

2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) **CVPR 2021**

Table of Contents

Message from the General and Program Chairs	clxvi
2021 Organizing Committee	clxvii
Area Chairs	clxix
Reviewers	clxxi

Session 01

Single-Stage Instance Shadow Detection With Bidirectional Relation Learning	1
<i>Tianyu Wang (The Chinese University of Hong Kong), Xiaowei Hu (The Chinese University of Hong Kong), Chi-Wing Fu (The Chinese University of Hong Kong), and Pheng-Ann Heng (The Chinese University of Hong Kong)</i>	
DeepMetaHandles: Learning Deformation Meta-Handles of 3D Meshes With Biharmonic Coordinates	12
<i>Minghua Liu (UCSD), Minhyuk Sung (KAIST), Radomir Mech (Adobe Systems Incorporated), and Hao Su (UCSD)</i>	
Learning Delaunay Surface Elements for Mesh Reconstruction	22
<i>Marie-Julie Rakotosaona (Ecole Polytechnique), Paul Guerrero (Adobe Research), Noam Aigerman (Adobe), Niloy J. Mitra (university college london), and Maks Ovsjanikov (Ecole polytechnique)</i>	
Fusing the Old with the New: Learning Relative Camera Pose with Geometry-Guided Uncertainty	32
<i>Bingbing Zhuang (NEC Labs) and Manmohan Chandraker (UC San Diego)</i>	
Skeleton Merger: An Unsupervised Aligned Keypoint Detector	43
<i>Ruoxi Shi (Shanghai Jiao Tong University), Zhengrong Xue (Shanghai Jiao Tong University), Yang You (Shanghai Jiao Tong University), and Cewu Lu (Shanghai Jiao Tong University)</i>	
Uncertainty Guided Collaborative Training for Weakly Supervised Temporal Action Detection	53
<i>Wenfei Yang (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China), Xiaoyuan Yu (Huawei Cloud), Tian Qi (Huawei Cloud & AI), Yongdong Zhang (University of Science and Technology of China), and Feng Wu (University of Science and Technology of China)</i>	

Pixel Codec Avatars	64
<i>Shugao Ma (Facebook), Tomas Simon (Facebook Reality Labs), Jason Saragih (Facebook), Dawei Wang (Facebook), Yuecheng Li (Facebook), Fernando De la Torre (Facebook), and Yaser Sheikh (Facebook Reality Labs)</i>	
HOTR: End-to-End Human-Object Interaction Detection With Transformers	74
<i>Bumsoo Kim (Korea University), Junhyun Lee (Korea University), Jaewoo Kang (Korea University), Eun-Sol Kim (Kakao Brain), and Hyunwoo J. Kim (Korea University)</i>	
Tuning IR-Cut Filter for Illumination-Aware Spectral Reconstruction From RGB	84
<i>Bo Sun (University of Southern California), Junchi Yan (Shanghai Jiao Tong University), Xiao Zhou (Hefei Normal University), and Yinqiang Zheng (The University of Tokyo)</i>	
DeFlow: Learning Complex Image Degradations From Unpaired Data With Conditional Flows	94
<i>Valentin Wolf (ETH Zurich), Andreas Lugmayr (ETH Zurich), Martin Danelljan (ETH Zurich), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)</i>	
AQD: Towards Accurate Quantized Object Detection	104
<i>Peng Chen (The University of Adelaide), Jing Liu (Monash University), Bohan Zhuang (Monash University), Mingkui Tan (South China University of Technology), and Chunhua Shen (University of Adelaide)</i>	
Privacy-Preserving Collaborative Learning With Automatic Transformation Search	114
<i>Wei Gao (Nanyang Technological University), Shangwei Guo (Chongqing University), Tianwei Zhang (Nanyang Technological University), Han Qiu (Telecom Paris), Yonggang Wen (Nanyang Technological University), and Yang Liu (Nanyang Technology University, Singapore)</i>	
