

2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2023)

**Waikoloa, Hawaii, USA
2-7 January 2023**

Pages 1-723



**IEEE Catalog Number: CFP23082-POD
ISBN: 978-1-6654-9347-5**

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

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

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

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

IEEE Catalog Number:	CFP23082-POD
ISBN (Print-On-Demand):	978-1-6654-9347-5
ISBN (Online):	978-1-6654-9346-8
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

2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) **WACV 2023**

Table of Contents

Message from the WACV 2023 General and Program Chairs	lxviii
WACV 2023 Organizing Committee	lxx
WACV 2023 Area Chairs	lxxi
Sponsors and Donors	lxxii

1A: Recognition

CNN2Graph: Building Graphs for Image Classification	1
<i>Vivek Trivedy (Temple University) and Longin Jan Latecki (Temple University)</i>	
Token Pooling in Vision Transformers for Image Classification	12
<i>Dmitrii Marin (U of Waterloo), Jen-Hao Rick Chang (Apple), Anurag Ranjan (Apple), Anish Prabhu (Apple), Mohammad Rastegari (Apple Inc), and Oncel Tuzel (Apple)</i>	
D2F2WOD: Learning Object Proposals for Weakly-Supervised Object Detection via Progressive Domain Adaptation	22
<i>Yuting Wang (Rutgers University), Ricardo Guerrero (Samsung AI Center Cambridge), and Vladimir Pavlovic (Rutgers University)</i>	
ML-Decoder: Scalable and Versatile Classification Head	32
<i>Gilad Sharir (Alibaba Group), Tal Ridnik (Alibaba), Emanuel Ben Baruch (Alibaba), Avi Ben-Cohen (Alibaba Group), and Asaf Noy (missing)</i>	
Large-Scale Open-Set Classification Protocols for ImageNet	42
<i>Andres Palechor (UZH), Annesha Bhounik (Adidas), and Manuel Günther (University of Zurich)</i>	
Composite Relationship Fields with Transformers for Scene Graph Generation	52
<i>George Adaimi (EPFL), David Mizrahi (EPFL), and Alexandre Alahi (EPFL)</i>	
CoKe: Contrastive Learning for Robust Keypoint Detection	65
<i>Yutong Bai (Johns Hopkins University), Angtian Wang (Johns Hopkins University), Adam Kortylewski (Max Planck Institute for Informatics), and Alan Yuille (Johns Hopkins University)</i>	
Towards Few-Annotation Learning for Object Detection: Are Transformer-Based Models more Efficient?	75
<i>Quentin Bouniot (CEALIST), Angélique Loesch (CEA LIST), Romaric Audigier (CEA LIST), and Amaury Habrard (University of St-Etienne Lab. H. Curien)</i>	

Scaling Novel Object Detection with Weakly Supervised Detection Transformers	85
<i>Tyler M Labonte (Georgia Institute of Technology), Yale Song (Facebook AI Research), Xin Wang (Microsoft Research), Vibhav Vineet (Microsoft), and Neel Joshi (Microsoft Research)</i>	
Dense Prediction with Attentive Feature Aggregation	97
<i>Yung-Hsu Yang (National Tsing Hua University), Thomas E Huang (ETH Zürich), Min Sun (NTHU), Samuel Rota Bulò (Facebook), Peter Kotschieder (Facebook), and Fisher Yu (ETH Zurich)</i>	
Boosting Vision Transformers for Image Retrieval	107
<i>Chull Hwan Song (Dealicious Inc.), Jooyoung Yoon (Dealicious Inc.), Shunghyun Choi (Dealicious Inc.), and Yannis Avrithis (IARAIathena RC)</i>	
Is Your Noise Correction Noisy? PLS: Robustness to Label Noise with Two Stage Detection	118
<i>Paul Albert (Insight Centre for Data Analytics (DCU)), Eric Arazo (Insight Centre for Data Analytics (DCU)), Tarun Krishna (DCU), Noel O'Connor (missing), and Kevin McGuinness (DCU)</i>	
Two-Level Data Augmentation for Calibrated Multi-View Detection	128
<i>Martin Engilberge (EPFL), Haixin Shi (EPFL), Zhiye Wang (EPFL), and Pascal Fua (EPFLSwitzerland)</i>	
TCAM: Temporal Class Activation Maps for Object Localization in Weakly-Labeled Unconstrained Videos	137
<i>Soufiane Belharbi (ÉTS Montreal), Ismail Ben Ayed (ETS Montreal), Luke Mccaffrey (McGill University), and Eric Granger (ETS Montreal)</i>	
LAVA: Label-Efficient Visual Learning and Adaptation	147
<i>Islam H Nassar (Monash University), Munawar Hayat (Monash University), Ehsan M Abbasnejad (The University of Adelaide), Hamid Rezaatofghi (Monash University), Mehrtash Harandi (Monash University), and Reza Haffari (Monash University Australia)</i>	
GEMS: Scene Expansion Using Generative Models of Graphs	157
<i>Rishi Agarwal (IIT Bombay), Tirupati Saketh Chandra (Indian Institute of Technology Bombay), Vaidehi Patil (Indian Institute of Technology Bombay), Aniruddha Mahapatra (Carnegie Mellon University), Kuldeep Kulkarni (Adobe Research), and Vishwa Vinay (Adobe Research)</i>	
Dynamic Mixture of Counter Network for Location-Agnostic Crowd Counting	167
<i>Mingjie Wang (Zhejiang Sci-Tech University), Hao Cai (Memorial University of Newfoundland), Yong Dai (Tencent AI Lab), and Minglun Gong (University of Guelph)</i>	

1B: Computational Photography

Simultaneous Acquisition of High Quality RGB Image and Polarization Information Using a Sparse Polarization Sensor	178
<i>Tepei Kurita (Sony Group Corporation), Yuhi Kondo (Sony Group Corporation), Legong Sun (Sony Group Corporation), and Yusuke Moriuchi (Sony Group Corporation)</i>	
DyStyle: Dynamic Neural Network for Multi-Attribute-Conditioned Style Editings	189
<i>Bingchuan Li (Bytedance), Shaofei Cai (bytedance), Wei Liu (Bytedance), Peng Zhang (Bytedance), Qian He (Bytedance), Miao Hua (Bytedance), and Zili Yi (ByteDance)</i>	

Lossy Image Compression with Quantized Hierarchical VAEs	198
<i>Zhihao Duan (Purdue University), Ming Lu (Nanjing University), Zhan Ma (Nanjing University), and Fengqing Maggie Zhu (Purdue UniversityUSA)</i>	
Keys to Better Image Inpainting: Structure and Texture Go Hand in Hand	208
<i>Jitesh Jain (IIT Roorkee ; Picsart), Yuqian Zhou (Adobe), Ning Yu (Salesforce Research), and Humphrey Shi (U of Oregon ; UIUC ; PAIR)</i>	
Frame Interpolation for Dynamic Scenes with Implicit Flow Encoding	218
<i>Pedro Figueiredo (Texas A&M University), Avinash Paliwal (Texas A&M University), and Nima Khademi Kalantari (Texas A&M University)</i>	
I See-Through You: A Framework for Removing Foreground Occlusion in Both Sparse and Dense Light Field Images	229
<i>Jiwan Hur (KAIST), Jae Young Lee (KAIST), Jaehyun Choi (KAIST), and Junmo Kim (KAIST)</i>	
Burst Reflection Removal Using Reflection Motion Aggregation Cues	239
<i>B H Pawan Prasad (Samsung Research), Green Rosh Ks (Samsung Research Institute Bangalore), Lokesh Boregowda (Samsung Research Institute Bangalore), and Kaushik Mitra (IIT Madras)</i>	
Line Search-Based Feature Transformation for Fast, Stable, and Tunable Content-Style Control in Photorealistic Style Transfer	249
<i>Tai-Yin Chiu (University of Texas) and Danna Gurari (University of Colorado Boulder)</i>	
Panoptic-Aware Image-to-Image Translation	259
<i>Liyun Zhang (Osaka University), Photchara Ratsamee (Osaka University), Bowen Wang (Osaka University), Zhaojie Luo (Osaka University), Yuki Uranishi (Osaka University), Manabu Higashida (Osaka University), and Haruo Takemura (Osaka University)</i>	
SimGlim: Simplifying Glimpse Based Active Visual Reconstruction	269
<i>Abhishek Jha (K.U. Leuven), Soroush Seifi (Toyota), and Tinne Tuytelaars (KU Leuven)</i>	
Evaluating Generative Networks Using Gaussian Mixtures of Image Features	279
<i>Lorenzo Luzi (Rice University), Carlos M Ortiz Marrero (Pacific Northwest National Laboratory), Nile N Wymar (PNNL), Richard Baraniuk (Rice University), and Michael Henry (Pacific Northwest National Laboratory)</i>	
More Control for Free! Image Synthesis with Semantic Diffusion Guidance	289
<i>Xihui Liu (UC Berkeley), Dong Huk Park (UC Berkeley), Samaneh Azadi (University of California Berkeley), Gong Zhang (University of Oregon), Arman Chopikyan (Picsart), Yuxiao Hu (Futurewei), Humphrey Shi (U of Oregon ; UIUC ; PAIR), Anna Rohrbach (UC Berkeley), and Trevor Darrell (UC Berkeley)</i>	
Placing Human Animations into 3D Scenes by Learning Interaction- and Geometry-Driven Keyframes	300
<i>James F Mullen (University of Maryland), Divya Kothandaraman (University of Maryland College Park), Dinesh Manocha (University of Maryland at College Park), and Aniket Bera (University of MarylandCollege Park)</i>	

Surface Normal Estimation from Optimized and Distributed Light Sources Using DNN-Based Photometric Stereo	311
<i>Takafumi Iwaguchi (Kyushu Univ.) and Hiroshi Kawasaki (Kyushu Univ.)</i>	
Interpolated SelectionConv for Spherical Images and Surfaces	321
<i>David M Hart (Brigham Young University), Michael Whitney (Brigham Young University), and Bryan S Morse (Brigham Young University)</i>	
RAST: Restorable Arbitrary Style Transfer via Multi-Restoration	331
<i>Yingnan Ma (University of Alberta), Chengqi Zhao (University of Alberta), Xudong Li (University of Alberta), and Anup Basu (University of AlbertaCanada)</i>	
On Quantizing Implicit Neural Representations	341
<i>Cameron Gordon (University of Adelaide), Shin-Fang Chng (The University of Adelaide), Lachlan Macdonald (Australian Institute for Machine Learning), and Simon Lucey (University of Adelaide)</i>	
Lightweight Video Denoising Using Aggregated Shifted Window Attention	351
<i>Lydia Lindner (Graz University of Technology), Alexander Effland (University of Bonn), Filip Ilic (Graz University of Technology), Thomas Pock (Graz University of Technology), and Erich Kobler (University Hospital Bonn)</i>	

1C: Domain Adaptation and Generalization

Federated Domain Generalization for Image Recognition via Cross-Client Style Transfer	361
<i>Junming Chen (HKUST), Meirui Jiang (The Chinese University of Hong Kong), Dou Qi (The Chinese University of Hong Kong), and Qifeng Chen (HKUST)</i>	
CTrGAN: Cycle Transformers GAN for Gait Transfer	371
<i>Shahar Mahpod (Ariel University), Noam Gaash (Ariel University), Hay Hoffman (Ariel university), and Gil Ben-Artzi (Ariel University)</i>	
SALAD: Source-Free Active Label-Agnostic Domain Adaptation for Classification, Segmentation and Detection	382
<i>Divya Kothandaraman (University of Maryland College Park), Sumit Shekhar (Adobe Research), Abhilasha Sancheti (Adobe Research), Manoj Ghuhan Arivazhagan (Carnegie Mellon University), Tripti Shukla (Adobe Research), and Dinesh Manocha (University of Maryland at College Park)</i>	
Backprop Induced Feature Weighting for Adversarial Domain Adaptation with Iterative Label Distribution Alignment	392
<i>Thomas Westfechtel (The University of Tokyo), Hao-Wei Yeh (The University of Tokyo), Qier Meng (Research Center for Advanced Science and TechnologyThe University of Tokyo), Yusuke Mukuta (The University of Tokyo), and Tatsuya Harada (The University of Tokyo / RIKEN)</i>	
Semi-Supervised Domain Adaptation with Auto-Encoder via Simultaneous Learning	402
<i>Md Mahmudur Rahman (University of Massachusetts Lowell), Rameswar Panda (MIT-IBM Watson AI Lab), and Mohammad Arif Ul Alam (University of Massachusetts Lowell)</i>	

Cross-Identity Video Motion Retargeting with Joint Transformation and Synthesis	412
<i>Haomiao Ni (Penn State University), Yihao Liu (Johns Hopkins University), Sharon Xiaolei Huang (The Pennsylvania State University), and Yuan Xue (Johns Hopkins University)</i>	
ConfMix: Unsupervised Domain Adaptation for Object Detection via Confidence-Based Mixing ...	423
<i>Giulio Mattolin (University of Trento), Luca Zanella (Fondazione Bruno Kessler), Elisa Ricci (University of Trento), and Yiming Wang (Fondazione Bruno Kessler)</i>	
Improving Diversity with Adversarially Learned Transformations for Domain Generalization	434
<i>Tejas Gokhale (Arizona State University), Rushil Anirudh (Lawrence Livermore National Laboratory), Jayaraman J. Thiagarajan (Lawrence Livermore National Laboratory), Bhavya Kailkhura (Lawrence Livermore National Laboratory), Chitta Baral (Arizona State University), and Yezhou Yang (Arizona State University)</i>	
Learning Across Domains and Devices: Style-Driven Source-Free Domain Adaptation in Clustered Federated Learning	444
<i>Donald Shenaj (University of Padova), Eros Fani (Politecnico Di Torino), Marco Toldo (University of Padova), Debora Caldarola (Politecnico di Torino), Antonio Tavera (Politecnico di Torino), Umberto Michieli (Samsung Research UK), Marco Ciccone (Politecnico di Torino), Pietro Zanuttigh (University of Padova), and Barbara Caputo (Politecnico di Torino)</i>	
CellTranspose: Few-Shot Domain Adaptation for Cellular Instance Segmentation	455
<i>Matthew R Keaton (West Virginia University), Ram J Zaveri (West Virginia University), and Gianfranco Doretto (West Virginia University)</i>	
CUDA-GHR: Controllable Unsupervised Domain Adaptation for Gaze and Head Redirection	467
<i>Swati Jindal (UCSC) and Xin Eric Wang (University of California Santa Cruz)</i>	
Towards Online Domain Adaptive Object Detection	478
<i>Vibashan Vs (Johns Hopkins University), Poojan B Oza (Johns Hopkins University), and Vishal Patel (Johns Hopkins University)</i>	
Domain Adaptive Video Semantic Segmentation via Cross-Domain Moving Object Mixing	489
<i>Kyusik Cho (Yonsei University), Suhyeon Lee (Yonsei University), Hongje Seong (Yonsei University), and Euntai Kim (Yonsei University)</i>	
Empirical Generalization Study: Unsupervised Domain Adaptation vs. Domain Generalization Methods for Semantic Segmentation in the Wild	499
<i>Fabrizio J Piva (Eindhoven University of Technology), Daan De Geus (Eindhoven University of Technology), and Gijs Dubbelman (Eindhoven University of Technology)</i>	
Intra-Source Style Augmentation for Improved Domain Generalization	509
<i>Yumeng Li (Bosch Center for Artificial Intelligence), Dan Zhang (Bosch Center for Artificial Intelligence), Margret Keuper (University of Mannheim), and Anna Khoreva (Bosch Center for Artificial Intelligence)</i>	
TVT: Transferable Vision Transformer for Unsupervised Domain Adaptation	520
<i>Jinyu Yang (Amazon), Jingjing Liu (Apple), Ning Xu (Kuaishou Technology), and Junzhou Huang (University of Texas at Arlington)</i>	

Learning Classifiers of Prototypes and Reciprocal Points for Universal Domain Adaptation	531
<i>Sungsu Hur (KAIST), Inkyu Shin (KAIST), Kwanyong Park (KAIST), Sanghyun Woo (KAIST), and In So Kweon (KAIST)</i>	
Auxiliary Task-Guided CycleGAN for Black-Box Model Domain Adaptation	541
<i>Michael Essich (Cognitive Systems GroupComputer Science DepartmentReutlingen University), Markus Rehmann (Reutlingen University), and Cristobal Curio (Reutlingen University)</i>	

2A: 3D Recognition

NeuralBF: Neural Bilateral Filtering for Top-Down Instance Segmentation on Point Clouds	551
<i>Weiwei Sun (University of British Columbia), Daniel Rebain (Google Inc.), Renjie Liao (University of British Columbia), Vladimir Tankovich (Google), Soroosh Yazdani (Google Inc.), Kwang Moo Yi (University of British Columbia), and Andrea Tagliasacchi (Google Brain and University of Toronto)</i>	
Mobile Robot Manipulation Using Pure Object Detection	561
<i>Brent Griffin (Agility Robotics)</i>	
SGPCR: Spherical Gaussian Point Cloud Representation and Its Application to Object Registration and Retrieval	572
<i>Driton Salihu (Technical University Munich) and Eckehard Steinbach (TUM)</i>	
GalA: Graphical Information Gain Based Attention Network for Weakly Supervised Point Cloud Semantic Segmentation	582
<i>Min Seok Lee (Korea University), Seok Woo Yang (Korea University), and Sung Won Han (Korea University)</i>	
PointInverter: Point Cloud Reconstruction and Editing via a Generative Model with Shape Priors	592
<i>Jaeyeon Kim (HKUST), Binh-Son Hua (VinAI Research), Thanh Nguyen (Deakin UniversityAustralia), and Sai-Kit Yeung (Hong Kong University of Science and Technology)</i>	
3D-SpLineNet: 3D Traffic Line Detection Using Parametric Spline Representations	602
<i>Maximilian M Pittner (Robert Bosch GmbH), Alexandru Condurache (Bosch), and Joel Janai (Robert Bosch GmbH)</i>	
Domain Adaptive Object Detection for Autonomous Driving under Foggy Weather	612
<i>Jinlong Li (Cleveland State University), Runsheng Xu (University of CaliforniaLos Angeles), Jin Ma (Cleveland State University), Qin Zou (Wuhan University), Jiaqi Ma (University of CaliforniaLos Angeles), and Hongkai Yu (Cleveland State University)</i>	
SAILOR: Scaling Anchors via Insights into Latent Object Representation	623
<i>Dušan Malić (Graz University of Technology), Christian Fruhwirth-Reisinger (Graz University of Technology), Horst Possegger (Graz University of Technology), and Horst Bischof (Graz University of Technology)</i>	
Class-Level Confidence Based 3D Semi-Supervised Learning	633
<i>Zhimin Chen (Clemson University), Longlong Jing (Waymo LLC), Yang Liang (City College of New York), Yingwei Li (Johns Hopkins University), and Bing Li (Clemson University)</i>	

MonoEdge: Monocular 3D Object Detection Using Local Perspectives	643
<i>Minghan Zhu (University of MichiganAnn Arbor), Lingting Ge (TuSimple), Panqu Wang (TuSimpleInc), and Huei Peng (University of Michigan)</i>	
Cross-Modality Feature Fusion Network for Few-Shot 3D Point Cloud Classification	653
<i>Minmin Yang (Syracuse University), Jiajing Chen (Syracuse University), and Senem Velipasalar (Syracuse University)</i>	
Dense Voxel Fusion for 3D Object Detection	663
<i>Anas Mahmoud (University of Toronto), Jordan Sir Kwang Hu (University of Toronto), and Steven L Waslander (University of Toronto)</i>	
Real-Time Concealed Weapon Detection on 3D Radar Images for Walk-Through Screening System	673
<i>Nagma S. Khan (NEC), Kazumine Ogura (NEC Corporation), Eric Cosatto (NEC Labs), and Masayuki Ariyoshi (NEC Corporation)</i>	
Resolving Class Imbalance for LiDAR-Based Object Detector by Dynamic Weight Average and Contextual Ground Truth Sampling	682
<i>Daeun Lee (Korea University) and Jinkyu Kim (Korea University)</i>	
Far3Det: Towards Far-Field 3D Detection	692
<i>Shubham Gupta (Carnegie mellon university), Jeet Kanjani (Carnegie Mellon University), Mengtian Li (Carnegie Mellon University), Francesco Ferroni (Argo AI), James Hays (Georgia Institute of TechnologyUSA), Deva Ramanan (Carnegie Mellon University), and Shu Kong (Carnegie Mellon University)</i>	

2B: Image & Scene Synthesis

UVCGAN: UNet Vision Transformer Cycle-Consistent GAN for Unpaired Image-to-Image Translation	702
<i>Dmitrii Torbunov (Brookhaven National Laboratory), Yi Huang (Brookhaven National Laboratory), Haiwang Yu (Brookhaven National Laboratory), Jin Huang (Brookhaven National Lab), Shinjae Yoo (Brookhaven National Laboratory), Meifeng Lin (Brookhaven Natinoal Laboratory), Brett Viren (Brookhaven National Lab), and Yihui Ren (Brookhaven National Laboratory)</i>	
Splatting-Based Synthesis for Video Frame Interpolation	713
<i>Simon Niklaus (Adobe Research), Ping Hu (Boston University), and Jiawen Chen (AdobeInc.)</i>	
CG-NeRF: Conditional Generative Neural Radiance Fields for 3D-Aware Image Synthesis	724
<i>Kyungmin Jo (Korea Advanced Institute of Science and Technology), Gyumin Shim (KAIST), Sanghun Jung (KAIST), Soyoung Yang (KAIST), and Jaegul Choo (Korea Advanced Institute of Science and Technology)</i>	
Spatially Multi-Conditional Image Generation	734
<i>Nikola Popović (ETH Zürich), Ritika Chakraborty (ETHZ), Danda Pani Paudel (ETH Zürich), Thomas Probst (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
WHFL: Wavelet-Domain High Frequency Loss for Sketch-to-Image Translation	744
<i>Min Woo Kim (Seoul National University) and Nam Ik Cho (Seoul National University)</i>	

DDNeRF: Depth Distribution Neural Radiance Fields	755
<i>David Dadon (Reichman University), Ohad Fried (IDC Herzliya), and Yacov Hel-Or (Reichman University)</i>	
Multi-Scale Contrastive Learning for Complex Scene Generation	764
<i>Hanbit Lee (Seoul National University), Youna Kim (Seoul National University), and Sang-Goo Lee (Seoul National University)</i>	
SIRA: Relightable Avatars from a Single Image	775
<i>Pol Caselles Rico (Universitat Politècnica de Catalunya), Eduard Ramon Maldonado (Universitat Politècnica de Catalunya), Jaime Garcia (Crisalix SA), Xavier Giro-I-Nieto (Universitat Politècnica de Catalunya), Francesc Moreno (IRI), and Gil Triginer Garces (Crisalix)</i>	
Learning Graph Variational Autoencoders with Constraints and Structured Priors for Conditional Indoor 3D Scene Generation	785
<i>Aditya Chattopadhyay (Johns Hopkins University), Xi Zhang (Amazon Inc.), David Wipf (Amazon), Himanshu Arora (Amazon), and Rene Vidal (Johns Hopkins UniversityUSA)</i>	
Beyond RGB: Scene-Property Synthesis with Neural Radiance Fields	795
<i>Mingtong Zhang (Tsinghua University), Shuhong Zheng (Peking University), Zhipeng Bao (Carnegie Mellon University), Martial Hebert (Carnegie Mellon School of Computer Science), and Yu-Xiong Wang (University of Illinois at Urbana-Champaign)</i>	
Vision Transformer for NeRF-Based View Synthesis from a Single Input Image	806
<i>Kai-En Lin (University of California San Diego), Yen-Chen Lin (MIT), Wei-Sheng Lai (Google), Tsung-Yi Lin (Nvidia Research), Yichang Shih (Google), and Ravi Ramamoorthi (University of California San Diego)</i>	
ScanNeRF: A Scalable Benchmark for Neural Radiance Fields	816
<i>Luca De Luigi (University of Bologna), Damiano Bolognini (Eyecan.ai), Federico Domeniconi (eyecan.ai), Daniele De Gregorio (eyecan.ai s.r.l.), Matteo Poggi (University of Bologna), and Luigi Di Stefano (University of Bologna)</i>	
Controllable 3D Generative Adversarial Face Model via Disentangling Shape and Appearance	826
<i>Fariborz Taherkhani (Carnegie Mellon University), Aashish Rai (Carnegie Mellon University), Quankai Gao (University of Southern California), Shaunak Srivastava (Carnegie Mellon University), Xuanbai Chen (Carnegie Mellon University), Fernando De La Torre (Carnegie Mellon), Steven L Song (Facebook), Aayush Prakash (Meta), and Daeil Kim (Meta)</i>	
Ev-NeRF: Event Based Neural Radiance Field	837
<i>Inwoo Hwang (Seoul National University), Junho Kim (Seoul National University), and Young Min Kim (Seoul National University)</i>	
Leveraging Off-the-Shelf Diffusion Model for Multi-Attribute Fashion Image Manipulation	848
<i>Chaerin Kong (Seoul National University), Donghyeon Jeon (Naver), Ohjoon Kwon (NAVER), and Nojun Kwak (Seoul National University)</i>	

