>Aamir Mustafa (University of Cambridge), Aliaksei Mikhailiuk (University of Cambridge), Dan Andrei Iliescu (University of Cambridge), Varun Babbar (University of Cambridge), and Rafał K. Mantiuk (Univ. Cambridge)</i>	
GANs Spatial Control via Inference-Time Adaptive Normalization	31
<i>Karin Jakoel (Technion), Liron Efraim (Technion), and Tamar Rott Shaham (Technion)</i>	
Single-Photon Camera Guided Extreme Dynamic Range Imaging	41
<i>Yuhao Liu (University of Wisconsin-Madison), Felipe Gutierrez-Barragan (University of Wisconsin-Madison), Atul Ingle (University of Wisconsin-Madison), Mohit Gupta (University of Wisconsin-Madison, USA), and Andreas Velten (University of Wisconsin - Madison)</i>	
Extracting Vignetting and Grain Filter Effects From Photos	52
<i>Abdelrahman Abdelhamed (Samsung), Jonghwa Yim (Samsung Electronics), Abhijith Punnappurath (Samsung AI Center Toronto), Michael S. Brown (York University), Jihwan Choe (Samsung), and Kihwan Kim (Samsung)</i>	
High Dynamic Range Imaging of Dynamic Scenes With Saturation Compensation but Without Explicit Motion Compensation	61
<i>Haesoo Chung (Seoul National University) and Nam Ik Cho (Seoul National University)</i>	

HERS Superpixels: Deep Affinity Learning for Hierarchical Entropy Rate Segmentation	72
<i>Hankui Peng (University of Cambridge), Angelica I. Aviles-Rivero (University of Cambridge), and Carola-Bibiane Schönlieb (Cambridge University)</i>	
Improving Single-Image Defocus Deblurring: How Dual-Pixel Images Help Through Multi-Task Learning	82
<i>Abdullah Abuolaim (York University), Mahmoud Afifi (Apple), and Michael S. Brown (York University)</i>	
Image Restoration by Deep Projected GSURE	91
<i>Shady Abu-Hussein (Tel Aviv University), Tom Tirer (Tel Aviv University), Se Young Chun (Seoul National University), Yonina C. Eldar, and Raja Giryes (Tel Aviv University)</i>	
Non-Blind Deblurring for Fluorescence: A Deformable Latent Space Approach With Kernel Parameterization	101
<i>Ziqiao Guan (Stony Brook University), Esther H. R. Tsai (Brookhaven National Laboratory), Xiaojing Huang (Brookhaven National Laboratory), Kevin G. Yager (Brookhaven National Laboratory), and Hong Qin (Stony Brook University)</i>	
Controlled GAN-Based Creature Synthesis via a Challenging Game Art Dataset – Addressing the Noise-Latent Trade-Off	110
<i>Vaibhav Vavilala (University of Illinois at Urbana-Champaign) and David Forsyth (Univeristy of Illinois at Urbana-Champaign)</i>	

Action and Pose

Hierarchical Modeling for Task Recognition and Action Segmentation in Weakly-Labeled Instructional Videos	120
<i>Reza Ghoddoosian (University of Texas at Arlington), Saif Sayed (University of Texas at Arlington), and Vassilis Athitsos (University of Texas at Arlington)</i>	
Quantified Facial Expressiveness for Affective Behavior Analytics	131
<i>Md Taufeeq Uddin (University of South Florida) and Shaun Canavan (University of South Florida)</i>	
Pose and Joint-Aware Action Recognition	141
<i>Anshul Shah (Johns Hopkins University), Shlok Mishra (University of Maryland College Park), Ankan Bansal (Amazon.com), Jun-Cheng Chen (Academia Sinica), Rama Chellappa (Johns Hopkins University), and Abhinav Shrivastava (University of Maryland)</i>	
Equine Pain Behavior Classification via Self-Supervised Disentangled Pose Representation	152
<i>Maheen Rashid (Univrses AB), Sofia Broomé (KTH Royal Institute of Technology), Katrina Ask (Swedish University of Agricultural Sciences), Elin Hernlund (Swedish University of Agricultural Sciences), Pia Haubro Andersen (Swedish University of Agricultural Sciences), Hedvig Kjellström (KTH Royal Institute of Technology), and Yong Jae Lee (University of California, Davis)</i>	

Domain Generalization Through Audio-Visual Relative Norm Alignment in First Person Action Recognition	163
<i>Mirco Planamente (Politecnico di Torino), Chiara Plizzari (Politecnico di Torino), Emanuele Alberti (Politecnico di Torino), and Barbara Caputo (Politecnico di Torino)</i>	
SSCAP: Self-Supervised Co-Occurrence Action Parsing for Unsupervised Temporal Action Segmentation	175
<i>Zhe Wang (UC-Irvine), Hao Chen (Amazon), Xinyu Li (Amazon), Chunhui Liu (Amazon), Yuanjun Xiong (Amazon), Joseph Tighe (Amazon), and Charless Fowlkes (UC Irvine)</i>	
Multi-Stream Dynamic Video Summarization	185
<i>Mohamed Elfeki (Microsoft), Liqiang Wang (University of Central Florida), and Ali Borji (MarkableAI)</i>	
SporeAgent: Reinforced Scene-Level Plausibility for Object Pose Refinement	196
<i>Dominik Bauer (TU Wien), Timothy Patten (University of Technology Sydney), and Markus Vincze (TU Wien)</i>	
A Structure-Aware Method for Direct Pose Estimation	205
<i>Hunter Blanton (University of Kentucky), Scott Workman (DZYNE Technologies), and Nathan Jacobs (University of Kentucky)</i>	

3D Computer Vision I

Stylizing 3D Scene via Implicit Representation and HyperNetwork	215
<i>Pei-Ze Chiang (National Chiao Tung University), Meng-Shiun Tsai (National Chiao Tung University), Hung-Yu Tseng (Facebook), Wei-Sheng Lai (Google), and Wei-Chen Chiu (National Chiao Tung University)</i>	
SIDE: Center-Based Stereo 3D Detector With Structure-Aware Instance Depth Estimation	225
<i>Xidong Peng (ShanghaiTech University), Xinge Zhu (The Chinese University of Hong Kong), Tai Wang (The Chinese University of Hong Kong), and Yuexin Ma (ShanghaiTech University)</i>	
Learning To Reconstruct 3D Non-Cuboid Room Layout From a Single RGB Image	235
<i>Cheng Yang (University of Electronic Science and Technology of China), Jia Zheng (Manycore Tech Inc. (Kujiale)), Xili Dai (University of California, Berkeley), Rui Tang (Kujiale.com), Yi Ma (UC Berkeley), and Xiaojun Yuan (University of Electronic Science and Technology of China)</i>	
Single-Shot Dense Active Stereo With Pixel-Wise Phase Estimation Based on Grid-Structure Using CNN and Correspondence Estimation Using GCN	245
<i>Ryo Furukawa (Hiroshima city univ.), Michihiro Mikamo (Hiroshima city univ.), Ryusuke Sagawa (AIST, Japan), and Hiroshi Kawasaki (Kyushu univ.)</i>	
EllipsoidNet: Ellipsoid Representation for Point Cloud Classification and Segmentation	256
<i>Yecheng Lyu (Worcester Polytechnic Institute), Xinming Huang (Worcester Polytechnic Institute), and Ziming Zhang (Worcester Polytechnic Institute)</i>	

What Makes for Effective Few-Shot Point Cloud Classification?	267
<i>Chuangguan Ye (Fudan University), Hongyuan Zhu (Astar), Yongbin Liao (Fudan University), Yanggang Zhang (Fudan University), Tao Chen (Fudan University), and Jiayuan Fan (Fudan University)</i>	
Mending Neural Implicit Modeling for 3D Vehicle Reconstruction in the Wild	277
<i>Shivam Duggal (Carnegie Mellon University), Zihao Wang (ByteDance AI-LAB), Wei-Chiu Ma (MIT), Sivabalan Manivasagam (University of Toronto), Justin Liang (Waabi), Shenlong Wang (Uber ATG, University of Toronto), and Raquel Urtasun (Uber ATG)</i>	
HybVIO: Pushing the Limits of Real-Time Visual-Inertial Odometry	287
<i>Otto Seiskari (Spectacular AI), Pekka Rantalankila (Spectacular AI), Juho Kannala (Aalto University, Finland), Jerry Ylilammi (Spectacular AI), Esa Rahtu (Tampere University), and Arno Solin (Aalto University)</i>	
Occlusion Resistant Network for 3D Face Reconstruction	297
<i>Hitika Tiwari (Indian Institute of Technology Kanpur & National Yang Ming Chiao Tung University), Vinod K. Kurmi (KU Leuven), K.S. Venkatesh (IIT Kanpur), and Yong-Sheng Chen (National Chiao Tung University)</i>	
StickyLocalization: Robust End-to-End Relocalization on Point Clouds Using Graph Neural Networks	307
<i>Kai Fischer (Valeo), Martin Simon (Valeo), Stefan Milz (Spleenlab GmbH), and Patrick Mäder (Technische Universität Ilmenau)</i>	
Symmetric-Light Photometric Stereo	317
<i>Kazuma Minami (Osaka university), Hiroaki Santo (Osaka University), Fumio Okura (Osaka University), and Yasuyuki Matsushita (Osaka University)</i>	
Lightweight Monocular Depth With a Novel Neural Architecture Search Method	326
<i>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)</i>	

Transfer, Few-shot, Semi-, and Un- Supervised Learning I

Knowledge Capture and Replay for Continual Learning	337
<i>Saisubramaniam Gopalakrishnan (Institute for Infocomm Research, A*STAR), Pranshu Ranjan Singh (Institute for Infocomm Research, A*STAR), Haytham Fayek (RMIT), Savitha Ramasamy (I2R A*STAR), and ArulMurugan Ambikapathi (ASTAR-Institute for Infocomm Research, A*STAR)</i>	
Meta-Learning for Multi-Label Few-Shot Classification	346
<i>Christian Simon (Australian National University), Piotr Koniusz (ANU College of Engineering and Computer Science), and Mehrtash Harandi (Monash University)</i>	
Cleaning Noisy Labels by Negative Ensemble Learning for Source-Free Unsupervised Domain Adaptation	356
<i>Waqar Ahmed (Istituto Italiano di Tecnologia), Pietro Morerio (Istituto Italiano di Tecnologia), and Vittorio Murino (Istituto Italiano di Tecnologia)</i>	

Learning Foreground-Background Segmentation From Improved Layered GANs	366
<i>Yu Yang (Tsinghua University), Hakan Bilen (University of Edinburgh), Qiran Zou (Tsinghua University), Wing Yin Cheung (Tsinghua University), and Xiangyang Ji (Tsinghua University)</i>	
To Miss-Attend Is to Misalign! Residual Self-Attentive Feature Alignment for Adapting Object Detectors	376
<i>Vaishnavi Khindkar (IIITH), Chetan Arora (Indian Institute of Technology Delhi), Vineeth N Balasubramanian (Indian Institute of Technology, Hyderabad), Anbumani Subramanian (IIIT-Hyderabad), Rohit Saluja (IIIT-Hyderabad), and C.V. Jawahar (IIIT-Hyderabad)</i>	
Contrast To Divide: Self-Supervised Pre-Training for Learning With Noisy Labels	387
<i>Evgenii Zheltonozhskii (Technion), Chaim Baskin (Technion), Avi Mendelson (Technion), Alex M. Bronstein (Technion), and Or Litany (NVIDIA)</i>	
PICA: Point-Wise Instance and Centroid Alignment Based Few-Shot Domain Adaptive Object Detection With Loose Annotations	398
<i>Chaoliang Zhong (Fujitsu R&D Center, Co., LTD, Beijing, China), Jie Wang (Fujitsu R&D Center, Co., LTD, Beijing, China), Cheng Feng (Fujitsu R&D Center, Co., LTD, Beijing, China), Ying Zhang (Fujitsu R&D Center, Co., LTD), Jun Sun (Fujitsu R&D Center Co., Ltd.), and Yasuto Yokota (Fujitsu LTD.)</i>	
Calibrating CNNs for Few-Shot Meta Learning	408
<i>Peng Yang (Baidu), Shaogang Ren (Baidu Research, USA), Yang Zhao (University at Buffalo), and Ping Li (Baidu Research)</i>	
FLUID: Few-Shot Self-Supervised Image Deraining	418
<i>Shyam Nandan Rai (IIIT-Hyderabad), Rohit Saluja (IIIT-Hyderabad), Chetan Arora (Indian Institute of Technology Delhi), Vineeth N Balasubramanian (Indian Institute of Technology, Hyderabad), Anbumani Subramanian (IIIT-Hyderabad), and C.V. Jawahar (IIIT-Hyderabad)</i>	
MaskSplit: Self-Supervised Meta-Learning for Few-Shot Semantic Segmentation	428
<i>Mustafa Sercan Amac (Hacettepe University), Ahmet Sencan (Middle East Technical University), Bugra Baran (METU), Nazli Ikizler-Cinbis (Hacettepe University), and Ramazan Gokberk Cinbis (METU)</i>	

Image/Video/Multimedia

All the Attention You Need: Global-Local, Spatial-Channel Attention for Image Retrieval	439
<i>Chull Hwan Song (Oddconcepts), Hye Joo Han (Oddconcepts), and Yannis Avrithis (Athena RC)</i>	
Hierarchical Proxy-Based Loss for Deep Metric Learning	449
<i>Zhibo Yang (Stony Brook University), Muhammet Bastan (Amazon), Xinliang Zhu (Amazon), Douglas Gray (Amazon), and Dimitris Samaras (Stony Brook University)</i>	
A Fast Partial Video Copy Detection Using KNN and Global Feature Database	459
<i>Weijun Tan (Linksprite Technologies), Hongwei Guo (Deepcam), and Rushuai Liu (Shenzhen Deepcam)</i>	

Learning With Label Noise for Image Retrieval by Selecting Interactions	468
<i>Sarah Ibrahim (University of Amsterdam), Arnaud Sors (Naver Labs Europe), Rafael Sampaio de Rezende (Naver Labs Europe), and Stéphane Clinchant (Naver Labs Europe)</i>	
Video and Text Matching With Conditioned Embeddings	478
<i>Ameen Ali (Tel Aviv University, Israel), Idan Schwartz (Technion), Tamir Hazan (Technion), and Lior Wolf (Tel Aviv University, Israel)</i>	
Re-Compose the Image by Evaluating the Crop on More Than Just a Score	488
<i>Yang Cheng (Purdue University), Qian Lin (HP), and Jan P. Allebach (Purdue University)</i>	
Discovering Underground Maps From Fashion	497
<i>Utkarsh Mall (Cornell University), Kavita Bala (Cornell University), Tamara Berg (Facebook AI Research), and Kristen Grauman (Facebook AI Research & UT Austin)</i>	
Strumming to the Beat: Audio-Conditioned Contrastive Video Textures	507
<i>Medhini Narasimhan (UC Berkeley), Shiry Ginosar (UC Berkeley), Andrew Owens (U Michigan), Alexei A. Efros (UC Berkeley), and Trevor Darrell (UC Berkeley)</i>	
MovingFashion: A Benchmark for the Video-To-Shop Challenge	517
<i>Marco Godi (University of Verona), Christian Joppi (University of Verona), Geri Skenderi (University of Verona), and Marco Cristani (University of Verona)</i>	
Challenges in Procedural Multimodal Machine Comprehension: A Novel Way To Benchmark ..	526
<i>Pritish Sahu (Rutgers University), Karan Sikka (SRI International), and Ajay Divakaran (SRI, USA)</i>	
Video Salient Object Detection via Contrastive Features and Attention Modules	536
<i>Yi-Wen Chen (University of California, Merced), Xiaojie Jin (NUS), Xiaohui Shen (ByteDance), and Ming-Hsuan Yang (University of California at Merced)</i>	