Rethinking and Improving the Robustness of Image Style Transfer	124
<i>Pei Wang (UC San Diego), Yijun Li (Adobe Research), and Nuno Vasconcelos (UC San Diego)</i>	
Style-Aware Normalized Loss for Improving Arbitrary Style Transfer	134
<i>Jiaxin Cheng (USC Information Sciences Institute), Ayush Jaiswal (Amazon.com Inc.), Yue Wu (Amazon.com Inc.), Pradeep Natarajan (Amazon.com Inc.), and Prem Natarajan (Amazon.com Inc.)</i>	
Faster Meta Update Strategy for Noise-Robust Deep Learning	144
<i>Youjiang Xu (Baidu), Linchao Zhu (University of Technology, Sydney), Lu Jiang (Google Research), and Yi Yang (UTS)</i>	
A Hyperbolic-to-Hyperbolic Graph Convolutional Network	154
<i>Jindou Dai (Beijing Institute of Technology), Yuwei Wu (Beijing Institute of Technology, (BIT), China), Zhi Gao (Beijing Institute of Technology), and Yunde Jia (Beijing Institute of Technology)</i>	
Quasi-Dense Similarity Learning for Multiple Object Tracking	164
<i>Jiangmiao Pang (Zhejiang University), Linlu Qiu (Georgia Institute of Technology), Xia Li (ETH Zurich), Haofeng Chen (Stanford University), Qi Li (Zhejiang University), Trevor Darrell (UC Berkeley), and Fisher Yu (ETH Zurich)</i>	
MetricOpt: Learning To Optimize Black-Box Evaluation Metrics	174
<i>Chen Huang (Apple), Shuangfei Zhai (Apple), Pengsheng Guo (Apple), and Josh Susskind (Apple)</i>	

Training Networks in Null Space of Feature Covariance for Continual Learning	184
<i>Shipeng Wang (Xi'an Jiaotong University), Xiaorong Li (Xi'an Jiaotong University), Jian Sun (Xi'an Jiaotong University), and Zongben Xu (Xi'an Jiaotong University)</i>	
Exponential Moving Average Normalization for Self-Supervised and Semi-Supervised Learning ...	194
<i>Zhaowei Cai (Amazon), Avinash Ravichandran (Amazon), Subhransu Maji (University of Massachusetts, Amherst), Charless Fowlkes (UC Irvine), Zhuowen Tu (UC San Diego), and Stefano Soatto (UCLA)</i>	
Learning To Recover 3D Scene Shape From a Single Image	204
<i>Wei Yin (University of Adelaide), Jianming Zhang (Adobe Research), Oliver Wang (Adobe Systems Inc), Simon Niklaus (Adobe Research), Long Mai (Adobe Research), Simon Chen (Adobe Research), and Chunhua Shen (University of Adelaide)</i>	
Fully Convolutional Networks for Panoptic Segmentation	214
<i>Yanwei Li (The Chinese University of Hong Kong), Hengshuang Zhao (University of Oxford), Xiaojuan Qi (The University of Hong Kong), Liwei Wang (Tencent), Zeming Li (Megvii(Face++) Inc), Jian Sun (Megvii Technology), and Jiaya Jia (Chinese University of Hong Kong)</i>	
Progressive Domain Expansion Network for Single Domain Generalization	224
<i>Lei Li (Institute of Computing Technology, Chinese), Ke Gao (Institute of Computing Technology, Chinese), Juan Cao (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Ziyao Huang (Institute of Computing Technology, ICT), Chinese Academy of Sciences, (CAS)), Yepeng Weng (Institute of Computing Technology, Chinese Academy of Sciences), Xiaoyue Mi (Institute of Computing Technology, Chinese Academy of Sciences), Zhengze Yu (Institute of Computing Technology, ICT), Chinese Academy of Sciences, (CAS)), Xiaoya Li (Institute of Computing Technology, Chinese Academy of Sciences), and Boyang Xia (Institute of Computing Technology, Chinese Academy of Science)</i>	