2C: Biometrics

Creating a Forensic Database of Shoeprints from Online Shoe-Tread Photos	858
<i>Samia Shafique (University of California Irvine), Bailey Kong (Ronin Institute), Shu Kong (Carnegie Mellon University), and Charless Fowlkes (UC Irvine)</i>	
Can Shadows Reveal Biometric Information?	869
<i>Safa Medin (Massachusetts Institute of Technology), Amir Weiss (Massachusetts Institute of Technology), Fredo Durand (MIT), William T. Freeman (MIT), and Gregory W Wornell (MIT)</i>	
Patch-Level Gaze Distribution Prediction for Gaze Following	880
<i>Qiaomu Miao (Stony Brook University), Minh Hoai (Stony Brook University), and Dimitris Samaras (Stony Brook University)</i>	
Searching Efficient Neural Architecture with Multi-Resolution Fusion Transformer for Appearance-Based Gaze Estimation	890
<i>Vikrant Nagpure (Honda Motor Co. Ltd.) and Kenji Okuma (Honda Motor Co. Ltd)</i>	
DeformIrisNet: An Identity-Preserving Model of Iris Texture Deformation	900
<i>Siamul Karim Khan (University of Notre Dame), Patrick Tinsley (University of Notre Dame), and Adam Czajka (University of Notre Dame)</i>	
Gait Recognition Using 3-D Human Body Shape Inference	909
<i>Haidong Zhu (University of Southern California), Zhaoheng Zheng (University of Southern California), and Ram Nevatia (U of Southern California)</i>	
CAST: Conditional Attribute Subsampling Toolkit for Fine-Grained Evaluation	919
<i>Wes J Robbins (University of Colorado Colorado Springs), Steven Zhou (University of Colorado Colorado Springs), Aman Bhatta (University of Notre Dame), Vitor Albiero (University of Notre Dame), Chad Mello (UCCS), Kevin Bowyer (University of Notre Dame), and Terrance Boulton (University of Colorado Colorado Springs)</i>	
Physically Plausible Animation of Human Upper Body from a Single Image	930
<i>Ziyuan Huang (National Taiwan University), Zhengping Zhou (Stanford University), Yung-Yu Chuang (National Taiwan University), Jiajun Wu (Stanford University), and Karen Liu (Stanford)</i>	
Online Adaptive Temporal Memory with Certainty Estimation for Human Trajectory Prediction ..	940
<i>Manh Huynh (UC Denver) and Gita Alaghband (University of Colorado Denver)</i>	
Context-Empowered Visual Attention Prediction in Pedestrian Scenarios	950
<i>Igor Vozniak (DFKI), Philipp Müller (DFKI GmbH), Lorena Hell (German Research Center for Artificial Intelligence), Nils Lipp (DFKI), Ahmed Abouelazm (Research Center for Artificial Intelligence (DFKI)), and Christian Mueller (DFKI)</i>	
Misclassifications of Contact Lens Iris PAD Algorithms: is It Gender Bias or Environmental Conditions?	961
<i>Akshay Agarwal (University at Buffalo), Nalini Ratha (SUNY Buffalo), Afzel Noore (Texas A&M University-Kingsville), Richa Singh (IIT Jodhpur), and Mayank Vatsa (IIT Jodhpur)</i>	

Synthetic Latent Fingerprint Generator	971
<i>Andre Wyzykowski (Michigan State University) and Anil Jain (Michigan State University)</i>	
UPAR: Unified Pedestrian Attribute Recognition and Person Retrieval	981
<i>Andreas Specker (Fraunhofer IOSB), Mickael Cormier (Fraunhofer IOSB Karlsruhe Germany), and Jürgen Beyerer (Fraunhofer IOSB)</i>	
Segmentation-Free Direct Iris Localization Networks	991
<i>Takahiro Toizumi (NEC corporation), Koichi Takahashi (NEC), and Masato Tsukada (University of Tsukuba)</i>	
THOR-Net: End-to-End Graformer-Based Realistic Two Hands and Object Reconstruction with Self-Supervision	1001
<i>Ahmed Tawfik Aboukhadra (DFKI), Jameel Malik (DFKI), Ahmed Elhayek (DFKI), Nadia Robertini (DFKI), and Didier Stricker (DFKI)</i>	

3A: Language & Vision

Fashion Image Retrieval with Text Feedback by Additive Attention Compositional Learning	1011
<i>Yuxin Tian (University of California Merced), Shawn Newsam (UC Merced), and Kofi A Boakye (Pinterest)</i>	
Cross-Modal Semantic Enhanced Interaction for Image-Sentence Retrieval	1022
<i>Xuri Ge (University of Glasgow), Fuhai Chen (Xiamen university), Songpei Xu (University of Glasgow), Fuxiang Tao (University of Glasgow), and Joemon M Jose (University of Glasgow)</i>	
Text-Guided Object Detector for Multi-Modal Video Question Answering	1032
<i>Ruoyue Shen (Tokyo Institute of Technology), Nakamasa Inoue (Tokyo Institute of Technology), and Koichi Shinoda (Tokyo Institute of Technology)</i>	
DRAMA: Joint Risk Localization and Captioning in Driving	1043
<i>Srikanth Malla (Honda Research Institute), Chiho Choi (Honda Research Institute US), Isht Dwivedi (Honda Research Institute USA), Joon Hee Choi (Samsung Semiconductor Inc), and Jiachen Li (Stanford University)</i>	
Interactive Image Manipulation with Complex Text Instructions	1053
<i>Ryugo Morita (Hosei University), Zhiqiang Zhang (Hosei University), Man M. Ho (Hosei University), and Jinjia Zhou (Hosei University)</i>	
InDiReCT: Language-Guided Zero-Shot Deep Metric Learning for Images	1063
<i>Konstantin Kobs (Universität Würzburg), Michael Steininger (University of Würzburg), and Andreas Hotho (Universität Würzburg)</i>	
Learning by Hallucinating: Vision-Language Pre-Training with Weak Supervision	1073
<i>Tzu-Jui Wang (Aalto University), Tomas Langer (Intuition Machines), Jorma Laaksonen (Aalto University), Heikki Arponen (Intuition Machines Inc.), and Bishop E Thomas (Intuition Machines)</i>	
Barlow Constrained Optimization for Visual Question Answering	1084
<i>Abhishek Jha (K.U. Leuven), Badri Narayana Patro (Microsoft), Luc Van Gool (KU Leuven), and Tinne Tuytelaars (KU Leuven)</i>	

A Priority Map for Vision-and-Language Navigation with Trajectory Plans and Feature-Location Cues	1094
<i>Jason Armitage (University of Zurich), Leonardo Impett (Cambridge University), and Rico Sennrich (University of Zurich)</i>	
Structure-Encoding Auxiliary Tasks for Improved Visual Representation in Vision-and-Language Navigation	1104
<i>Chia-Wen Kuo (Georgia Institute of Technology), Chih-Yao Ma (Facebook), Judy Hoffman (Georgia Tech), and Zsolt Kira (Georgia Institute of Technology)</i>	
Dense but Efficient VideoQA for Intricate Compositional Reasoning	1114
<i>Jihyeon Lee (Kakao Brain), Woo Young Kang (Kakao Brain), and Eun-Sol Kim (Hanyang University)</i>	
Switching to Discriminative Image Captioning by Relieving a Bottleneck of Reinforcement Learning	1124
<i>Ukyo Honda (Nara Institute of Science and Technology / RIKEN Center for Advanced Intelligence Project), Taro Watanabe (Nara Institute of Science and Technology), and Yuji Matsumoto (RIKEN Center for Advanced Intelligence Project)</i>	
NAPReg: Nouns as Proxies Regularization for Semantically Aware Cross-Modal Embeddings	1135
<i>Bhavin Jawade (University at Buffalo), Deen D Mohan (University at Buffalo), Naji Mohamed Ali (University at Buffalo), Srirangaraj Setlur (University at BuffaloSUNY), and Venu Govindaraju (University at BuffaloSUNY)</i>	
Image-Text Pre-Training for Logo Recognition	1145
<i>Mark Hubenthal (Amazon Inc.) and Suren Kumar (Amazon Inc.)</i>	
VLC-BERT: Visual Question Answering with Contextualized Commonsense Knowledge	1155
<i>Sahithya Ravi (University of British Columbia), Aditya Chinchure (University of British Columbia), Leonid Sigal (University of British Columbia), Renjie Liao (University of British Columbia), and Vered Shwartz (University of British Columbia)</i>	

3B: 3D Vision

TVCalib: Camera Calibration for Sports Field Registration in Soccer	1166
<i>Jonas Theiner (L3S Research CenterLeibniz Universität Hannover) and Ralph Everth (TIB - Leibniz Information Center for Science and Technology)</i>	
3D Change Localization and Captioning from Dynamic Scans of Indoor Scenes	1176
<i>Yue Qiu (National Institute of Advanced Industrial Science and Technology (AIST)), Shintaro Yamamoto (Toshiba), Ryosuke Yamada (National Institute of Advanced Industrial Science and Technology (AIST); Tokyo Denki University (TDU)), Ryota Suzuki (Saitama University), Hirokatsu Kataoka (National Institute of Advanced Industrial Science and Technology (AIST)), Kenji Iwata (National Institute of Advanced Industrial Science and Technology (AIST)), and Yutaka Satoh (National Institute of Advanced Industrial Science and Technology (AIST))</i>	

Adaptive Feature Fusion for Cooperative Perception Using LiDAR Point Clouds	1186
<i>Donghao Qiao (Queen's University) and Farhana H Zulkernine (Queen's University)</i>	
Centroid Distance Keypoint Detector for Colored Point Clouds	1196
<i>Hanzhe Teng (University of CaliforniaRiverside), Dimitrios Chatziparaschis (University of CaliforniaRiverside), Xinyue Kan (University of CaliforniaRiverside), Amit K. Roy-Chowdhury (University of CaliforniaRiverside), and Konstantinos Karydis (University of California Riverside)</i>	
PP4AV: A Benchmarking Dataset for Privacy-Preserving Autonomous Driving	1206
<i>Linh Khac Trinh (VinFast), Phuong Pham (VinFast), Hoang Trinh (VinFast), Nguyen Bach (VinFast), Dung T Nguyen (FPT SoftwareAIC), Giang Duy Nguyen (VinFast), and Huy Q Nguyen (VinFast)</i>	
Self-Supervised Correspondence Estimation via Multiview Registration	1216
<i>Mohamed El Banani (University of Michigan), Ignacio Rocco (Facebook AI Research), David Novotny (Facebook AI Research), Andrea Vedaldi (University of Oxford / Facebook AI Research), Natalia Neverova (Facebook AI Research), Justin Johnson (University of Michigan), and Ben Graham (Facebook Research)</i>	
Seg&Struct: The Interplay between Part Segmentation and Structure Inference for 3D Shape Parsing	1226
<i>Jeonghyun Kim (KAIST), Kaichun Mo (NVIDIA), Minhyuk Sung (KAIST), and Woontack Woo (KAIST)</i>	
Compressing Explicit Voxel Grid Representations: Fast NeRFs Become Also Small	1236
<i>Chenxi Deng (Telecom Paris) and Enzo Tartaglione (T��l��com Paris - Institut Polytechnique de Paris)</i>	
Nearest Neighbors Meet Deep Neural Networks for Point Cloud Analysis	1246
<i>Renrui Zhang (Shanghai AI Lab), Liuhui Wang (Peking University), Ziyu Guo (Peking University), and Jianbo Shi (University of Pennsylvania)</i>	
Generative Range Imaging for Learning Scene Priors of 3D LiDAR Data	1256
<i>Kazuto Nakashima (Kyushu University), Yumi Iwashita (NASA Jet Propulsion Laboratory), and Ryo Kurazume (Kyushu University)</i>	
Rebalancing Gradient to Improve Self-Supervised Co-Training of Depth, Odometry and Optical Flow Predictions	1267
<i>Marwane Hariat (N/A), Antoine Manzanera (ENSTA Paris), and David Filliat (ENSTA Paris)</i>	
RSF: Optimizing Rigid Scene Flow from 3D Point Clouds without Labels	1277
<i>David Deng (UC Berkeley) and Avidesh Zakhori (UC Berkeley)</i>	
Improving the Robustness of Point Convolution on K-Nearest Neighbor Neighborhoods with a Viewpoint-Invariant Coordinate Transform	1287
<i>Xingyi Li (Oregon State University), Wenxuan Wu (Oregon State University), Xiaoli Fern (Oregon State University), and Li Fuxin (Oregon State University)</i>	
PIDS: Joint Point Interaction-Dimension Search for 3D Point Cloud	1298
<i>Tunhou Zhang (Duke University), Mingyuan Ma (Duke University), Feng Yan (University of Houston), Hai Li (Duke University), and Yiran Chen (Duke University)</i>	

3C: Adversarial Learning, Attacks, Defenses, and Deepfakes

Leveraging Local Patch Differences in Multi-Object Scenes for Generative Adversarial Attacks	1308
<i>Abhishek Aich (University of California, Riverside), Shasha Li (University of California, Riverside), M. Salman Asif (University of California, Riverside), Chengyu Song (University of California, Riverside), Srikanth V. Krishnamurthy (University of California, Riverside), and Amit K. Roy-Chowdhury (University of California, Riverside)</i>	
Motif Mining: Finding and Summarizing Remixed Image Content	1319
<i>William Theisen (University of Notre Dame), Daniel Gonzalez Cedre (University of Notre Dame), Zachariah Carmichael (University of Notre Dame), Daniel Moreira (University of Notre Dame), Tim Weninger (University of Notre Dame), and Walter Scheirer (University of Notre Dame)</i>	
DeepPrivacy2: Towards Realistic Full-Body Anonymization	1329
<i>Håkon Hukkelås (Norwegian University of Science and Technology) and Frank Lindseth (NTNU)</i>	
A Continual Deepfake Detection Benchmark: Dataset, Methods, and Essentials	1339
<i>Chuangqiao Li (ETH Zurich), Zhiwu Huang (Singapore Management University), Danda Pani Paudel (ETH Zürich), Yabin Wang (Xi'an Jiaotong University), Mohamad Shahbazi (ETH Zurich), Xiaopeng Hong (Harbin Institute of Technology), and Luc Van Gool (ETH Zurich)</i>	
Task Agnostic and Post-Hoc Unseen Distribution Detection	1350
<i>Radhika Dua (KAIST), Seongjun Yang (KAIST), Yixuan Li (University of Wisconsin-Madison), and Edward Choi (KAIST)</i>	
Closer Look at the Transferability of Adversarial Examples: How They Fool Different Models Differently	1360
<i>Futa Kai Waseda (The University of Tokyo), Sosuke Nishikawa (The University of Tokyo), Trung-Nghia Le (University of ScienceVNU-HCM), Huy Hong Nguyen (National Institute of Informatics), and Isao Echizen (National Institute of Informatics)</i>	
My Face My Choice: Privacy Enhancing Deepfakes for Social Media Anonymization	1369
<i>Umur A Ciftci (State University of New York at Binghamton), Gokturk Yuksek (State University of New York at Binghamton), and Ilke Demir (Intel Corporation)</i>	
Do Adaptive Active Attacks Pose Greater Risk Than Static Attacks?	1380
<i>Nathan Drenkow (Johns Hopkins University Applied Physics Laboratory), Max Lennon (JHU/APL), I-Jeng Wang (Johns Hopkins University), and Philippe Burlina (JHU/APL/CS/SOM)</i>	

4A: Neural Architecture Search, Representation Learning, and Explainable Models

Neural Weight Search for Scalable Task Incremental Learning	1390
<i>Jian Jiang (King's college london) and Oya Celiktutan (King's College London)</i>	
Toward Edge-Efficient Dense Predictions with Synergistic Multi-Task Neural Architecture Search	1400
<i>Thanh M Vu (University of North Carolina at Chapel Hill), Yanqi Zhou (Google), Chunfeng Wen (Google), Yueqi Li (X, The Moonshot Factory), and Jan-Michael Frahm (UNC-Chapel Hill)</i>	
Addressing Feature Suppression in Unsupervised Visual Representations	1411
<i>Tianhong Li (MIT), Lijie Fan (MIT), Yuan Yuan (MIT), Hao He (Massachusetts Institute of Technology), Yonglong Tian (MIT), Rogerio Feris (MIT-IBM Watson AI Lab/IBM Research), Piotr Indyk (MIT), and Dina Katabi (Massachusetts Institute of Technology)</i>	
A Protocol for Evaluating Model Interpretation Methods from Visual Explanations	1421
<i>Hamed Behzadi-Khormouji (University of Antwerp/pimec-IDLab) and Jose Oramas (University of Antwerp/pimec-IDLab)</i>	
Realistic Full-Body Anonymization with Surface-Guided GANs	1430
<i>Håkon Hukkelås (Norwegian University of Science and Technology), Morten Smebye (Capra Consulting AS), Rudolf Mester (Norwegian University of Science and Technology), and Frank Lindseth (NTNU)</i>	
Global-Local Self-Distillation for Visual Representation Learning	1441
<i>Tim Lebaillly (KU Leuven) and Tinne Tuytelaars (KU Leuven)</i>	
Large-to-Small Image Resolution Asymmetry in Deep Metric Learning	1451
<i>Pavel Suma (Czech Technical University in Prague) and Giorgos Tolias (Czech Technical University in Prague)</i>	
Learning How to MIMIC: Using Model Explanations to Guide Deep Learning Training	1461
<i>Matthew S Watson (Durham University), Bashar Awwad Shiekh Hasan (Amazon), and Noura Al Moubayed (University of Durham)</i>	
The Box Size Confidence Bias Harms Your Object Detector	1471
<i>Johannes Gilg (Technical University of Munich), Torben Teepe (Technical University of Munich), Fabian Herzog (Technical University of Munich), and Gerhard Rigoll (Institute for Human-Machine Communication/TU Munich/Germany)</i>	
ProtoSeg: Interpretable Semantic Segmentation with Prototypical Parts	1481
<i>Mikołaj Sacha (Jagiellonian University), Dawid Damian Rymarczyk (Jagiellonian University), Łukasz Struski (Jagiellonian University), Jacek Tabor (Jagiellonian University), and Bartosz Zieliński (Jagiellonian University)</i>	
FreeREA: Training-Free Evolution-Based Architecture Search	1493
<i>Niccolò Cavagnero (Politecnico di Torino), Luca Robbiano (Politecnico di Torino), Barbara Caputo (Politecnico di Torino), and Giuseppe Averta (Politecnico di Torino)</i>	
SVD-NAS: Coupling Low-Rank Approximation and Neural Architecture Search	1503
<i>Zheven Yu (Imperial College London) and Christos-Savvas Bouganis (Imperial College London)</i>	

Visually Explaining 3D-CNN Predictions for Video Classification with an Adaptive Occlusion Sensitivity Analysis	1513
<i>Tomoki Uchiyama (University of Tsukuba), Naoya Sogi (NEC corporation), Koichiro Niinuma (FUJITSU RESEARCH OF AMERICA), and Kazuhiro Fukui (University of Tsukuba)</i>	
Orthogonal Transforms for Learning Invariant Representations in Equivariant Neural Networks	1523
<i>Jaspreet Singh (Punjabi University Patiala), Chandan Singh (punjabi university patiala), and Ankur Rana (Punjabi University)</i>	
Representation Disentanglement in Generative Models with Contrastive Learning	1531
<i>Shentong Mo (Carnegie Mellon University), Zhun Sun (Tohoku University), and Chao Li (RIKEN)</i>	
Calibrating Deep Neural Networks Using Explicit Regularisation and Dynamic Data Pruning	1541
<i>Rishabh Patra (APPCAIRBITS Pilani KK Birla Goa Campus), Ramya Hebbalaguppe (TCS Research), Tirtharaj Dash (BITS PilaniGoa Campus), Gautam Shroff (Tata Consultancy Services Ltd.), and Lovekesh Vig (Innovation LabsTata Consultancy Services Limited)</i>	
Patch-Based Privacy Preserving Neural Network for Vision Tasks	1550
<i>Mitsuhiro Mabuchi (Toyota Motor Corporation) and Tetsuya Ishikawa (Woven Core)</i>	

4B: Tracking & Reidentification

PreViTS: Contrastive Pretraining with Video Tracking Supervision	1560
<i>Brian Chen (Columbia University), Ramprasaath R. Selvaraju (Salesforce Research), Shih-Fu Chang (Columbia University), Juan Carlos Niebles (Salesforce & Stanford University), and Nikhil Naik (MIT)</i>	
Efficient Visual Tracking with Exemplar Transformers	1571
<i>Philippe Blatter (ETH Zurich), Menelaos Kanakis (ETH Zurich), Martin Danelljan (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
Multi-View Tracking Using Weakly Supervised Human Motion Prediction	1582
<i>Martin Engilberge (EPFL), Weizhe Liu (EPFL), and Pascal Fua (EPFLSwitzerland)</i>	
Planar Object Tracking via Weighted Optical Flow	1593
<i>Jonáš Šerých (CTU in Prague) and Jiri Matas (Czech Technical UniversityPrague)</i>	
Feature Disentanglement Learning with Switching and Aggregation for Video-Based Person Re-Identification	1603
<i>Minjung Kim (Yonsei University), Myeongah Cho (Yonsei University), and Sangyoun Lee (Yonsei University)</i>	
Body Part-Based Representation Learning for Occluded Person Re-Identification	1613
<i>Vladimir Somers (UCLouvain), Christophe De Vleeschouwer (Université Catholique de Louvain), and Alexandre Alahi (EPFL)</i>	

Camera Alignment and Weighted Contrastive Learning for Domain Adaptation in Video Person ReID	1624
<i>Djebril Mekhazni (Ecole de Technologie Supérieure), Maximilien F Dufau (Efrei Paris), Christian Desrosiers (ETSCanada), Marco Pedersoli (Ecole de technologie supérieure), and Eric Granger (ETS Montreal)</i>	
MEVID: Multi-View Extended Videos with Identities for Video Person Re-Identification	1634
<i>Daniel Davila (Kitware Inc), Dawei Du (KitwareInc.), Bryon C Lewis (Kitware), Christopher Funk (KitwareInc), Joseph B Van Pelt (KitwareInc.), Roderic Collins (Kitware Inc.), Kellie N Corona (Kitware Inc.), Matt S Brown (KitwareInc.), Scott McCloskey (Kitware), Anthony Hoogs (Kitware), and Brian S Clipp (Kitware)</i>	
Unsupervised 4D LiDAR Moving Object Segmentation in Stationary Settings with Multivariate Occupancy Time Series	1644
<i>Thomas Kreutz (TU Darmstadt), Max Mühlhäuser (Technische Universität Darmstadt), and Alejandro Sanchez Guinea (TU Darmstadt)</i>	
AttTrack: Online Deep Attention Transfer for Multi-Object Tracking	1654
<i>Keivan Nalaie (McMaster University) and Rong Zheng (McMaster University)</i>	
Multi-Frame Attention with Feature-Level Warping for Drone Crowd Tracking	1664
<i>Takanori Asanomi (Kyushu University), Kazuya Nishimura (Kyushu University), and Ryoma Bise (Kyushu University)</i>	
BURST: A Benchmark for Unifying Object Recognition, Segmentation and Tracking in Video	1674
<i>Ali Athar (RWTH Aachen), Jonathon Luiten (RWTH Aachen University), Paul Voigtlaender (Google), Tarasha Khurana, Achal Dave (Carnegie Mellon University), Bastian Leibe (RWTH Aachen University), and Deva Ramanan (Carnegie Mellon University)</i>	
Gallery Filter Network for Person Search	1684
<i>Lucas W Jaffe (UC Berkeley) and Avidesh Zakhori (University of California Berkeley)</i>	