Language, Text and Documents

Multimodal Learning Using Optimal Transport for Sarcasm and Humor Detection	546
<i>Shraman Pramanick (Johns Hopkins University), Aniket Roy (Johns Hopkins University), and Vishal M. Patel (Johns Hopkins University)</i>	
DeepPatent: Large Scale Patent Drawing Recognition and Retrieval	557
<i>Michal Kucer (Los Alamos National Laboratory), Diane Oyen (Los Alamos National Lab), Juan Castorena (Old Dominion University), and Jian Wu (LANL)</i>	
Typenet: Towards Camera Enabled Touch Typing on Flat Surfaces Through Self-Refinement ...	567
<i>Ben Maman (Tel Aviv University) and Amit Bermano (Tel-Aviv University)</i>	
Less Can Be More: Sound Source Localization With a Classification Model	577
<i>Arda Senocak (KAIST), Hyeonggon Ryu (KAIST), Junsik Kim (Harvard), and In So Kweon (KAIST)</i>	

SeeTek: Very Large-Scale Open-Set Logo Recognition With Text-Aware Metric Learning	587
<i>Chenge Li (Amazon.com Inc), István Fehérvári (Amazon), Xiaonan Zhao (Amazon), Ives Macedo (Amazon), and Srikar Appalaraju (Amazon)</i>	
SAC: Semantic Attention Composition for Text-Conditioned Image Retrieval	597
<i>Surgan Jandial (Media and Data Science Research Lab, Adobe), Pinkesh Badjatiya (Adobe Systems), Pranit Chawla (IIT Kharagpur), Ayush Chopra (MIT), Mausoom Sarkar (Adobe), and Balaji Krishnamurthy (Media and Data Science Research Lab, Adobe)</i>	
3DRefTransformer: Fine-Grained Object Identification in Real-World Scenes Using Natural Language	607
<i>Ahmed Abdelreheem (KAUST), Ujjwal Upadhyay (KAUST), Ivan Skorokhodov (KAUST), Rawan Al Yahya (KAUST), Jun Chen (King Abdullah University of Science and Technology), and Mohamed Elhoseiny (KAUST)</i>	

3D Computer Vision II

Monocular Depth Estimation With Adaptive Geometric Attention	617
<i>Taher Naderi (The university of Tennessee at Knoxville), Amir Sadovnik (The University of Tennessee), Jason Hayward (University of Tennessee-Knoxville), and Hairong Qi (University of Tennessee-Knoxville)</i>	
Shadow Art Revisited: A Differentiable Rendering Based Approach	628
<i>Kaustubh Sadekar (Indian Institute of Technology Gandhinagar), Ashish Tiwari (Indian Institute of Technology Gandhinagar), and Shanmuganathan Raman (Indian Institute of Technology (IIT) Gandhinagar)</i>	
Deep Two-Stream Video Inference for Human Body Pose and Shape Estimation	637
<i>Ziwen Li (OPPO Research Institute), Bo Xu (OPPO Research Institute), Han Huang (OPPO Research Institute), Cheng Lu (XPENG), and Yandong Guo (OPPO Research Institute)</i>	
RGL-NET: A Recurrent Graph Learning Framework for Progressive Part Assembly	647
<i>Abhinav Narayan (Indian Institute of Technology Gandhinagar (IITGN)), Rajendra Nagar (Indian Institute of Technology Jodhpur), and Shanmuganathan Raman (Indian Institute of Technology (IIT) Gandhinagar)</i>	
Spatial-Temporal Transformer for 3D Point Cloud Sequences	657
<i>Yimin Wei (Sun Yat-Sen University), Hao Liu (Sun Yat-Sen University), Tingting Xie (Queen Mary University of London), Qihong Ke (The University of Melbourne), and Yulan Guo (Sun Yat-sen University)</i>	
SBEVNet: End-to-End Deep Stereo Layout Estimation	667
<i>Divam Gupta (Carnegie Mellon University), Wei Pu (National Robotics Engineering Center), Trenton Tabor (Carnegie Mellon University), and Jeff Schneider (missing)</i>	
MobileStereoNet: Towards Lightweight Deep Networks for Stereo Matching	677
<i>Faranak Shamsafar (University of Tuebingen), Samuel Woerz (University of Tübingen), Rafia Rahim (University of Tuebingen), and Andreas Zell (University of Tuebingen)</i>	

Data Augmented 3D Semantic Scene Completion With 2D Segmentation Priors	687
<i>Aloisio Dourado (University of Brasilia), Frederico Guth (University of Brasilia), and Teofilo de Campos (Universidade de Brasilia)</i>	
3D Modeling Beneath Ground: Plant Root Detection and Reconstruction Based on Ground-Penetrating Radar	697
<i>Yawen Lu (Rochester Institute of Technology) and Guoyu Lu (Rochester Institute of Technology)</i>	
An Experimental Comparison of Multi-View Stereo Approaches on Satellite Images	707
<i>Alvaro Gómez (Universidad de la República - Facultad de Ingeniería), Gregory Randall (ENS Paris - Saclay), Gabriele Facciolo (Ecole Normale Supérieure Paris-Saclay), and Rafael Grompone von Gioi (Universidad de la República, Facultad de Ingeniería)</i>	
Creating and Reenacting Controllable 3D Humans With Differentiable Rendering	717
<i>Thiago L. Gomes (Universidade Federal de Ouro Preto), Thiago M. Coutinho (Universidade Federal de Minas Gerais), Rafael Azevedo (Universidade Federal de Minas Gerais), Renato Martins (Université Bourgogne Franche-Comté), and Erickson R. Nascimento (UFMG)</i>	
Transferable 3D Adversarial Textures Using End-to-End Optimization	727
<i>Camilo Pestana (The University of Western Australia), Naveed Akhtar (The University of Western Australia), Nazanin Rahnavard (University of Central Florida), Mubarak Shah (University of Central Florida), and Ajmal Mian (University of Western Australia)</i>	

Action and Behavior Recognition

Spatiotemporal Initialization for 3D CNNs With Generated Motion Patterns	737
<i>Hirokatsu Kataoka (National Institute of Advanced Industrial Science and Technology (AIST)), Kensho Hara (Tokyo Institute of Technology), Ryusuke Hayashi (National Institute of Advanced Industrial Science and Technology (AIST)), Eisuke Yamagata (National Institute of Advanced Industrial Science and Technology (AIST)), and Nakamasa Inoue (Tokyo Institute of Technology)</i>	
MUGL: Large Scale Multi Person Conditional Action Generation With Locomotion	747
<i>Shubh Maheshwari (IIIT Hyderabad), Debtanu Gupta (IIIT Hyderabad), and Ravi Kiran Sarvadevabhatla (IIIT Hyderabad)</i>	
Busy-Quiet Video Disentangling for Video Classification	756
<i>Guoxi Huang (The University of York) and Adrian G. Bors (University of York)</i>	
Contextual Proposal Network for Action Localization	766
<i>He-Yen Hsieh (Academia Sinica), Ding-Jie Chen (Academia Sinica), and Tyng-Luh Liu (Academia Sinica)</i>	
Multi-Level Attentive Adversarial Learning With Temporal Dilation for Unsupervised Video Domain Adaptation	776
<i>Peipeng Chen (Sun Yat-sen University), Yuan Gao (School of Data and Computer Science, Sun Yat-sen University), and Andy J. Ma (Sun Yat-sen University)</i>	
MM-ViT: Multi-Modal Video Transformer for Compressed Video Action Recognition	786
<i>Jiawei Chen (InnoPeak Technology) and Chiu Man Ho (OPPO US R&D)</i>	

PERF-Net: Pose Empowered RGB-Flow Net	798
<i>Yinxiao Li (Google), Zhichao Lu (Google), Xuehan Xiong (Google), and Jonathan Huang (Google)</i>	
Action Anticipation Using Latent Goal Learning	808
<i>Debaditya Roy (Agency for Science, Technology and Research) and Basura Fernando (Agency for Science, Technology and Research, A*STAR, Singapore)</i>	
Leaky Gated Cross-Attention for Weakly Supervised Multi-Modal Temporal Action Localization.....	817
<i>Jun-Tae Lee (Qualcomm AI Research), Sungrack Yun (Qualcomm AI Research), and Mihir Jain (Qualcomm AI Research)</i>	
NUTA: Non-Uniform Temporal Aggregation for Action Recognition	827
<i>Xinyu Li (Amazon), Chunhui Liu (Amazon), Bing Shuai (Amazon), Yi Zhu (Amazon), Hao Chen (Amazon), and Joseph Tighe (Amazon)</i>	
Skeleton-DML: Deep Metric Learning for Skeleton-Based One-Shot Action Recognition	837
<i>Raphael Memmesheimer (University of Koblenz-Landau), Simon Häring (University of Koblenz-Landau), Nick Theisen (University of Koblenz-Landau), and Dietrich Paulus (missing)</i>	
Self-Supervised Video Representation Learning With Cross-Stream Prototypical Contrasting ..	846
<i>Martine Toering (University of Amsterdam), Ioannis Gatopoulos (University of Amsterdam), Maarten Stol (BrainCreators), and Vincent Tao Hu (University of Amsterdam)</i>	

Computational Photography and Image Processing II

Novel-View Synthesis of Human Tourist Photos	857
<i>Jonathan Freer (University of Birmingham), Kwang Moo Yi (University of British Columbia), Wei Jiang (University of British Columbia), Jongwon Choi (Chung-Ang University), and Hyung Jin Chang (University of Birmingham)</i>	
Deep Online Fused Video Stabilization	865
<i>Zhenmei Shi (University of Wisconsin Madison), Fuhao Shi (Google), Wei-Sheng Lai (Google), Chia-Kai Liang (Google), and Yingyu Liang (University of Wisconsin Madison)</i>	
Normalizing Flow as a Flexible Fidelity Objective for Photo-Realistic Super-Resolution	874
<i>Andreas Lugmayr (ETH Zurich), Martin Danelljan (ETH Zurich), Fisher Yu (ETH Zurich), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)</i>	
Deep Feature Prior Guided Face Deblurring	884
<i>Soo Hyun Jung (Ajou University), Tae Bok Lee (Ajou University), and Yong Seok Heo (Ajou University)</i>	
Deep Photo Scan: Semi-Supervised Learning for Dealing With the Real-World Degradation in Smartphone Photo Scanning	894
<i>Man M. Ho (Hosei University) and Jinjia Zhou (Hosei University)</i>	
Learning Color Representations for Low-Light Image Enhancement	904
<i>Bomi Kim (KAIST), Sunhyeok Lee (KAIST), Nahyun Kim (KAIST), Donggon Jang (KAIST), and Dae-Shik Kim (KAIST)</i>	

DAQ: Channel-Wise Distribution-Aware Quantization for Deep Image Super-Resolution Networks. 913
Cheemun Hong (Seoul National University), Heewon Kim (Seoul National University), Sungyong Baik (Seoul National University), Junghun Oh (Seoul National University), and Kyoung Mu Lee (Seoul National University)

Supervised Compression for Resource-Constrained Edge Computing Systems 923
Yoshitomo Matsubara (University of California, Irvine), Ruihan Yang (University of California Irvine), Marco Levorato (University of California, Irvine), and Stephan Mandt (University of California, Irvine)

Auto White-Balance Correction for Mixed-Illuminant Scenes 934
Mahmoud Afifi (Apple), Marcus A. Brubaker (York University), and Michael S. Brown (York University)

Adversarial Open Domain Adaptation for Sketch-to-Photo Synthesis 944
Xiaoyu Xiang (Purdue University), Ding Liu (Bytedance), Xiao Yang (Bytedance AI Lab), Yiheng Zhu (ByteDance AI Lab), Xiaohui Shen (ByteDance AI Lab), and Jan P. Allebach (Purdue University)

AE-StyleGAN: Improved Training of Style-Based Auto-Encoders 955
Ligong Han (Rutgers University), Sri Harsha Musunuri (Rutgers University, New Brunswick), Martin Renqiang Min (NEC Labs America-Princeton), Ruijiang Gao (University of Texas at Austin), Yu Tian (Rutgers), and Dimitris Metaxas (Rutgers)

Late-Resizing: A Simple but Effective Sketch Extraction Strategy for Improving Generalization of Line-Art Colorization 965
Dohyun Kim (NAVER WEBTOON Corp.), Dajung Je (Naver Webtoon Corporation), Kwangjin Lee (NAVER WEBTOON Corp.), Moohyun Kim (NAVER WEBTOON Corp.), and Han Kim (NAVER WEBTOON Corp.)