Sketch, Ground, and Refine: Top-Down Dense Video Captioning	234
<i>Chaorui Deng (University of Adelaide), Shizhe Chen (INRIA), Da Chen (Alibaba Group), Yuan He (Alibaba Group), and Qi Wu (University of Adelaide)</i>	
Shared Cross-Modal Trajectory Prediction for Autonomous Driving	244
<i>Chiho Choi (Honda Research Institute US), Joon Hee Choi (Sungkyunkwan University), Jiachen Li (University of California, Berkeley), and Srikanth Malla (Honda Research Institute)</i>	
Glancing at the Patch: Anomaly Localization With Global and Local Feature Comparison	254
<i>Shenzhi Wang (SenseTime), Liwei Wu (SenseTime Research), Lei Cui (Tsinghua University), and Yujun Shen (Dept. of IE, CUHK)</i>	
RankDetNet: Delving Into Ranking Constraints for Object Detection	264
<i>Ji Liu (xilinx beijing), Dong Li (Xilinx), Rongzhang Zheng (Xilinx Inc.), Lu Tian (Xilinx, Inc.), and Yi Shan (Xilinx)</i>	
GAIA: A Transfer Learning System of Object Detection That Fits Your Needs	274
<i>Xingyuan Bu (SenseTime Group Limited), Junran Peng (Chinese Academy of Sciences), Junjie Yan (Sensetime Group Limited), Tieniu Tan (NLPR, China), and Zhaoxiang Zhang (Chinese Academy of Sciences, China)</i>	

Primitive Representation Learning for Scene Text Recognition	284
<i>Ruijie Yan (Tsinghua University), Liangrui Peng (Tsinghua University), Shanyu Xiao (Tsinghua University), and Gang Yao (Tsinghua University)</i>	
Keep Your Eyes on the Lane: Real-Time Attention-Guided Lane Detection	294
<i>Lucas Tabelini (UFES), Rodrigo Berriel (UFES), Thiago M. Paixão (UFES), Claudine Badue (UFES), Alberto F. De Souza (UFES), and Thiago Oliveira-Santos (UFES)</i>	
OTA: Optimal Transport Assignment for Object Detection	303
<i>Zheng Ge (Waseda University), Songtao Liu (MEGVII), Zeming Li (Megvii(Face++) Inc), Osamu Yoshie (Waseda University), and Jian Sun (Megvii Technology)</i>	
StickyPillars: Robust and Efficient Feature Matching on Point Clouds Using Graph Neural Networks	313
<i>Kai Fischer (Valeo), Martin Simon (Valeo), Florian Ölsner Ölsner (Spleenlab), Stefan Milz (Spleenlab.ai / Ilmenau University), Horst-Michael Groß (Ilmenau University of Technology, Neuroinformatics and Cognitive Robotics Lab), and Patrick Mäder (Technische Universität Ilmenau)</i>	
Semantic Scene Completion via Integrating Instances and Scene In-the-Loop	324
<i>Yingjie Cai (CUHK), Xuesong Chen (The Chinese University of Hong Kong), Chao Zhang (Samsung Telecommunication Research Institute), Kwan-Yee Lin (SenseTime Research), Xiaogang Wang (Chinese University of Hong Kong, Hong Kong), and Hongsheng Li (Chinese University of Hong Kong)</i>	
Holistic 3D Human and Scene Mesh Estimation From Single View Images	334
<i>Zhenzhen Weng (Stanford University) and Serena Yeung (Stanford University)</i>	
Point Cloud Upsampling via Disentangled Refinement	344
<i>Ruihui Li (The Chinese University of Hong Kong), Xianzhi Li (The Chinese University of Hong Kong), Pheng-Ann Heng (The Chinese University of Hong Kong), and Chi-Wing Fu (The Chinese University of Hong Kong)</i>	
DyCo3D: Robust Instance Segmentation of 3D Point Clouds Through Dynamic Convolution	354
<i>Tong He (University of Adelaide), Chunhua Shen (University of Adelaide), and Anton van den Hengel (University of Adelaide)</i>	