4C: Enhancement, Restoration, and Super-Resolution

HIME: Efficient Headshot Image Super-Resolution with Multiple Exemplars	1694
<i>Xiaoyu Xiang (Meta Platforms Inc.), Jon Morton (Facebook), Fitsum Reda (Google), Lucas D Young (Facebook), Federico Perazzi (FacebookInc.), Rakesh Ranjan (Facebook), Amit Kumar (Facebook), Andrea Colaco (Google), and Jan Allebach (Purdue University)</i>	
Fine-Context Shadow Detection Using Shadow Removal	1705
<i>Jeya Maria Jose Valanarasu (Johns Hopkins University) and Vishal Patel (Johns Hopkins University)</i>	
Robust Real-World Image Enhancement Based on Multi-Exposure LDR Images	1715
<i>Haoyu Ren (Zeku Inc), Yi Fan (Zeku Inc), and Hsilin Huang (Zeku)</i>	

Pik-Fix: Restoring and Colorizing Old Photos	1724
<i>Runsheng Xu (University of CaliforniaLos Angeles), Zhengzhong Tu (University of Texas at Austin), Yuanqi Du (Cornell University), Xiaoyu Dong (Northwestern University), Jinlong Li (Cleveland State University), Zibo Meng (Innopeak Technology), Jiaqi Ma (University of CaliforniaLos Angeles), Alan Bovik (University of Texas at Austin), and Hongkai Yu (Cleveland State University)</i>	
Fast Online Video Super-Resolution with Deformable Attention Pyramid	1735
<i>Dario Fuoli (ETH Zurich), Martin Danelljan (ETH Zurich), Radu Timofte (University of Wurzburg & ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
Style-Guided Inference of Transformer for High-Resolution Image Synthesis	1745
<i>Jonghwa Yim (Vision AI LabAI CenterNCSoft) and Minjae Kim (Vision AI LabAI CenterNCSoft)</i>	
PSENet: Progressive Self-Enhancement Network for Unsupervised Extreme-Light Image Enhancement	1756
<i>Hue Nguyen (VinAI Research), Diep Tran (VinAI Research), Khoi Nguyen (VinAI Research), and Rang Nguyen (VinAI Research)</i>	
Cross-Resolution Flow Propagation for Foveated Video Super-Resolution	1766
<i>Eugene Eu Tzuan Lee (National Yang Ming Chiao Tung University), Lien-Feng Hsu (National Chiao Tung University), Evan Chen (National Chiao Tung University), and Cy Lee (National Chiao Tung University)</i>	
GeoFill: Reference-Based Image Inpainting with Better Geometric Understanding	1776
<i>Yunhan Zhao (University of CaliforniaIrvine), Connelly Barnes (Adobe), Yuqian Zhou (Adobe), Eli Shechtman (Adobe ResearchUS), Sohrab Amirghodsi (Adobe Research), and Charless Fowlkes (UC Irvine)</i>	
iColoriT: Towards Propagating Local Hints to the Right Region in Interactive Colorization by Leveraging Vision Transformer	1787
<i>Jooyeol Yun (Korea Advanced Institute of Science and Technology), Sanghyeon Lee (Korea Advanced Institute of Science and Technology), Minho Park (Korea Advanced Institute of Science and Technology), and Jaegul Choo (Korea Advanced Institute of Science and Technology)</i>	
Deep Model-Based Super-Resolution with Non-Uniform Blur	1797
<i>Charles Laroche (GoPro), Andres Almansa (CNRS & Université de Paris), and Matias Tassano (Meta Inc)</i>	
SHARDS: Efficient Shadow Removal Using Dual Stage Network for High-Resolution Images	1809
<i>Mrinmoy Sen (Samsung R&D Institute India - Bangalore), Sai Pradyumna Chermala (Samsung R&D Institute IndiaBangalore), Nazrinbanu Nurmohammad Nagori (Samsung Research India Bangalore Pvt Ltd), Venkat Peddigari (Samsung Research Institute Bangalore), Praful Mathur (Samsung R&D Institute Bangalore), B H Pawan Prasad (Samsung Research), and Moonhwan Jeong (Samsung Electronics)</i>	
Guiding Users to Where to Give Color Hints for Efficient Interactive Sketch Colorization via Unsupervised Region Prioritization	1818
<i>Younghin Cho (Korea Advanced Institute of Science and Technology), Junsoo Lee (NAVER WEBTOON Ltd.), Soyoung Yang (KAIST), Juntae Kim (Korea University), Yeojeong Park (KAIST), Haneol Lee (UNIST), Mohammad Azam Khan (KAIST), Daesik Kim (Naver webtoon), and Jaegul Choo (Korea Advanced Institute of Science and Technology)</i>	

Efficient Reference-Based Video Super-Resolution (ERVSR): Single Reference Image is All You Need	1828
<i>Youngrae Kim (KAIST), Jinsu Lim (KAIST), Hoonhee Cho (KAIST), Minji Lee (KAIST), Dongman Lee (KAIST), Kuk-Jin Yoon (KAIST), and Ho-Jin Choi (KAIST)</i>	
Efficient Flow-Guided Multi-Frame De-Fencing	1838
<i>Stavros Tsogkas (Samsung AI Centre Toronto), Fengjia Zhang (Samsung), Allan D Jepson (Samsung Toronto AIC), and Alex Levinshtein (Samsung AI Centre Toronto)</i>	
Perceptual Image Enhancement for Smartphone Real-Time Applications	1848
<i>Marcos V. Conde (University of Würzburg), Florin-Alexandru Vasluianu (Computer Vision LabUniversity of Würzburg), Javier Vazquez-Corral (Autonomous University of Barcelona), and Radu Timofte (University of Würzburg & ETH Zurich)</i>	

5A: Computer Vision for Healthcare

Expert-Defined Keywords Improve Interpretability of Retinal Image Captioning	1859
<i>Ting-Wei Wu (Georgia Institute of Technology), Jia-Hong Huang (University of Amsterdam), Joseph Lin (UCLA), and Marcel Worring (University of Amsterdam)</i>	
Diffeomorphic Image Registration with Neural Velocity Field	1869
<i>Kun Han (University of California Irvine), Shanlin Sun (University of CaliforniaIrvine), Xiangyi Yan (University of CaliforniaIrvine), Chenyu You (Yale University), Hao Tang (University of California Irvine), Junayed A Naushad (UCI), Haoyu Ma (University of CaliforniaIrvine), Deying Kong (University of California Irvine), and Xiaohui Xie (University of CaliforniaIrvine)</i>	
ATCON: Attention Consistency for Vision Models	1880
<i>Ali Mirzazadeh (Georgia Tech), Florian Dubost (Stanford University), Maxwell Pike (Stanford University), Krish Maniar (Stanford), Max Zuo (Georgia Institute of Technology), Christopher Lee-Messer (Stanford University), and Daniel Rubin (missing)</i>	
Semi-Supervised Learning for Sparsely-Labeled Sequential Data: Application to Healthcare Video Processing	1890
<i>Florian Dubost (Stanford University), Erin Hong (Stanford University), Siyi Tang (Stanford University), Nandita Bhaskhar (Stanford University), Christopher Lee-Messer (Stanford University), and Daniel Rubin (missing)</i>	
Analysis of Master Vein Attacks on Finger Vein Recognition Systems	1900
<i>Huy Hong Nguyen (National Institute of Informatics), Trung-Nghia Le (University of ScienceVNU-HCM), Junichi Yamagishi (National Institute of Informatics), and Isao Echizen (National Institute of Informatics)</i>	

Computer Vision to the Rescue: Infant Postural Symmetry Estimation from Incongruent Annotations	1909
<i>Xiaofei Huang (Northeastern University), Michael Wan (The Roux Institute at Northeastern University), Lingfei Luan (Northeastern University), Bethany Tunik (Independent Physical Therapist), and Sarah Ostadabbas (Northeastern University)</i>	
VSGD-Net: Virtual Staining Guided Melanocyte Detection on Histopathological Images	1918
<i>Kechun Liu (University of Washington), Beibin Li (Microsoft Research), Wenjun Wu (University of Washington), Caitlin May (Dermatopathology Northwest), Oliver Chang (University of Washington), Stevan Knezevich (Pathology Associates), Lisa Reisch (University of Washington), Joann Elmore (University of California Los Angeles), and Linda Shapiro (University of Washington)</i>	
PINER: Prior-Informed Implicit Neural Representation Learning for Test-Time Adaptation in Sparse-View CT Reconstruction	1928
<i>Bowen Song (Stanford University), Liye Shen (Stanford University), and Lei Xing (Stanford University)</i>	
Robust and Efficient Alignment of Calcium Imaging Data through Simultaneous Low Rank and Sparse Decomposition	1938
<i>Junmo Cho (KAIST), Seungjae Han (KAIST), Eun-Seo Cho (KAIST), Kijung Shin (KAIST), and Young-Gyu Yoon (KAIST)</i>	
MRI Imputation Based on Fused Index- and Intensity-Registration	1948
<i>Jiyeon Shin (Seoul National University) and Jungwoo Lee (Seoul National University)</i>	
DBCE: A Saliency Method for Medical Deep Learning through Anatomically-Consistent Free-Form Deformations	1958
<i>Joshua Peters (The University of Queensland), Leo Lebrat (CSIRO), Rodrigo Santa Cruz (CSIRO), Aaron M Nicolson (CSIRO), Gregg R Belous (CSIRO), Salamata Konate (CSIRO), Parnesh Raniga (Australian e-Health Research Centre), Vincent Dore (CSIRO), Pierrick Bourgeat (CSIRO), Jurgen Fripp (Australian e-Health Research Centre), Clinton Fookes (Queensland University of Technology), and Olivier Salvado (CSIRO)</i>	
Masked Image Modeling Advances 3D Medical Image Analysis	1969
<i>Zekai Chen (Bristol Myers Squibb), Devansh Agarwal (Bristol Myers Squibb), Kshitij Aggarwal (Bristol Myers Squibb), Wiem Safta (Bristol Myers Squibb), Mariann Micsinai-Balan (BMS), and Kevin Brown (Bristol Myers Squibb)</i>	
EmbryosFormer: Deformable Transformer and Collaborative Encoding-Decoding for Embryos Stage Development Classification	1980
<i>Tien-Phat Nguyen (University of Science VNU-HCM), Trong-Thang Pham (ohn von Neumann Institute Ho Chi Minh City), Tri C Nguyen (HOPE Research Center My Duc Hospital Ho Chi Minh City), Hieu Le Xuan (Australian National University), Dung P. Nguyen (My Duc Phu Nhuan Hospital), Hau Thi My Lam (Olea Fertility Vinmec Central Park International Hospital), Phong Nguyen (FPT Software), Jennifer M Fowler (Arkansas NSF EPSCoR), Minh-Triet Tran (University of Science VNU-HCM), and Ngan Le (University of Arkansas)</i>	

Performer: A Novel PPG-to-ECG Reconstruction Transformer for a Digital Biomarker of Cardiovascular Disease Detection	1990
<i>Ella Lan (The Harker School)</i>	
A Morphology Focused Diffusion Probabilistic Model for Synthesis of Histopathology Images	1999
<i>Puria Azadi Moghadam (The University of British Columbia), Sanne Van Dalen (Eindhoven University of Technology), Karina Martin (The University of British Columbia), Jochen Lennerz (Massachusetts General Hospital/Harvard Medical School), Stephen Yip (BC Cancer Agency), Hossein Farahani (The University of British Columbia), and Ali Bashashati (The University of British Columbia)</i>	

5B: Video Analysis & Optical Flow

Contrastive Losses Are Natural Criteria for Unsupervised Video Summarization	2009
<i>Zongshang Pang (Osaka University), Yuta Nakashima (Osaka University), Mayu Otani (CyberAgent), and Hajime Nagahara (Osaka University)</i>	
M-FUSE: Multi-Frame Fusion for Scene Flow Estimation	2019
<i>Lukas Mehl (University of Stuttgart), Azin Jahedi (Universität Stuttgart), Jenny Schmalfluss (University of Stuttgart), and Andrés Bruhn (University of Stuttgart)</i>	
BoxMask: Revisiting Bounding Box Supervision for Video Object Detection	2029
<i>Khurram Azeem Hashmi (Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH), A. A. Pagani (DFKI), Didier Stricker (DFKI), and Muhammad Zeshan Afzal (Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH)</i>	
Lightweight Network for Video Motion Magnification	2040
<i>Jasdeep Singh (cvpr labiiit ropar), Subrahmanyam Murala (IIT Ropar), and G Sankara Raju Kosuru (IIT Ropar)</i>	
TinyHD: Efficient Video Saliency Prediction with Heterogeneous Decoders Using Hierarchical Maps Distillation	2050
<i>Feiyan Hu (Dublin City University), Simone Palazzo (University of Catania), Federica Proietto Salanitri (University of Catania), Giovanni Bellitto (University of Catania), Morteza Moradi (University of Catania), Concetto Spampinato (University of Catania), and Kevin McGuinness (DCU)</i>	
BrightFlow: Brightness-Change-Aware Unsupervised Learning of Optical Flow	2060
<i>Rémi Marsal (CEA), Florian Chabot (CEA), Angélique Loesch (CEA LIST), and Hichem Sahbi (Sorbonne University)</i>	
FLAVR: Flow-Agnostic Video Representations for Fast Frame Interpolation	2070
<i>Tarun Kalluri (UC San Diego), Deepak Pathak (Carnegie Mellon University), Manmohan Chandraker (UC San Diego), and Du Tran (Facebook AI)</i>	
MovieCLIP: Visual Scene Recognition in Movies	2082
<i>Digbalay Bose (University of Southern California), Rajat Hebbar (University of Southern California), Krishna Somandepalli (Google Research), Haoyang Zhang (University of Southern California), Yin Cui (Google), Kree Cole-Mclaughlin (Google), Huisheng Wang (Google), and Shrikanth Narayanan (USC)</i>	

Neural Implicit Representations for Physical Parameter Inference from a Single Video	2092
<i>Florian Hoffherr (Technical University of Munich), Lukas Koestler (Technical University of Munich), Florian Bernard (University of Bonn), and Daniel Cremers (TU Munich)</i>	
Improving Saliency Models' Predictions of the Next Fixation with Humans' Intrinsic Cost of Gaze Shifts	2103
<i>Florian Kadner (Technical University Darmstadt), Tobias Thomas (Technical University Darmstadt), David Hoppe (Technical University Darmstadt), and Constantin Rothkopf (TU Darmstadt)</i>	
Match Cutting: Finding Cuts with Smooth Visual Transitions	2114
<i>Boris Chen (Netflix), Amir Ziai (Netflix), Rebecca S. Tucker (Netflix), and Yuchen Xie (Netflix)</i>	
TTTFlow: Unsupervised Test-Time Training with Normalizing Flow	2125
<i>David Osowiecki (Ecole de Technologie Supérieure), Gustavo A Vargas Hakim (École de Technologie Supérieure), Mehrdad Noori (École de Technologie Supérieure), Milad Cheraghalikhani (École de Technologie Supérieure), Ismail Ben Ayed (ETS Montreal), and Christian Desrosiers (ETSCanada)</i>	
Weakly-Supervised Optical Flow Estimation for Time-of-Flight	2134
<i>Michael Schelling (Ulm University - Institute of Media Informatics), Pedro Hermosilla Casajus (Ulm University), and Timo Ropinski (Ulm University)</i>	
Meta-Learning for Adaptation of Deep Optical Flow Networks	2144
<i>Chaerin Min (Hanyang University), Tae Hyun Kim (Hanyang Univeristy), and Jongwoo Lim (Hanyang University)</i>	
Fine-Grained Affordance Annotation for Egocentric Hand-Object Interaction Videos	2154
<i>Zecheng Yu (The University of Tokyo), Yifei Huang (The University of Tokyo), Ryosuke Furuta (The University of Tokyo), Takuma Yagi (The University of Tokyo), Yusuke Goutsu (The University of Tokyo), and Yoichi Sato (University of Tokyo)</i>	

5C: Vision, Audio, and Other Modalities

Dissecting Deep Metric Learning Losses for Image-Text Retrieval	2163
<i>Hong Xuan (Microsoft) and Xi Stephen Chen (Microsoft)</i>	
Content-Based Music-Image Retrieval Using Self- and Cross-Modal Feature Embedding Memory	2173
<i>Takayuki Nakatsuka (National Institute of Advanced Industrial Science and Technology (AIST)), Masahiro Hamasaki (National Institute of Advanced Industrial Science and Technology (AIST)), and Masataka Goto (National Institute of Advanced Industrial Science and Technology (AIST))</i>	
Complementary Cues from Audio Help Combat Noise in Weakly-Supervised Object Detection ...	2184
<i>Cagri Gungor (University of Pittsburgh) and Adriana Kovashka (University of Pittsburgh)</i>	

Multimodal Multi-Head Convolutional Attention with Various Kernel Sizes for Medical Image Super-Resolution	2194
<i>Mariana-Iuliana Georgescu (University of Bucharest), Radu Tudor Ionescu (University of Bucharest), Andreea Iuliana D Miron (University of Medicine and Pharmacy Carol Davila), Olivian Savencu (University of Medicine and Pharmacy Carol Davila), Nicolae Catalin Ristea (University Politehnica of Bucharest), Nicolae Verga (Carol Davila University of Medicine and Pharmacy), and Fahad Shahbaz Khan (MBZUAI)</i>	
AudioViewer: Learning to Visualize Sounds	2205
<i>Chunjin Song (UBC), Yuchi Zhang (The University of British Columbia), Willis Peng (The University of British Columbia), Parmis Mohaghegh (University of British Columbia), Bastian Wandt (University of British Columbia), and Helge Rhodin (UBC)</i>	
Towards MOOCs for Lipreading: Using Synthetic Talking Heads to Train Humans in Lipreading at Scale	2216
<i>Aditya Agarwal (IIIT Hyderabad), Bipasha Sen (International Institute of Information Technology Hyderabad), Rudrabha Mukhopadhyay (IIIT Hyderabad), Vinay Namboodiri (University of Bath), and C.V. Jawahar (IIIT-Hyderabad)</i>	
Relaxing Contrastiveness in Multimodal Representation Learning	2226
<i>Zudi Lin (Harvard University), Erhan Bas (Amazon), Kunwar Y Singh (Amazon), Gurumurthy Swaminathan (Amazon), and Rahul Bhotika (Amazon)</i>	
Event-Specific Audio-Visual Fusion Layers: A Simple and New Perspective on Video Understanding	2236
<i>Arda Senocak (KAIST), Junsik Kim (Harvard University), Tae-Hyun Oh (POSTECH), Dingzeyu Li (Adobe Research), and In So Kweon (KAIST)</i>	
BirdSoundsDenoising: Deep Visual Audio Denoising for Bird Sounds	2247
<i>Youshan Zhang (Yeshiva University) and Jialu Li (Cornell University)</i>	
Audio-Visual Efficient Conformer for Robust Speech Recognition	2257
<i>Maxime Burchi (University of Würzburg) and Radu Timofte (University of Würzburg & ETH Zurich)</i>	
Recipe2Video: Synthesizing Personalized Videos from Recipe Texts	2267
<i>Prateksha Udhayan (Adobe Research), Suryateja Bv (Avanti Fellows), Parth S Laturia (Morgan Stanley), Dev G Chauhan (IIT Kanpur), Darshan Khandelwal (Adobe Research), Stefano Petrangeli (Adobe), and Balaji Vasan Srinivasan (Adobe Research)</i>	
Hear the Flow: Optical Flow-Based Self-Supervised Visual Sound Source Localization	2277
<i>Dennis Fedorishin (University at Buffalo), Deen D Mohan (University at Buffalo), Bhavin Jawade (University at Buffalo), Srirangaraj Setlur (University at BuffaloSUNY), and Venu Govindaraju (University at BuffaloSUNY)</i>	

6A: Machine Learning Architectures & Algorithms 1

Instance-Dependent Noisy Label Learning via Graphical Modelling	2287
<i>Arpit Garg (School of Computer ScienceThe University of Adelaide), Cuong Cao Nguyen (The University of Adelaide), Rafael Felix (The University of Adelaide), Thanh-Toan Do (Monash University), and Gustavo Carneiro (University of Adelaide)</i>	
Composite Learning for Robust and Effective Dense Predictions	2298
<i>Menelaos Kanakis (ETH Zurich), Thomas E Huang (ETH Zürich), David Brüggemann (ETH Zurich), Fisher Yu (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
Improving Multi-Fidelity Optimization with a Recurring Learning Rate for Hyperparameter Tuning	2308
<i>Hyunjae Lee (Lunit Inc.), Gi-Hyeon Lee (Lunit), Junhwan Kim (Lunit), Sungjun Cho (Lunit), Dohyun Kim (OnePredict), and Donggeun Yoo (Lunit)</i>	
Understanding the Role of Mixup in Knowledge Distillation: An Empirical Study	2318
<i>Hongjun Choi (Arizona State University), Eunsom Jeon (Arizona State University), Ankita Shukla (ASU), and Pavan Turaga (Arizona State University)</i>	
Cross-Task Attention Mechanism for Dense Multi-Task Learning	2328
<i>Ivan Lopes (Inria), Tuan-Hung Vu (Valeo.ai), and Raoul De Charette (Inria)</i>	
Continual Learning with Dependency Preserving Hypernetworks	2338
<i>Srikar Chandra Dupati (IIT Hyderabad), Sakshi Varshney (IIT Hyderabad), P. K. Srijith (IIT Hyderabad), and Sunil Gupta (Deakin UniversityAustralia)</i>	
FLOAT: Fast Learnable Once-for-All Adversarial Training for Tunable Trade-Off between Accuracy and Robustness	2348
<i>Souvik Kundu (University of Southern California), Sairam Sundaresan (Intel AI Lab), Massoud Pedram (University of Southern California), and Peter A. Beerel (University of Southern California)</i>	
Online Knowledge Distillation for Multi-Task Learning	2358
<i>Geethu M Jacob (Rakuten Institute of Technology), Vishal Agarwal (rakuten), and Bjorn Stenger (Rakuten Institute of Technology)</i>	
HyperPosePDF – Hypernetworks Predicting the Probability Distribution on SO(3)	2368
<i>Timon Höfer (University of Tuebingen), Benjamin Kiefer (University of Tuebingen), Martin Messmer (Universität Tübingen), and Andreas Zell (University of Tuebingen)</i>	
Improving Pixel-Level Contrastive Learning by Leveraging Exogenous Depth Information	2379
<i>Ahmed Ben Saad (ENS Paris Saclay), Kristina Prokopetc (AI LabSchlumberger), Josselin Jk Kherroubi (Schlumberger), Axel Davy (HGH Infrared Systems), Adrien Courtois (ENS Paris-Saclay), and Gabriele Facciolo (ENS Paris - Saclay)</i>	
What Can We Learn by Predicting Accuracy?	2389
<i>Olivier Risser-Maroux (Université de Paris) and Benjamin Chamand (IRIT)</i>	

AdvisIL – A Class-Incremental Learning Advisor	2399
<i>Eva Feillet (CEA LIST), Grégoire Petit (CEA), Adrian Popescu (CEA LIST), Céline Hudelot (CentraleSupélec), and Marina Reyboz (Univ. Grenoble AlpesCEALISTF-38000 GrenobleFrance)</i>	
Searching for Robust Binary Neural Networks via Bimodal Parameter Perturbation	2409
<i>Daehyun Ahn (POSTECH), Hyungjun Kim (POSTECH), Taesu Kim (POSTECH), Eunhyeok Park (POSTECH), and Jae-Joon Kim (Seoul National University)</i>	
RIFT: Disentangled Unsupervised Image Translation via Restricted Information Flow	2419
<i>Ben Usman (Boston University), Dina Bashkirova (Boston University), and Kate Saenko (Boston University)</i>	
Enabling ISPless Low-Power Computer Vision	2429
<i>Gourav Datta (University of Southern California), Zeyu Liu (University of Southern California), Zihan Yin (USC), Linyu Sun (University of Southern California), Akhilesh Jaiswal (University of Southern California), and Peter A. Beerel (University of Southern California)</i>	