Transfer Learning

Hierarchically Decoupled Spatial-Temporal Contrast for Self-Supervised Video Representation Learning 975
Zehua Zhang (Indiana University) and David Crandall (Indiana University)

Boosting Contrastive Self-Supervised Learning With False Negative Cancellation 986
Tri Huynh (Google), Simon Kornblith (Google Brain), Matthew R. Walter (Toyota Technological Institute at Chicago), Michael Maire (University of Chicago), and Maryam Khademi (Google)

Self-Supervised Learning of Domain Invariant Features for Depth Estimation 997
Hiroyasu Akada (KAUST, Keio University), Shariq Farooq Bhat (KAUST), Ibraheem Alhashim (National Center for Artificial Intelligence (NCAI), Saudi Data and Artificial Intelligence Authority (SDAIA), Riyadh, Kingdom of Saudi Arabia), and Peter Wonka (KAUST)

Adversarial Branch Architecture Search for Unsupervised Domain Adaptation	1008
<i>Luca Robbiano (Politecnico di Torino), Muhammad Rameez Ur Rahman (Sapienza University of Rome), Fabio Galasso (Sapienza University), Barbara Caputo (Politecnico di Torino), and Fabio Maria Carlucci (Huawei Noah's Ark Lab)</i>	
Forgery Detection by Internal Positional Learning of Demosaicing Traces	1019
<i>Quentin Bammey (École Normale Supérieure Paris-Saclay), Rafael Grompone von Gioi (Ecole Normale Supérieure Paris-Saclay), and Jean-Michel Morel (CMLA, ENS Cachan (Paris-Saclay))</i>	
Distance-Based Hyperspherical Classification for Multi-Source Open-Set Domain Adaptation	1030
<i>Silvia Bucci (Italian Institute of Technology), Francesco Cappio Borlino (Politecnico di Torino), Barbara Caputo (Politecnico di Torino), and Tatiana Tommasi (Politecnico di Torino)</i>	
Few-Shot Weakly-Supervised Object Detection via Directional Statistics	1040
<i>Amirreza Shaban (University of Washington), Amir Rahimi (Australian National University), Thalaiyasingam Ajanthan (ANU), Byron Boots (University of Washington), and Richard Hartley (Australian National University, Australia)</i>	
Self-Supervised Pretraining Improves Self-Supervised Pretraining	1050
<i>Colorado J Reed (University of California, Berkeley), Xiangyu Yue (University of California, Berkeley), Ani Nrusimha (UC Berkeley), Sayna Ebrahimi (UC Berkeley), Vivek Vijaykumar (Georgia Tech), Richard Mao (UC Berkeley), Bo Li (Microsoft Research Asia), Shanghang Zhang (UC Berkeley), Devin Guillory (UC Berkeley), Sean Metzger (UC Berkeley), Kurt Keutzer (EECS, UC Berkeley), and Trevor Darrell (UC Berkeley)</i>	
SC-UDA: Style and Content Gaps Aware Unsupervised Domain Adaptation for Object Detection ...	1061
<i>Fuxun Yu (George Mason University), Di Wang (Microsoft), Yinpeng Chen (Microsoft), Nikolaos Karianakis (Microsoft), Tong Shen (Microsoft), Pei Yu (Microsoft), Dimitrios Lymberopoulos (Microsoft), Sidi Lu (Wayne State University), Weisong Shi (Wayne State University), and Xiang Chen (George Mason University)</i>	
Coupled Training for Multi-Source Domain Adaptation	1071
<i>Ohad Amosy (Bar Ilan University) and Gal Chechik (Bar Ilan University)</i>	
Federated Multi-Target Domain Adaptation	1081
<i>Chun-Han Yao (University of California at Merced), Boqing Gong (Google), Hang Qi (Google), Yin Cui (Google), Yukun Zhu (Google Inc.), and Ming-Hsuan Yang (University of California at Merced)</i>	
Unsupervised Learning for Human Sensing Using Radio Signals	1091
<i>Tianhong Li (MIT), Lijie Fan (MIT), Yuan Yuan (MIT), and Dina Katabi (Massachusetts Institute of Technology)</i>	
Few-Shot Object Detection by Attending to Per-Sample-Prototype	1101
<i>Hojun Lee (Seoul National University), Myunggi Lee (Seoul National University), and Nojun Kwak (Seoul National University)</i>	

Estimating Image Depth in the Comics Domain	1111
<i>Deblina Bhattacharjee (EPFL), Martin Everaert (EPFL), Mathieu Salzmann (EPFL), and Sabine Süsstrunk (EPFL)</i>	

Biometrics

A Deep Insight Into Measuring Face Image Utility With General and Face-Specific Image Quality Metrics	1121
<i>Biyang Fu (Fraunhofer IGD), Cong Chen (Fraunhofer IGD), Olaf Henniger (Fraunhofer IGD), and Naser Damer (Fraunhofer IGD)</i>	
Learnable Multi-Level Frequency Decomposition and Hierarchical Attention Mechanism for Generalized Face Presentation Attack Detection	1131
<i>Meiling Fang (Fraunhofer Institute for Computer Graphics Research IGD), Naser Damer (Fraunhofer IGD), Florian Kirchbuchner (Fraunhofer Institute for Computer Graphics Research IGD), and Arjan Kuijper (Fraunhofer Institute for Computer Graphics Research IGD and Mathematical and Applied Visual Computing group, TU Darmstadt)</i>	
Attribute-Based Deep Periocular Recognition: Leveraging Soft Biometrics to Improve Periocular Recognition	1141
<i>Veeru Talreja (West Virginia University), Nasser M. Nasrabadi (West Virginia University), and Matthew C. Valenti (West Virginia University)</i>	
Face Verification With Challenging Imposters and Diversified Demographics	1151
<i>Adrian Popescu (CEA LIST), Liviu-Daniel Ștefan (University Politehnica of Bucharest), Jérôme Deshayes-Chossart (CEA), and Bogdan Ionescu (University Politehnica of Bucharest, Romania)</i>	
MTGLS: Multi-Task Gaze Estimation With Limited Supervision	1161
<i>Shreya Ghosh (Monash University), Munawar Hayat (Monash University), Abhinav Dhall (Monash University), and Jarrod Knibbe (University of Melbourne)</i>	
Complete Face Recovery GAN: Unsupervised Joint Face Rotation and De-Occlusion From a Single-View Image	1173
<i>Yeong-Joon Ju (Korea University), Gun-Hee Lee (Korea University), Jung-Ho Hong (Korea University), and Seong-Whan Lee (Korea University)</i>	
Matching and Recovering 3D People From Multiple Views	1184
<i>Alejandro Perez-Yus (Institut de Robòtica i Informàtica Industrial, CSIC-UPC) and Antonio Agudo (Institut de Robotica i Informatica Industrial, CSIC-UPC)</i>	
On Black-Box Explanation for Face Verification	1194
<i>Domingo Mery (Universidad Catolica de Chile) and Bernardita Morris (Universidad Catolica de Chile)</i>	
LwPosr: Lightweight Efficient Fine Grained Head Pose Estimation	1204
<i>Naina Dhingra (ETH Zurich)</i>	

Mobile Based Human Identification Using Forehead Creases: Application and Assessment Under COVID-19 Masked Face Scenarios	1215
<i>Rohit Bharadwaj (Birla Institute of Technology & Science Pilani), Gaurav Jaswal (Indian Institute of Technology Delhi), Aditya Nigam (IIT mandi), and Kamlesh Tiwari (BITS Pilani)</i>	
3DFaceFill: An Analysis-by-Synthesis Approach To Face Completion	1224
<i>Rahul Dey (Michigan State University) and Vishnu Naresh Boddeti (Michigan State University)</i>	
Disentangled Representation With Dual-Stage Feature Learning for Face Anti-Spoofing	1234
<i>Yu-Chun Wang (National Tsing Hua University), Chien-Yi Wang (Microsoft), and Shang-Hong Lai (Microsoft)</i>	
Joint Classification and Trajectory Regression of Online Handwriting Using a Multi-Task Learning Approach	1244
<i>Felix Ott (Fraunhofer IIS), David Rügamer (LMU Munich), Lucas Heublein (Fraunhofer IIS), Bernd Bischl (LMU Munich), and Christopher Mutschler (Fraunhofer IIS)</i>	
Human-Aided Saliency Maps Improve Generalization of Deep Learning	1255
<i>Aidan Boyd (University of Notre Dame), Kevin W. Bowyer (University of Notre Dame), and Adam Czajka (University of Notre Dame)</i>	

Object Detection and Recognition

ImVoxelNet: Image to Voxels Projection for Monocular and Multi-View General-Purpose 3D Object Detection	1265
<i>Danila Rukhovich (Samsung AI Center Moscow), Anna Vorontsova (Samsung AI Center), and Anton Konushin (Lomonosov Moscow State University)</i>	
Fusion Point Pruning for Optimized 2D Object Detection With Radar-Camera Fusion	1275
<i>Lukas Stäcker (Opel Automobile GmbH), Philipp Heidenreich (Opel Automobile GmbH), Jason Rambach (DFKI), and Didier Stricker (DFKI)</i>	
Detecting Tear Gas Canisters With Limited Training Data	1283
<i>Ashwin D'Cruz (Forensic Architecture), Christopher Tegho (Forensic Architecture), Sean Greaves (Forensic Architecture), and Lachlan Kermode (Forensic Architecture)</i>	
Online Continual Learning via Candidates Voting	1292
<i>Jiangpeng He (Purdue University) and Fengqing Zhu (Purdue University, USA)</i>	
A Semi-Supervised Generalized VAE Framework for Abnormality Detection Using One-Class Classification	1302
<i>Renuka Sharma (IITB), Satvik Mashkaria (IITB), and Suyash P. Awate (Indian Institute of Technology (IIT) Bombay)</i>	
YOLO-ReT: Towards High Accuracy Real-Time Object Detection on Edge GPUs	1311
<i>Prakhar Ganesh (National University of Singapore), Yao Chen (Advanced Digital Sciences Center), Yin Yang (Hamad bin Khalifa University), Deming Chen (University of Illinois at Urbana-Champaign), and Marianne Winslett (CS Illinois)</i>	

Improving Object Detection by Label Assignment Distillation	1322
<i>Chuong H. Nguyen (CyberCore AI), Thuy C. Nguyen (Cybercore AI), Tuan N. Tang (CyberCore AI), and Nam L.H. Phan (CyberCore AI)</i>	
Robustly Recognizing Irregular Scene Text by Rectifying Principle Irregularities	1332
<i>Changsheng Xu (Fudan University), Yang Wang (Tongji University), Fan Bai (Fudan University), Jihong Guan (Tongji University), and Shuigeng Zhou (Fudan University)</i>	
Multi-Head Deep Metric Learning Using Global and Local Representations	1340
<i>Mohammad K. Ebrahimpour (OV Labs), Gang Qian (ObjectVideo), and Allison Beach (ObjectVideo)</i>	
Extractive Knowledge Distillation	1350
<i>Takumi Kobayashi (National Institute of Advanced Industrial Science and Technology)</i>	
Recursive Contour-Saliency Blending Network for Accurate Salient Object Detection	1360
<i>Yun Yi Ke (Nanyang Technological University) and Takahiro Tsubono (Open8 Singapore Pte. Ltd)</i>	
Adversarial Robustness of Deep Sensor Fusion Models	1371
<i>Shaojie Wang (Washington University in St. Louis), Tong Wu (Washington University in St. Louis), Ayan Chakrabarti (Washington University in St. Louis), and Yevgeniy Vorobeychik (Washington University in St. Louis)</i>	
WEPDToF: A Dataset and Benchmark Algorithms for In-the-Wild People Detection and Tracking From Overhead Fisheye Cameras	1381
<i>Ozan Tezcan (Boston University), Zhihao Duan (Purdue University), Mertcan Cokbas (Boston University), Prakash Ishwar (Boston University), and Janusz Konrad (Boston University, Massachusetts)</i>	
A Context-Enriched Satellite Imagery Dataset and an Approach for Parking Lot Detection	1391
<i>Yifang Yin (A*STAR), Wenmiao Hu (National University of Singapore), An Tran (Grabtaxi Holdings), Hannes Kruppa (Grab Holdings Inc.), Roger Zimmermann (NUS), and See-Kiong Ng (National University of Singapore)</i>	
Densely-Packed Object Detection via Hard Negative-Aware Anchor Attention	1401
<i>Sungmin Cho (Chung-Ang Univ., Korea), Jinwook Paeng (Chung-Ang Univ., Korea), and Junseok Kwon (Chung-Ang Univ., Korea)</i>	

Remote Sensing and Vision Systems

Robust 3D Garment Digitization From Monocular 2D Images for 3D Virtual Try-On Systems ..	1411
<i>Sahib Majithia (Myntra Design Pvt. Ltd.), Sandeep N. Parameswaran (Myntra Designs Pvt. Ltd.), Sadbhavana Babar (Myntra Designs Pvt. Ltd.), Vikram Garg (Myntra Designs Pvt. Ltd.), Astitva Srivastava (IIITH), and Avinash Sharma (CVIT, IIIT-Hyderabad)</i>	
PPCD-GAN: Progressive Pruning and Class-Aware Distillation for Large-Scale Conditional GANs Compression	1422
<i>Duc Minh Vo (The University of Tokyo), Akihiro Sugimoto (NII), and Hideki Nakayama (The University of Tokyo)</i>	

Billion-Scale Pretraining With Vision Transformers for Multi-Task Visual Representations	1431
<i>Josh Beal (Pinterest), Hao-Yu Wu (Pinterest.com), Dong Huk Park (UC Berkeley), Andrew Zhai (Pinterest, Inc.), and Dmitry Kislyuk (Pinterest)</i>	
Multi-Task Classification of Sewer Pipe Defects and Properties Using a Cross-Task Graph Neural Network Decoder	1441
<i>Joakim Bruslund Haurum (Aalborg University), Meysam Madadi (CVC), Sergio Escalera (Computer Vision Center (UAB) & University of Barcelona), and Thomas B. Moeslund (Aalborg University)</i>	
Pixel-Level Bijective Matching for Video Object Segmentation	1453
<i>Suhwan Cho (Yonsei University), Heansung Lee (Yonsei University), Minjung Kim (Yonsei University), Sungjun Jang (Yonsei University), and Sangyoun Lee (Yonsei University)</i>	
Extraction of Positional Player Data From Broadcast Soccer Videos	1463
<i>Jonas Theiner (L3S Research Center, Leibniz Universität Hannover), Wolfgang Gritz (Leibniz Information Centre for Science and Technology (TIB)), Eric Müller-Budack (TIB - Leibniz Information Centre for Science and Technology), Robert Rein (German Sports University Cologne), Daniel Memmert (German Sports University Cologne), and Ralph Ewerth (TIB - Leibniz Information Center for Science and Technology)</i>	
Interpretable Semantic Photo Geolocation	1474
<i>Jonas Theiner (L3S Research Center, Leibniz Universität Hannover), Eric Müller-Budack (TIB - Leibniz Information Centre for Science and Technology), and Ralph Ewerth (TIB - Leibniz Information Center for Science and Technology)</i>	
Active Learning for Improved Semi-Supervised Semantic Segmentation in Satellite Images ...	1485
<i>Shasvat Desai (Orbital Insight) and Debasmita Ghose (Yale University)</i>	
Lane-Level Street Map Extraction From Aerial Imagery	1496
<i>Songtao He (Massachusetts Institute of Technology) and Hari Balakrishnan (Massachusetts Institute of Technology)</i>	

Explainable AI

Discrete Neural Representations for Explainable Anomaly Detection	1506
<i>Stanislaw Szymanowicz (University of Cambridge), James Charles (Cambridge University), and Roberto Cipolla (University of Cambridge)</i>	
Auditing Saliency Cropping Algorithms	1515
<i>Abeba Birhane (University College Dublin), Vinay Uday Prabhu (UnifyID AI Labs), and John Whaley (UnifyID Inc)</i>	
Agree To Disagree: When Deep Learning Models With Identical Architectures Produce Distinct Explanations	1524
<i>Matthew Watson (Durham University), Bashar Awwad Shiekh Hasan (Amazon), and Noura Al Moubayed (University of Durham)</i>	
Visualizing Paired Image Similarity in Transformer Networks	1534
<i>Samuel Black (Temple University), Abby Stylianou (Saint Louis University), Robert Pless (George Washington University), and Richard Souvenir (Temple University)</i>	

X-MIR: EXplainable Medical Image Retrieval	1544
<i>Brian Hu (Kitware Inc.), Bhavan Vasu (Kitware), and Anthony Hoogs (Kitware)</i>	
Uncertainty Learning Towards Unsupervised Deformable Medical Image Registration	1555
<i>Xuan Gong (University at Buffalo), Luckyson Khaidem (University at Buffalo), Wentao Zhu (Kuaishou), Baochang Zhang (Beihang University), and David Doermann (University at Buffalo)</i>	
How Good Is Your Explanation? Algorithmic Stability Measures To Assess the Quality of Explanations for Deep Neural Networks	1565
<i>Thomas Fel (IRT Saint Exupery), David Vigouroux (IRT Saint Exupery), Rémi Cadène (LIP6), and Thomas Serre (Brown University)</i>	
SWAG-V: Explanations for Video Using Superpixels Weighted by Average Gradients	1576
<i>Thomas Hartley (Cardiff University), Kirill Sidorov (Cardiff University), Christopher Willis (BAE Systems AI Labs), and David Marshall (Cardiff University)</i>	