HCRF-Flow: Scene Flow From Point Clouds With Continuous High-Order CRFs and Position-Aware Flow Embedding	364
<i>Ruibo Li (Nanyang Technological University), Guosheng Lin (Nanyang Technological University), Tong He (University of Adelaide), Fayao Liu (Institute for Infocomm Research, A*STAR), and Chunhua Shen (University of Adelaide)</i>	
Iso-Points: Optimizing Neural Implicit Surfaces With Hybrid Representations	374
<i>Wang Yifan (ETH Zurich), Shihao Wu (ETH Zurich), Cengiz Öztireli Öztireli (University of Cambridge, Google), and Olga Sorkine-Hornung (ETH Zurich)</i>	

Fast Sinkhorn Filters: Using Matrix Scaling for Non-Rigid Shape Correspondence With Functional Maps	384
<i>Gautam Pai (École Polytechnique), Jing Ren (KAUST), Simone Melzi (Sapienza University of Rome), Peter Wonka (KAUST), and Maks Ovsjanikov (Ecole polytechnique)</i>	
Globally Optimal Relative Pose Estimation With Gravity Prior	394
<i>Yaqing Ding (Nanjing University of Science and Technology), Daniel Barath (ETH Zürich), Jian Yang (Nanjing University of Science and Technology), Hui Kong (Nanjing University of Science and Technology), and Zuzana Kukelova (Czech Technical University in Prague)</i>	
Discovering Relationships Between Object Categories via Universal Canonical Maps	404
<i>Natalia Neverova (Facebook AI Research), Artsiom Sanakoyeu (Heidelberg University), Patrick Labatut (Facebook AI Research), David Novotny (Facebook AI Research), and Andrea Vedaldi (University of Oxford / Facebook AI Research)</i>	
Neural Reprojection Error: Merging Feature Learning and Camera Pose Estimation	414
<i>Hugo Germain (Ecole des Ponts ParisTech), Vincent Lepetit (Ecole des Ponts ParisTech), and Guillaume Bourmaud (Université de Bordeaux)</i>	
Rotation-Only Bundle Adjustment	424
<i>Seong Hun Lee (University of Zaragoza) and Javier Civera (Universidad de Zaragoza)</i>	
PAUL: Procrustean Autoencoder for Unsupervised Lifting	434
<i>Chaoyang Wang (Carnegie Mellon University) and Simon Lucey (CMU)</i>	
Robust Multimodal Vehicle Detection in Foggy Weather Using Complementary Lidar and Radar Signals	444
<i>Kun Qian (UCSD), Shilin Zhu (UCSD), Xinyu Zhang (UCSD), and Li Erran Li (Amazon / Columbia University)</i>	
Depth-Conditioned Dynamic Message Propagation for Monocular 3D Object Detection	454
<i>Li Wang (Fudan University), Liang Du (Fudan University), Xiaoqing Ye (baidu), Yanwei Fu (Fudan University), Guodong Guo (Baidu), Xiangyang Xue (Fudan University), Jianfeng Feng (Fudan University), and Li Zhang (University of Oxford)</i>	
Actor-Context-Actor Relation Network for Spatio-Temporal Action Localization	464
<i>Junting Pan (The Chinese University of Hong Kong), Siyu Chen (Peking University), Mike Zheng Shou (Facebook Research), Yu Liu (The Chinese University of Hong Kong), Jing Shao (Sensetime), and Hongsheng Li (Chinese University of Hong Kong)</i>	
Temporal-Relational CrossTransformers for Few-Shot Action Recognition	475
<i>Toby Perrett (University of Bristol), Alessandro Masullo (University of Bristol), Tilo Burghardt (University of Bristol), Majid Mirmehdi (University of Bristol), and Dima Damen (University of Bristol)</i>	

Temporal Context Aggregation Network for Temporal Action Proposal Refinement	485