6B: Low-Level Vision, Compression, Few-Shot Learning

On the Importance of Denoising When Learning to Compress Images	2439
<i>Benoit Brummer (intoPIX University of Louvain) and Christophe De Vleeschouwer (Université Catholique de Louvain)</i>	
End-to-End Single-Frame Image Signal Processing for High Dynamic Range Scenes	2448
<i>Khanh Quoc Dinh (Samsung Research) and Kwang Pyo Choi (Samsung Electronics)</i>	
No Reference Opinion Unaware Quality Assessment of Authentically Distorted Images	2458
<i>Nithin C Babu (Indian Institute of ScienceBangalore), Vignesh Kannan (Indian Institute of Science), and Rajiv Soundararajan (Indian Institute of Science)</i>	
HyperShot: Few-Shot Learning by Kernel HyperNetworks	2468
<i>Marcin Sendera (Jagiellonian University), Marcin Przewięźlikowski (Jagiellonian University), Konrad Karanowski (Wroclaw University of Science and Technology), Maciej Zieba (Wroclaw University of Science and Technology), Jacek Tabor (Jagiellonian University), and Przemysław Spurek (Jagiellonian University)</i>	
Contrastive Knowledge-Augmented Meta-Learning for Few-Shot Classification	2478
<i>Rakshith Subramanyam (Arizona State University), Mark Heimann (Lawrence Livermore), T.S. Jayram (LLNL), Rushil Anirudh (Lawrence Livermore National Laboratory), and Jayaraman J. Thiagarajan (Lawrence Livermore National Laboratory)</i>	
Few-Shot Medical Image Segmentation with Cycle-Resemblance Attention	2487
<i>Hao Ding (Illinois Institute of Technology), Changchang Sun (Illinois Institute of Technology), Hao Tang (ETH Zurich), Dawen Cai (University of Michigan), and Yan Yan (Illinois Institute of Technology)</i>	
Neural Distributed Image Compression with Cross-Attention Feature Alignment	2497
<i>Nitish Mital (Imperial College London), Ezgi Ozyilkan (New York University), Ali Garjani (EPFL), and Deniz Gunduz (Imperial College London)</i>	

MASTAF: A Model-Agnostic Spatio-Temporal Attention Fusion Network for Few-Shot Video Classification	2507
<i>Xin Liu (University of CaliforniaDavis), Huanle Zhang (University of CaliforniaDavis), Hamed Pirsiavash (University of California Davis), and Xin Liu (University of California)</i>	
Pixel-Wise Prediction Based Visual Odometry via Uncertainty Estimation	2517
<i>Hao-Wei Chen (National Tsing Hua University), Ting-Hsuan Liao (National Tsing Hua University), Hsuan-Kung Yang (National Tsing Hua University), and Chun-Yi Lee (National Tsing Hua University)</i>	
Universal Deep Image Compression via Content-Adaptive Optimization with Adapters	2528
<i>Koki Tsubota (The University of Tokyo), Hiroaki Akutsu (HitachiLtd.), and Kiyoharu Aizawa (The University of Tokyo)</i>	
Language-Free Training for Zero-Shot Video Grounding	2538
<i>Dahye Kim (Yonsei University), Jungin Park (Yonsei University), Jiyoung Lee (NAVER AI LAB), Seongheon Park (Yonsei Univ.), and Kwanghoon Sohn (Yonsei Univ.)</i>	
Separating Partially-Polarized Diffuse and Specular Reflection Components under Unpolarized Light Sources	2548
<i>Soma Kajiyama (Kyushu Institute of Technology), Taihe Piao (Kyushu Institute of Technology), Ryo Kawahara (Kyushu Institute of Technology), and Takahiro Okabe (Kyushu Institute of Technology)</i>	
Elimination of Non-Novel Segments at Multi-Scale for Few-Shot Segmentation	2558
<i>Alper Kayabaşı (METU), Gülin Tüfekci (Aselsan), and İlkan Uluşoy (METU)</i>	
An Unified Framework for Language Guided Image Completion	2567
<i>Jihyun Kim (Sogang University), Seong-Hun Jeong (Pukyong National University), Kyeongbo Kong (Pukyong National University), and Suk-Ju Kang (Sogang University)</i>	

6C: Anomaly & Out-of-Distribution Detection

Cross-Domain Video Anomaly Detection without Target Domain Adaptation	2578
<i>Abhishek Aich (University of CaliforniaRiverside), Kuan-Chuan Peng (Mitsubishi Electric Research Laboratories (MERL)), and Amit K. Roy-Chowdhury (University of CaliforniaRiverside)</i>	
Asymmetric Student-Teacher Networks for Industrial Anomaly Detection	2591
<i>Marco Rudolph (Leibniz University Hannover), Tom Wehrbein (Leibniz University Hannover), Bodo Rosenhahn (Leibniz University Hannover), and Bastian Wandt (University of British Columbia)</i>	
Heatmap-Based Out-of-Distribution Detection	2602
<i>Julia Hornauer (Ulm University) and Vasileios Belagiannis (Friedrich-Alexander-Universität Erlangen-Nürnberg)</i>	
Anomaly Detection in 3D Point Clouds Using Deep Geometric Descriptors	2612
<i>Paul Bergmann (MVTec Software GmbH) and David Sattlegger (MVTec Software GmbH)</i>	

Training Auxiliary Prototypical Classifiers for Explainable Anomaly Detection in Medical Image Segmentation	2623
<i>Wonwoo Cho (Korea Advanced Institute of Science and Technology), Jeonghoon Park (Korea Advanced Institute of Science and Technology), and Jaegul Choo (Korea Advanced Institute of Science and Technology)</i>	
Bi-Directional Frame Interpolation for Unsupervised Video Anomaly Detection	2633
<i>Hanqiu Deng (University of Alberta), Zhaoxiang Zhang (University of Alberta), Shihao Zou (University of Alberta), and Xingyu Li (University of Alberta)</i>	
Hyperdimensional Feature Fusion for Out-of-Distribution Detection	2643
<i>Samuel James Wilson (Queensland University of Technology), Tobias Fischer (Queensland University of Technology), Niko Suenderhauf (Queensland University of Technology), and Feras Dayoub (Queensland University of Technology)</i>	
Towards Interpretable Video Anomaly Detection	2654
<i>Keval Doshi (University of South Florida) and Yasin Yilmaz (University of South Florida)</i>	
Normality Guided Multiple Instance Learning for Weakly Supervised Video Anomaly Detection	2664
<i>Seongheon Park (Yonsei Univ.), Hanjae Kim (Yonsei Univ.), Minsu Kim (Yonsei Univ.), Dahye Kim (Yonsei University), and Kwanghoon Sohn (Yonsei Univ.)</i>	
Mutual Learning for Long-Tailed Recognition	2674
<i>Changhwa Park (42dot Inc.), Junho Yim (LG Energy Solution), and Eunji Jun (Hyundai Motor Group)</i>	

7A: Self-Supervised Learning

Representation Recovering for Self-Supervised Pre-Training on Medical Images	2684
<i>Xiangyi Yan (University of CaliforniaIrvine), Junayed A Naushad (UCI), Shanlin Sun (University of CaliforniaIrvine), Kun Han (University of California Irvine), Hao Tang (University of California Irvine), Deying Kong (University of California Irvine), Haoyu Ma (University of CaliforniaIrvine), Chenyu You (Yale University), and Xiaohui Xie (University of CaliforniaIrvine)</i>	
Self-Supervised Pyramid Representation Learning for Multi-Label Visual Analysis and Beyond...	2695
<i>Cheng-Yen Hsieh (National Taiwan University), Chih-Jung Chang (Stanford University), Fu-En Yang (National Taiwan University), and Yu-Chiang Frank Wang (National Taiwan University)</i>	
Similarity Contrastive Estimation for Self-Supervised Soft Contrastive Learning	2705
<i>Julien Denize (CEA), Jaonary Rabarisoa (CEA), Astrid Orcesi (CEA LIST), Romain Hérault (Normandy Université / INSA de Rouen / LITIS), and Stéphane Canu (INSA Rouen Normandie)</i>	

Magnification Prior: A Self-Supervised Method for Learning Representations on Breast Cancer Histopathological Images	2716
<i>Prakash Chandra Chhipa (Luleå University of Technology), Richa Upadhyay (Luleå University of Technology), Gustav Grund Pihlgren (Luleå University of Technology), Rajkumar Saini (Luleå tekniska universitetLuleåSweden), Seiichi Uchida (Kyushu University), and Marcus Liwicki (Luleå University of Technology)</i>	
Motion Aware Self-Supervision for Generic Event Boundary Detection	2727
<i>Ayush K. Rai (Insight SFI Centre for Data Analytics, Dublin City University (DCU)), Tarun Krishna (Insight SFI Centre for Data Analytics, Dublin City University (DCU)), Julia Dietlmeier (Insight SFI Centre for Data Analytics, Dublin City University (DCU)), Kevin Mcguinness (Insight SFI Centre for Data Analytics, Dublin City University (DCU)), Alan Smeaton (Insight SFI Centre for Data Analytics, Dublin City University (DCU)), and Noel O'Connor (Insight SFI Centre for Data Analytics, Dublin City University (DCU))</i>	
Improving Predicate Representation in Scene Graph Generation by Self-Supervised Learning	2739
<i>So Hasegawa (Fujitsu Limited), Masayuki Hiromoto (Fujitsu Limited), Akira Nakagawa (Fujitsu Laboratories Ltd.), and Yuhei Umeda (Fujitsu)</i>	
An Embedding-Dynamic Approach to Self-Supervised Learning	2749
<i>Suhong Moon (UC Berkeley), Domas Buracas (UC Berkeley), Seunghyun Park (Clova AI ResearchNAVER Corp.), Jinkyu Kim (Korea University), and John F Canny (UC Berkeley)</i>	
TeST: Test-Time Self-Training under Distribution Shift	2758
<i>Samarth Sinha (University of TorontoVector Institute), Peter Gehler (Amazon), Francesco Locatello (Amazon Lablet), and Bernt Schiele (MPI Informatics)</i>	
SSSD: Self-Supervised Self Distillation	2769
<i>Wei-Chi Chen (National Cheng Kung University) and Wei-Ta Chu (National Cheng Kung University)</i>	
Multi-Level Contrastive Learning for Self-Supervised Vision Transformers	2777
<i>Shentong Mo (Carnegie Mellon University), Zhun Sun (Tohoku University), and Chao Li (RIKEN)</i>	
Self-Supervised 2D/3D Registration for X-Ray to CT Image Fusion	2787
<i>Srikrishna Jaganathan (Pattern Recognition LabFAU Erlangen-Nuremberg), Maximilian Kukla (Uni Bamberg), Jian Wang (Siemens Healthcare GmbH), Karthik Shetty (Friedrich-Alexander-Universität Erlangen-Nürnberg), and Andreas K Maier (Pattern Recognition LabFAU Erlangen-Nuremberg)</i>	
FUSSL: Fuzzy Uncertain Self Supervised Learning	2798
<i>Salman Mohamadi (West Virginia University), Gianfranco Doretto (West Virginia University), and Donald Adjeroh (West Virginia University)</i>	
Rethinking Rotation in Self-Supervised Contrastive Learning: Adaptive Positive or Negative Data Augmentation	2808
<i>Atsuyuki Miyai (The University of Tokyo), Qing Yu (The University of Tokyo), Daiki Ikami (NTT Corporation), Go Irie (NTT Corporation), and Kiyoharu Aizawa (The University of Tokyo)</i>	

Self-Supervised Distilled Learning for Multi-Modal Misinformation Identification	2818
<i>Michael Mu (State University of New York at Buffalo), Sreyasee Das Bhattacharjee (State University of New York at Buffalo), and Junsong Yuan (State University of New York at BuffaloUSA)</i>	
Self-Distilled Self-Supervised Representation Learning	2828
<i>Jiho Jang (Seoul National University), Kiyoon Yoo (Seoul National University), Seonhoon Kim (Coupang), Chaerin Kong (Seoul National University), Jangho Kim (NAVER WEBTOON), and Nojun Kwak (Seoul National University)</i>	

7B: Pose & Place Recognition

TransVLAD: Multi-Scale Attention-Based Global Descriptors for Visual Geo-Localization	2839
<i>Yifan Xu (Shanghai JiaoTong University), Pourya Shamsolmoali (East China Normal University), Eric Granger (ETS Montreal), Claire Nicodeme (SNCF), Laurent Gardes (SNCF), and Jie Yang (Shanghai Jiao Tong University)</i>	
Learnable Human Mesh Triangulation for 3D Human Pose and Shape Estimation	2849
<i>Sungho Chun (Kwangwoon University), Sungbum Park (NCSOFT), and Ju Yong Chang (Kwangwoon University)</i>	
COPE: End-to-End Trainable Constant Runtime Object Pose Estimation	2859
<i>Stefan Thalhammer (ACIN TU Wien), Timothy Patten (University of Technology Sydney), and Markus Vincze (TU Wien)</i>	
ElliPose: Stereoscopic 3D Human Pose Estimation by Fitting Ellipsoids	2870
<i>Christian A Grund (AISC GmbH), Julian Tanke (University of Bonn), and Jürgen Gall (University of Bonn)</i>	
Partially Calibrated Semi-Generalized Pose from Hybrid Point Correspondences	2881
<i>Snehal Bhayani (University of OuluFinland), Viktor Larsson (Lund University), Torsten Sattler (Czech Technical University in Prague), Janne Heikkila (University of OuluFinland), and Zuzana Kukelova (Czech Technical University in Prague)</i>	
ImPosing: Implicit Pose Encoding for Efficient Visual Localization	2891
<i>Arthur Moreau (Mines ParisTech - PSL University), Thomas Gilles (Mines ParisTech), Nathan Piasco (Huawei Technologies France), Dzmitry Tsishkou (Huawei Technologies), Bogdan Stanciulescu (Mines ParisTech), and Arnaud De La Fortelle (MINES ParisTech - PSL University)</i>	
Uplift and Upsample: Efficient 3D Human Pose Estimation with Uplifting Transformers	2902
<i>Moritz Einfalt (University of Augsburg), Katja Ludwig (Universität Augsburg), and Rainer Lienhart (Universität AugsburgGermany)</i>	
Cross-View Image Sequence Geo-Localization	2913
<i>Xiaohan Zhang (University of Vermont), Waqas Sultani (Information Technology University), and Safwan Wshah (University of Vermont)</i>	
CameraPose: Weakly-Supervised Monocular 3D Human Pose Estimation by Leveraging In-the-Wild 2D Annotations	2923
<i>Cheng-Yen Yang (University of Washington), Jiajia Luo (Amazon), Lu Xia (Amazon), Yuyin Sun (Amazon), Ke Zhang (Amazon), Nan Qiao (Amazon), Zhongyu Jiang (University of Washington), Jenq-Neng Hwang (University of Washington), and Cheng-Hao Kuo (Amazon)</i>	

Image-Free Domain Generalization via CLIP for 3D Hand Pose Estimation	2933
<i>Seongyeon Lee (UNIST), Hansoo Park (UNIST), Dong Uk Kim (UNIST), Jihyeon Kim (UNIST), Muhammadjon Boboev (UNIST), and Seungryul Baek (UNIST)</i>	
Benchmarking Visual Localization for Autonomous Navigation	2944
<i>Lauri Suomela (Tampere University), Jussi Kalliola (Tampere University), Atakan Dag (Tampere University), Harry M Edelman (Tampere University / Turku University of Applied Sciences), and Joni-Kristian Kamarainen (Tampere University)</i>	
Learning 3D Human Pose Estimation from Dozens of Datasets Using a Geometry-Aware Autoencoder to Bridge between Skeleton Formats	2955
<i>István Sárándi (RWTH Aachen University), Alexander Hermans (RWTH Aachen University), and Bastian Leibe (RWTH Aachen University)</i>	
3D GAN Inversion with Pose Optimization	2966
<i>Jaehoon Ko (Korea University), Kyusun Cho (Korea University), Daewon Choi (Korea University), Kwangrok Ryoo (Korea University), and Seungryong Kim (Korea University)</i>	
Marker-Removal Networks to Collect Precise 3D Hand Data for RGB-Based Estimation and Its Application in Piano	2976
<i>Erwin Wu (Tokyo Institute of Technology), Hayato Nishioka (Sony Computer Science Laboratories Inc.), Shinichi Furuya (Sony Computer Science Laboratories Inc.), and Hideki Koike (Tokyo Institute of Technology)</i>	
Vis2Rec: A Large-Scale Visual Dataset for Visit Recommendation	2986
<i>Michaël Soumm (CEA LIST), Adrian Popescu (CEA LIST), and Bertrand Delezoide (CEA LIST)</i>	
MixVPR: Feature Mixing for Visual Place Recognition	2997
<i>Amar Ali-Bey (Laval University), Brahim Chaib-Draa (Laval University), and Philippe Giguere (Laval University)</i>	

7C: Depth Estimation & 3D Reconstruction

CountNet3D: A 3D Computer Vision Approach to Infer Counts of Occluded Objects	3007
<i>Porter Jenkins (Brigham Young University), Kyle W Armstrong (Delicious Ai), Stephen W Nelson (Brigham Young University), Siddhesh Gotad (DeliciousAI), James S Jenkins (DeliciousAi), Wade Wilkey (Delicious AI), and Tanner A Watts (University of Utah)</i>	
Temporally Consistent Online Depth Estimation in Dynamic Scenes	3017
<i>Zhaoshuo Li (Johns Hopkins University), Wei Ye (Facebook), Dilin Wang (Facebook), Francis Creighton (Johns Hopkins University), Russ Taylor (Johns Hopkins University), Ganesh Venkatesh (Facebook), and Mathias Unberath (Johns Hopkins University)</i>	
Uncertainty-Aware Interactive LiDAR Sampling for Deep Depth Completion	3027
<i>Kensuke Taguchi (Kyocera Corporation), Shogo Morita (Kyocera Corporation), Yusuke Hayashi (Kyocera Corporation), Wataru Imaeda (Kyocera Corp.), and Hironobu Fujiyoshi (Chubu University)</i>	

nLMVS-Net: Deep Non-Lambertian Multi-View Stereo	3036
<i>Kohei Yamashita (Kyoto University), Yuto Enyo (Kyoto University), Shohei Nobuhara (Kyoto University), and Ko Nishino (Kyoto University)</i>	
Wiener Guided DIP for Unsupervised Blind Image Deconvolution	3046
<i>Gustav Bredell (ETH Zürich), Ertunc Erdil (ETH Zurich), Bruno Weber (University of Zurich), and Ender Konukoglu (ETH Zurich)</i>	
360MVSNet: Deep Multi-View Stereo Network with 360° Images for Indoor Scene Reconstruction.....	3056
<i>Ching-Ya Chiu (National Taiwan University), Yu-Ting Wu (National Taipei University), I-Chao Shen (The University of Tokyo), and Yung-Yu Chuang (National Taiwan University)</i>	
Fast Differentiable Transient Rendering for Non-Line-of-Sight Reconstruction	3066
<i>Markus Plack (University of Bonn), Clara Callenberg (University of Bonn), Monika Schneider (University of Bonn), and Matthias B Hullin (University of Bonn)</i>	
MonoDVPS: A Self-Supervised Monocular Depth Estimation Approach to Depth-Aware Video Panoptic Segmentation	3076
<i>Andra Petrovai (Technical University of Cluj-Napoca) and Sergiu Nedeveschi (Technical University of Cluj-Napoca)</i>	
DELS-MVS: Deep Epipolar Line Search for Multi-View Stereo	3086
<i>Christian Sormann (Graz University of Technology), Emanuele Santellani (Technical University of Graz), Mattia Rossi (SONY Europe B.V.), Andreas Kuhn (SONY Europe), and Friedrich Fraundorfer (Graz University of Technology)</i>	
Probabilistic Volumetric Fusion for Dense Monocular SLAM	3096
<i>Antoni Rosinol (MIT), John Leonard (MIT), and Luca Carlone (Massachusetts Institute of Technology)</i>	
High-Quality RGB-D Reconstruction via Multi-View Uncalibrated Photometric Stereo and Gradient-SDF	3105
<i>Lu Sang (Technical University of Munich), Bjoern Haefner (Technical University of Munich), Xingxing Zuo (Technical University of Munich), and Daniel Cremers (TU Munich)</i>	
High-Resolution Depth Estimation for 360° Panoramas through Perspective and Panoramic Depth Images Registration	3115
<i>Chi-Han Peng (National Yang Ming Chiao Tung University) and Jiayao Zhang (KAUST)</i>	
Multi-View Photometric Stereo Revisited	3125
<i>Berk Kaya (ETH Zurich), Suryansh Kumar (ETH Zurich), Carlos Oliveira (CVL Zurich), Vittorio Ferrari (Google Research), and Luc Van Gool (ETH Zurich)</i>	
Automated Line Labelling: Dataset for Contour Detection and 3D Reconstruction	3135
<i>Hari Santhanam (University of Pennsylvania), Nehal Doiphode (University of Pennsylvania), and Jianbo Shi (University of Pennsylvania)</i>	
Anisotropic Multi-Scale Graph Convolutional Network for Dense Shape Correspondence	3145
<i>Mohammad Farazi (Arizona State University), Wenhui Zhu (Arizona State University), Zhangsihao Yang (Arizona State University), and Yalin Wang (Arizona State University)</i>	

Automatically Annotating Indoor Images with CAD Models via RGB-D Scans	3155
<i>Stefan Ainetter (Graz University of Technology), Sinisa Stekovic (Graz University of Technology), Friedrich Fraundorfer (Graz University of Technology), and Vincent Lepetit (Ecole des Ponts ParisTech)</i>	

8A: Segmentation

Intra-Batch Supervision for Panoptic Segmentation on High-Resolution Images	3164
<i>Daan De Geus (Eindhoven University of Technology) and Gijs Dubbelman (Eindhoven University of Technology)</i>	
Refign: Align and Refine for Adaptation of Semantic Segmentation to Adverse Conditions	3173
<i>David Brüggenmann (ETH Zurich), Christos Sakaridis (ETH Zurich), Prune Truong (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
Weakly Supervised Cell-Instance Segmentation with Two Types of Weak Labels by Single Instance Pasting	3184
<i>Kazuya Nishimura (Kyushu University) and Ryoma Bise (Kyushu University)</i>	
Attribution-Aware Weight Transfer: A Warm-Start Initialization for Class-Incremental Semantic Segmentation	3194
<i>Dipam Goswami (BITS Pilani), René Schuster (DFKI), Joost Van De Weijer (Computer Vision Center), and Didier Stricker (DFKI)</i>	
Semantic Segmentation of Degraded Images Using Layer-Wise Feature Adjustor	3204
<i>Kazuki Endo (Teikyo Heisei University), Masayuki Tanaka (Tokyo Institute of Technology), and Masatoshi Okutomi (Tokyo Institute of Technology)</i>	
Automated Detection of Label Errors in Semantic Segmentation Datasets via Deep Learning and Uncertainty Quantification	3213
<i>Matthias Rottmann (University of Wuppertal) and Marco Reese (University of Wuppertal)</i>	
Full Contextual Attention for Multi-Resolution Transformers in Semantic Segmentation	3223
<i>Loic Themyr (CNAM), Clément Rambour (Cnam), Nicolas Thome (CNAMParis), Toby Collins (IRCAD), and Alexandre Hostettler (IRCAD France)</i>	
WSNet: Towards an Effective Method for Wound Image Segmentation	3233
<i>Subba Reddy Oota (IIIT Hyderabad), Vijay Rowtula (IIIT Hyderabad), Shahid Saleem Mohammed (Woundtech), Mingsun Liu (Woundtech), and Manish Gupta (Microsoft)</i>	
Autoencoder-Based Background Reconstruction and Foreground Segmentation with Background Noise Estimation	3243
<i>Bruno E. Sauvalle (MINES ParisTech) and Arnaud De La Fortelle (MINES ParisTech - PSL University)</i>	
LoopDA: Constructing Self-Loops to Adapt Nighttime Semantic Segmentation	3255
<i>Fengyi Shen (Technical University of Munich), Zador Pataki (ETH Zurich), Akhil Gurram (Universitat Autònoma de Barcelona), Ziyuan Liu (Huawei group), He Wang (Peking University), and Alois C. Knoll (Robotics and Embedded Systems)</i>	

Unsupervised Multi-Object Segmentation Using Attention and Soft-Argmax	3266
<i>Bruno E. Sauvalle (MINES ParisTech) and Arnaud De La Fortelle (MINES ParisTech - PSL University)</i>	
Image Segmentation-Based Unsupervised Multiple Objects Discovery	3276
<i>Sandra Kara (CEA), Hejer Ammar (CEA), Florian Chabot (CEA), and Quoc-Cuong Pham (CEA)</i>	
X-Align: Cross-Modal Cross-View Alignment for Bird’s-Eye-View Segmentation	3286
<i>Shubhankar Borse (Qualcomm AI Research), Marvin Klingner (Technische Universität Braunschweig), Varun Ravi Kumar (Qualcomm), Hong Cai (Qualcomm AI Research), Abdulaziz Almuzairee (University of California San Diego), Senthil Yogamani (Qualcomm), and Fatih Porikli (Qualcomm AI Research)</i>	