Transfer, Few-shot, Semi-, and Un- Supervised Learning II

SEGA: Semantic Guided Attention on Visual Prototype for Few-Shot Learning	1586
<i>Fengyuan Yang (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)</i>	
Enhancing Few-Shot Image Classification With Unlabelled Examples	1597
<i>Peyman Bateni (University of British Columbia), Jarred Barber (Charles River Analytics), Jan-Willem van de Meent (University of Amsterdam), and Frank Wood (University of British Columbia)</i>	
A Pixel-Level Meta-Learner for Weakly Supervised Few-Shot Semantic Segmentation	1607
<i>Yuan-Hao Lee (National Taiwan University), Fu-En Yang (National Taiwan University), and Yu-Chiang Frank Wang (National Taiwan University)</i>	
COCOA: Context-Conditional Adaptation for Recognizing Unseen Classes in Unseen Domains	1618
<i>Puneet Mangla (IIT Hyderabad), Shivam Chandhok (Indian Institute of Technology, Hyderabad), Vineeth N Balasubramanian (Indian Institute of Technology, Hyderabad), and Fahad Shahbaz Khan (MBZUAI)</i>	
Meta-Meta Classification for One-Shot Learning	1628
<i>Arkabandhu Chowdhury (Amazon.com), Dipak Chaudhari (The University of Texas at Austin), Swarat Chaudhuri (The University of Texas at Austin), and Chris Jermaine (Rice University)</i>	
Transfer Learning for Pose Estimation of Illustrated Characters	1638
<i>Shuhong Chen (University of Maryland - College Park) and Matthias Zwicker (University of Maryland)</i>	
From Node To Graph: Joint Reasoning on Visual-Semantic Relational Graph for Zero-Shot Detection	1648
<i>Hui Nie (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)</i>	

Mutual Learning of Joint and Separate Domain Alignments for Multi-Source Domain Adaptation...	1658
<i>Yuanyuan Xu (Institute of Computing Technology, Chinese Academy of Sciences), Meina Kan (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)</i>	
Generalized Clustering and Multi-Manifold Learning With Geometric Structure Preservation	1668
<i>Lirong Wu (Westlake University), Zicheng Liu (Westlake University), Jun Xia (Westlake University and Zhejiang University), Zelin Zang (Westlake University), Siyuan Li (Westlake University), and Stan Z. Li (Westlake University)</i>	
Trading-Off Information Modalities in Zero-Shot Classification	1677
<i>Jorge Sánchez (CONICET) and Matías Molina (National University of Córdoba)</i>	
Semi-Supervised Domain Adaptation via Sample-to-Sample Self-Distillation	1686
<i>Jeongbeen Yoon (POSTECH), Dahyun Kang (POSTECH), and Minsu Cho (POSTECH)</i>	

Medical Imaging / Imaging for Bioinformatics / Biological and Cell Microscopy

Towards Durability Estimation of Bioprosthetic Heart Valves via Motion Symmetry Analysis ..	1696
<i>Maryam Alizadeh (University of Victoria), Melissa Cote (University of Victoria), and Alexandra Branzan Albu (University of Victoria)</i>	
Co-Net: A Collaborative Region-Contour-Driven Network for Fine-to-Finer Medical Image Segmentation	1706
<i>Anran Liu (Xi'an Jiaotong University), Xiangsheng Huang (Institute of Automation Chinese Academy of Sciences), Tong Li (Xi'an Jiaotong University), and Pengcheng Ma (Institute of Automation Chinese Academy of Sciences)</i>	
Biomass Prediction With 3D Point Clouds From LiDAR	1716
<i>Liyuan Pan (The Australian National University), Liu Liu (ANU), Anthony G. Condon (CSIRO), Gonzalo M. Estavillo (CSIRO), Robert A. Coe (The Commonwealth Scientific and Industrial Research Organisation (CSIRO)), Geoff Bull (High Resolution Plant Phenomics Centre - CSIRO Agriculture and Food), Eric A. Stone (The Australian National University), Lars Petersson (Data61/CSIRO), and Vivien Rolland (CSIRO)</i>	
Dynamic CNNs Using Uncertainty To Overcome Domain Generalization for Surgical Instrument Localization	1727
<i>Markus Philipp (Karlsruhe Institute of Technology / Carl Zeiss Meditec AG), Anna Alperovich (Carl Zeiss AG), Marielena Gutt-Will (University Hospital of Bern), Andrea Mathis (University Hospital of Bern), Stefan Saur (Carl Zeiss Meditec AG), Andreas Raabe (University Hospital of Bern), and Franziska Mathis-Ullrich (Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology)</i>	

Self-Supervised Generative Style Transfer for One-Shot Medical Image Segmentation	1737
<i>Devavrat Tomar (Swiss Federal Institute of Technology Lausanne), Behzad Bozorgtabar (EPFL), Manana Lortkipanidze (EPFL), Guillaume Vray (EPFL), Mohammad Saeed Rad (École Polytechnique Fédérale de Lausanne), and Jean-Philippe Thiran (École Polytechnique Fédérale de Lausanne)</i>	
UNETR: Transformers for 3D Medical Image Segmentation	1748
<i>Ali Hatamizadeh (NVIDIA Corporation), Yucheng Tang (Vanderbilt University), Vishwesh Nath (NVIDIA), Dong Yang (NVIDIA Corporation), Andriy Myronenko (NVIDIA), Bennett Landman (Vanderbilt University), Holger R. Roth (NVIDIA), and Daguang Xu (NVIDIA Corporation)</i>	
Consistent Cell Tracking in Multi-Frames With Spatio-Temporal Context by Object-Level Warping Loss	1759
<i>Junya Hayashida (Kyushu University), Kazuya Nishimura (Kyushu University), and Ryoma Bise (Kyushu University)</i>	
Semi-Supervised Semantic Segmentation of Vessel Images Using Leaking Perturbations	1769
<i>Jinyong Hou (University of Otago), Xuejie Ding (University of Otago), and Jeremiah D. Deng (University of Otago, New Zealand)</i>	
Compressed Sensing MRI Reconstruction With Co-VeGAN: Complex-Valued Generative Adversarial Network	1779
<i>Bhavya Vasudeva (ISI Kolkata), Puneesh Deora (ISI Kolkata), Saumik Bhattacharya (Indian Institute of Technology Kharagpur), and Pyari Mohan Pradhan (IIT Roorkee)</i>	
Knowledge-Augmented Contrastive Learning for Abnormality Classification and Localization in Chest X-Rays With Radiomics Using a Feedback Loop	1789
<i>Yan Han (UT Austin), Chongyan Chen (UT Austin), Ahmed Tewfik (Electrical and Computer Engineering at The University of Texas at Austin), Benjamin Glicksberg (Icahn School of Medicine at Mount Sinai), Ying Ding (University of Texas at Austin), Yifan Peng (Weill Cornell Medicine), and Zhangyang Wang (University of Texas at Austin)</i>	
T-Net: A Resource-Constrained Tiny Convolutional Neural Network for Medical Image Segmentation	1799
<i>Tariq M. Khan (Deakin University), Antonio Robles-Kelly (Deakin University), and Syed S. Naqvi (COMSATS University Islamabad)</i>	

Vision for Robotics and Other Applications

Automated Defect Inspection in Reverse Engineering of Integrated Circuits	1809
<i>Ann-Christin Bette (TU Munich), Patrick Brus (Infineon), Gabor Balazs (Technical University of Munich), Matthias Ludwig (TU Munich), and Alois Knoll (Robotics and Embedded Systems)</i>	
CFLOW-AD: Real-Time Unsupervised Anomaly Detection With Localization via Conditional Normalizing Flows	1819
<i>Denis Gudovskiy (Panasonic), Shun Ishizaka (Panasonic Corporation), and Kazuki Kozuka (Panasonic Corporation)</i>	
Fully Convolutional Cross-Scale-Flows for Image-Based Defect Detection	1829
<i>Marco Rudolph (Leibniz University Hannover), Tom Wehrbein (Leibniz University Hannover), Bodo Rosenhahn (Leibniz University Hannover), and Bastian Wandt (University of British Columbia)</i>	

Network Generalization Prediction for Safety Critical Tasks in Novel Operating Domains	1839
<i>Molly O'Brien (Johns Hopkins University), Mike Medoff (exida), Julia Bukowski (Villanova University), and Gregory D. Hager (The Johns Hopkins University)</i>	
CoordiNet: Uncertainty-Aware Pose Regressor for Reliable Vehicle Localization	1848
<i>Arthur Moreau (Mines ParisTech - PSL University), Nathan Piasco (Huawei France), Dzmitry Tsishkou (Huawei France), Bogdan Stanculescu (Mines ParisTech), and Arnaud de La Fortelle (Mines ParisTech)</i>	
SIGNAV: Semantically-Informed GPS-Denied Navigation and Mapping in Visually-Degraded Environments	1858
<i>Alex Krasner (SRI International), Mikhail Sizintsev (SRI International), Abhinav Rajvanshi (SRI International), Han-Pang Chiu (SRI International), Niluthpol Mithun (SRI International), Kevin Kaighn (SRI International), Philip Miller (SRI International), Ryan Villamil (SRI International), and Supun Samarasekera (SRI International)</i>	
Self-Supervised Domain Adaptation for Visual Navigation With Global Map Consistency	1868
<i>Eun Sun Lee (Seoul National University), Junho Kim (Seoul National University), and Young Min Kim (Seoul National University)</i>	
FT-DeepNets: Fault-Tolerant Convolutional Neural Networks With Kernel-Based Duplication	1878
<i>Iljoo Baek (Carnegie Mellon University), Wei Chen (Carnegie Mellon University), Zhihao Zhu (Carnegie Mellon University), Soheil Samii (GM Motors R&D), and Raj Rajkumar (Carnegie Mellon University)</i>	
Learning Maritime Obstacle Detection From Weak Annotations by Scaffolding	1888
<i>Lojze Žust (University of Ljubljana) and Matej Kristan (University of Ljubljana)</i>	
Siamese Transformer Pyramid Networks for Real-Time UAV Tracking	1898
<i>Daitao Xing (New York University), Nikolaos Evangeliou (New York University Abu Dhabi), Athanasios Tsoukalas (New York University Abu Dhabi), and Anthony Tzes (New York University Abu Dhabi)</i>	
FastAno: Fast Anomaly Detection via Spatio-Temporal Patch Transformation	1908
<i>Chaewon Park (Yonsei University), MyeongAh Cho (Yonsei University), Minhyeok Lee (Yonsei University), and Sangyoun Lee (Yonsei University)</i>	

Segmentation

Temporally Stable Video Segmentation Without Video Annotations	1919
<i>Aharon Azulay (Lightricks, HUJI), Tavi Halperin (Lightricks), Orestis Vantzos (Lightricks), Nadav Bornstein (Lightricks), and Ofir Bibi (Lightricks)</i>	
D2Conv3D: Dynamic Dilated Convolutions for Object Segmentation in Videos	1929
<i>Christian Schmidt (RWTH Aachen), Ali Athar (RWTH Aachen), Sabarinath Mahadevan (RWTH Aachen University), and Bastian Leibe (RWTH Aachen University)</i>	
Single-Shot Path Integrated Panoptic Segmentation	1939
<i>Sukjun Hwang (Yonsei University), Seoung Wug Oh (Adobe Research), and Seon Joo Kim (Yonsei University)</i>	

Robust Lane Detection via Expanded Self Attention	1949
<i>Minhyeok Lee (Yonsei University), Junhyeop Lee (Yonsei University), Dogyoon Lee (Yonsei University), Woojin Kim (Yonsei University), Sangwon Hwang (Yonsei University), and Sangyoun Lee (Yonsei University)</i>	
Pixel-by-Pixel Cross-Domain Alignment for Few-Shot Semantic Segmentation	1959
<i>Antonio Tavera (Politecnico di Torino), Fabio Cermelli (Politecnico di Torino), Carlo Masone (Istituto Italiano di Tecnologia), and Barbara Caputo (Politecnico di Torino)</i>	
VCSeg: Virtual Camera Adaptation for Road Segmentation	1969
<i>Gong Cheng (York University) and James H. Elder (York University)</i>	
Maximizing Cosine Similarity Between Spatial Features for Unsupervised Domain Adaptation in Semantic Segmentation	1979
<i>Inseop Chung (Seoul National University), Daesik Kim (Naver webtoon), and Nojun Kwak (Seoul National University)</i>	
Hyper-Convolution Networks for Biomedical Image Segmentation	1989
<i>Tianyu Ma (Cornell University), Adrian V. Dalca (MIT), and Mert R. Sabuncu (Cornell)</i>	
Plugging Self-Supervised Monocular Depth Into Unsupervised Domain Adaptation for Semantic Segmentation	1999
<i>Adriano Cardace (University of Bologna), Luca De Luigi (University of Bologna), Pierluigi Zama Ramirez (University of Bologna), Samuele Salti (University of Bologna), and Luigi Di Stefano (University of Bologna)</i>	
Shallow Features Guide Unsupervised Domain Adaptation for Semantic Segmentation at Class Boundaries	2010
<i>Adriano Cardace (University of Bologna), Pierluigi Zama Ramirez (University of Bologna), Samuele Salti (University of Bologna), and Luigi Di Stefano (University of Bologna)</i>	

Transfer, Few-shot, Semi-, and Un- Supervised Learning III

Bayesian Uncertainty and Expected Gradient Length – Regression: Two Sides of the Same Coin?	2021
<i>Megh Shukla (Mercedes-Benz R&D India)</i>	
The Hitchhiker’s Guide to Prior-Shift Adaptation	2031
<i>Tomáš Šipka (CTU FEE), Milan Šulc (Czech Technical University in Prague), and Jiří Matas (CMP CTU FEE)</i>	
Ortho-Shot: Low Displacement Rank Regularization With Data Augmentation for Few-Shot Learning	2040
<i>Uche Osahor (West Virginia University) and Nasser M. Nasrabadi (West Virginia University)</i>	
Tensor Feature Hallucination for Few-Shot Learning	2050
<i>Michalis Lazarou (Imperial College London), Tania Stathaki (Imperial College London), and Yannis Avrithis (Athena RC)</i>	