<i>Zhiwu Qing (Huazhong University of Science and Technology), Haisheng Su (SenseTime Group Limited), Weihao Gan (SenseTime Group Limited), Dongliang Wang (SenseTime Group Limited), Wei Wu (SenseTime Group Limited), Xiang Wang (Huazhong University of Science and Technology), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Junjie Yan (Sensetime Group Limited), Changxin Gao (Huazhong University of Science and Technology), and Nong Sang (Huazhong University of Science and Technology)</i>	
Affordance Transfer Learning for Human-Object Interaction Detection	495
<i>Zhi Hou (The University of Sydney), Baosheng Yu (The University of Sydney), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Xiaojiang Peng (Shenzhen Technology University), and Dacheng Tao (The University of Sydney)</i>	
Achieving Robustness in Classification Using Optimal Transport With Hinge Regularization	505
<i>Mathieu Serrurier (Université Toulouse Paul Sabatier IRIT), Franck Mamalet (IRT Saint-Exupery), Alberto González-Sanz (Université Toulouse Paul Sabatier Institut de Mathématiques de Toulouse), Thibaut Boissin (IRT), Jean-Michel Loubes (Université Toulouse Paul Sabatier Institut de Mathématiques de Toulouse), and Eustasio del Barrio (IMUVA)</i>	
Over-the-Air Adversarial Flickering Attacks Against Video Recognition Networks	515
<i>Roi Pony (Technion – Israel Institute of Technology), Itay Naeh (Rafael - Advanced Defense Systems Ltd.), and Shie Mannor (Technion)</i>	
Deep Dual Consecutive Network for Human Pose Estimation	525
<i>Zhenguang Liu (Zhejiang University), Haoming Chen (Zhejiang Gongshang University), Runyang Feng (Zhejiang Gongshang University), Shuang Wu (Nanyang Technological University), Shouling Ji (Zhejiang University), Bailin Yang (Zhejiang Gongshang University), and Xun Wang (Zhejiang Gongshang University)</i>	
StereoPIFu: Depth Aware Clothed Human Digitization via Stereo Vision	535
<i>Yang Hong (University of Science and Technology of China), Juyong Zhang (University of Science and Technology of China), Boyi Jiang (University of Science and Technology of China), Yudong Guo (University of Science and Technology of China), Ligang Liu (University of Science and Technology of China), and Hujun Bao (Zhejiang University)</i>	
Body Meshes as Points	546
<i>Jianfeng Zhang (NUS), Dongdong Yu (ByteDance), Jun Hao Liew (NUS), Xuecheng Nie (YiTu Inc.), and Jiashi Feng (NUS)</i>	
Cross-View Cross-Scene Multi-View Crowd Counting	557
<i>Qi Zhang (City University of Hong Kong, Hong Kong), Wei Lin (Northwestern Polytechnical University, Center for OPTical IMagery Analysis and Learning), and Antoni B. Chan (City University of Hong Kong, Hong Kong)</i>	
GANmut: Learning Interpretable Conditional Space for Gamut of Emotions	568
<i>Stefano d’Apolito (ETH Zurich), Danda Pani Paudel (ETH Zürich), Zhiwu Huang (ETH Zurich), Andrés Romero (ETH Zürich), and Luc Van Gool (ETH Zurich)</i>	

Consistent Instance False Positive Improves Fairness in Face Recognition	578
<i>Xingkun Xu (Tencent), Yuge Huang (Tencent YouTu), Pengcheng Shen (Tencent), Shaoxin Li (Tencent), Jilin Li (Tencent), Feiyue Huang (Tencent), Yong Li (Nanjing University of Science and Technology), and Zhen Cui (Nanjing University of Science and Technology)</i>	
Neural Feature Search for RGB-Infrared Person Re-Identification	587