8B: Action Recognition

Modality Mixer for Multi-Modal Action Recognition	3297
<i>Sumin Lee (KAIST), Sangmin Woo (KAIST), Yeonju Park (KAIST), Muhammad Adi Nugroho (KAIST), and Changick Kim (KAIST)</i>	
Fine-Grained Activities of People Worldwide	3307
<i>Jeffrey Byrne (Visym Labs), Gregory Castanon (Systems & Technology Research), Zhongheng Li (Systems and Technology Research), and Gil Ettinger (STR)</i>	
STAR-Transformer: A Spatio-Temporal Cross Attention Transformer for Human Action Recognition	3319
<i>Dasom Ahn (Keimyung University), Sangwon Kim (Keimyung University), Hyunsu Hong (Difine), and Byoung Chul Ko (Keimyung University)</i>	
Holistic Interaction Transformer Network for Action Detection	3329
<i>Gueter Josmy Faure (National Tsing Hua University), Min-Hung Chen (Microsoft), and Shang-Hong Lai (National Tsing Hua University)</i>	
VirtualHome Action Genome: A Simulated Spatio-Temporal Scene Graph Dataset with Consistent Relationship Labels	3340
<i>Yue Qiu (National Institute of Advanced Industrial Science and Technology (AIST)), Yoshiki Nagasaki (National Institute of Advanced Industrial Science and Technology (AIST); Keio University), Kensho Hara (National Institute of Advanced Industrial Science and Technology (AIST)), Hirokatsu Kataoka (National Institute of Advanced Industrial Science and Technology (AIST)), Ryota Suzuki (Saitama University), Kenji Iwata (National Institute of Advanced Industrial Science and Technology (AIST)), and Yutaka Satoh (National Institute of Advanced Industrial Science and Technology (AIST))</i>	
Stop or Forward: Dynamic Layer Skipping for Efficient Action Recognition	3350
<i>Jonghyeon Seon (Seoul National University), Jaedong Hwang (MIT), Jonghwan Mun (Kakao Brain), and Bohyung Han (Seoul National University)</i>	
Reconstructing Humpty Dumpty: Multi-Feature Graph Autoencoder for Open Set Action Recognition	3360
<i>Dawei Du (KitwareInc.), Ameya Shringi (Kitware Inc.), Anthony Hoogs (Kitware), and Christopher Funk (KitwareInc)</i>	

Multi-View Action Recognition Using Contrastive Learning	3370
<i>Ketul Shah (Johns Hopkins University), Anshul Shah (Johns Hopkins University), Chun Pong Lau (Johns Hopkins University), Celso De Melo (Army Research Laboratory), and Rama Chellappa (Johns Hopkins University)</i>	
Multimodal Vision Transformers with Forced Attention for Behavior Analysis	3381
<i>Tanay Agrawal (INRIA), Michal Balazsia (INRIA), Philipp Müller (DFKI GmbH), and Francois Bremond (Inria Sophia AntipolisFrance)</i>	
Efficient Skeleton-Based Action Recognition via Joint-Mapping Strategies	3392
<i>Min-Seok Kang (KAKAO Enterprise), Dongoh Kang (KAKAO Enterprise), and Hansaem Kim (KAKAO Enterprise)</i>	
GliTr: Glimpse Transformers with Spatiotemporal Consistency for Online Action Prediction	3402
<i>Samrudhdhi B. Rangrej (McGill University), Kevin J Liang (Meta), Tal Hassner (Facebook AI), and James J. Clark (McGill University)</i>	

8C: Face Modeling and Recognition

Harnessing Unrecognizable Faces for Improving Face Recognition	3413
<i>Siqi Deng (Amazon), Yuanjun Xiong (Amazon), Meng Wang (Amazon), Wei Xia (Amazon), and Stefano Soatto (AWS Amazon ML)</i>	
AT-DDPM: Restoring Faces Degraded by Atmospheric Turbulence Using Denoising Diffusion Probabilistic Models	3423
<i>Nithin Gopalakrishnan Nair (Johns Hopkins University), Kangfu Mei (Johns Hopkins University), and Vishal Patel (Johns Hopkins University)</i>	
IFQA: Interpretable Face Quality Assessment	3433
<i>Byungho Jo (Mayo Clinic), Donghyeon Cho (Chungnam National University), In Kyu Park (Inha University), and Sungeun Hong (Inha University)</i>	
FaceDancer: Pose- and Occlusion-Aware High Fidelity Face Swapping	3443
<i>Felix Rosberg (Berge Consulting), Eren Erdal Aksoy (Halmstad University), Cristofer Englund (RISE), and Fernando Alonso-Fernandez (Halmstad University)</i>	
Fine Gaze Redirection Learning with Gaze Hardness-Aware Transformation	3453
<i>Sangjin Park (Inha University), Daeha Kim (Inha University), and Byung Cheol Song (Inha University)</i>	
Scaling Neural Face Synthesis to High FPS and Low Latency by Neural Caching	3463
<i>Frank Yu (University of British Columbia), Sidney Fels (University of British Columbia), and Helge Rhodin (UBC)</i>	
QMagFace: Simple and Accurate Quality-Aware Face Recognition	3473
<i>Philipp Terhörst (Paderborn University), Malte Ihfeld (Technische Universität Darmstadt), Marco Huber (Fraunhofer IGD), Naser Damer (Fraunhofer IGD), Florian Kirchbuchner (Fraunhofer Institute for Computer Graphics Research IGD), Kiran Raja (NTNU), and Arjan Kuijper (Fraunhofer Institute for Computer Graphics Research IGD and Mathematical and Applied Visual Computing groupTU Darmstadt)</i>	

FaceOff: A Video-to-Video Face Swapping System	3484
<i>Aditya Agarwal (IIIT Hyderabad), Bipasha Sen (International Institute of Information Technology Hyderabad), Rudrabha Mukhopadhyay (IIIT Hyderabad), Vinay Namboodiri (University of Bath), and C.V. Jawahar (IIIT-Hyderabad)</i>	
Weakly Supervised Face Naming with Symmetry-Enhanced Contrastive Loss	3494
<i>Tingyu Qu (KU Leuven), Tinne Tuytelaars (KU Leuven), and Sien Moens (KU Leuven)</i>	
Mesh-Tension Driven Expression-Based Wrinkles for Synthetic Faces	3504
<i>Chirag A Raman (Delft University of Technology), Charlie Hewitt (Microsoft), Erroll W Wood (Microsoft), and Tadas Baltrusaitis (Microsoft)</i>	
DigiFace-1M: 1 Million Digital Face Images for Face Recognition	3515
<i>Gwangbin Bae (University of Cambridge), Martin De La Gorce (microsoft), Tadas Baltrusaitis (Microsoft), Charlie Hewitt (Microsoft), Dong Chen (Microsoft Research Asia), Julien Valentin (Microsoft), Roberto Cipolla (University of Cambridge), and Jingjing Shen (Microsoft)</i>	
3DMM-RF: Convolutional Radiance Fields for 3D Face Modeling	3525
<i>Efstathios Galanakis (Imperial College London), Baris Gecer (Huawei), Alexandros Lattas (Imperial College London), and Stefanos Zafeiriou (Imperial College London)</i>	
Unifying Margin-Based Softmax Losses in Face Recognition	3537
<i>Yang Zhang (Yahoo Research), Simao Herdade (Yahoo Research), Kapil Thadani (Yahoo Research), Eric M Dodds (Block), Jack Culpepper (Yahoo Research), and Yueh-Ning Ku (ByteDance)</i>	
FastSwap: A Lightweight One-Stage Framework for Real-Time Face Swapping	3547
<i>Sahng-Min Yoo (KAIST), Tae-Min Choi (KAIST), Jae-Woo Choi (KAIST), and Jong-Hwan Kim (KAIST)</i>	

9A: Text & Documents, Medical Imaging, Retail Applications

Knowing What to Label for Few Shot Microscopy Image Cell Segmentation	3557
<i>Youssef Dawoud (Universität Ulm), Arij Bouazizi (Ulm University), Katharina Ernst (Ulm University Medical Center), Gustavo Carneiro (University of Adelaide), and Vasileios Belagiannis (Friedrich-Alexander-Universität Erlangen-Nürnberg)</i>	
A Deep Neural Framework to Detect Individual Advertisement (Ad) from Videos	3567
<i>Zongyi Liu (Amazon.com)</i>	
Delving into Masked Autoencoders for Multi-Label Thorax Disease Classification	3577
<i>Junfei Xiao (Johns Hopkins University), Yutong Bai (Johns Hopkins University), Alan Yuille (Johns Hopkins University), and Zongwei Zhou (Johns Hopkins University)</i>	
OutfitTransformer: Learning Outfit Representations for Fashion Recommendation	3590
<i>Rohan Sarkar (Purdue University), Navaneeth Bodla (University of Maryland), Mariya Vasileva (Amazon), Yenliang Lin (Amazon), Anurag Beniwal (Amazon), Alan Lu (Amazon), and Gerard Medioni (USC)</i>	

LayerDoc: Layer-Wise Extraction of Spatial Hierarchical Structure in Visually-Rich Documents	3599
<i>Puneet Mathur (University of Maryland), Rajiv Jain (Adobe Research), Ashutosh Mehra (Adobe Inc), Jiuxiang Gu (Adobe Research), Franck Dernoncourt (Adobe Research), Anandhavelu N (Adobe Research), Quan Tran (Adobe Research), Verena Kaynig-Fittkau (Adobe), Ani Nenkova (Adobe Research), Dinesh Manocha (University of Maryland at College Park), and Vlad I Morariu (Adobe Research)</i>	
Color Recommendation for Vector Graphic Documents Based on Multi-Palette Representation	3610
<i>Qianru Qiu (CyberAgentInc.), Xueting Wang (CyberAgentInc.), Mayu Otani (CyberAgent), and Yuki Iwazaki (CyberAgentInc.)</i>	
Probabilistic Integration of Object Level Annotations in Chest X-Ray Classification	3619
<i>Tom J Van Sonsbeek (University of Amsterdam), Xiantong Zhen (University of Amsterdam), Dwarikanath Mahapatra (Inception Institute of Artificial Intelligence), and Marcel Worring (University of Amsterdam)</i>	
D-Extract: Extracting Dimensional Attributes from Product Images	3630
<i>Pushpendu Ghosh (Amazon), Nancy X.R. Wang (Amazon), and Promod Yenigalla (Amazon)</i>	
Generative Colorization of Structured Mobile Web Pages	3639
<i>Kotaro Kikuchi (CyberAgent), Naoto Inoue (CyberAgent), Mayu Otani (CyberAgent), Edgar Simo-Serra (Waseda University), and Kota Yamaguchi (CyberAgent)</i>	
The Fully Convolutional Transformer for Medical Image Segmentation	3649
<i>Athanasios Tragakis (University of Glasgow), Chaitanya Kaul (University of Glasgow), Roderick Murray-Smith (University of Glasgow), and Dirk Husmeier (University of Glasgow)</i>	
Multi-Scale Cell-Based Layout Representation for Document Understanding	3659
<i>Yuzhi Shi (Chubu University), Mijung Kim (Rakuten Institute of Technology), and Yeongnam Chae (Rakuten Institute of Technology)</i>	
Efficient Few-Shot Learning for Pixel-Precise Handwritten Document Layout Analysis	3669
<i>Axel De Nardin (Università di Udine), Silvia Zottin (University of Udine), Matteo Paier (University of Udine), Gian Luca Foresti (University of UdineItaly), Emanuela Colombi (Università di Udine), and Claudio Piciarelli (University of UdineItaly)</i>	
OCR-VQGAN: Taming Text-Within-Image Generation	3678
<i>Juan Antonio Rodríguez Garcia (Pompeu Fabra University), David Vazquez (ServiceNow), Issam Hadj Laradji (ServiceNow), Marco Pedersoli (École de technologie supérieure), and Pau Rodriguez (Element AI)</i>	

9B: Remote Sensing, Agriculture and Biology, Embedded & Real-Time, Few-Shot Learning

Tracking Growth and Decay of Plant Roots in Minirhizotron Images	3688
<i>Alexander Gillert (Fraunhofer IGD), Bo Peters (Universität Greifswald), Uwe Freiherr Von Lukas (Fraunhofer IGD), Juergen Kreyling (Universität of Greifswald), and Gesche Blume-Werry (Umeå University)</i>	

Handling Image and Label Resolution Mismatch in Remote Sensing	3698
<i>Scott Workman (DZYNE Technologies), Armin Hadzic (DZYNE Technologies Inc.), and M. Usman Rafique (Kitware Inc.)</i>	
ARUBA: An Architecture-Agnostic Balanced Loss for Aerial Object Detection	3708
<i>Rebbapragada V C Sairam (Indian Institute of TechnologyHyderabad), Monish K Keswani (Czech Technical University), Uttaran Sinha (IIT Hyderabad), Nishit Shah (IIT Hyderabad), and Vineeth N Balasubramanian (Indian Institute of TechnologyHyderabad)</i>	
The CropAndWeed Dataset: A Multi-Modal Learning Approach for Efficient Crop and Weed Manipulation	3718
<i>Daniel Steininger (AIT Austrian Institute of Technology), Andreas Trondl (AIT), Gerardus Croonen (AITAustrian Institute of Technology GmbH), Julia Simon (AIT Austrian Institute of Technology), and Verena Widhalm (AIT Austrian Institute of Technology)</i>	
DSTrans: Dual-Stream Transformer for Hyperspectral Image Restoration	3728
<i>Dabing Yu (Hohai university), Qingwu Li (Hohai University), Xiaolin Wang (Hohai University), Zhang Zhiliang (Hohai university), Yixi Qian (Hohai university), and Xu Chang (Hohai University)</i>	
Learning Few-Shot Segmentation from Bounding Box Annotations	3739
<i>Byeolyi Han (Georgia Institute of Technology) and Tae-Hyun Oh (POSTECH)</i>	
Towards Discriminative and Transferable One-Stage Few-Shot Object Detectors	3749
<i>Karim Guirguis (Robert Bosch Corporate Research), Mohamed Abdelsamad (Universität Stuttgart), George Eskandar (University of Stuttgart), Ahmed M Hendawy (University of Stuttgart), Matthias Kayser (Robert Bosch Corporate Research), Bin Yang (University of Stuttgart), and Jürgen Beyerer (Fraunhofer IOSB)</i>	
Learning Incoherent Light Emission Steering from Metasurfaces Using Generative Models	3759
<i>Prasad P Iyer (Sandia National Labs), Saaketh Desai (Sandia National Laboratories), Sadhvikas J Addamane (Sandia National Laboratories), Remi Dingreville (Sandia National Laboratories), and Igal Brener (Sandia National Labs)</i>	
Transformers for Recognition in Overhead Imagery: A Reality Check	3767
<i>Francesco Luzi (Duke University), Aneesh Gupta (Duke University), Leslie Collins (Duke University), Kyle Bradbury (Duke University), and Jordan Malof (University of Montana)</i>	
Wavelength-Aware 2D Convolutions for Hyperspectral Imaging	3777
<i>Leon Varga (Universiteat Tuebingen), Martin Messmer (Universität Tübingen), Nuri H Benbarka (University of Tübingen), and Andreas Zell (University of Tuebingen)</i>	
Semantic Segmentation in Aerial Imagery Using Multi-Level Contrastive Learning with Local Consistency	3787
<i>Maofeng Tang (University of Tennessee), Konstantinos Georgiou (University Of TennesseeKnoxville), Hairong Qi (University of Tennessee-Knoxville), Cody Champion (AFS), and Marc Bosch (AFS)</i>	
Are Straight-Through Gradients and Soft-Thresholding All You Need for Sparse Training?	3797
<i>Antoine Vanderschueren (UCLouvain) and Christophe De Vleeschouwer (Université Catholique de Louvain)</i>	

LCS: Learning Compressible Subspaces for Efficient, Adaptive, Real-Time Network Compression at Inference Time	3807
<i>Elvis Nunez (UCLA), Maxwell C Horton (AppleXnor.Ai and University of Washington), Anish J Prabhu (Apple), Anurag Ranjan (Apple), Ali Farhadi (University of Washington), Allen Institute for AI (Apple), and Mohammad Rastegari (University of Washington)</i>	
Learning Attention Propagation for Compositional Zero-Shot Learning	3817
<i>Muhammad Gul Zain Ali Ali Khan (Deutsches Forschungszentrum für künstliche Intelligenz (DFKI) GmbH), Muhammad Ferjad Naeem (ETH Zürich), Luc Van Gool (ETH Zurich), A. A. Pagani (DFKI), Didier Stricker (DFKI), and Muhammad Zeshan Afzal (Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH)</i>	
Computer Vision for International Border Legibility	3827
<i>Trevor A Ortega (Western Washington University), Thomas J Nelson (Western Washington University), Skyler B Crane (Western Washington University), Josh David Myers-Dean (University of Colorado Boulder), and Scott Wehrwein (Western Washington University)</i>	

9C: Machine Learning Architectures & Algorithms 2

SONGs: Self-Organizing Neural Graphs	3837
<i>Lukasz Struski (Jagiellonian University), Tomasz Danel (Jagiellonian University), Marek Smieja, Jacek Tabor (Jagiellonian University), and Bartosz Zieliński (Jagiellonian University)</i>	
Learning Lightweight Neural Networks via Channel-Split Recurrent Convolution	3847
<i>Guojun Wu (Bytedance), Xin Zhang (Worcester Polytechnic Institute), Ziming Zhang (Worcester Polytechnic Institute), Yanhua Li (Worcester Polytechnic Institute USA), Xun Zhou (University of Iowa), Christopher Brinton (Purdue University), and Zhenming Liu (College of William & Mary)</i>	
SPIQ: Data-Free Per-Channel Static Input Quantization	3858
<i>Edouard Yvinec (Datakalab), Arnaud Dapogny (Pierre and Marie Curie University (UPMC)), Matthieu Cord (Sorbonne University), and Kevin Bailly (UPMC)</i>	
Spatial Consistency Loss for Training Multi-Label Classifiers from Single-Label Annotations	3868
<i>Thomas Verelst (KU Leuven), Paul Rubenstein (MPI for Intelligent Systems Tuebingen), Marcin Eichner (Apple), Tinne Tuytelaars (KU Leuven), and Maxim Berman (Apple)</i>	
CORL: Compositional Representation Learning for Few-Shot Classification	3879
<i>Ju He (Johns Hopkins University), Adam Kortylewski (Max Planck Institute for Informatics), and Alan Yuille (Johns Hopkins University)</i>	
Compact and Optimal Deep Learning with Recurrent Parameter Generators	3889
<i>Jiayun Wang (UC Berkeley / ICSI), Yubei Chen (University of California Berkeley), Stella X Yu (UC Berkeley / ICSI), Brian Cheung (UC Berkeley), and Yann Lecun (Facebook)</i>	

FeTrIL: Feature Translation for Exemplar-Free Class-Incremental Learning	3900
<i>Grégoire Petit (CEA), Adrian Popescu (CEA LIST), Hugo Schindler (CEA), David Picard (ENPC), and Bertrand Delezoide (CEA LIST)</i>	
Gradient-Based Quantification of Epistemic Uncertainty for Deep Object Detectors	3910
<i>Tobias Riedlinger (University of Wuppertal), Matthias Rottmann (University of Wuppertal), Marius Schubert (University of Wuppertal), and Hanno Gottschalk (University of Wuppertal)</i>	
Adaptive Sample Selection for Robust Learning under Label Noise	3921
<i>Deep Patel (Indian Institute of Science) and P. S. Sastry (Indian Institute of Science)</i>	
Randomness is the Root of All Evil: more Reliable Evaluation of Deep Active Learning	3932
<i>Yilin Ji (Karlsruhe Institute of Technology), Daniel Kaestner (Karlsruhe Institute of Technology), Oliver Wirth (Karlsruhe Institute of Technology), and Christian Wressnegger (Karlsruhe Institute of Technology)</i>	
PatchDropout: Economizing Vision Transformers Using Patch Dropout	3942
<i>Yue Liu (KTH Royal Institute of Technology), Christos Matsoukas (KTH Royal Institute of Technology), Fredrik Strand (Karolinska Institutet), Hossein Azizpour (Royal Institute of Technology (KTH)), and Kevin Smith (KTH Royal Institute of Technology)</i>	
PRN: Panoptic Refinement Network	3952
<i>Bo Sun (Adobe Inc.), Jason Kuen (Adobe Research), Zhe Lin (Adobe Research), Philippos Mordohai (Stevens Institute of Technology), and Simon Chen (Adobe Research)</i>	
Self-Attentive Pooling for Efficient Deep Learning	3963
<i>Fang Chen (University of Southern California), Gourav Datta (University of Southern California), Souvik Kundu (University of Southern California), and Peter A. Bearel (University of Southern California)</i>	
Accumulated Trivial Attention Matters in Vision Transformers on Small Datasets	3973
<i>Xiangyu Chen (The University of Kansas), Qinghao Hu (Institute of AutomationChinese Academy of Sciences), Kaidong Li (University of Kansas), Cuncong Zhong (University of Kansas), and Guanghui Wang (Ryerson University)</i>	

Virtual: Recognition

The Change You Want to See	3982
<i>Ragav Sachdeva (University of Oxford) and Andrew Zisserman (University of Oxford)</i>	
Ancestor Search: Generalized Open Set Recognition via Hyperbolic Side Information Learning....	3992
<i>Xiwen Dengxiong (Rochester Institute of Technology) and Yu Kong (Michigan State University)</i>	

Trans4Map: Revisiting Holistic Bird’s-Eye-View Mapping from Egocentric Images to Allocentric Semantics with Vision Transformers	4002
<i>Chang Chen (Karlsruhe Institute of Technology), Jiaming Zhang (Karlsruhe Institute of Technology), Kailun Yang (Karlsruhe Institute of Technology), Kunyu Peng (KIT), and Rainer Stiefelhagen (Karlsruhe Institute of Technology)</i>	
More Knowledge, Less Bias: Unbiasing Scene Graph Generation with Explicit Ontological Adjustment	4012
<i>Zhanwen Chen (University of Virginia), Saed Rezayi (University of Georgia), and Sheng Li (University of Virginia)</i>	
AFPSNet: Multi-Class Part Parsing Based on Scaled Attention and Feature Fusion	4022
<i>Njuod Alsudays (Cardiff University), Jing Wu (Cardiff University), Yu-Kun Lai (Cardiff University), and Ze Ji (Cardiff University)</i>	
Foreground Guidance and Multi-Layer Feature Fusion for Unsupervised Object Discovery with Transformers	4032
<i>Zhiwei Lin (Peking University), Zengyu Yang (Beijing University of Technology), and Yongtao Wang (Peking University)</i>	

Virtual: Computational Photography

Adaptively-Realistic Image Generation from Stroke and Sketch with Diffusion Model	4043
<i>Shin-I Cheng (National Chiao Tung University), Yu-Jie Chen (National Chiao Tung University), Wei-Chen Chiu (National Chiao Tung University), Hung-Yu Tseng (Facebook), and Hsin-Ying Lee (Snap Inc)</i>	
Single-Image HDR Reconstruction by Multi-Exposure Generation	4052
<i>Phuoc-Hieu Le (VinAI Research), Quynh T Le (VinAI Research), Rang Nguyen (VinAI Research), and Binh-Son Hua (VinAI Research)</i>	
Dynamic Neural Portraits	4062
<i>Michail C Doukas (Imperial College London), Stylianos Ploumpis (Huawei Technologies Co. Ltd), and Stefanos Zafeiriou (Imperial College London)</i>	
Is Bigger Always Better? An Empirical Study on Efficient Architectures for Style Transfer and Beyond	4073
<i>Jie An (University of Rochester), Tao Li (Peking University), Haozhi Huang (Tencent), Jinwen Ma (Peking University), and Jiebo Luo (U. Rochester)</i>	
SLI-pSp: Injecting Multi-Scale Spatial Layout in pSp	4084
<i>Aradhya Mathur (IIITD), Anish Madan (Indraprastha Institute of Information Technology Delhi), and Ojaswa Sharma (IIITD)</i>	
Semi-Supervised Learning for Low-Light Image Restoration through Quality Assisted Pseudo-Labeling	4094
<i>Sameer Malik (Indian Institute Of Science) and Rajiv Soundararajan (Indian Institute of Science)</i>	