HierMatch: Leveraging Label Hierarchies for Improving Semi-Supervised Learning	2061
<i>Ashima Garg (IIIT Delhi), Shaurya Bagga (IIIT Delhi), Yashvardhan Singh (IIIT Delhi), and Saket Anand (Indraprastha Institute of Information Technology Delhi)</i>	
Identifying Wrongly Predicted Samples: A Method for Active Learning	2071
<i>Rahaf Aljundi (Toyota Motor Europe), Nikolay Chumerin (Toyota Motor Europe), and Daniel Olmeda Reino (Toyota Motor Europe)</i>	
Multi-Domain Incremental Learning for Semantic Segmentation	2080
<i>Prachi Garg (International Institute of Information Technology (IIIT)), Rohit Saluja (IIIT-Hyderabad), Vineeth N Balasubramanian (Indian Institute of Technology, Hyderabad), Chetan Arora (Indian Institute of Technology Delhi), Anbumani Subramanian (IIIT-Hyderabad), and C.V. Jawahar (IIIT-Hyderabad)</i>	
Few-Shot Open-Set Recognition of Hyperspectral Images With Outlier Calibration Network ..	2091
<i>Debabrata Pal (Honeywell), Valay Bundele (IIT Bombay), Renuka Sharma (IIT Bombay), Biplab Banerjee (Indian Institute of Technology, Bombay), and Yogananda Jeppu (Honeywell)</i>	
Multi-Motion and Appearance Self-Supervised Moving Object Detection	2101
<i>Fan Yang (United Imaging Intelligence), Srikrishna Karanam (United Imaging Intelligence), Meng Zheng (United Imaging Intelligence), Terrence Chen (United Imaging Intelligence), Haibin Ling (Stony Brook University), and Ziyang Wu (United Imaging Intelligence)</i>	
Masking Modalities for Cross-Modal Video Retrieval	2111
<i>Valentin Gabeur (Google), Arsha Nagrani (Oxford University), Chen Sun (Google), Karteek Alahari (Inria), and Cordelia Schmid (Google)</i>	
Single Source One Shot Reenactment Using Weighted Motion From Paired Feature Points ...	2121
<i>Soumya Tripathy (Tampere University of Technology), Juho Kannala (Aalto University, Finland), and Esa Rahtu (Tampere University)</i>	

Multimedia and Other Applications

Sign Language Translation With Hierarchical Spatio-Temporal Graph Neural Network	2131
<i>Jichao Kan (University of Sydney), Kun Hu (The University of Sydney), Markus Hagenbuchner (The University of Wollongong, Australia), Ah Chung Tsoi (University of Wollongong), Mohammed Bennamoun (University of Western Australia), and Zhiyong Wang (The University of Sydney)</i>	
Transductive Weakly-Supervised Player Detection Using Soccer Broadcast Videos	2141
<i>Chris Andrew Gadge (International Institute of Information Technology, Hyderabad) and C.V. Jawahar (IIIT-Hyderabad)</i>	
Beyond Mono to Binaural: Generating Binaural Audio From Mono Audio With Depth and Cross Modal Attention	2151
<i>Kranti Kumar Parida (IIT Kanpur), Siddharth Srivastava (IIT Delhi), and Gaurav Sharma (IIT Kanpur)</i>	
Unsupervised Sounding Object Localization With Bottom-Up and Top-Down Attention	2161
<i>Jiayin Shi (Shanghai Jiao Tong University) and Chao Ma (Shanghai Jiao Tong University)</i>	

Visually Guided Sound Source Separation and Localization Using Self-Supervised Motion Representations	2171
<i>Lingyu Zhu (Tampere University) and Esa Rahtu (Tampere University)</i>	
V-SlowFast Network for Efficient Visual Sound Separation	2182
<i>Lingyu Zhu (Tampere University) and Esa Rahtu (Tampere University)</i>	
CrossLocate: Cross-Modal Large-Scale Visual Geo-Localization in Natural Environments Using Rendered Modalities	2193
<i>Jan Tomešek (Brno University of Technology), Martin Čadík (Brno University of Technology), and Jan Brejcha (Brno University of Technology)</i>	
C-VTON: Context-Driven Image-Based Virtual Try-On Network	2203
<i>Benjamin Fele (University of Ljubljana), Ajda Lampe (University of Ljubljana), Peter Peer (University of Ljubljana), and Vitomir Struc (University of Ljubljana)</i>	
Predicting Levels of Household Electricity Consumption in Low-Access Settings	2213
<i>Simone Fobi (Columbia University), Joel Mugenyi (University of Massachusetts, Amherst), Nathaniel J. Williams (Rochester Institute of Technology), Vijay Modi (Columbia University), and Jay Taneja (University of Massachusetts)</i>	
Occlusion-Robust Object Pose Estimation With Holistic Representation	2223
<i>Bo Chen (The University of Adelaide), Tat-Jun Chin (The University of Adelaide), and Marius Klimavicius (Blackswan Technologies)</i>	
Dual-Head Contrastive Domain Adaptation for Video Action Recognition	2234
<i>Victor G. Turrisi da Costa (University of Trento), Giacomo Zara (University of Trento), Paolo Rota (University of Trento), Thiago Oliveira-Santos (UFES), Nicu Sebe (University of Trento), Vittorio Murino (Istituto Italiano di Tecnologia), and Elisa Ricci (University of Trento)</i>	
Towards Class-Oriented Poisoning Attacks Against Neural Networks	2244
<i>Bingyin Zhao (Clemson University) and Yingjie Lao (Clemson University)</i>	

3D Computer Vision III

Tensor-Based Non-Rigid Structure From Motion	2254
<i>Stella Graßhof (IT University of Copenhagen) and Sami Sebastian Brandt (IT University of Copenhagen)</i>	
Registration of Human Point Set Using Automatic Key Point Detection and Region-Aware Features	2264
<i>Amar Maharjan (University of North Texas) and Xiaohui Yuan (University of North Texas)</i>	
Style Agnostic 3D Reconstruction via Adversarial Style Transfer	2273
<i>Felix Petersen (University of Konstanz), Bastian Goldluecke (University of Konstanz), Oliver Deussen (University of Konstanz), and Hilde Kuehne (University of Frankfurt)</i>	

Seeing Implicit Neural Representations As Fourier Series	2283
<i>Nuri Benbarka (University of Tübingen), Timon Höfer (University of Tuebingen), Hamd ul-Moqueet Riaz (University of Tuebingen), and Andreas Zell (University of Tuebingen)</i>	
M3DETR: Multi-Representation, Multi-Scale, Mutual-Relation 3D Object Detection With Transformers	2293
<i>Tianrui Guan (University of Maryland, College Park), Jun Wang (The University of Maryland, College Park), Shiyi Lan (University of Maryland), Rohan Chandra (University of Maryland), Zuxuan Wu (UMD), Larry Davis (University of Maryland), and Dinesh Manocha (University of Maryland at College Park)</i>	
Neural Architecture Search for Efficient Uncalibrated Deep Photometric Stereo	2304
<i>Francesco Sarno (ETH), Suryansh Kumar (ETH Zurich), Berk Kaya (ETH Zurich), Zhiwu Huang (Singapore Management University), Vittorio Ferrari (Google Research), and Luc Van Gool (ETH Zurich)</i>	
Shape-Coded ArUco: Fiducial Marker for Bridging 2D and 3D Modalities	2315
<i>Lilika Makabe (Osaka University), Hiroaki Santo (Osaka University), Fumio Okura (Osaka University), and Yasuyuki Matsushita (Osaka University)</i>	
Modeling Dynamic Target Deformation in Camera Calibration	2325
<i>Annika Hagemann (Robert Bosch GmbH), Moritz Knorr (Robert Bosch GmbH), and Christoph Stiller (Institute of Measurement and Control Systems, Karlsruhe Institute of Technology (KIT))</i>	
Information Bottlenecked Variational Autoencoder for Disentangled 3D Facial Expression Modelling	2334
<i>Hao Sun (University of York), Nick Pears (University of York, UK), and Yajie Gu (University of York)</i>	
Mesh Convolutional Autoencoder for Semi-Regular Meshes of Different Sizes	2344
<i>Sara Hahner (Fraunhofer SCAI) and Jochen Garcke (Fraunhofer SCAI)</i>	

Deep Learning I

Auto-X3D: Ultra-Efficient Video Understanding via Finer-Grained Neural Architecture Search.	2354
<i>Yifan Jiang (University of Texas at Austin), Xinyu Gong (University of Texas at Austin), Junru Wu (Texas A&M University), Humphrey Shi (U of Oregon; UIUC), Zhicheng Yan (Facebook AI), and Zhangyang Wang (University of Texas at Austin)</i>	
EdgeConv With Attention Module for Monocular Depth Estimation	2364
<i>Minhyeok Lee (Yonsei University), Sangwon Hwang (Yonsei University), Chaewon Park (Yonsei University), and Sangyoun Lee (Yonsei University)</i>	
On the Maximum Radius of Polynomial Lens Distortion	2374
<i>Matthew J. Leotta (Kitware), David Russell (Kitware), and Andrew Matrai (Kitware)</i>	

Non-Semantic Evaluation of Image Forensics Tools: Methodology and Database	2383
<i>Quentin Bamme</i> (École Normale Supérieure Paris-Saclay), <i>Tina Nikoukhah</i> (Centre Borelli, ENS Paris-Saclay), <i>Marina Gardella</i> (École normale supérieure Paris-Saclay), <i>Rafael Grompone von Gioi</i> (Ecole Normale Supérieure Paris-Saclay), <i>Miguel Colom</i> (ENS Paris-Saclay), and <i>Jean-Michel Morel</i> (CMLA, ENS Cachan (Paris-Saclay))	
Addressing Out-of-Distribution Label Noise in Webly-Labelled Data	2393
<i>Paul Albert</i> (Insight Centre for Data Analytics (DCU)), <i>Diego Ortego</i> (Insight Centre for Data Analytics (DCU)), <i>Eric Arazo</i> (Insight Centre for Data Analytics (DCU)), <i>Noel E. O'Connor</i> (Dublin City University (DCU)), and <i>Kevin McGuinness</i> (Insight Centre for Data Analytics)	
SpectraNet: Learned Recognition of Artificial Satellites From High Contrast Spectroscopic Imagery	2403
<i>J. Zachary Gazak</i> (United States Space Force), <i>Ian McQuaid</i> (KBR), <i>Ryan Swindle</i> (United States Space Force), <i>Matthew Phelps</i> (United States Space Force), and <i>Justin Fletcher</i> (United States Space Force)	
Improving Fractal Pre-Training	2412
<i>Connor Anderson</i> (Brigham Young University) and <i>Ryan Farrell</i> (Brigham Young University)	
Multi-Domain Semantic Segmentation With Overlapping Labels	2422
<i>Petra Bevandić</i> (University of Zagreb), <i>Marin Oršić</i> (University of Zagreb), <i>Ivan Grubišić</i> (University of Zagreb), <i>Josip Šarić</i> (University of Zagreb), and <i>Siniša Šegvić</i> (University of Zagreb)	
GraN-GAN: Piecewise Gradient Normalization for Generative Adversarial Networks	2432
<i>Vineeth S. Bhaskara</i> (Samsung AI Center Toronto), <i>Tristan Aumentado-Armstrong</i> (Samsung Toronto AIC), <i>Allan D. Jepson</i> (Samsung Toronto AIC), and <i>Alex Levinshstein</i> (Samsung AI Centre Toronto)	

Vision and Language / Document Analysis

Co-Segmentation Aided Two-Stream Architecture for Video Captioning	2442
<i>Jayesh Vaidya</i> (Indian Institute of Technology, Madras), <i>Arulkumar Subramaniam</i> (Indian Institute of Technology Madras), and <i>Anurag Mittal</i> (Indian Institute of Technology Madras)	
Improve Image Captioning by Estimating the Gazing Patterns From the Caption	2453
<i>Rehab Alahmadi</i> (George Washington University) and <i>James Hahn</i> (The George Washington University)	
GraDual: Graph-Based Dual-Modal Representation for Image-Text Matching	2463
<i>Siqu Long</i> (University of Sydney), <i>Soyeon Caren Han</i> (University of Sydney), <i>Xiaojun Wan</i> (Peking University), and <i>Josiah Poon</i> (The University of Sydney)	
Let There Be a Clock on the Beach: Reducing Object Hallucination in Image Captioning	2473
<i>Ali Furkan Biten</i> (Computer Vision Center), <i>Lluís Gómez</i> (Universitat Autònoma de Barcelona), and <i>Dimosthenis Karatzas</i> (Computer Vision Centre)	

Is an Image Worth Five Sentences? A New Look Into Semantics for Image-Text Matching	2483
<i>Ali Furkan Biten (Computer Vision Center), Andrés Mafla (Computer Vision Centre), Lluís Gómez (Universitat Autònoma de Barcelona), and Dimosthenis Karatzas (Computer Vision Centre)</i>	
Variational Stacked Local Attention Networks for Diverse Video Captioning	2493
<i>Tonmoay Deb (Independent University, Bangladesh), Akib Sadmanee (Independent University, Bangladesh), Kishor Kumar Bhaumik (Independent University Bangladesh), Amin Ahsan Ali (Independent University Bangladesh), M Ashraful Amin (Independent University, Bangladesh), and A K M Mahbubur Rahman (Independent University, Bangladesh)</i>	
QUALIFIER: Question-Guided Self-Attentive Multimodal Fusion Network for Audio Visual Scene-Aware Dialog	2503
<i>Muchao Ye (The Pennsylvania State University), Quanzeng You (Microsoft Azure Computer Vision), and Fenglong Ma (Pennsylvania State University)</i>	
An Investigation of Critical Issues in Bias Mitigation Techniques	2512
<i>Robik Shrestha (Rochester Institute of Technology), Kushal Kafle (Adobe Research), and Christopher Kanan (Rochester Institute of Technology)</i>	
Natural Language Video Moment Localization Through Query-Controlled Temporal Convolution ..	2524
<i>Lingyu Zhang (Rensselaer Polytechnic Institute) and Richard J. Radke (Rensselaer Polytechnic Institute)</i>	
Post-OCR Paragraph Recognition by Graph Convolutional Networks	2533
<i>Renshen Wang (Google Research), Yasuhisa Fujii (Google Inc.), and Ashok C. Popat (Google Research)</i>	
Visual Understanding of Complex Table Structures From Document Images	2543
<i>Sachin Raja (IIIT Hyderabad), Ajoy Mondal (IIIT Hyderabad), and C.V. Jawahar (IIIT-Hyderabad)</i>	
Parsing Line Chart Images Using Linear Programming	2553
<i>Hajime Kato (Rakuten Institute of Technology), Mitsuru Nakazawa (Rakuten Institute of Technology, Rakuten Group, Inc.), Hsuan-Kung Yang (National Tsing Hua University), Mark Chen (The University of Tokyo), and Björn Stenger (Rakuten Institute of Technology)</i>	
One-Shot Compositional Data Generation for Low Resource Handwritten Text Recognition ..	2563
<i>Mohamed Ali Souibgui (Computer Vision Center, Barcelona, Spain), Ali Furkan Biten (Computer Vision Center), Sounak Dey (Computer Vision Center), Alicia Fornés (CVC), Yousri Kessentini (MIRACL Laboratory), Lluís Gómez (Universitat Autònoma de Barcelona), Dimosthenis Karatzas (Computer Vision Centre), and Josep Lladós (Computer Vision Center, Barcelona)</i>	
Efficient Counterfactual Debiasing for Visual Question Answering	2572
<i>Camila Kolling (PUCRS), Martin More (PUCRS), Nathan Gavenski (PUCRS), Eduardo Pooch (PUCRS), Otávio Parraga (Pucrs), and Rodrigo C. Barros (PUCRS)</i>	