<i>Yehansen Chen (China University of Geosciences), Lin Wan (China University of Geosciences), Zhihang Li (Institute of Automation, Chinese Academy of Science), Qianyan Jing (China University of Geosciences), and Zongyuan Sun (China University of Geosciences)</i>	
Coarse-To-Fine Person Re-Identification With Auxiliary-Domain Classification and Second-Order Information Bottleneck	598
<i>Anguo Zhang (Fuzhou University), Yueming Gao (Fuzhou University), Yuzhen Niu (Fuzhou University), Wenxi Liu (Fuzhou University), and Yongcheng Zhou (Fuzhou University)</i>	
EvDistill: Asynchronous Events To End-Task Learning via Bidirectional Reconstruction-Guided Cross-Modal Knowledge Distillation	608
<i>Lin Wang (KAIST), Yujeong Chae (KAIST), Sung-Hoon Yoon (KAIST), Tae-Kyun Kim (Imperial College London), and Kuk-Jin Yoon (KAIST)</i>	
iVPF: Numerical Invertible Volume Preserving Flow for Efficient Lossless Compression	620
<i>Shifeng Zhang (Huawei Noah's Ark Lab), Chen Zhang (Huawei Noah's Ark Lab), Ning Kang (Huawei Noah's Ark Lab), and Zhenguo Li (Huawei Noah's Ark Lab)</i>	
Deep Convolutional Dictionary Learning for Image Denoising	630
<i>Hongyi Zheng (The Hong Kong Polytechnic University), Hongwei Yong (The Hong Kong Polytechnic University), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)</i>	
Semi-Supervised Video Deraining With Dynamical Rain Generator	642
<i>Zongsheng Yue (Xi'an Jiaotong University), Jianwen Xie (Baidu Research), Qian Zhao (Xi'an Jiaotong University), and Deyu Meng (Xi'an Jiaotong University)</i>	
PPR10K: A Large-Scale Portrait Photo Retouching Dataset With Human-Region Mask and Group-Level Consistency	653
<i>Jie Liang (The Hong Kong Polytechnic University), Hui Zeng (The Hong Kong Polytechnic University), Miaomiao Cui (Alibaba-inc), Xuansong Xie (Alibaba), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)</i>	
Removing Diffraction Image Artifacts in Under-Display Camera via Dynamic Skip Connection Network	662
<i>Ruicheng Feng (Nanyang Technological University), Chongyi Li (Nanyang Technological University), Huaijin Chen (SenseBrain Technology Ltd.), Shuai Li (Sensebrain), Chen Change Loy (Nanyang Technological University), and Jinwei Gu (SenseBrain)</i>	
GAN Prior Embedded Network for Blind Face Restoration in the Wild	672
<i>Tao Yang (Alibaba), Peiran Ren (Alibaba), Xuansong Xie (Alibaba), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)</i>	

Polarimetric Normal Stereo	682
<i>Yoshiki Fukao (Kyoto University), Ryo Kawahara (Kyoto University), Shohei Nobuhara (Kyoto University), and Ko Nishino (Kyoto University)</i>	
Practical Single-Image Super-Resolution Using Look-Up Table	691
<i>Younghyun Jo (Yonsei University) and Seon Joo Kim (Yonsei University)</i>	
Deep Learning in Latent Space for Video Prediction and Compression	701
<i>Bowen Liu (University of Michigan - Ann Arbor), Yu Chen (University of Michigan), Shiyu Liu (University of Michigan - Ann Arbor), and Hun-Seok Kim (Nil)</i>	
Debiased Subjective Assessment of Real-World Image Enhancement	711
<i>Peibei Cao (City University of Hong Kong), Zhangyang Wang (University of Texas at Austin), and Kede Ma (City University of Hong Kong)</i>	
BABEL: Bodies, Action and Behavior With English Labels	722