Virtual: Domain Adaptation and Generalization

Contrastive Learning of Semantic Concepts for Open-Set Cross-Domain Retrieval	4104
<i>Aishwarya Agarwal (Indian Institute of Technology Bombay), Srikrishna Karanam (Adobe Research), Balaji Vasan Srinivasan (Adobe Research), and Biplob Banerjee (Indian Institute of Technology Bombay)</i>	
Generative Alignment of Posterior Probabilities for Source-Free Domain Adaptation	4114
<i>Sachin Chhabra (Arizona State University), Hemanth Venkateswara (Arizona State University), and Baoxin Li (Arizona State University)</i>	
FFM: Injecting Out-of-Domain Knowledge via Factorized Frequency Modification	4124
<i>Zijian Wang (University of Queensland), Yadan Luo (The University of Queensland), Zi Helen Huang (University of Queensland), and Mahsa Baktashmotlagh (University of Queensland)</i>	
Exploiting Instance-Based Mixed Sampling via Auxiliary Source Domain Supervision for Domain-Adaptive Action Detection	4134
<i>Yifan Lu (ETH Zurich), Gurkirt Singh (ETH Zurich), Suman Saha (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
Center-Aware Adversarial Augmentation for Single Domain Generalization	4146
<i>Tianle Chen (University of Queensland), Mahsa Baktashmotlagh (University of Queensland), Zijian Wang (University of Queensland), and Mathieu Salzmann (EPFL)</i>	
Self-Distillation for Unsupervised 3D Domain Adaptation	4155
<i>Adriano Cardace (University of Bologna), Riccardo Spezialetti (University of Bologna), Pierluigi Zama Ramirez (University of Bologna), Samuele Salti (University of Bologna), and Luigi Di Stefano (University of Bologna)</i>	
CoNMix for Source-Free Single and Multi-Target Domain Adaptation	4167
<i>Vikash Kumar (Indian Institute of Science), Rohit Lal (Indian Institute of Science), Himanshu Patil (Indian Institute of Science), and Anirban Chakraborty (Indian Institute of Science)</i>	
Domain Adaptation Using Self-Training with Mixup for One-Stage Object Detection	4178
<i> jitender Kumar Maurya (Toshiba software India limited), Keyur Ranipa (Toshiba Software India Pvt Ltd), Osamu Yamaguchi (Toshiba Corporation), Tomoyuki Shibata (Toshiba Corporation), and Daisuke Kobayashi (Toshiba)</i>	
Recur, Attend or Convolve? On Whether Temporal Modeling Matters for Cross-Domain Robustness in Action Recognition	4188
<i>Sofia Broomé (KTH Royal Institute of Technology), Ernest J Pokropek (KTH Royal Institute of Technology), Boyu Li (KTH Royal Institute of Technology), and Hedvig Kjellström (KTH Royal Institute of Technology)</i>	
Select, Label, and Mix: Learning Discriminative Invariant Feature Representations for Partial Domain Adaptation	4199
<i>Aadarsh Sahoo (MIT-IBM Watson AI Lab), Rameswar Panda (MIT-IBM Watson AI Lab), Rogerio Feris (MIT-IBM Watson AI Lab/IBM Research), Kate Saenko (Boston University), and Abir Das (IIT Kharagpur)</i>	
Learning Style Subspaces for Controllable Unpaired Domain Translation	4209
<i>Gaurav Bhatt (The University of British Columbia) and Vineeth N Balasubramanian (Indian Institute of Technology Hyderabad)</i>	

Virtual: 3D Recognition

TransPillars: Coarse-To-Fine Aggregation for Multi-Frame 3D Object Detection	4219
<i>Zhipeng Luo (Nanyang Technological University), Gongjie Zhang (Nanyang Technological University), Changqing Zhou (Nanyang Technological University), Tianrui Liu (Nanyang Technological University), Shijian Lu (Nanyang Technological University), and Liang Pan (Nanyang Technological University)</i>	
Fractal Projection Forest: Fast and Explainable Point Cloud Classifier	4229
<i>Hanxiao Tan (TU Dortmund University)</i>	
Li3DeTr: A LiDAR Based 3D Detection Transformer	4239
<i>Gopi Krishna Erabati (University of Coimbra) and Helder Araujo (University of Coimbra)</i>	
ImpDet: Exploring Implicit Fields for 3D Object Detection	4249
<i>Xuelin Qian (Fudan University), Li Wang (Fudan University), Yi Zhu (Amazon), Li Zhang (Fudan University), Yanwei Fu (Fudan University), and Xiangyang Xue (Fudan University)</i>	
Weakly-Supervised Point Cloud Instance Segmentation with Geometric Priors	4260
<i>Heming Du (Australian National University), Xin Yu (University of Technology Sydney), Farookh Hussain (University of Technology Sydney), Mohammad Ali Armin (CSIRO(Data61)), Lars Petersson (Data61/CSIRO), and Weihao Li (Data61CSIRO)</i>	
Multivariate Probabilistic Monocular 3D Object Detection	4270
<i>Xuepeng Shi (Imperial College London), Zhixiang Chen (University of Sheffield), and Tae-Kyun Kim (Imperial College London)</i>	
Learning to Detect 3D Lanes by Shape Matching and Embedding	4280
<i>Ruixin Liu (Xi'an Jiaotong University), Zhihao Guan (Xi'an Jiaotong University), Zejian Yuan (Xi'an Jiaotong University), Ao Liu (Tencent), Tong Zhou (Tencent), Tang Kun (Tencent), Erlong Li (Tencent), Chao Zheng (Tencent), and Shuqi Mei (Tencent)</i>	

Virtual: Image & Scene Synthesis

DSAG: A Scalable Deep Framework for Action-Conditioned Multi-Actor Full Body Motion Synthesis	4289
<i>Debtanu Gupta (IIIT Hyderabad), Shubh Maheshwari (IIIT Hyderabad), Sai Shashank Kalakonda (International Institute of Information Technology Hyderabad), Manasvi V Vaidyula (International Institute of Information Technology Hyderabad), and Ravi Kiran Sarvadevabhatla (IIIT Hyderabad)</i>	
Self-Improving Multiplane-To-Layer Images for Novel View Synthesis	4298
<i>Pavel Ilich Solovev (Samsung), Taras Khakhulin (Skolkovo Institute of Science and Technology), and Denis Korzhenkov (Yandex Research)</i>	
SketchInverter: Multi-Class Sketch-Based Image Generation via GAN Inversion	4308
<i>Zirui An (Beihang University), Jingbo Yu (BeiHang University), Runtao Liu (Johns Hopkins University), Chuang Wang (Beihang University), and Qian Yu (Beihang University)</i>	

Recovering Fine Details for Neural Implicit Surface Reconstruction	4319
<i>Decai Chen (Fraunhofer Heinrich Hertz Institute), Peng Zhang (Fraunhofer Heinrich Hertz Institute), Ingo Feldmann (Fraunhofer HHI), Oliver Schreer (Fraunhofer Heinrich-Hertz-Institute), and Peter Eisert (Fraunhofer Heinrich Hertz Institute)</i>	
Control-NeRF: Editable Feature Volumes for Scene Rendering and Manipulation	4329
<i>Verica Lazova (Max Planck Institute for Informatics), Vladimir Guzov (University of Tuebingen), Kyle B Olszewski (Snap Inc.), Sergey Tulyakov (Snap Inc), and Gerard Pons-Moll (University of Tübingen)</i>	

Virtual: Biometrics

Split to Learn: Gradient Split for Multi-Task Human Image Analysis	4340
<i>Weijian Deng (Australian National University), Yumin Suh (NEC Labs America), Xiang Yu (NEC Labs), Masoud Faraki (NEC Labs), Liang Zheng (Australian National University), and Manmohan Chandraker (UC San Diego)</i>	
A Suspect Identification Framework Using Contrastive Relevance Feedback	4350
<i>Devansh Gupta (Indraprastha Institute of Information Technology Delhi), Aditya Saini (IIIT Delhi), Sarthak Bhagat (IIIT-Delhi), Shagun Uppal (IIIT-Delhi), Drishti Bhasin (IIT Roorkee), Rishi Raj Raj Jain (Indraprastha Institute of Information Technology Delhi), Ponnuram Kumaraguru (IIIT Hyderabad), and Rajiv Ratn Shah (IIIT Delhi)</i>	
Towards a Framework for Privacy-Preserving Pedestrian Analysis	4359
<i>Anil Kunchala (Technological University Dublin), Melanie Bouroche (Trinity College Dublin), and Bianca Schoen-Phelan (Technical University Dublin)</i>	

Virtual: Language & Vision

Guiding Visual Question Answering with Attention Priors	4370
<i>Thao Minh Le (Deakin University), Vuong Le (Deakin University), Sunil Gupta (Deakin University Australia), Svetha Venkatesh (Deakin University), and Truyen Tran (Deakin University)</i>	
Grounding Scene Graphs on Natural Images via Visio-Lingual Message Passing	4380
<i>Aditay Tripathi (Indian Institute of Science), Anand Mishra (Indian Institute of Technology Jodhpur), and Anirban Chakraborty (Indian Institute of Science)</i>	
K-VQG: Knowledge-Aware Visual Question Generation for Common-Sense Acquisition	4390
<i>Kohei Uehara (The University of Tokyo) and Tatsuya Harada (The University of Tokyo / RIKEN)</i>	
Perceiver-VL: Efficient Vision-and-Language Modeling with Iterative Latent Attention	4399
<i>Zineng Tang (UNC at Chapel Hill), Jaemin Cho (UNC Chapel Hill), Jie Lei (UNC Chapel Hill), and Mohit Bansal (University of North Carolina at Chapel Hill)</i>	

Text and Image Guided 3D Avatar Generation and Manipulation	4410
<i>Zehranaz Canfes (Bogazici University), M. Furkan Atasoy (Bogazici University), Alara Dirik (Bogazici University), and Pinar Yanardag (Bogazici University)</i>	
More Than Just Attention: Improving Cross-Modal Attentions with Contrastive Constraints for Image-Text Matching	4421
<i>Yuxiao Chen (Rutgers University), Jianbo Yuan (Bytedance), Long Zhao (Google Research), Tianlang Chen (University of Rochester), Rui Luo (AMAZON COM SERVICES INC), Larry Davis (AMAZON COM SERVICES INC), and Dimitris N. Metaxas (Rutgers)</i>	
Watching the News: Towards VideoQA Models That Can Read	4430
<i>Soumya Shamarao Jahagirdar (International Institute of Information Technology Hyderabad), Minesh Mathew (CVITIIIT-Hyderabad), Dimosthenis Karatzas (Computer Vision Centre), and C.V. Jawahar (IIIT-Hyderabad)</i>	
How to Practice VQA on a Resource-Limited Target Domain	4440
<i>Mingda Zhang (Google Inc), Rebecca Hua (University of Pittsburgh), and Adriana Kovashka (University of Pittsburgh)</i>	
Arbitrary Style Guidance for Enhanced Diffusion-Based Text-to-Image Generation	4450
<i>Zhihong Pan (Baidu Research (USA)), Xin Zhou (Baidu USA), and Hao Tian (BaiduInc)</i>	

Virtual: 3D Vision

GarSim: Particle Based Neural Garment Simulator	4461
<i>Lokender Tiwari (TCS Research) and Brojeshwar Bhowmick (Tata Consultancy Services)</i>	
IDD-3D: Indian Driving Dataset for 3D Unstructured Road Scenes	4471
<i>Shubham Dokania (IIIT Hyderabad), A. H. Abdul Hafez (Hasan Kalyoncu University), Anbumani Subramanian (IIIT-Hyderabad), Manmohan Chandraker (UC San Diego), and C.V. Jawahar (IIIT-Hyderabad)</i>	
Fast and Accurate: Video Enhancement Using Sparse Depth	4481
<i>Yu Feng (University of Rochester), Patrick Hansen (Tenstorrent), Paul Whatmough (Arm Research), Guoyu Lu (Rochester Institute of Technology), and Yuhao Zhu (University of Rochester)</i>	
Complementary Bi-Directional Feature Compression for Indoor 360° Semantic Segmentation with Self-Distillation	4490
<i>Zishuo Zheng (Beijing Jiaotong University), Chunyu Lin (Beijing Jiaotong University), Lang Nie (Beijing Jiaotong University), Kang Liao (Beijing Jiaotong University), Zhijie Shen (Beijing Jiaotong University), and Yao Zhao (Beijing Jiaotong University)</i>	
Overlap-Guided Gaussian Mixture Models for Point Cloud Registration	4500
<i>Guofeng Mei (University of Technology Sydney), Fabio Poiesi (Fondazione Bruno Kessler), Cristiano Saltori (University of Trento), Jian Zhang (UTS), Elisa Ricci (University of Trento), and Nicu Sebe (University of Trento)</i>	

3D Neural Sculpting (3DNS): Editing Neural Signed Distance Functions	4510
<i>Petros Tzathas (National Technical University of Athens), Petros Maragos (National Technical University of Athens), and Anastasios Roussos (Institute of Computer Science Foundation for Research and Technology Hellas)</i>	
Sim2real Transfer Learning for Point Cloud Segmentation: An Industrial Application Case on Autonomous Disassembly	4520
<i>Chengzhi Wu (Karlsruhe Institute of Technology), Xuelei Bi (KIT), Julius Pfrommer (Fraunhofer IOSB), Alexander Cebulla (Karlsruhe Institute of Technology), Simon Mangold (wbk Institute of Production Science), and Jürgen Beyerer (Fraunhofer IOSB)</i>	

Virtual: Adversarial Learning, Attacks, Defenses, and Deepfakes

Robustness of Trajectory Prediction Models under Map-Based Attacks	4530
<i>Zhihao Zheng (Lehigh University), Xiaowen Ying (Lehigh University), Zhen Yao (Lehigh University), and Mooi Choo Chuah (Lehigh University)</i>	
Inducing Data Amplification Using Auxiliary Datasets in Adversarial Training	4540
<i>Saehyung Lee (Seoul National University) and Hyungyu Lee (Seoul National University)</i>	
Certified Defense for Content Based Image Retrieval	4550
<i>Kazuya Kakizaki (NEC Corporation / University of Tsukuba), Kazuto Fukuchi (University of Tsukuba / RIKEN AIP), and Jun Sakuma (University of Tsukuba / RIKEN AIP)</i>	
Phantom Sponges: Exploiting Non-Maximum Suppression to Attack Deep Object Detectors	4560
<i>Avishag Shapira (The Open University), Alon Zolfi (Ben-Gurion University of the Negev), Luca Demetrio (Università degli Studi di Cagliari), Battista Biggio (University of Cagliari Italy), and Asaf Shabtai (Ben-Gurion University of the Negev)</i>	
Explainability-Aware One Point Attack for Point Cloud Neural Networks	4570
<i>Hanxiao Tan (TU Dortmund University) and Helena Kotthaus (TU Dortmund)</i>	
Image Completion with Heterogeneously Filtered Spectral Hints	4580
<i>Xingqian Xu (UIUC), Shant Navasardyan (Picsart Inc.), Vahram Tadevosyan (Picsart), Andranik Sargsyan (Picsart), Yadong Mu (Peking University), and Humphrey Shi (U of Oregon ; UIUC ; PAIR)</i>	
Proactive Deepfake Defence via Identity Watermarking	4591
<i>Yuan Zhao (University of Technology Sydney), Bo Liu (University of Technology Sydney (UTS)), Ming Ding (Data61CSIRO), Baoping Liu (UTS), Tianqing Zhu (University of Technology Sydney), and Xin Yu (University of Technology Sydney)</i>	
Avoiding Lingering in Learning Active Recognition by Adversarial Disturbance	4601
<i>Lei Fan (Northwestern University) and Ying Wu (Northwestern University)</i>	
DE-CROP: Data-Efficient Certified Robustness for Pretrained Classifiers	4611
<i>Gaurav Kumar Nayak (Indian Institute of Science Bangalore), Ruchit Rawal (Indian Institute of Science), and Anirban Chakraborty (Indian Institute of Science)</i>	

PatchZero: Defending Against Adversarial Patch Attacks by Detecting and Zeroing the Patch	4621
<i>Ke Xu (University of Southern California), Yao Xiao (University of Southern California), Zhaozheng Zheng (University of Southern California), Kaijie Cai (University of Southern California), and Ram Nevatia (U of Southern California)</i>	
CFL-Net: Image Forgery Localization Using Contrastive Learning	4631
<i>Fahim Faisal Niloy (Agency LabIndependent University Bangladesh), Kishor Kumar Bhaumik (Sungkyunkwan University), and Simon S Woo (Sungkyunkwan University (SKKU))</i>	
Indirect Adversarial Losses via an Intermediate Distribution for Training GANs	4641
<i>Rui Yang (The University of Tokyo), Duc Minh Vo (The University of Tokyo), and Hideki Nakayama (The University of Tokyo)</i>	
Adversarial Robustness in Discontinuous Spaces via Alternating Sampling & Descent	4651
<i>Rahul M V (Stanford University), Eric Wong (MIT), and Zico Kolter (Carnegie Mellon University)</i>	
RANCER: Non-Axis Aligned Anisotropic Certification with Randomized Smoothing	4661
<i>Taras Rumezhak (Ukrainian Catholic University), Francisco Girbal Eiras (University of Oxford), Philip Torr (University of Oxford), and Adel Bibi (University of Oxford)</i>	
Adversarial Local Distribution Regularization for Knowledge Distillation	4670
<i>Thanh Nguyen-Duc (Monash University), Trung Le (Monash University), He Zhao (CSIRO's Data61), Jianfei Cai (Monash University), and Dinh Phung (Monash University)</i>	
TI2Net: Temporal Identity Inconsistency Network for Deepfake Detection	4680
<i>Baoping Liu (UTS), Bo Liu (University of Technology Sydney (UTS)), Ming Ding (Data61CSIRO), Tianqing Zhu (University of Technology Sydney), and Xin Yu (University of Technology Sydney)</i>	
Interpreting Disparate Privacy-Utility Tradeoff in Adversarial Learning via Attribute Correlation	4690
<i>Likun Zhang (Institute of Information EngineeringCASChina), Yahong Chen (Institute of Information EngineeringChinese Academy of Sciences), Ang Li (Duke University), Binghui Wang (Illinois Institute of Technology), Yiran Chen (Duke University), Fenghua Li (Institute of Information EngineeringCASChina), Jin Cao (Xidian University), and Ben Niu (Institute of Information EngineeringCASChina)</i>	
Watch Those Words: Video Falsification Detection Using Word-Conditioned Facial Motion	4699
<i>Shruti Agarwal (MIT), Liwen Hu (Pinscreen), Evonne Ng (UC Berkeley), Trevor Darrell (UC Berkeley), Hao Li (Pinscreen), and Anna Rohrbach (UC Berkeley)</i>	

Virtual: Neural Architecture Search, Representation Learning, and Explainable Models

Augmentation by Counterfactual Explanation – Fixing an Overconfident Classifier	4709
<i>Sumedha Singla (University of Pittsburgh), Nihal Murali (University of Pittsburgh), Forough Arabshahi (Facebook), Sofia Triantafyllou (University of Pittsburgh Department of Biomedical Informatics), and Kayhan Batmanghelich (University of Pittsburgh)</i>	

Fantastic Style Channels and Where to Find Them: A Submodular Framework for Discovering Diverse Directions in GANs	4720
<i>Enis Simsar (TU Munich), Umut Kocasari (Bogazici University), Ezgi Gulperi Er (Bogazici University), and Pinar Yanardag (Bogazici University)</i>	
Visualizing Global Explanations of Point Cloud DNNs	4730
<i>Hanxiao Tan (TU Dortmund University)</i>	
Revisiting Training-Free NAS Metrics: An Efficient Training-Based Method	4740
<i>Taojiannan Yang (University of Central Florida), Linjie Yang (ByteDance AI Lab), Xiaojie Jin (Bytedance Inc. USA), and Chen Chen (University of Central Florida)</i>	
Encouraging Disentangled and Convex Representation with Controllable Interpolation Regularization	4750
<i>Yunhao Ge (University of Southern California), Zhi Xu (University of Southern California), Yao Xiao (University of Illinois at Urbana-Champaign), Gan Xin (University of Southern California), Yunkui Pang (University of North Carolina Chapel Hill), and Laurent Itti (University of Southern California)</i>	
RNAS-MER: A Refined Neural Architecture Search with Hybrid Spatiotemporal Operations for Micro-Expression Recognition	4759
<i>Monu Verma (MNIT), Priyanka Lubal (MNIT), Santosh Kumar Vipparthi (Indian Institute of Technology Guwahati), and Mohamed Abdel-Mottaleb (University of Miami)</i>	
Concept Correlation and Its Effects on Concept-Based Models	4769
<i>Lena Heidemann (Fraunhofer IKS), Maureen Monnet (Fraunhofer IKS), and Karsten Roscher (Fraunhofer IKS)</i>	

Virtual: Tracking & Reidentification

Graph-Based Self-Learning for Robust Person Re-Identification	4778
<i>Yuqiao Xian (Sun Yat-sen University), Jinrui Yang (Sun Yat-Sen University), Fufu Yu (Tencent Youtu), Jun Zhang (Tencent), and Xing Sun (Shopee)</i>	
Hard to Track Objects with Irregular Motions and Similar Appearances? Make It Easier by Buffering the Matching Space	4788
<i>Fan Yang (Fujitsu Research), Shigeyuki Odashima (Fujitsu Research), Shoichi Masui (Fujitsu Research), and Shan Jiang (Fujitsu Research)</i>	
Back to MLP: A Simple Baseline for Human Motion Prediction	4798
<i>Wen Guo (INRIA), Yuming Du (École des Ponts ParisTech), Xi Shen (Tencent AI Lab), Vincent Lepetit (Ecole des Ponts ParisTech), Xavier Alameda-Pineda (INRIA), and Francesc Moreno (IRI)</i>	
SAT: Scale-Augmented Transformer for Person Search	4809
<i>Mustansar Fiaz (MBZUAI), Hisham Cholakkal (MBZUAI), Rao Muhammad Anwer (MBZUAI/AALTO), and Fahad Shahbaz Khan (MBZUAI)</i>	
HOOT: Heavy Occlusions in Object Tracking Benchmark	4819
<i>Gozde Sahin (University of Southern CaliforniaUSA) and Laurent Itti (University of Southern California)</i>	

Relation Preserving Triplet Mining for Stabilising the Triplet Loss in Re-Identification Systems	4829
<i>Adhiraj Ghosh (University of Tübingen), Kuruparan Shanmugalingam (Singapore Management University), and Wen-Yan Lin (SMU)</i>	
Detection Recovery in Online Multi-Object Tracking with Sparse Graph Tracker	4839
<i>Jeongseok Hyun (HKUST), Myunggu Kang (Clova AI VideoNAVER Corp.), Dongyoon Wee (Clova AI VideoNAVER Corp.), and Dit-Yan Yeung (HKUST)</i>	
MMPTRACK: Large-Scale Densely Annotated Multi-Camera Multiple People Tracking Benchmark	4849
<i>Xiaotian Han (Microsoft), Quanzeng You (Microsoft Azure Computer Vision), Chunyu Wang (Microsoft Research asia), Zhizheng Zhang (Microsoft Research), Peng Chu (Microsoft), Houdong Hu (Microsoft AI and Research), Jiang Wang (Microsoft), and Zicheng Liu (Microsoft)</i>	
TransMOT: Spatial-Temporal Graph Transformer for Multiple Object Tracking	4859
<i>Peng Chu (Microsoft), Jiang Wang (Microsoft), Quanzeng You (Microsoft Azure Computer Vision), Haibin Ling (Stony Brook University), and Zicheng Liu (Microsoft)</i>	
Bent & Broken Bicycles: Leveraging Synthetic Data for Damaged Object Re-Identification	4870
<i>Luca Piano (Politecnico di Torino), Filippo Gabriele Praticò (Politecnico di Torino), Alessandro Sebastian S Russo (Politecnico di Torino), Lorenzo Lanari (Politecnico di Torino), Lia Morra (Politecnico di Torino), and Fabrizio Lamberti (Politecnico di Torino)</i>	