InfographicVQA	2582
<i>Minesh Mathew (CVIT, IIIT-Hyderabad), Viraj Bagal (Indian Institute of Science Education and Research, Pune), Rubèn Tito (Computer Vision Center), Dimosthenis Karatzas (Computer Vision Centre), Ernest Valveny (Universitat Autònoma de Barcelona), and C.V. Jawahar (IIIT-Hyderabad)</i>	

Segmentation, Tracking, and Scene Understanding

Mixed-Dual-Head Meets Box Priors: A Robust Framework for Semi-Supervised Segmentation	2592
<i>Chenshu Chen (Hikvision Research Institute), Tao Liu (Hikvision Research Institute), Wenming Tan (Hikvision Research Institute), and Shiliang Pu (Hikvision Research Institute)</i>	
Self-Supervised Test-Time Adaptation on Video Data	2603
<i>Fatemeh Azimi (TU Kaiserslautern), Sebastian Palacio (DFKI), Federico Raue (DFKI), Jörn Hees (DFKI), Luca Bertinetto (University of Oxford), and Andreas Dengel (DFKI GmbH)</i>	
Modeling Aleatoric Uncertainty for Camouflaged Object Detection	2613
<i>Jiawei Liu (Australian National University), Jing Zhang (Australian National University), and Nick Barnes (ANU)</i>	
Perceptual Consistency in Video Segmentation	2623
<i>Yizhe Zhang (Qualcomm AI Research), Shubhankar Borse (Qualcomm AI Research), Hong Cai (Qualcomm AI Research), Ying Wang (Innopeak Technology Inc.), Ning Bi (Qualcomm), Xiaoyun Jiang (Qualcomm), and Fatih Porikli (Qualcomm AI Research)</i>	
AuxAdapt: Stable and Efficient Test-Time Adaptation for Temporally Consistent Video Semantic Segmentation	2633
<i>Yizhe Zhang (Qualcomm AI Research), Shubhankar Borse (Qualcomm AI Research), Hong Cai (Qualcomm AI Research), and Fatih Porikli (Qualcomm AI Research)</i>	
Time-Space Transformers for Video Panoptic Segmentation	2643
<i>Andra Petrovai (Technical University of Cluj-Napoca) and Sergiu Nedevschi (Technical University of Cluj-Napoca)</i>	
Inferring the Class Conditional Response Map for Weakly Supervised Semantic Segmentation	2653
<i>Weixuan Sun (Australian National University), Jing Zhang (Australian National University), and Nick Barnes (ANU)</i>	
Semi-Supervised Multi-Task Learning for Semantics and Depth	2663
<i>Yufeng Wang (Beihang University), Yi-Hsuan Tsai (Phiar Technologies), Wei-Chih Hung (Waymo), Wenrui Ding (Beihang University), Shuo Liu (Beihang University), and Ming-Hsuan Yang (University of California at Merced)</i>	
Compensation Tracker: Reprocessing Lost Object for Multi-Object Tracking	2673
<i>Zhibo Zou (Chongqing University of Posts and Telecommunications), Junjie Huang (Chongqing University of Posts and Telecommunications), and Ping Luo (Chongqing University of Posts and Telecommunications)</i>	

Hole-Robust Wireframe Detection	2684
<i>Naejin Kong (Samsung), Kiwoong Park (Samsung Electronics), and Harshith Goka (Samsung Research)</i>	
Global Assists Local: Effective Aerial Representations for Field of View Constrained Image Geo-Localization	2694
<i>Royston Rodrigues (NEC) and Masahiro Tani (NEC)</i>	
Dynamic Iterative Refinement for Efficient 3D Hand Pose Estimation	2703
<i>John Yang (Seoul National University), Yash Bhargat (Qualcomm AI Research), Simyung Chang (Qualcomm Korea YH), Fatih Porikli (Qualcomm AI Research), and Nojun Kwak (Seoul National University)</i>	
FASST: Fast Attention Based Single-Stage Segmentation Net for Real-Time Instance Segmentation	2714
<i>Yuan Cheng (Shanghai Jiao Tong University), Rui Lin (The University of Hong Kong), Peining Zhen (Shanghai Jiao Tong University), Tianshu Hou (Shanghai Jiao Tong University), Chiu Wa Ng (En:ai Ltd.), Hai-Bao Chen (Shanghai Jiao Tong University), Hao Yu (Southern University of Science and Technology), and Ngai Wong (The University of Hong Kong)</i>	
RLSS: A Deep Reinforcement Learning Algorithm for Sequential Scene Generation	2723
<i>Azimkhon Ostonov (KAUST), Peter Wonka (KAUST), and Dominik L. Michels (KAUST)</i>	
Learning Temporal Video Procedure Segmentation From an Automatically Collected Large Dataset	2733
<i>Lei Ji (Microsoft), Chenfei Wu (Microsoft), Daisy Zhou (Microsoft), Kun Yan (Beihang University), Edward Cui (Microsoft), Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences), and Nan Duan (Microsoft Research)</i>	

Adversarial Methods, Biometrics and Face Processing

Geometry-Aware Hierarchical Bayesian Learning on Manifolds	2743
<i>Yonghui Fan (Arizona State University) and Yalin Wang (Arizona State University)</i>	
Generative Adversarial Graph Convolutional Networks for Human Action Synthesis	2753
<i>Bruno Degardin (University of Beira Interior), João Neves (Tomiworld), Vasco Lopes (NOVA Lincs, Universidade da Beira Interior), João Brito (IT: Instituto de Telecomunicações - University of Beira Interior), Ehsan Yaghoubi (University of Beira Interior, Covilha), and Hugo Proença (U-Beira Interior)</i>	
A Riemannian Framework for Analysis of Human Body Surface	2763
<i>Emery Pierson (Université de Lille), Mohamed Daoudi (IMT Lille Douai), and Alice-Barbara Tumpach (Universite Lille)</i>	
Detection and Localization of Facial Expression Manipulations	2773
<i>Ghazal Mazaheri (University of California, Riverside) and Amit K. Roy-Chowdhury (University of California, Riverside)</i>	
Generalized Facial Manipulation Detection With Edge Region Feature Extraction	2784
<i>Dong-Keon Kim (Sungkyunkwan University) and Kwang-Su Kim (Sungkyunkwan University)</i>	

Digital and Physical-World Attacks on Remote Pulse Detection	2795
<i>Jeremy Speth (University of Notre Dame), Nathan Vance (University of Notre Dame), Patrick Flynn (University of Notre Dame), Kevin W. Bowyer (University of Notre Dame), and Adam Czajka (University of Notre Dame)</i>	
Unsupervised Robust Domain Adaptation Without Source Data	2805
<i>Peshal Agarwal (ETH Zürich), Danda Pani Paudel (ETH Zürich), Jan-Nico Zaech (ETH Zürich), and Luc Van Gool (ETH Zurich)</i>	
Attack Agnostic Detection of Adversarial Examples via Random Subspace Analysis	2815
<i>Nathan Drenkow (Johns Hopkins University Applied Physics Laboratory), Neil Fendley (JHU/APL), and Philippe Burlina (JHU/APL/CS/SOM)</i>	
Evaluating the Robustness of Semantic Segmentation for Autonomous Driving Against Real-World Adversarial Patch Attacks	2826
<i>Federico Nesti (Scuola Superiore Sant'Anna), Giulio Rossolini (Scuola Superiore Sant'Anna), Saasha Nair (Scuola Superiore Sant'Anna), Alessandro Biondi (Scuola Superiore Sant'Anna), and Giorgio Buttazzo (Scuola Superiore Sant'Anna)</i>	
ADC: Adversarial Attacks Against Object Detection That Evade Context Consistency Checks ..	2836
<i>Mingjun Yin (University of California, Riverside), Shasha Li (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)</i>	
Semantically Stealthy Adversarial Attacks Against Segmentation Models	2846
<i>Zhenhua Chen (Indiana University), Chuhua Wang (Indiana Univeristy, Bloomington), and David Crandall (Indiana University)</i>	
One-Class Learned Encoder-Decoder Network With Adversarial Context Masking for Novelty Detection	2856
<i>John Taylor Jewell (Western University), Vahid Reza Khazaie (Western University), and Yalda Mohsenzadeh (The University of Western Ontario)</i>	

Deep Learning, Diverse Applications in Computer Vision

Fair and Accurate Age Prediction Using Distribution Aware Data Curation and Augmentation	2867
<i>Yushi Cao (Nanyang Technological University), David Berend (Nanyang Technological University), Palina Tolmach (Nanyang Technological University), Guy Amit (Ben-Gurion University of the Negev), Moshe Levy (Ben-Gurion University of the Negev), Yang Liu (Nanyang Technology University, Singapore), Asaf Shabtai (Ben-Gurion University of the Negev), and Yuval Elovici (Ben-Gurion University of the Negev)</i>	
BiHPF: Bilateral High-Pass Filters for Robust Deepfake Detection	2878
<i>Yonghyun Jeong (Samsung SDS), Doyeon Kim (SamsungSDS), Seungjai Min (Samsung SDS), Seongho Joe (Samsung SDS), Youngjune Gwon (Samsung SDS), and Jongwon Choi (Chung-Ang University)</i>	
How and What To Learn: Taxonomizing Self-Supervised Learning for 3D Action Recognition .	2888
<i>Amor Ben Tanfous (ANITI), Aimen Zerroug (ANITI), Drew Linsley (Brown University), and Thomas Serre (Brown University)</i>	

Cross-Modal Adversarial Reprogramming	2898
<i>Paarth Neekhara (UCSD), Shehzeen Hussain (UCSD), Jinglong Du (UCSD), Shlomo Dubnov (UC San Diego), Farinaz Koushanfar (UCSD), and Julian McAuley (UCSD)</i>	
The Untapped Potential of Off-the-Shelf Convolutional Neural Networks	2907
<i>Matthew Inkawhich (Duke University), Nathan Inkawhich (Duke University), Eric Davis (SRC, Inc.), Hai Li (Duke University), and Yiran Chen (Duke University)</i>	
MisConv: Convolutional Neural Networks for Missing Data	2917
<i>Marcin Przewiężlikowski (Jagiellonian University), Marek Śmieja, Łukasz Struski (Jagiellonian University), and Jacek Tabor (Jagiellonian University)</i>	
Surrogate Model-Based Explainability Methods for Point Cloud NNs	2927
<i>Hanxiao Tan (TU Dortmund University) and Helena Kotthaus (TU Dortmund)</i>	
AttWalk: Attentive Cross-Walks for Deep Mesh Analysis	2937
<i>Ran Ben Izhak (Technion), Alon Lahav (Technion), and Ayellet Tal (Technion)</i>	
Self-Supervised Knowledge Transfer via Loosely Supervised Auxiliary Tasks	2947
<i>Seungbum Hong (Hutom), Jihun Yoon (Hutom), Min-Kook Choi (KAIST), and Junmo Kim (Hutom)</i>	
Sandwich Batch Normalization: A Drop-In Replacement for Feature Distribution Heterogeneity.....	2957
<i>Xinyu Gong (University of Texas at Austin), Wuyang Chen (University of Texas at Austin), Tianlong Chen (University of Texas at Austin), and Zhangyang Wang (University of Texas at Austin)</i>	
In-Field Phenotyping Based on Crop Leaf and Plant Instance Segmentation	2968
<i>Jan Weyler (University of Bonn), Federico Magistri (University of Bonn), Peter Seitz (Robert Bosch GmbH), Jens Behley (University of Bonn), and Cyrill Stachniss (University of Bonn)</i>	

Applications and Systems

PhotoWCT2: Compact Autoencoder for Photorealistic Style Transfer Resulting From Blockwise Training and Skip Connections of High-Frequency Residuals	2978
<i>Tai-Yin Chiu (University of Texas) and Danna Gurari (University of Colorado Boulder)</i>	
MAPS: Multimodal Attention for Product Similarity	2988
<i>Nilotpai Das (Amazon), Aniket Joshi (International Institute of Information Technology, Hyderabad), Promod Yenigalla (Amazon), and Gourav Agrwal (Amazon India)</i>	
mToFNet: Object Anti-Spoofing With Mobile Time-of-Flight Data	2997
<i>Yonghyun Jeong (Samsung SDS), Doyeon Kim (SamsungSDS), Jaehyeon Lee (Samsung SDS), Minki Hong (Samsung SDS), Solbi Hwang (Samsung sds), and Jongwon Choi (Chung-Ang University)</i>	
A Modular and Unified Framework for Detecting and Localizing Video Anomalies	3007
<i>Keval Doshi (University of South Florida) and Yasin Yilmaz (University of South Florida)</i>	

PROVES: Establishing Image Provenance Using Semantic Signatures	3017
<i>Mingyang Xie (Washington University in St. Louis), Manav Kulshrestha (University of Massachusetts), Shaojie Wang (Washington University in St. Louis), Jinghan Yang (Washington University in St. Louis), Ayan Chakrabarti (Washington University in St. Louis), Ning Zhang (Washington University in St. Louis), and Yevgeniy Vorobeychik (Washington University in St. Louis)</i>	
Multi-Branch Neural Networks for Video Anomaly Detection in Adverse Lighting and Weather Conditions	3027
<i>Sam Leroux (Ghent University - IMEC), Bo Li (Ghent University - imec), and Pieter Simoens (Ghent University - imec)</i>	
Rethinking Video Anomaly Detection – A Continual Learning Approach	3036
<i>Keval Doshi (University of South Florida) and Yasin Yilmaz (University of South Florida)</i>	
LEAD: Self-Supervised Landmark Estimation by Aligning Distributions of Feature Similarity ...	3046
<i>Tejan Karmali (Indian Institute of Science, Bengaluru), Abhinav Atrishi (Indian Institute of Science, Bengaluru), Sai Sree Harsha (National Institute of Technology, Karnataka, Surathkal), Susmit Agrawal (Indian Institute of Science), Varun Jampani (Google), and R. Venkatesh Babu (Indian Institute of Science)</i>	
Pro-CCaps: Progressively Teaching Colourisation to Capsules	3056
<i>Rita Pucci (University of Udine), Christian Micheloni (University of Udine, Italy), Gian Luca Foresti (University of Udine, Italy), and Niki Martinel (University of Udine)</i>	
Multi-Scale Patch-Based Representation Learning for Image Anomaly Detection and Segmentation	3065
<i>Chin-Chia Tsai (National Tsing Hua University), Tsung-Hsuan Wu (National Tsing Hua University), and Shang-Hong Lai (National Tsing Hua University)</i>	
Short-Term Solar Irradiance Prediction From Sky Images With a Clear Sky Model	3074
<i>Huiyu Gao (The Australian National University) and Miaomiao Liu (The Australian National University)</i>	
Intelligent Camera Selection Decisions for Target Tracking in a Camera Network	3083
<i>Anil Sharma (IIIT-Delhi), Saket Anand (Indraprastha Institute of Information Technology Delhi), and Sanjit K Kaul (IIIT-Delhi)</i>	