<i>Abhinanda R. Punnakal (Max Planck Institute for Intelligent Systems), Arjun Chandrasekaran (Max Planck Institute for Intelligent Systems), Nikos Athanasiou (Max Planck Institute for Intelligent Systems), Alejandra Quirós-Ramírez (Max Planck Institute for Intelligent Systems, Tuebingen), and Michael J. Black (Max Planck Institute for Intelligent Systems)</i>	
Rethinking Channel Dimensions for Efficient Model Design	732
<i>Dongyoon Han (NAVER AI LAB), Sangdoon Yun (NAVER AI LAB), Byeongho Heo (NAVER AI LAB), and YoungJoon Yoo (Clova AI Research, NAVER Corp.)</i>	
Automated Log-Scale Quantization for Low-Cost Deep Neural Networks	742
<i>Sangyun Oh (UNIST), Hyeonuk Sim (UNIST), Sugil Lee (UNIST), and Jongeun Lee (UNIST)</i>	
ProSelfLC: Progressive Self Label Correction for Training Robust Deep Neural Networks	752
<i>Xinshao Wang (University of Oxford), Yang Hua (Queen's University Belfast), Elyor Kodirov (Anyvision), David A. Clifton (University of Oxford), and Neil M. Robertson (Queen's University Belfast)</i>	
The Lottery Ticket Hypothesis for Object Recognition	762
<i>Sharath Girish (University of Maryland, College Park), Shishira R Maiya (University of Maryland), Kamal Gupta (University of Maryland), Hao Chen (University of Maryland), Larry S. Davis (University of Maryland), and Abhinav Shrivastava (University of Maryland)</i>	
Spatial-Phase Shallow Learning: Rethinking Face Forgery Detection in Frequency Domain	772
<i>Honggu Liu (University of Science and Technology of China), Xiaodan Li (Alibaba Group), Wenbo Zhou (University of Science and Technology of China), Yuefeng Chen (Alibaba Group), Yuan He (Alibaba Group), Hui Xue (Alibaba), Weiming Zhang (University of Science and Technology of China), and Nenghai Yu (University of Science and Technology of China)</i>	
Transformer Interpretability Beyond Attention Visualization	782
<i>Hila Chefer (Tel Aviv University), Shir Gur (Tel Aviv University), and Lior Wolf (Tel Aviv University, Israel)</i>	
Mixed-Privacy Forgetting in Deep Networks	792
<i>Aditya Golatkar (University of California, Los Angeles), Alessandro Achille (Amazon Web Services), Avinash Ravichandran (Amazon), Marzia Polito (Amazon Web Services), and Stefano Soatto (UCLA)</i>	

CoSMo: Content-Style Modulation for Image Retrieval With Text Feedback	802
<i>Seungmin Lee (Seoul National University), Dongwan Kim (Seoul National University), and Bohyung Han (Seoul National University)</i>	
Discrimination-Aware Mechanism for Fine-Grained Representation Learning	813
<i>Furong Xu (Ant Group), Meng Wang (Ant group), Wei Zhang (Ant Financial Services Group), Yuan Cheng (Alibaba Group), and Wei Chu (Ant Group)</i>	
Dual Contradistinctive Generative Autoencoder	823
<i>Gaurav Parmar (UC San Diego), Dacheng Li (University of California, San Diego), Kwonjoon Lee (UC San Diego), and Zhuowen Tu (UC San Diego)</i>	
Cross-Modal Contrastive Learning for Text-to-Image Generation	833
<i>Han Zhang (Google), Jing Yu Koh (Google Research), Jason Baldridge (Google Inc.), Honglak Lee (Google / U. Michigan), and Yinfei Yang (Google Research)</i>	
Bridging the Visual Gap: Wide-Range Image Blending	843
<i>Chia-Ni Lu (National Chiao Tung University), Ya-Chu Chang (National Chiao Tung University), and Wei-Chen Chiu (National Chiao Tung University)</i>	