Virtual: Enhancement, Restoration, and Super-Resolution

Kernel-Aware Burst Blind Super-Resolution	4881
<i>Wenyi Lian (Uppsala University) and Shanglian Peng (CUIT)</i>	
Modeling the Lighting in Scenes as Style for Auto White-Balance Correction	4892
<i>Furkan Osman Kınlı (Özyeğin University), Doğa Yılmaz (Özyeğin University), Barış Özcan (Özyeğin University), and Furkan Kirac (Ozyegin University)</i>	
Real-Time Restoration of Dark Stereo Images	4903
<i>Mohit Lamba (Indian Institute of Technology Madras), Venkata Anoop Suhas Kumar Morisetty (IIT Madras), and Kaushik Mitra (IIT Madras)</i>	
LRA&LDRA: Rethinking Residual Predictions for Efficient Shadow Detection and Removal	4914
<i>Mehmet Kerim Yücel (Samsung R&D UK), Valia Dimaridou (CERTH-ITI), Bruno Manganelli (Samsung Research UK), Mete Ozay (Samsung Research UK), Anastasios Drosou (CERTH-ITI), and Albert Saà-Garriga (Samsung R&D UK)</i>	
Single Image Super-Resolution via a Dual Interactive Implicit Neural Network	4925
<i>Quan Nguyen (University of Texas at Arlington) and William Beksi (University of Texas at Arlington)</i>	
Joint Video Rolling Shutter Correction and Super-Resolution	4935
<i>Akash Gupta (University of CaliforniaRiverside), Sudhir K Singh (Vimaan Robotics), and Amit K. Roy-Chowdhury (University of CaliforniaRiverside)</i>	

Enriched CNN-Transformer Feature Aggregation Networks for Super-Resolution	4945
<i>Jinsu Yoo (Hanyang University), Taehoon Kim (LG AI Research), Sihaeng Lee (LG AI Research), Seung Hwan Kim (LG AI Research), Honglak Lee (LG AI Research), and Tae Hyun Kim (Hanyang Univeristy)</i>	

Virtual: Computer Vision for Healthcare

DSFormer: A Dual-Domain Self-Supervised Transformer for Accelerated Multi-Contrast MRI Reconstruction	4955
<i>Bo Zhou (Yale University), Neel Dey (New York University), Jo Schlemper (Hyperfine), Seyed Sadegh Mohseni Salehi (Hyperfine Research), Chi Liu (Yale University), James S Duncan (Yale University), and Michal Sofka (Hyperfine Research)</i>	
RADIANT: Better rPPG Estimation Using Signal Embeddings and Transformer	4965
<i>Anup Kumar Gupta (Indian Institute of Technology Indore), Rupesh Kumar (Indian Institute of Technology Indore), Lokendra Birla (IIT Indore), and Puneet Gupta (IIT Indore)</i>	
Attend Who is Weak: Pruning-Assisted Medical Image Localization under Sophisticated and Implicit Imbalances	4976
<i>Ajay Kumar Jaiswal (UT Austin), Tianlong Chen (University of Texas at Austin), Justin F Rousseau (Dell Medical School at University of Texas at Austin), Yifan Peng (Weill Cornell Medicine), Ying Ding (University of Texas at Austin), and Zhangyang Wang (University of Texas at Austin)</i>	
HoechtGAN: Virtual Lymphocyte Staining Using Generative Adversarial Networks	4986
<i>Georg Wölflein (University of St Andrews), In Hwa Um (University of St Andrews), David J Harrison (university of st andrews), and Ognjen Arandjelovic (University of St Andrews)</i>	
EfficientPhys: Enabling Simple, Fast and Accurate Camera-Based Cardiac Measurement	4997
<i>Xin Liu (University of Washington), Brian Hill (UCLA), Ziheng Jiang (University of Washington and OctoML), Shwetak Patel (University of Washington), and Daniel Mcduff (Microsoft Research)</i>	
Improving Deep Facial Phenotyping for Ultra-Rare Disorder Verification Using Model Ensembles	5007
<i>Alexander Hustinx (Institute for Genomic Statistics and Bioinformatics), Fabio Hellmann (University of Augsburg), Ömer Sümer (University of Augsburg), Behnam Javanmardi (Institute for Genomic Statistics and Bioinformatics), Elisabeth André (University of AugsburgGermany), Peter Krawitz (Institute for Genomic Statistics and Bioinformatics), and Tzung-Chien Hsieh (Institute for Genomic Statistics and Bioinformatics)</i>	
ALPINE: Improving Remote Heart Rate Estimation Using Contrastive Learning	5018
<i>Lokendra Birla (IIT Indore), Sneha Shukla (Indian Institute of Technology Indore), Anup Kumar Gupta (Indian Institute of Technology Indore), and Puneet Gupta (IIT Indore)</i>	

Exemplar Guided Deep Neural Network for Spatial Transcriptomics Analysis of Gene Expression Prediction	5028
<i>Yan Yang (The Australian National University), Md Zakir Hossain (The Australian National University), Eric A Stone (The Australian National University), and Shafin Rahman (North South University)</i>	

Virtual: Video Analysis & Optical Flow

Enhanced Bi-Directional Motion Estimation for Video Frame Interpolation	5038
<i>Xin Jin (Samsung Research), Longhai Wu (Samsung Electronics (China) R & D centre), Shen Guotao (Srcn), Chen Youxin (Samsung Electronics (China) R&D Center), Jie Chen (Software R&D CenterNanjingsamsung electronicsChina), Jayoon Koo (Samsung Electronics Co. Ltd), and Cheul-Hee Hahm (Samsung Electronics Co. Ltd)</i>	
Dance Style Transfer with Cross-Modal Transformer	5047
<i>Wenjie Yin (KTH Royal Institute of Technology), Hang Yin (KTH), Kim Baraka (Free University of Amsterdam), Danica Kragic (KTH Royal Institute of Technology), and Marten Bjorkman (KTH)</i>	
MFCFlow: A Motion Feature Compensated Multi-Frame Recurrent Network for Optical Flow Estimation	5057
<i>Yonghu Chen (SIMIT), Dongchen Zhu (SIMIT), Wenjun Shi (SIMIT), Guanghui Zhang (SIMIT), Tianyu Zhang (SIMIT), Xiaolin Zhang (SIMIT), and Jiamao Li (SIMIT)</i>	
GlobalFlowNet: Video Stabilization Using Deep Distilled Global Motion Estimates	5067
<i>Jerin Geo James (Indian Institute of Technology Bombay), Devansh Jain (IIT Bombay), and Ajit Rajwade (IIT Bombay)</i>	
Towards Equivariant Optical Flow Estimation with Deep Learning	5077
<i>Stefano Savian (Istituto Italiano di Tecnologia), Pietro Morerio (Istituto Italiano di Tecnologia), Alessio Del Bue (Istituto Italiano di Tecnologia (IIT)), Andrea A Janes (Free University of Bozen/Bolzano), and Tammam Tillo (Indraprastha Institute of Information Technology - Delhi)</i>	
Skew-Robust Human-Object Interactions in Videos	5087
<i>Apoorva Agarwal (IIT Bombay), Rishabh Dabral (IIT Bombay), Arjun Jain (Indian Institute Of Technology Bombay), and Ganesh Ramakrishnan (IIT Bombay)</i>	
Video Joint Denoising and Demosaicing with Recurrent CNNs	5097
<i>Valéry Dewil (Centre Borelli), Adrien Courtois (ENS Paris-Saclay), Mariano Rodriguez (ENS Paris-Saclay), Thibaud Ehret (Centre BorelliENS Paris-Saclay), Nicola Brandonisio (Huawei Technologies France), Denis Bujoreanu (Huawei Technologies France), Gabriele Facciolo (ENS Paris - Saclay), and Pablo Arias (ENS Paris-Saclay)</i>	
Video Object Matting via Hierarchical Space-Time Semantic Guidance	5109
<i>Yumeng Wang (Northwestern Polytechnical University), Bo Xu (OPPO Research Institute), Ziwen Li (OPPO Research Institute), Han Huang (OPPO Research Institute), Cheng Lu (XPENG), and Yandong Guo (OPPO Research Institute)</i>	

Exploiting Long-Term Dependencies for Generating Dynamic Scene Graphs	5119
<i>Shengyu Feng (Carnegie Mellon University), Hesham Mostafa (Intel Corporation), Marcel Nassar (Intel Corporation), Somdeb Majumdar (Intel Labs), and Subarna Tripathi (Intel Labs)</i>	
Treating Motion as Option to Reduce Motion Dependency in Unsupervised Video Object Segmentation	5129
<i>Suhwan Cho (Yonsei University), Minhyeok Lee (Yonsei University), Seunghoon Lee (Yonsei University), Chaewon Park (Yonsei University), Donghyeong Kim (Yonsei University), and Sangyoun Lee (Yonsei University)</i>	
DCVNet: Dilated Cost Volume Networks for Fast Optical Flow	5139
<i>Huaizu Jiang (Northeastern University) and Erik Learned-Miller (University of Massachusetts Amherst)</i>	

Virtual: Vision, Audio, and Other Modalities

AVE-CLIP: AudioCLIP-Based Multi-Window Temporal Transformer for Audio Visual Event Localization	5147
<i>Tanvir Mahmud (The University of Texas at Austin) and Diana Marculescu (The University of Texas at Austin)</i>	
SeCo: Separating Unknown Musical Visual Sounds with Consistency Guidance	5157
<i>Xinchi Zhou (The University of Sydney), Dongzhan Zhou (The University of Sydney), Wanli Ouyang (The University of Sydney), Hang Zhou (The Chinese University of Hong Kong), and Di Hu (Renmin University of China)</i>	
Audio-Visual Face Reenactment	5167
<i>Madhav Agarwal (IIIT-Hyderabad), Rudrabha Mukhopadhyay (IIIT Hyderabad), Vinay Namboodiri (University of Bath), and C.V. Jawahar (IIIT-Hyderabad)</i>	
LiveSeg: Unsupervised Multimodal Temporal Segmentation of Long Livestream Videos	5177
<i>Jielin Qiu (Carnegie Mellon University), Franck Deroncourt (Adobe Research), Trung Bui (Adobe Research), Zhaowen Wang (Adobe Research), Ding Zhao (Carnegie Mellon University), and Hailin Jin (Adobe Research)</i>	
Exploiting Visual Context Semantics for Sound Source Localization	5188
<i>Xinchi Zhou (The University of Sydney), Dongzhan Zhou (The University of Sydney), Di Hu (Renmin University of China), Hang Zhou (The Chinese University of Hong Kong), and Wanli Ouyang (The University of Sydney)</i>	
Towards Generating Ultra-High Resolution Talking-Face Videos with Lip Synchronization	5198
<i>Anchit Gupta (IIIT Hyderabad), Rudrabha Mukhopadhyay (IIIT Hyderabad), Sindhu B Hegde (International Institute of Information Technology (IIIT) Hyderabad), Faizan Farooq Khan (IIIT Hyderabad), Vinay Namboodiri (University of Bath), and C.V. Jawahar (IIIT-Hyderabad)</i>	
Towards Disturbance-Free Visual Mobile Manipulation	5208
<i>Tianwei Ni (Université de Montréal), Kiana Ehsani (Allen Institute for Artificial Intelligence), Luca Weihs (Allen Institute for Artificial Intelligence), and Jordi Salvador (Allen Institute for AI)</i>	

Unsupervised Audio-Visual Lecture Segmentation	5221
<i>Darshan Singh S (IIIT Hyderabad), Anchit Gupta (IIIT Hyderabad), C.V. Jawahar (IIIT-Hyderabad), and Makarand Tapaswi (Wadhvani AIIIT Hyderabad)</i>	
GAFNet: A Global Fourier Self Attention Based Novel Network for Multi-Modal Downstream Tasks	5231
<i>Onkar K Susladkar (Vishwakarma Institute of Information TechnologyPune), Gayatri S Deshmukh (Vishwakarma Institute of Information Technology), Dhruv Makwana (Independent Researcher), Sparsh Mittal (IIT Roorkee), Sai Chandra Teja R (Independent Researcher), and Rekha Singhal (TCS)</i>	
MT-DETR: Robust End-to-End Multimodal Detection with Confidence Fusion	5241
<i>Shih-Yun Chu (National Taiwan University) and Ming-Sui Lee (National Taiwan University)</i>	

Virtual: Machine Learning Architectures & Algorithms 1

Hyperspherical Quantization: Toward Smaller and more Accurate Models	5251
<i>Dan Liu (McGill University), Xi Chen (Huawei Noah's Ark Lab; McGill University), Chen Ma (City University of Hong Kong), and Xue Liu (McGill University)</i>	
Saliency Guided Experience Packing for Replay in Continual Learning	5262
<i>Gobinda Saha (Purdue University) and Kaushik Roy (Purdue University)</i>	
AdaNorm: Adaptive Gradient Norm Correction Based Optimizer for CNNs	5273
<i>Shiv Ram Dubey (Indian Institute of Information TechnologyAllahabad), Satish Kumar Singh (IIIT Alahabad), and Bidyut Baran Chaudhuri (ISI Kolkata)</i>	
Spike-Based Anytime Perception	5283
<i>Matthew Dutson (University of Wisconsin-Madison), Yin Li (University of Wisconsin-Madison), and Mohit Gupta (University of Wisconsin-MadisonUSA)</i>	
Meta-OLE: Meta-Learned Orthogonal Low-Rank Embedding	5294
<i>Ze Wang (Purdue University), Yue Lu (Harvard UniversityUSA), and Qiang Qiu (Purdue University)</i>	
GEMS: Generating Efficient Meta-Subnets	5304
<i>Varad Pimpalkhute (University of Massachusetts Amherst), Shruti Kunde (TCS), and Rekha Singhal (TCS)</i>	
SERF: Towards Better Training of Deep Neural Networks Using Log-Softplus ERror Activation Function	5313
<i>Sayan Nag (University of Toronto), Mayukh Bhattacharyya (Stony Brook University), Anuraag Mukherjee (Indian Institute of Science Education and ResearchMohali), and Rohit Kundu (University of CaliforniaRiverside)</i>	
Learning Latent Structural Relations with Message Passing Prior	5323
<i>Shaogang Ren (Texas A&M University), Hongliang Fei (Baidu Research), Dingcheng Li (Baidu Inc), and Ping Li (Baidu)</i>	

Bootstrapping the Relationship between Images and Their Clean and Noisy Labels	5333
<i>Brandon Smart (Australian Institute of Machine Learning) and Gustavo Carneiro (University of Adelaide)</i>	

Virtual: Low-Level Vision, Compression, Few-Shot Learning

Boosting Neural Video Codecs by Exploiting Hierarchical Redundancy	5344
<i>Reza Pourreza (Qualcomm), Hoang Le (Qualcomm AI), Amir Said (Qualcomm TechnologiesInc.), Guillaume Sautiere (Qualcomm AI Research), and Auke Wiggers (Qualcomm AI Research)</i>	
A Neural Video Codec with Spatial Rate-Distortion Control	5354
<i>Noor Fathima (Qualcomm Technologies Netherlands B.V), Jens Petersen (Qualcomm AI Research), Guillaume Sautiere (Qualcomm AI Research), Auke Wiggers (Qualcomm AI Research), and Reza Pourreza (Qualcomm)</i>	
Burst Vision Using Single-Photon Cameras	5364
<i>Sizhuo Ma (University of Wisconsin-Madison), Paul Mos (AQUA EPFL), Edoardo Charbon (EPFL), and Mohit Gupta (University of Wisconsin-MadisonUSA)</i>	
Few-Shot Object Detection via Improved Classification Features	5375
<i>Xinyu Jiang (Tongji University), Zhengjia Li (Alibaba), Maoqing Tian (SenseTime Group Limited), Jianbo Liu (The Chinese University of Hong Kong), Shuai Yi (SenseTime Group Limited), and Duoqian Miao (Tongji University)</i>	
EventPoint: Self-Supervised Interest Point Detection and Description for Event-Based Camera	5385
<i>Ze Huang (Xiamen University), Li Sun (Univeristy of Sheffield), Cheng Zhao (University of Oxford), Song Li (Infomatics SchoolXiamen University), and Songzhi Su (Xiamen University)</i>	
Sim2RealVS: A New Benchmark for Video Stabilization with a Strong Baseline	5395
<i>Qi Rao (University of Technology Sydney), Xin Yu (University of Technology Sydney), Shant Navasardyan (Picsart Inc.), and Humphrey Shi (U of Oregon; UIUC; PAIR)</i>	
Effective Invertible Arbitrary Image Rescaling	5405
<i>Zhihong Pan (Baidu Research (USA)), Baopu Li (Oracle Health and AI), Dongliang He (Baidu), Wenhao Wu (Baidu), and Errui Ding (Baidu Inc.)</i>	
Self-Attention Message Passing for Contrastive Few-Shot Learning	5415
<i>Ojas Shirekar (TU Delft), Hadi Jamali-Rad (Shell & TU Delft), and Anuj R Singh (Delft University of Technology)</i>	
One-Shot Doc Snippet Detection: Powering Search in Document Beyond Text	5426
<i>Abhinav Java (AdobeMDSR Labs), Shripad V Deshmukh (AdobeMDSR Labs), Milan Aggarwal (Adobe), Surgan Jandial (MDSR LabsAdobe), Mausoom Sarkar (Adobe), and Balaji Krishnamurthy (missing)</i>	
Semantic Guided Latent Parts Embedding for Few-Shot Learning	5436
<i>Fengyuan Yang (Institute of Computing TechnologyChinese Academy of Sciences), Ruiping Wang (Institute of Computing TechnologyChinese Academy of Sciences), and Xilin Chen (Institute of Computing TechnologyChinese Academy of Sciences)</i>	

Event-Based RGB Sensing with Structured Light	5447
<i>Seyed Ehsan Marjani Bajestani (Polytechnique Montreal) and Giovanni Beltrame (Polytechnique Montreal)</i>	
Discrete Cosin TransFormer: Image Modeling from Frequency Domain	5457
<i>Xinyu Li (ByteDance), Yanyi Zhang (Rutgers University), Jianbo Yuan (Bytedance), Hanlin Lu (ByteDance Inc.), and Yibo Zhu (ByteDance)</i>	

Virtual: Anomaly & Out-of-Distribution Detection

Anomaly Clustering: Grouping Images into Coherent Clusters of Anomaly Types	5468
<i>Kihyuk Sohn (Google), Jinsung Yoon (Google), Chun-Liang Li (Google), Chen-Yu Lee (Google), and Tomas Pfister (Google)</i>	
Image-Consistent Detection of Road Anomalies as Unpredictable Patches	5480
<i>Tomas Vojir (Czech Technical UniversityPrague) and Jiri Matas (CMP CTU FEE)</i>	
GLAD: A Global-to-Local Anomaly Detector	5490
<i>Aitor A Artola (Centre BorelliENS Paris-Saclay), Yannis Kolodziej (Visionairy), Jean-Michel Morel (Centre Borelli ENS Paris-Saclay), and Thibaud Ehret (Centre BorelliENS Paris-Saclay)</i>	
No Shifted Augmentations (NSA): Compact Distributions for Robust Self-Supervised Anomaly Detection	5500
<i>Mohamed Yousef (Intuition Machines Inc.), Marcel Ackermann (IntuitionMachines), Unmesh Kurup (Intuition Machines), and Tom Bishop (Glass ImagingInc.)</i>	
Out-of-Distribution Detection via Frequency-Regularized Generative Models	5510
<i>Mu Cai (University of Wisconsin-Madison) and Yixuan Li (University of Wisconsin-Madison)</i>	
Mixture Outlier Exposure: Towards Out-of-Distribution Detection in Fine-Grained Environments	5520
<i>Jingyang Zhang (Duke University), Nathan Inkawhich (Air Force Research Laboratory), Randolph W Linderman (Duke University), Yiran Chen (Duke University), and Hai Li (Duke University)</i>	
DyAnNet: A Scene Dynamicity Guided Self-Trained Video Anomaly Detection Network	5530
<i>Kamalakar Vijay Thakare (Indian Institute of TechnologyBhubaneswar), Yash Raghuwanshi (IIT Bhubaneswar), Debi Prosad Dogra (IIT Bhubaneswar), Heeseung Choi (Korea Institute of Science and Technology (KIST)), and Ig-Jae Kim (KIST)</i>	
Out-of-Distribution Detection with Reconstruction Error and Typicality-Based Penalty	5540
<i>Genki Osada (LINE Corporation), Tsubasa Takahashi (LINE Corporation), Budrul Ahsan (IBM Japan), and Takashi Nishide (University of Tsukuba)</i>	
Zero-Shot Versus Many-Shot: Unsupervised Texture Anomaly Detection	5553
<i>Toshimichi Aota (Sanoh Industrial Co.Ltd.), Lloyd Tzer Tong Teh (Sanoh Industrial Co.Ltd.), and Takayuki Okatani (Tohoku University/RIKEN AIP)</i>	

Virtual: Self-Supervised Learning

ViewCLR: Learning Self-Supervised Video Representation for Unseen Viewpoints	5562
<i>Srijan Das (University of North Carolina at Charlotte) and Michael S Ryoo (Stony Brook/Google)</i>	
Progressive Video Summarization via Multimodal Self-Supervised Learning	5573
<i>Haopeng Li (The University of Melbourne), Qihong Ke (Monash University), Mingming Gong (University of Melbourne), and Tom Drummond (University of Melbourne)</i>	
Self-Supervised Learning with Masked Image Modeling for Teeth Numbering, Detection of Dental Restorations, and Instance Segmentation in Dental Panoramic Radiographs	5583
<i>Amani H Almalki (Temple University) and Longin Jan Latecki (Temple University)</i>	
Cooperative Self-Training for Multi-Target Adaptive Semantic Segmentation	5593
<i>Yangsong Zhang (Telecom Paris), Subhankar Roy (University of Trento), Hongtao Lu (Shanghai Jiao Tong University), Elisa Ricci (University of Trento), and Stéphane Lathuilière (Telecom-Paris)</i>	
Self Supervised Low Dose Computed Tomography Image Denoising Using Invertible Network Exploiting Inter Slice Congruence	5603
<i>Sutanu Bera (Indian Institute of Technology Kharagpur) and Prabir Kumar Biswas (IIT Khargpur)</i>	
Self-Supervised Learning with Local Contrastive Loss for Detection and Semantic Segmentation	5613
<i>Ashraful Islam (Nvidia), Benjamin Lundell (Microsoft), Harpreet Sawhney (Microsoft), Sudipta Sinha (Microsoft), Peter Morales (Microsoft), and Richard Radke (Rensselaer Polytechnic Institute)</i>	
Self-Supervised Clustering Based on Manifold Learning and Graph Convolutional Networks	5623
<i>Leonardo T Lopes (UNESP) and Daniel Pedronette (UNESP)</i>	
Unifying Distribution Alignment as a Loss for Imbalanced Semi-Supervised Learning	5633
<i>Justin Lazarow (UC San Diego), Kihyuk Sohn (Google), Chen-Yu Lee (Google), Chun-Liang Li (Google), Zizhao Zhang (Google), and Tomas Pfister (Google)</i>	
Accelerating Self-Supervised Learning via Efficient Training Strategies	5643
<i>Mustafa Taha Kocyigit (University of Edinburgh), Timothy Hospedales (Edinburgh University), and Hakan Bilen (University of Edinburgh)</i>	

Virtual: Pose & Place Recognition

ETR: An Efficient Transformer for Re-Ranking in Visual Place Recognition	5654
<i>Hao Zhang (Fudan University), Xin Chen (Fudan University), Heming Jing (Fudan University), Yingbin Zheng (Videt Technology), Yuan Wu (Fudan University ChinaPR), and Cheng Jin (Fudan University)</i>	
HandGCNFormer: A Novel Topology-Aware Transformer Network for 3D Hand Pose Estimation	5664
<i>Yintong Wang (SIMIT), Lili Chen (SIMIT), Jiamao Li (SIMIT), and Xiaolin Zhang (SIMIT)</i>	

SD-Pose: Structural Discrepancy Aware Category-Level 6D Object Pose Estimation	5674
<i>Guowei Li (SIMIT), Dongchen Zhu (SIMIT), Guanghui Zhang (SIMIT), Wenjun Shi (SIMIT), Tianyu Zhang (SIMIT), Xiaolin Zhang (SIMIT), and Jiamao Li (SIMIT)</i>	
Rethinking the Data Annotation Process for Multi-View 3D Pose Estimation with Active Learning and Self-Training	5684
<i>Qi Feng (Meta), Kun He (Facebook Reality Labs), He Wen (Facebook Reality Labs), Cem Keskin (Facebook), and Yuting Ye (Facebook Reality Labs)</i>	
Self-Supervised Relative Pose with Homography Model-Fitting in the Loop	5694
<i>Bruce Muller (University of York) and William Smith (University of York)</i>	
HuPR: A Benchmark for Human Pose Estimation Using Millimeter Wave Radar	5704
<i>Shih-Po Lee (National Chiao Tung University), Niraj Prakash Kini (National Yang Ming Chiao Tung University), Wen-Hsiao Peng (National Chiao Tung University), Ching-Wen Ma (National Yang Ming Chiao Tung University), and Jenq-Neng Hwang (University of Washington)</i>	
Kinematic-Aware Hierarchical Attention Network for Human Pose Estimation in Videos	5714
<i>Kyung-Min Jin (Korea University), Byoungsung Lim (Korea University), Gun-Hee Lee (Korea University), Tae-Kyung Kang (Korea University), and Seong-Whan Lee (Korea University)</i>	
Interacting Hand-Object Pose Estimation via Dense Mutual Attention	5724
<i>Rong Wang (The Australian National University), Wei Mao (the Australian National University), and Hongdong Li (Australian National UniversityAustralia)</i>	
CRT-6D: Fast 6D Object Pose Estimation with Cascaded Refinement Transformers	5735
<i>Pedro Castro (Imperial College London) and Tae-Kyun Kim (Imperial College London)</i>	