Computational Photography, Image and Video Synthesis

Revealing Disocclusions in Temporal View Synthesis Through Infilling Vector Prediction	3093
<i>Vijayalakshmi Kanchana (Indian Institute of Science), Nagabhushan Somraj (Indian Institute of Science), Suraj Yadwad (Indian Institute of Science), and Rajiv Soundararajan (Indian Institute of Science)</i>	
Pose-Guided Generative Adversarial Net for Novel View Action Synthesis	3103
<i>Xianhang Li (University of California, Santa Cruz), Junhao Zhang (National University of Singapore), Kunchang Li (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Shruti Vyas (University of Central Florida), and Yogesh S. Rawat (University of Central Florida)</i>	

Facial Attribute Transformers for Precise and Robust Makeup Transfer	3113
<i>Zhaoyi Wan (U. Rochester), Haoran Chen (Mybank, Ant Group), Jie An (University of Rochester), Wentao Jiang (Beihang University), Cong Yao (Alibaba DAMO Academy), and Jiebo Luo (U. Rochester)</i>	
Enhanced Correlation Matching Based Video Frame Interpolation	3123
<i>Sungho Lee (Samsung Research), Narae Choi (Samsung Electronics), and Woong Il Choi (Samsung)</i>	
Robust High-Resolution Video Matting With Temporal Guidance	3132
<i>Shanchuan Lin (University of Washington), Linjie Yang (ByteDance AI Lab), Imran Saleemi (ByteDance), and Soumyadip Sengupta (University of Washington)</i>	
Tailor Me: An Editing Network for Fashion Attribute Shape Manipulation	3142
<i>Youngjoong Kwon (The University of North Carolina at Chapel Hill), Stefano Petrangeli (Adobe), Dahun Kim (KAIST), Haoliang Wang (Adobe Research), Viswanathan Swaminathan (Adobe), and Henry Fuchs (unc)</i>	
Fast and Efficient Restoration of Extremely Dark Light Fields	3152
<i>Mohit Lamba (Indian Institute of Technology Madras) and Kaushik Mitra (IIT Madras)</i>	
S2FGAN: Semantically Aware Interactive Sketch-To-Face Translation	3162
<i>Yan Yang (The Australian National University), Md Zakir Hossain (The Australian National University), Tom Gedeon (The Australian National University), and Shafin Rahman (North South University)</i>	
Resolution-Robust Large Mask Inpainting With Fourier Convolutions	3172
<i>Roman Suvorov (Samsung AI Center Moscow), Elizaveta Logacheva (Samsung AI Center, Moscow), Anton Mashikhin (Samsung), Anastasia Remizova (Skoltech Institute of Science and Technology), Arsenii Ashukha (Samsung AI Center), Aleksei Silvestrov (Samsung AI Center Moscow), Naejin Kong (Samsung), Harshith Goka (Samsung Research), Kiwoong Park (Samsung Electronics), and Victor Lempitsky (Samsung)</i>	
CharacterGAN: Few-Shot Keypoint Character Animation and Reposing	3183
<i>Tobias Hinz (Adobe Research), Matthew Fisher (Adobe Research), Oliver Wang (Adobe Systems Inc), Eli Shechtman (Adobe Research, US), and Stefan Wermter (University of Hamburg)</i>	
Fast Nonlinear Image Unblending	3193
<i>Daichi Horita (The University of Tokyo), Kiyoharu Aizawa (The University of Tokyo), Ryohei Suzuki (Preferred Networks, Inc.), Taizan Yonetsuji (Preferred Networks, Inc.), and Huachun Zhu (Preferred Networks)</i>	
Measuring Hidden Bias Within Face Recognition via Racial Phenotypes	3202
<i>Seyma Yucer (Durham University), Furkan Tektas (Bubo AI), Noura Al Moubayed (University of Durham), and Toby P. Breckon (Durham University)</i>	

Imaging for Medical and Bioinformatics

- Weakly Supervised Branch Network With Template Mask for Classifying Masses in 3D Automated Breast Ultrasound 3212
Daekyung Kim (Seoul National University), Chang-Mo Nam (Monitor Corporation), Haesol Park (KIST), Mijung Jang (Seoul National University Bundang Hospital), and Kyong Joon Lee (SNUBH)
- Weakly-Supervised Convolutional Neural Networks for Vessel Segmentation in Cerebral Angiography 3220
Arvind Vepa (University of California, Los Angeles), Andrew Choi (University of California, Los Angeles), Noor Nakhaei (University of California, Los Angeles), Wonjun Lee (University of California, Los Angeles), Noah Stier (University of California, Santa Barbara), Andrew Vu (University of California, Los Angeles), Greyson Jenkins (Pepperdine University), Xiaoyan Yang (University of California, Los Angeles), Manjot Shergill (University of California, Los Angeles), Moira Desphy (University of California, Los Angeles), Kevin Delao (University of California, Los Angeles), Mia Levy (University of California, Los Angeles), Cristopher Garduno (University of California, Los Angeles), Lacy Nelson (University of California, Los Angeles), Wandu Liu (University of California, Los Angeles), Fan Hung (University of California, Los Angeles), and Fabien Scalzo (University of California, Los Angeles)
- METGAN: Generative Tumour Inpainting and Modality Synthesis in Light Sheet Microscopy .. 3230
Izabela Horvath (TUM), Johannes Paetzold (TUM), Oliver Schoppe (Technische Universität München), Rami Al-Maskari (TUM), Ivan Ezhov (TUM), Suprosanna Shit (TUM), Hongwei Li (Technical University of Munich), Ali Ertürk (Institute for Stroke and Dementia Research, LMU Munich), and Bjoern Menze (TUM)
- TA-Net: Topology-Aware Network for Gland Segmentation 3241
Haotian Wang (University of Idaho), Min Xian (University of Idaho), and Aleksandar Vakanski (University of Idaho)
- Non-Local Attention Improves Description Generation for Retinal Images 3250
Jia-Hong Huang (University of Amsterdam), Ting-Wei Wu (Georgia Institute of Technology), C.-H. Huck Yang (Georgia Institute of Technology), Zenglin Shi (University of Amsterdam), I-Hung Lin (Department of Ophthalmology, Tri-Service General Hospital, National Defense Medical Center), Jesper Tegner (KAUST), and Marcel Worring (University of Amsterdam)
- Weakly Supervised Learning for Joint Image Denoising and Protein Localization in Cryo-Electron Microscopy 3260
Qinwen Huang (Duke University), Ye Zhou (Duke University), Hsuan-Fu Liu (Duke University), and Alberto Bartesaghi (Duke University)
- AFTer-UNet: Axial Fusion Transformer UNet for Medical Image Segmentation 3270
Xiangyi Yan (University of California, Irvine), Hao Tang (University of California Irvine), Shanlin Sun (University of California, Irvine), Haoyu Ma (University of California, Irvine), Deying Kong (University of California, Irvine), and Xiaohui Xie (University of California, Irvine)

Recognition I

Danish Fungi 2020 – Not Just Another Image Recognition Dataset	3281
<i>Lukáš Pícek (University of West Bohemia), Milan Sulc (Czech Technical University in Prague), Jiří Matas (CMP CTU FEE), Thomas S. Jeppesen (University of Copenhagen), Jacob Heilmann-Clausen (GBIF), Thomas Læssøe (Department of Biology/Globe Institute, University of Copenhagen), and Tobias Frøslev (GLOBE Institute)</i>	
Multi-View Fusion of Sensor Data for Improved Perception and Prediction in Autonomous Driving	3292
<i>Sudeep Fadau (Aurora), Shreyash Pandey (Aurora), Darshan Hegde (Dexterity Inc), Yi Shi (Aurora), Fang-Chieh Chou (Aurora), Nemanja Djuric (Aurora Innovation), and Carlos Vallespi-Gonzalez (Uber)</i>	
Learned Event-Based Visual Perception for Improved Space Object Detection	3301
<i>Nikolaus Salvatore (Centauri Corporation) and Justin Fletcher (AFSPC)</i>	
Dataset Knowledge Transfer for Class-Incremental Learning Without Memory	3311
<i>Habib Slim (CEA-LIST), Eden Belouadah (CEA LIST), Adrian Popescu (CEA LIST), and Darian Onchis (West University of Timisoara)</i>	
Learning to Weight Filter Groups for Robust Classification	3321
<i>Siyang Yuan (Duke University), Yitong Li (Apple Inc), Dong Wang (Duke University), Ke Bai (Duke University), Lawrence Carin (Duke University; KAUST, Saudi Arabia), and David Carlson (Duke University)</i>	
Low-Cost Multispectral Scene Analysis With Modality Distillation	3331
<i>Heng Zhang (Univ Rennes 1), Elisa Fromont (Université Rennes 1, IRISA/INRIA rba), Sébastien Lefèvre (Université de Bretagne Sud / IRISA), and Bruno Avignon (Atermes)</i>	
HHP-Net: A Light Heteroscedastic Neural Network for Head Pose Estimation With Uncertainty	3341
<i>Giorgio Cantarini (MaLGA-DIBRIS, Università degli Studi di Genova), Federico Figari Tomenotti (MaLGA-DIBRIS, Università degli Studi di Genova), Nicoletta Noceti (MaLGA-DIBRIS, Università degli Studi di Genova), and Francesca Odone (University of Genova, Italy)</i>	
Batch Normalization Tells You Which Filter Is Important	3351
<i>Junghun Oh (Seoul National University), Heewon Kim (Seoul National University), Sungyong Baik (Seoul National University), Cheeun Hong (Seoul National University), and Kyoung Mu Lee (Seoul National University)</i>	
Novel Ensemble Diversification Methods for Open-Set Scenarios	3361
<i>Miriam Farber (Amazon.com), Roman Goldenberg (Google), George Leifman (Google), and Gal Novich (Amazon)</i>	
Measuring Representation of Race, Gender, and Age in Children’s Books: Face Detection and Feature Classification in Illustrated Images	3371
<i>Teodora Szasz (The University of Chicago), Emileigh Harrison (University of Chicago), Ping-Jung Liu (The University of Chicago), Ping-Chang Lin (University of Chicago), Hakizumwami Biral Runesha (University of Chicago), and Anjali Adukia (The University of Chicago)</i>	

Vision for Robotics / Deep Learning

- CeyMo: See More on Roads – A Novel Benchmark Dataset for Road Marking Detection 3381
Oshada Jayasinghe (University of Moratuwa), Sahan Hemachandra (University of Moratuwa), Damith Anhettigama (University of Moratuwa), Shenali Kariyawasam (University of Moratuwa), Ranga Rodrigo (University of Moratuwa), and Peshala Jayasekara (University of Moratuwa)
- Towards Active Vision for Action Localization With Reactive Control and Predictive Learning 3391
Shubham Trehan (Oklahoma State University) and Sathyanarayanan N. Aakur (OK State)
- ForeSI: Success-Aware Visual Navigation Agent 3401
Mahdi Kazemi Moghaddam (University of Adelaide), Ehsan Abbasnejad (The University of Adelaide), Qi Wu (University of Adelaide), Javen Qinfeng Shi (University of Adelaide), and Anton Van Den Hengel (University of Adelaide)
- DG-Labeler and DGL-MOTS Dataset: Boost the Autonomous Driving Perception 3411
Yiming Cui (University of Florida), Zhiwen Cao (Purdue University), Yixin Xie (UTEP), Xingyu Jiang (Purdue University), Feng Tao (University of Texas at San Antonio), Yingjie Victor Chen (Purdue University), Lin Li (UTEP), and Dongfang Liu (Rochester Institute of Technology)
- TricubeNet: 2D Kernel-Based Object Representation for Weakly-Occluded Oriented Object Detection 3421
Beomyoung Kim (NAVER), Janghyeon Lee (LG AI Research), Sihaeng Lee (LG AI Research), Doyeon Kim (KAIST), and Junmo Kim (KAIST)
- AirCamRTM: Enhancing Vehicle Detection for Efficient Aerial Camera-Based Road Traffic Monitoring 3431
Rafael Makrigiorgis (KIOS COE, UCY), Nicolas Hadjittoouli (University of Cyprus), Christos Kyrkou (KIOS CoE, University of Cyprus), and Theocharis Theocharides (University of Cyprus)
- StyleMC: Multi-Channel Based Fast Text-Guided Image Generation and Manipulation 3441
Umut Kocasari (Bogazici University), Alara Dirik (Bogazici University), Mert Tiftikli (Bogazici University), and Pinar Yanardag (Bogazici University)
- Self-Guidance: Improve Deep Neural Network Generalization via Knowledge Distillation 3451
Zhenzhu Zheng (University of Delaware) and Xi Peng (University of Delaware)
- REFICS: A Step Towards Linking Vision With Hardware Assurance 3461
Ronald Wilson (University of Florida), Hangwei Lu (Florida Institute for Cybersecurity (FICS) Research, University of Florida), Mengdi Zhu (University of Florida), Domenic Forte (University of Florida), and Damon L. Woodard (University of Florida)

Data InInstance Prior (DISP) in Generative Adversarial Networks	3471
<i>Puneet Mangla (IIT Hyderabad), Nupur Kumari (Carnegie Mellon University), Mayank Singh (Adobe Systems), Balaji Krishnamurthy (missing), and Vineeth N. Balasubramanian (Indian Institute of Technology Hyderabad)</i>	

Deep Learning - Architectures, Training, and Inference

Online Knowledge Distillation by Temporal-Spatial Boosting	3482
<i>Chengcheng Li (The University of Tennessee, Knoxville), Zi Wang (UTK), and Hairong Qi (University of Tennessee-Knoxville)</i>	
Multi-Dimensional Dynamic Model Compression for Efficient Image Super-Resolution	3492
<i>Zejiang Hou (Princeton University) and Sun-Yuan Kung (Princeton University)</i>	
Contextual Gradient Scaling for Few-Shot Learning	3503
<i>Sanghyuk Lee (Inha University), Seunghyun Lee (Inha University), and Byung Cheol Song (Inha University)</i>	
Channel Pruning via Lookahead Search Guided Reinforcement Learning	3513
<i>Zi Wang (UTK) and Chengcheng Li (The University of Tennessee, Knoxville)</i>	
Progressive Automatic Design of Search Space for One-Shot Neural Architecture Search	3525
<i>Xin Xia (Bytedance Inc), Xuefeng Xiao (ByteDance Inc), Xing Wang (Bytedance AI Lab), and Min Zheng (ByteDance)</i>	
Meta Approach to Data Augmentation Optimization	3535
<i>Ryuichiro Hataya (The University of Tokyo), Jan Zdenek (The University of Tokyo), Kazuki Yoshizoe (Kyushu University), and Hideki Nakayama (The University of Tokyo)</i>	
Approximate Neural Architecture Search via Operation Distribution Learning	3545
<i>Xingchen Wan (Oxford), Binxin Ru (University of Oxford), Pedro M. Esparana (Huawei), and Fabio Maria Carlucci (Huawei Noah's Ark Lab)</i>	
Resource-Efficient Hybrid X-Formers for Vision	3555
<i>Pranav Jeevan (IIT Bombay) and Amit Sethi (Indian Institute of Technology Bombay)</i>	
Leveraging Test-Time Consensus Prediction for Robustness Against Unseen Noise	3564
<i>Anindya Sarkar (IIT Hyderabad), Anirban Sarkar (IIT Hyderabad), and Vineeth N Balasubramanian (Indian Institute of Technology, Hyderabad)</i>	
Latent Reweighting, an Almost Free Improvement for GANs	3574
<i>Thibaut Issenhuth (Criteo), Ugo Tanielian (UPMC-Criteo), David Picard (ENPC), and J�er�mie Mary (Criteo AI Lab)</i>	
Towards a Robust Differentiable Architecture Search Under Label Noise	3584
<i>Christian Simon (Australian National University), Piotr Koniusz (ANU College of Engineering and Computer Science), Lars Petersson (Data61/CSIRO), Yan Han (The Australian National University), and Mehrtash Harandi (Monash University)</i>	