Virtual: Depth Estimation & 3D Reconstruction

Meta-Auxiliary Learning for Future Depth Prediction in Videos	5745
<i>Huan Liu (McMaster University), Zhixiang Chi (Huawei Noah's Ark Laboratory), Yuanhao Yu (Huawei Noah's Ark Laboratory), Yang Wang (Concordia University), Jun Chen (McMaster University), and Jin Tang (Huawei Noah's Ark Laboratory)</i>	
X-NeRF: Explicit Neural Radiance Field for Multi-Scene 360° Insufficient RGB-D Views	5755
<i>Haoyi Zhu (Shanghai Jiao Tong University)</i>	
Self-Supervised Monocular Depth Estimation: Solving the Edge-Fattening Problem	5765
<i>Xingyu Chen (Peking University), Ruonan Zhang (Peking Universityshenzhen graduate school), Ji Jiang (Peking University), Yan Wang (Peking University), Ge Li (SECEShenzhen Graduate SchoolPeking University), and Thomas H Li (Advanced Institute of Information TechnologyPeking University)</i>	
PointNeuron: 3D Neuron Reconstruction via Geometry and Topology Learning of Point Clouds .	5776
<i>Runkai Zhao (University of Sydney), Heng Wang (The University of Sydney), Chaoyi Zhang (University of Sydney), and Weidong Cai (University of Sydney)</i>	

Self-Supervised Monocular Depth Estimation from Thermal Images via Adversarial Multi-Spectral Adaptation	5787
<i>Ukcheol Shin (KAIST), Kwanyong Park (KAIST), Byeong-Uk Lee (KAIST), Kyunghyun Lee (KAIST), and In So Kweon (KAIST)</i>	
Frequency-Aware Self-Supervised Monocular Depth Estimation	5797
<i>Xingyu Chen (Peking University), Thomas H Li (Advanced Institute of Information TechnologyPeking University), Ruonan Zhang (Peking Universityshenzhen graduate school), and Ge Li (SECEShenzhen Graduate SchoolPeking University)</i>	
SIUNet: Sparsity Invariant U-Net for Edge-Aware Depth Completion	5807
<i>Avinash Nittur Ramesh (Fraunhofer FHR), Fabio Giovanneschi (Fraunhofer FHR), and Maria Antonia Gonzalez Huici (Fraunhofer FHR)</i>	
Semantics-Depth-Symbiosis: Deeply Coupled Semi-Supervised Learning of Semantics and Depth	5817
<i>Nitin Bansal (OPPO US Research Center), Pan Ji (OPPO US Research Center), Junsong Yuan (State University of New York at BuffaloUSA), and Yi Xu (OPPO US Research Center)</i>	
Expansion of Visual Hints for Improved Generalization in Stereo Matching	5829
<i>Andrea Pilzer (NVIDIA), Yuxin Hou (Aalto University), Niki A Loppi (NVIDIA), Arno Solin (Aalto University), and Juho Kannala (Aalto UniversityFinland)</i>	
Heightfields for Efficient Scene Reconstruction for AR	5839
<i>Jamie Watson (Niantic), Sara Vicente (Niantic), Oisin Mac Aodha (University of Edinburgh), Clement Ljc Godard (Niantic), Gabriel Brostow (University College London), and Michael Firman (Niantic)</i>	
Attention Attention Everywhere: Monocular Depth Prediction with Skip Attention	5850
<i>Ashutosh Agarwal (IIT Delhi) and Chetan Arora (Indian Institute of Technology Delhi)</i>	
Sparsity Agnostic Depth Completion	5860
<i>Andrea Conti (University of Bologna), Matteo Poggi (University of Bologna), and Stefano Mattoccia (University of Bologna)</i>	

Virtual: Segmentation

Human-in-the-Loop Video Semantic Segmentation Auto-Annotation	5870
<i>Nan Qiao (Amazon), Yuyin Sun (Amazon), Chong Liu (UC Santa Barbara), Lu Xia (Amazon), Jiajia Luo (Amazon), Ke Zhang (Amazon), and Cheng-Hao Kuo (Amazon)</i>	
A Simple and Powerful Global Optimization for Unsupervised Video Object Segmentation	5881
<i>Georgy Ponimatkin (Ecole des Ponts ParisTech), Nermin Samet (Ecole des Ponts ParisTech), Yang Xiao (École des ponts ParisTech), Yuming Du (École des Ponts ParisTech), Renaud Marlet (Ecole des Ponts ParisTech), and Vincent Lepetit (Ecole des Ponts ParisTech)</i>	

Reducing Annotation Effort by Identifying and Labeling Contextually Diverse Classes for Semantic Segmentation under Domain Shift	5893
<i>Sharat Agarwal (Indraprastha Institute of Information Technology Delhi), Saket Anand (Indraprastha Institute of Information Technology Delhi), and Chetan Arora (Indian Institute of Technology Delhi)</i>	
Pruning-Guided Curriculum Learning for Semi-Supervised Semantic Segmentation	5903
<i>Heejo Kong (Korea university), Gun-Hee Lee (Korea University), Suneung Kim (Korea university), and Seong-Whan Lee (Korea University)</i>	
Unsupervised Video Object Segmentation via Prototype Memory Network	5913
<i>Minhyeok Lee (Yonsei University), Suhwan Cho (Yonsei University), Seunghoon Lee (Yonsei University), Chaewon Park (Yonsei University), and Sangyoun Lee (Yonsei University)</i>	
BEVSegFormer: Bird's Eye View Semantic Segmentation from Arbitrary Camera Rigs	5924
<i>Lang Peng (NullMax), Zhirong Chen (Nullmax), Zhangjie Fu (nullmax.ai), Pengpeng Liang (Zhengzhou University), and Erkang Cheng (Nullmax)</i>	
SCTS: Instance Segmentation of Single Cells Using a Transformer-Based Semantic-Aware Model and Space-Filling Augmentation	5933
<i>Yating Zhou (Institute of Automation Chinese Academy of Sciences), Wenjing Li (Chinese Academy of Sciences), and Ge Yang (National Laboratory of Pattern Recognition Institute of Automation Chinese Academy of Sciences)</i>	
Single Stage Weakly Supervised Semantic Segmentation of Complex Scenes	5943
<i>Peri Akiva (Rutgers University) and Kristin Dana (Rutgers University)</i>	
Semantic Segmentation with Active Semi-Supervised Learning	5955
<i>Aneesh Rangnekar (Rochester Institute of Technology), Christopher Kanan (University of Rochester), and Matthew J Hoffman (Rochester Institute of Technology)</i>	
Urban Scene Semantic Segmentation with Low-Cost Coarse Annotation	5967
<i>Anurag Das (Max-Planck-Institut für Informatik Saarbrücken), Yongqin Xian (ETH Zurich), Yang He (Amazon), Zeynep Akata (University of Tübingen), and Bernt Schiele (MPI Informatics)</i>	

Virtual: Action Recognition

Ego-Vehicle Action Recognition Based on Semi-Supervised Contrastive Learning	5977
<i>Chihiro Noguchi (Toyota Motor Corporation) and Toshihiro Tanizawa (Toyota Motor Corporation)</i>	
A Simple and Efficient Pipeline to Build an End-to-End Spatial-Temporal Action Detector	5988
<i>Lin Sui (Nanjing University), Chen-Lin Zhang (4ParadigmInc), Lixin Gu (dataelemInc), and Feng Han (DataElemInc)</i>	
Spatio-Temporal Action Detection under Large Motion	5998
<i>Gurkirt Singh (ETH Zurich), Vasileios Choutas (ETH Zurich), Suman Saha (ETH Zurich), Fisher Yu (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	

FAN-Trans: Online Knowledge Distillation for Facial Action Unit Detection	6008
<i>Jing Yang (University of Nottingham), Jie Shen (Imperial College London), Yiming Lin (Imperial College London / Meta), Yordan Hristov (independent), and Maja Pantic (Imperial College London/ Facebook)</i>	
Temporal Feature Enhancement Dilated Convolution Network for Weakly-Supervised Temporal Action Localization	6017
<i>Jianxiong Zhou (Northwestern University) and Ying Wu (Northwestern University)</i>	
Adaptive Local-Component-Aware Graph Convolutional Network for One-Shot Skeleton-Based Action Recognition	6027
<i>Anqi Zhu (The University of Melbourne), QiuHong Ke (Monash University), Mingming Gong (University of Melbourne), and James Bailey (The University of Melbourne)</i>	
Intention-Conditioned Long-Term Human Egocentric Action Anticipation	6037
<i>Esteve Valls Mascaro (TU Wien), Hyemin Ahn (Ulsan National Institute of Science and Technology), and Dongheui Lee (Technische Universität Wien (TU Wien))</i>	
Action-Aware Masking Network with Group-Based Attention for Temporal Action Localization .	6047
<i>Tae-Kyung Kang (Korea University), Gun-Hee Lee (Korea University), Kyung-Min Jin (Korea University), and Seong-Whan Lee (Korea University)</i>	
Anticipative Feature Fusion Transformer for Multi-Modal Action Anticipation	6057
<i>Zeyun Zhong (Karlsruhe Institute of Technology), David Schneider (Karlsruhe Institute of Technology), Michael Voit (Fraunhofer IOSB), Rainer Stiefelhagen (Karlsruhe Institute of Technology), and Jürgen Beyerer (Fraunhofer IOSB)</i>	

Virtual: Face Modeling and Recognition

Nested Deformable Multi-Head Attention for Facial Image Inpainting	6067
<i>Shruti S Phutke (Indian Institute of Technology Ropar) and Subrahmanyam Murala (IIT Ropar)</i>	
Uncertainty-Aware Label Distribution Learning for Facial Expression Recognition	6077
<i>Nhat Minh Le (AIOZ), Khanh Duy Nguyen (University of ScienceVNU-HCM), Quang Duy Tran (AIOZ), Erman Tjiputra (AIOZ), Bac Le (University of ScienceVNU-HCM), and Anh Nguyen (University of Liverpool)</i>	
Domain Invariant Vision Transformer Learning for Face Anti-Spoofing	6087
<i>Chen-Hao Liao (National Taiwan University), Wen-Cheng Chen (National Cheng Kung University), Hsuantung Liu (E.SUN Financial Holding Co.Ltd.), Yi-Ren Yeh (National Kaohsiung Normal University), Min-Chun Hu (National Tsing Hua University), and Chu-Song Chen (National Taiwan University)</i>	
CYBORG: Blending Human Saliency into the Loss Improves Deep Learning-Based Synthetic Face Detection	6097
<i>Aidan Boyd (University of Notre Dame), Patrick Tinsley (University of Notre Dame), Kevin Bowyer (University of Notre Dame), and Adam Czajka (University of Notre Dame)</i>	

ReEnFP: Detail-Preserving Face Reconstruction by Encoding Facial Priors	6107
<i>Yasheng Sun (Tokyo Institute of Technology), Jiangke Lin (NetEase Fuxi AI Lab), Hang Zhou (The Chinese University of Hong Kong), Zhiliang Xu (Baidu Inc.), Dongliang He (Baidu), and Hideki Koike (Tokyo Institute of Technology)</i>	
A Quality Aware Sample-to-Sample Comparison for Face Recognition	6118
<i>Mohammad Saeed Ebrahimi Saadabadi (West Virginia University), Sahar Rahimi Malakshan (West Virginia university), Ali Zafari (West Virginia University), Moktari Mostofa (West Virginia University), and Nasser Nasrabadi (West Virginia University)</i>	

Virtual: Text & Documents, Medical Imaging, Retail Applications

Treatment Learning Causal Transformer for Noisy Image Classification	6128
<i>Chao-Han Huck Yang (Georgia Institute of Technology), I-Te Hung (Columbia University), Yi-Chieh Liu (Georgia Institute of Technology), and Pin-Yu Chen (IBM Research)</i>	
Modeling Stroke Mask for End-to-End Text Erasing	6140
<i>Xiangcheng Du (Fudan University), Zhao Zhou (Fudan University), Yingbin Zheng (Videt Technology), Tianlong Ma (East China Normal University), Xingjiao Wu (East China Normal University), and Cheng Jin (Fudan University)</i>	
Cut-Paste Consistency Learning for Semi-Supervised Lesion Segmentation	6149
<i>Boon Peng Yap (Nanyang Technological UniversitySingapore) and Bk Ng (NTU)</i>	
ScoreNet: Learning Non-Uniform Attention and Augmentation for Transformer-Based Histopathological Image Classification	6159
<i>Thomas Stegmüller (EPFL), Behzad Bozorgtabar (EPFL), Antoine Spahr (EPFL), and Jean-Philippe Thiran (École Polytechnique Fédérale de Lausanne)</i>	
Seq-UPS: Sequential Uncertainty-Aware Pseudo-Label Selection for Semi-Supervised Text Recognition	6169
<i>Gaurav Patel (Purdue University), Jan Allebach (Purdue University), and Qiang Qiu (Purdue University)</i>	
From Forks to Forceps: A New Framework for Instance Segmentation of Surgical Instruments	6180
<i>Britty Baby (IIT Delhi), Daksh Thapar (Indian Institute of TechnologyMandi), Mustafa E Chasmai (Indian Institute of Technology Delhi), Tamajit Banerjee (Indian Institute of Technology Delhi), Kunal Dargan (Indian Institute of Technology Delhi), Ashish Suri (AIIMSDelhi), Subhashis Banerjee (IIT Delhi), and Chetan Arora (Indian Institute of Technology Delhi)</i>	

HiFormer: Hierarchical Multi-Scale Representations Using Transformers for Medical Image Segmentation	6191
<i>Moein Heidari (Iran University of Science and Technology), Amirhossein Kazerouni (Iran University of Science and Technology), Milad Soltany (Iran University of Science and Technology), Reza Azad (RWTH University), Ehsan Khodapanah Aghdam (IUST), Julien Cohen-Adad (NeuroPoly LabInstitute of Biomedical EngineeringPolytechnique Montreal), and Dorit Merhof (RWTH Aachen University)</i>	
LineEX: Data Extraction from Scientific Line Charts	6202
<i>Shivasankaran V P (IIT Gandhinagar), Muhammad Yusuf Hassan (IIT Gandhinagar), and Mayank Singh (IIT Gandhinagar)</i>	
Medical Image Segmentation via Cascaded Attention Decoding	6211
<i>Md Mostafijur Rahman (The University of Texas at Austin) and Radu Marculescu (The University of Texas at Austin)</i>	

Virtual: Remote Sensing, Agriculture and Biology, Embedded & Real-Time, Few-Shot Learning

MFFN: Multi-View Feature Fusion Network for Camouflaged Object Detection	6221
<i>Dehua Zheng Zheng (Huazhong University of Science and Technology), Xiaochen Zheng (ETH Zurich), Laurence T. Yang (St. Francis Xavier University), Yuan Gao (Huazhong University of Science and Technology), Chenlu Zhu (Hubei Chutian Expressway Digital Technology Co.LtdChina.), and Yiheng Ruan (Hubei Chutian Expressway Digital Technology Co.LtdChina.)</i>	
Semantics Guided Contrastive Learning of Transformers for Zero-Shot Temporal Activity Detection	6232
<i>Sayak Nag (University of California Riverside), Orpaz Goldstein (Edgecast), and Amit K. Roy-Chowdhury (University of CaliforniaRiverside)</i>	
OpenEarthMap: A Benchmark Dataset for Global High-Resolution Land Cover Mapping	6243
<i>Junshi Xia (RIKEN), Naoto Yokoya (The University of Tokyo), Bruno Adriano (RIKEN Center for Advanced Intelligence Project (AIP)), and Clifford Broni-Bediako (RIKEN)</i>	
Few-Shot Learning of Compact Models via Task-Specific Meta Distillation	6254
<i>Yong Wu (Shanghai University), Shekhor Chanda (University of Manitoba), Mehrdad Hosseinzadeh (University of Manitoba; Huawei Technologies Canada), Zhi Liu (Shanghai UniversityChina), and Yang Wang (Concordia University)</i>	
SSFE-Net: Self-Supervised Feature Enhancement for Ultra-Fine-Grained Few-Shot Class Incremental Learning	6264
<i>Zicheng Pan (Griffith University), Xiaohan Yu (Griffith University), Miaohua Zhang (Griffith University), and Yongsheng Gao (Griffith University)</i>	
One-Shot Synthesis of Images and Segmentation Masks	6274
<i>Vadim Sushko (Bosch Center for Artificial Intelligence), Dan Zhang (Bosch Center for Artificial Intelligence), Jürgen Gall (University of Bonn), and Anna Khoreva (Bosch Center for Artificial Intelligence)</i>	

MORGAN: Meta-Learning-Based Few-Shot Open-Set Recognition via Generative Adversarial Network	6284
<i>Debabrata Pal (Honeywell), Shirsha Bose (Technical University of Munich), Biplab Banerjee (Indian Institute of TechnologyBombay), and Yogananda Jeppu (Honeywell)</i>	
Aerial Image Dehazing with Attentive Deformable Transformers	6294
<i>Ashutosh C Kulkarni (Indian Institute of TechnologyRopar) and Subrahmanyam Murala (IIT Ropar)</i>	
Few-Shot Object Counting with Similarity-Aware Feature Enhancement	6304
<i>Zhiyuan You (Shanghai Jiao Tong University), Kai Yang (SenseTime Research), Wenhan Luo (Sun Yat-sen University), Xin Lu (SenseTime Research), Lei Cui (Tsinghua University), and Xinyi Le (Shanghai Jiao Tong University)</i>	
Aggregating Bilateral Attention for Few-Shot Instance Localization	6314
<i>He-Yen Hsieh (Academia Sinica), Ding-Jie Chen (Academia Sinica), Cheng-Wei Chang (Academia Sinica), and Tyng-Luh Liu (Academia Sinica)</i>	
Deep Learning Methodology for Early Detection and Outbreak Prediction of Invasive Species Growth	6324
<i>Nathan Elias (Liberal Arts and Science Academy (LASA))</i>	
Improving the Pair Selection and the Model Fusion Steps of Satellite Multi-View Stereo Pipelines	6333
<i>Alvaro Gomez (Universidad de la República - Facultad de Ingeniería), Gabriele Facciolo (ENS Paris - Saclay), Rafael Grompone Von Gioi (Centre BorelliENS Paris-Saclay), and Gregory Randall (Universidad de la RepúblicaFacultad de Ingeniería)</i>	
GAF-Net: Improving the Performance of Remote Sensing Image Fusion Using Novel Global Self and Cross Attention Learning	6343
<i>Ankit Jha (Indian Institute of Technology Bombay), Shirsha Bose (Technical University of Munich), and Biplab Banerjee (Indian Institute of TechnologyBombay)</i>	
Hyperblock Floating Point: Generalised Quantization Scheme for Gradient and Inference Computation	6353
<i>Marcelo Gennari Do Nascimento (Unviersity of Oxford), Roger Fawcett (Intel), Martin Langhammer (Intel Corporation), and Victor Adrian Prisacariu (University of Oxford)</i>	
Self-Pair: Synthesizing Changes from Single Source for Object Change Detection in Remote Sensing Imagery	6363
<i>Minseok Seo (si-analytics), Hakjin Lee (SI Analytics), Yongjin Jeon (SI Analytics), and Junghoon Seo (SI Analytics)</i>	
Jointly Learning Band Selection and Filter Array Design for Hyperspectral Imaging	6373
<i>Ke Li (ETH Zurich), Dengxin Dai (MPI for Informatics), and Luc Van Gool (ETH Zurich)</i>	
Computer Vision for Ocean Eddy Detection in Infrared Imagery	6384
<i>Evangelos Moschos (École PolytechniqueAmphitrite), Alisa Kugusheva (Ecole PolytechniqueAmphitrite), Paul Coste (IMT AtlantiqueAmphitrite), and Alexandre Stegner (Ecole PolytechniqueCNRSAmphitrite)</i>	

ROMA: Run-Time Object Detection to Maximize Real-Time Accuracy	6394
<i>Junkyu Lee (Queen's University Belfast), Blesson Varghese (University of St Andrews), and Hans Vandierendonck (Queen's University Belfast)</i>	
AnoLeaf: Unsupervised Leaf Disease Segmentation via Structurally Robust Generative Inpainting	6404
<i>Swati Bhugra (IIT Delhi), Vinay Kaushik (IIT Delhi), Amit Gupta (DView AI), Brejesh Lall (IIT Delhi), and Santanu Chaudhury (IITDelhiIndia)</i>	
LAB: Learnable Activation Binarizer for Binary Neural Networks	6414
<i>Sieger Falkena (Delft university of Technology), Hadi Jamali-Rad (Shell & TU Delft), and Jan C Van Gemert (Delft University of Technology)</i>	

Virtual: Machine Learning Architectures & Algorithms 2

Collaborative Multi-Teacher Knowledge Distillation for Learning Low Bit-Width Deep Neural Networks	6424
<i>Cuong Pham (Monash University), Tuan Na Hoang (Bytedance), and Thanh-Toan Do (Monash University)</i>	
Difficulty-Net: Learning to Predict Difficulty for Long-Tailed Recognition	6433
<i>Saptarshi Sinha (Hitachi CRL) and Hiroki Ohashi (Hitachi Ltd)</i>	
SD-Conv: Towards the Parameter-Efficiency of Dynamic Convolution	6443
<i>Shwai He (JD Explore Academy), Chenbo Jiang (Nanjing University of Science and Technology Zijin College), Daize Dong (University of Electronic Science and Technology of China), and Liang Ding (JD Explore Academy)</i>	
Dynamic Re-Weighting for Long-Tailed Semi-Supervised Learning	6453
<i>Hanyu Peng (Baidu Research), Weiguo Pian (University of Luxembourg), Mingming Sun (Baidu Research), and Ping Li (Baidu Research)</i>	
Couplformer: Rethinking Vision Transformer with Coupling Attention	6464
<i>Hai Lan (Fujian Institute of Research on the Structure of MatterChinese Academy of Sciences), Xihao Wang (Technical University of Munich), Hao Shen (fortiss GmbH), Peidong Liang (Fujian (Quanzhou) HIT Research Institute of Engineering and Technology), and Xian Wei (ECNU)</i>	
Do Pre-Trained Models Benefit Equally in Continual Learning?	6474
<i>Kuan-Ying Lee (University of Illinois at Urbana-Champaign), Yuanyi Zhong (University of Illinois at Urbana-Champaign), and Yu-Xiong Wang (University of Illinois at Urbana-Champaign)</i>	
Performance Comparison of DVS Data Spatial Downscaling Methods Using Spiking Neural Networks	6483
<i>Amélie Gruel (I3S / CNRS), Jean Martinet (I3SUniv. Cote d'AzurCNRS), Bernabe Linares-Barranco (Instituto de Microelectronica de Sevilla (IMSE-CNM); CSIC and Univ. of Sevilla), and Teresa Serrano-Gotarredona (Instituto de Microelectronica de Sevilla (IMSE-CNM); CSIC and Univ. of Sevilla)</i>	

Pushing the Efficiency Limit Using Structured Sparse Convolutions	6492
<i>Vinay Kumar Verma (Duke University), Nikhil Mehta (Duke University), Shijing Si (Ping An Technology (Shenzhen) Co. Ltd), Ricardo Henao (Duke University), and Lawrence Carin Duke (CS)</i>	
Dataset Condensation with Distribution Matching	6503
<i>Bo Zhao (The University of Edinburgh) and Hakan Bilen (University of Edinburgh)</i>	
Mapping DNN Embedding Manifolds for Network Generalization Prediction	6513
<i>Molly O'Brien (The Johns Hopkins University), Brett Wolfinger (The Johns Hopkins University), Julia Bukowski (Villanova University), Mathias Unberath (Johns Hopkins University), Aria Pezeshk (missing), and Gregory D. Hager (The Johns Hopkins University)</i>	
Federated Learning for Commercial Image Sources	6523
<i>Shreyansh Jain (IIIT Delhi) and Koteswar Rao Jerripothula (IIIT Delhi)</i>	

Author Index