EZCrop: Energy-Zoned Channels for Robust Output Pruning	3595
<i>Rui Lin (The University of Hong Kong), Jie Ran (HKU), Dongpeng Wang (umechk), King Hung Chiu (umechk), and Ngai Wong (The University of Hong Kong)</i>	

Deep Learning II

PRECODE – A Generic Model Extension To Prevent Deep Gradient Leakage	3605
<i>Daniel Scheliga (Technische Universität Ilmenau), Patrick Mäder (Technische Universität Ilmenau), and Marco Seeland (TU Ilmenau)</i>	
S2-MLP: Spatial-Shift MLP Architecture for Vision	3615
<i>Tan Yu (Baidu Research), Xu Li (Baidu Research), Yunfeng Cai (Baidu Research), Mingming Sun (Baidu Research), and Ping Li (Baidu Research)</i>	
Preventing Catastrophic Forgetting and Distribution Mismatch in Knowledge Distillation via Synthetic Data	3625
<i>Kuluhan Binici (National University of Singapore), Nam Trung Pham (Institute For Infocomm Research), Tulika Mitra (National University of Singapore), and Karianto Leman (i2r.a-star.edu.sg)</i>	
FalCon: Fine-Grained Feature Map Sparsity Computing With Decomposed Convolutions for Inference Optimization	3634
<i>Zirui Xu (George Mason University), Fuxun Yu (George Mason University), Chenxi Liu (George Mason University), Zhe Wu (Comcast Research Lab), Hongcheng Wang (Comcast), and Xiang Chen (George Mason University)</i>	
Model Compression Using Optimal Transport	3645
<i>Sahas Lohit (Mitsubishi Electric Research Laboratories) and Michael Jones (MERL)</i>	
Sharing Decoders: Network Fission for Multi-Task Pixel Prediction	3655
<i>Steven Hickson (Google), Karthik Raveendran (Google), and Irfan Essa (Google)</i>	
Hessian-Aware Pruning and Optimal Neural Implant	3665
<i>Shixing Yu (Peking University), Zhewei Yao (University of California, Berkeley), Amir Gholami (UC Berkeley), Zhen Dong (UC Berkeley), Sehoon Kim (University of California, Berkeley), Michael W. Mahoney (UC Berkeley), and Kurt Keutzer (EECS, UC Berkeley)</i>	
Latent to Latent: A Learned Mapper for Identity Preserving Editing of Multiple Face Attributes in StyleGAN-Generated Images	3677
<i>Siavash Khodadadeh (University of Central Florida), Shabnam Ghadar (Adobe), Saeid Motiian (Adobe), Wei-An Lin (Adobe), Ladislau Bölöni (University of Central Florida), and Ratheesh Kalarot (Adobe)</i>	

Recognition II

SeaDronesSee: A Maritime Benchmark for Detecting Humans in Open Water	3686
<i>Leon Amadeus Varga (University of Tuebingen), Benjamin Kiefer (Universität Tübingen), Martin Messmer (Universiteat Tuebingen), and Andreas Zell (University of Tuebingen)</i>	

Meta-UDA: Unsupervised Domain Adaptive Thermal Object Detection Using Meta-Learning .	3697
<i>Vibashan VS (Johns Hopkins University), Domenick Poster (WVU), Suya You (US Army Research Laboratory), Shuowen Hu (ARL), and Vishal M. Patel (Johns Hopkins University)</i>	
Class-Balanced Active Learning for Image Classification	3707
<i>Javad Zolfaghari Bengar (Computer Vision Center), Joost van de Weijer (Computer Vision Center), Laura Lopez Fuentes (Universitat de les Illes Balears), and Bogdan Raducanu (Computer Version Center)</i>	
Single Image Object Counting and Localizing Using Active-Learning	3717
<i>Inbar Huberman-Spiegelglas (The Hebrew University of Jerusalem) and Raanan Fattal (Hebrew University of Jerusalem, Israel)</i>	
F-CAM: Full Resolution Class Activation Maps via Guided Parametric Upscaling	3727
<i>Soufiane Belharbi (ÉTS Montreal), Aydin Sarraf (Ericsson), Marco Pedersoli (École de technologie supérieure), Ismail Ben Ayed (ETS Montreal), Luke McCaffrey (McGill University), and Eric Granger (ETS Montreal)</i>	
Learning To Generate the Unknowns as a Remedy to the Open-Set Domain Shift	3737
<i>Mahsa Baktashmotlagh (University of Queensland), Tianle Chen (University of Queensland), and Mathieu Salzmann (EPFL)</i>	
Fast-CLOCs: Fast Camera-LiDAR Object Candidates Fusion for 3D Object Detection	3747
<i>Su Pang (Michigan State University), Daniel Morris (Michigan State University), and Hayder Radha (Michigan State University)</i>	
Learable Adaptive Cosine Estimator (LACE) for Image Classification	3757
<i>Joshua Peebles (University of Florida), Connor H. McCurley (University of Florida), Sarah Walker (University), Dylan Stewart (University of Florida), and Alina Zare (University of Florida)</i>	
Self-Supervised Shape Alignment for Sports Field Registration	3768
<i>Feng Shi (SPORTLOGIQ), Paul Marchwica (Sportlogiq), Juan Camilo Gamboa Higuera (SPORTLOGiQ), Michael Jamieson (Sportlogiq), Mehrsan Javan (SPORTLOGiQ Inc., McGill University), and Parthipan Siva (Sportlogiq)</i>	
Inpaint2Learn: A Self-Supervised Framework for Affordance Learning	3778
<i>Lingzhi Zhang (University of Pennsylvania), Weiyu Du (University of Pennsylvania), Shenghao Zhou (University of Pennsylvania), Jiancong Wang (University of Pennsylvania), and Jianbo Shi (University of Pennsylvania)</i>	

Explainability, Robustness and Ethics in Vision

DAD: Data-Free Adversarial Defense at Test Time	3788
<i>Gaurav Kumar Nayak (Indian Institute of Science, Bangalore), Ruchit Rawal (Indian Institute of Science), and Anirban Chakraborty (Indian Institute of Science)</i>	

Physical Adversarial Attacks on an Aerial Imagery Object Detector	3798
<i>Andrew Du (The University of Adelaide), Bo Chen (The University of Adelaide), Tat-Jun Chin (The University of Adelaide), Yee Wei Law (University of South Australia), Michele Sasdelli (The University of Adelaide), Ramesh Rajasegaran (The University of Adelaide), and Dillon Campbell (The University of Adelaide)</i>	
Geometrically Adaptive Dictionary Attack on Face Recognition	3809
<i>Junyoung Byun (KAIST), Hyojun Go (KAIST), and Changick Kim (KAIST)</i>	
On the Effectiveness of Small Input Noise for Defending Against Query-Based Black-Box Attacks	3819
<i>Junyoung Byun (KAIST), Hyojun Go (KAIST), and Changick Kim (KAIST)</i>	
REGroup: Rank-Aggregating Ensemble of Generative Classifiers for Robust Predictions	3829
<i>Lokender Tiwari (TCS Research), Anish Madan (Indraprastha Institute of Information Technology, Delhi), Saket Anand (Indraprastha Institute of Information Technology Delhi), and Subhashis Banerjee (IIT Delhi)</i>	
Generative Adversarial Attack on Ensemble Clustering	3839
<i>Chetan Kumar (University of Massachusetts Dartmouth), Deepak Kumar (University of Massachusetts Dartmouth), and Ming Shao (University of Massachusetts Dartmouth)</i>	
Adversarial Semantic Hallucination for Domain Generalized Semantic Segmentation	3849
<i>Gabriel Tjio (Institute of High Performance Computing, A*STAR), Ping Liu (IHPC), Joey Tianyi Zhou (IHPC, A*STAR), and Rick Siow Mong Goh (Institute of High Performance Computing, A*STAR)</i>	
Unveiling Real-Life Effects of Online Photo Sharing	3859
<i>Van-Khoa Nguyen (Bach Khoa University), Adrian Popescu (CEA LIST), and Jérôme Deshayes-Chossart (CEA)</i>	
Reconstructing Training Data From Diverse ML Models by Ensemble Inversion	3870
<i>Qian Wang (Apple) and Daniel Kurz (Apple)</i>	
Evaluating and Mitigating Bias in Image Classifiers: A Causal Perspective Using Counterfactuals	3879
<i>Saloni Dash (Microsoft Research), Vineeth N Balasubramanian (Indian Institute of Technology, Hyderabad), and Amit Sharma (Microsoft Research)</i>	
Fair Visual Recognition in Limited Data Regime Using Self-Supervision and Self-Distillation	3889
<i>Pratik Mazumder (Indian Institute of Technology, Kanpur), Pravendra Singh (IIT Roorkee), and Vinay P. Namboodiri (University of Bath)</i>	
Does Data Repair Lead to Fair Models? Curating Contextually Fair Data To Reduce Model Bias.....	3898
<i>Sharat Agarwal (Indraprastha Institute of Information Technology, Delhi), Sumanyu Muku (Indian Institute of Technology Delhi), Saket Anand (Indraprastha Institute of Information Technology Delhi), and Chetan Arora (Indian Institute of Technology Delhi)</i>	

Generating and Controlling Diversity in Image Search	3908
<i>Md. Mehrab Tanjim (University of California, San Diego), Ritwik Sinha (Adobe Research), Krishna Kumar Singh (Adobe Research), Sridhar Mahadevan (Adobe Research), David Arbour (Adobe Research), Moumita Sinha (Adobe), and Garrison W. Cottrell (University of California, San Diego)</i>	
PoP-Net: Pose Over Parts Network for Multi-Person 3D Pose Estimation From a Depth Image	3917
<i>Yuliang Guo (Robert Bosch LLC), Zhong Li (OPPO US Research Center), Zekun Li (OPPO US Research Center), Xiangyu Du (OPPO US Research Center), Shuxue Quan (OPPO US Research Center), and Yi Xu (OPPO US Research Center)</i>	
Improving Model Generalization by Agreement of Learned Representations From Data Augmentation	3927
<i>Rowel Atienza (University of the Philippines)</i>	

Low-Level Vision / Image Processing

Event-Based Kilohertz Eye Tracking Using Coded Differential Lighting	3937
<i>Timo Stoffregen (Monash University), Hossein Daraei (Facebook), Clare Robinson (Facebook), and Alexander Fix (Facebook Reality Labs)</i>	
MEGAN: Memory Enhanced Graph Attention Network for Space-Time Video Super-Resolution	3946
<i>Chenyu You (Yale University), Lianyi Han (Tencent America), Aosong Feng (Yale University), Ruihan Zhao (UT Austin), Hui Tang (Tencent Medical AI Lab), and Wei Fan (Tencent)</i>	
Single Image Deraining Network With Rain Embedding Consistency and Layered LSTM	3957
<i>Yizhou Li (Tokyo Institute of Technology), Yusuke Monno (Tokyo Institute of Technology), and Masatoshi Okutomi (Tokyo Institute of Technology)</i>	
Neural Radiance Fields Approach to Deep Multi-View Photometric Stereo	3967
<i>Berk Kaya (ETH Zurich), Suryansh Kumar (ETH Zurich), Francesco Sarno (ETH), Vittorio Ferrari (Google Research), and Luc Van Gool (ETH Zurich)</i>	
Detail Preserving Residual Feature Pyramid Modules for Optical Flow	3980
<i>Libo Long (University of Ottawa) and Jochen Lang (University of Ottawa)</i>	
No-Reference Image Quality Assessment via Transformers, Relative Ranking, and Self-Consistency	3989
<i>S. Alireza Golestaneh (Carnegie Mellon University/ Bosch AI), Saba Dadsetan (University of Pittsburgh), Kris M. Kitani (CMU), and Alireza Golestaneh (Bosch Center for AI and Research)</i>	
Learning From the CNN-Based Compressed Domain	4000
<i>Zhenzhen Wang (Nanyang technological University), Minghai Qin (Western Digital Research), and Yen-Kuang Chen (Alibaba DAMO Academy)</i>	

MoESR: Blind Super-Resolution Using Kernel-Aware Mixture of Experts	4009
<i>Mohammad Emad (Eindhoven University of Technology), Maurice Peemen (Thermo Fisher Scientific), and Henk Corporaal (TU Eindhoven)</i>	
edge-SR: Super-Resolution for the Masses	4019
<i>Pablo Navarrete Michelini (BOE Technology Group Co., Ltd.), Yunhua Lu (BOE Technology Group Co., Ltd.), and Xingqun Jiang (BOE Technology Group Co., Ltd.)</i>	
Image-Adaptive Hint Generation via Vision Transformer for Outpainting	4029
<i>Daehyeon Kong (Sogang University), Kyeongbo Kong (Pukyong National University), Kyunghun Kim (Sogang University), Sung-Jun Min (Sogang University), and Suk-Ju Kang (Sogang University)</i>	
Hyperspectral Image Super-Resolution With RGB Image Super-Resolution as an Auxiliary Task	4039
<i>Ke Li (ETH Zurich), Dengxin Dai (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
Deep Optimization Prior for THz Model Parameter Estimation	4049
<i>Tak Ming Wong (University of Siegen), Hartmut Bauermeister (University of Siegen), Matthias Kahl (University of Siegen), Peter Haring Bolívar (University of Siegen), Michael Möller (University of Siegen), and Andreas Kolb (University of Siegen)</i>	
Geometry-Inspired Top-K Adversarial Perturbations	4059
<i>Nurislam Tursynbek (Skolkovo Institute of Science and Technology), Aleksandr Petiushko (MSU; Huawei), and Ivan Oseledets (Skolkovo Institute of Science and Technology)</i>	
Nonnegative Low-Rank Tensor Completion via Dual Formulation With Applications to Image and Video Completion	4069
<i>Tanmay Kumar Sinha (International Institute of Information Technology, Hyderabad), Jayadev Naram (International Institute of Information Technology, Hyderabad), and Pawan Kumar (IIIT, Hyderabad)</i>	
PredStereo: An Accurate Real-Time Stereo Vision System	4078
<i>Diksha Moolchandani (IIT Delhi), Nivedita Shrivastava (IIT Delhi), Anshul Kumar (IIT Delhi), and Smruti R. Sarangi (IIT Delhi)</i>	

Author Index