Exploiting Spatial Dimensions of Latent in GAN for Real-Time Image Editing	852
<i>Hyunsu Kim (NAVER AI Lab), Yunjey Choi (NAVER AI Lab), Junho Kim (NAVER AI Lab), Sungjoo Yoo (Seoul National University), and Youngjung Uh (Yonsei University)</i>	
ArtFlow: Unbiased Image Style Transfer via Reversible Neural Flows	862
<i>Jie An (University of Rochester), Siyu Huang (Nanyang Technological University), Yibing Song (Tencent), Dejing Dou (Baidu), Wei Liu (Tencent), and Jiebo Luo (U. Rochester)</i>	
DualAST: Dual Style-Learning Networks for Artistic Style Transfer	872
<i>Haibo Chen (Zhejiang University), Lei Zhao (Zhejiang University), Zhizhong Wang (Zhejiang University), Huiming Zhang (Zhejiang University), Zhiwen Zuo (Zhejiang University), Ailin Li (College of Computer Science and Technology, Zhejiang University), Wei Xing (Zhejiang University), and Dongming Lu (Zhejiang University)</i>	
Single-Shot Freestyle Dance Reenactment	882
<i>Oran Gafni (Facebook AI Research), Oron Ashual (Facebook AI Research), and Lior Wolf (Tel Aviv University, Israel)</i>	
SceneGen: Learning To Generate Realistic Traffic Scenes	892
<i>Shuhan Tan (Sun Yat-Sen University), Kelvin Wong (Uber Advanced Technologies Group, University of Toronto), Shenlong Wang (Uber ATG, University of Toronto), Sivabalan Manivasagam (University of Toronto), Mengye Ren (University of Toronto), and Raquel Urtasun (Uber ATG)</i>	
Learning Semantic-Aware Dynamics for Video Prediction	902
<i>Xinzhu Bei (UCLA), Yanchao Yang (Stanford University), and Stefano Soatto (UCLA)</i>	
Reinforced Attention for Few-Shot Learning and Beyond	913
<i>Jie Hong (Australian National University), Pengfei Fang (The Australian National University), Weihao Li (Data61, CSIRO), Tong Zhang (The Australian National University), Christian Simon (Australian National University), Mehrtash Harandi (Monash University), and Lars Petersson (Data61/CSIRO)</i>	

Fast and Accurate Model Scaling	924
<i>Piotr Dollár (FAIR), Mannat Singh (Facebook AI Research), and Ross Girshick (FAIR)</i>	
Multi-Label Learning From Single Positive Labels	933
<i>Elijah Cole (Caltech), Oisin Mac Aodha (University of Edinburgh), Titouan Lorieul (Inria), Pietro Perona (California Institute of Technology), Dan Morris (Microsoft), and Nebojsa Jojic (Microsoft Research)</i>	
Contrastive Learning Based Hybrid Networks for Long-Tailed Image Classification	943
<i>Peng Wang (University of Wollongong), Kai Han (University of Bristol), Xiu-Shen Wei (Nanjing University of Science and Technology), Lei Zhang (Northwestern Polytechnical University), and Lei Wang (University of Wollongong, Australia)</i>	
Learning Graph Embeddings for Compositional Zero-Shot Learning	953
<i>Muhammad Ferjad Naeem (ETH Zürich), Yongqin Xian (Max Planck Institute Informatics), Federico Tombari (Google, TU Munich), and Zeynep Akata (University of Tübingen)</i>	
Multispectral Photometric Stereo for Spatially-Varying Spectral Reflectances: A Well Posed Problem?	963
<i>Heng Guo (Osaka University), Fumio Okura (Osaka University), Boxin Shi (Peking University), Takuya Funatomi (Nara Institute of Science and Technology), Yasuhiro Mukaigawa (NAIST), and Yasuyuki Matsushita (Osaka University)</i>	
LiBR: A Practical Bayesian Approach to Adversarial Detection	972
<i>Zhijie Deng (Tsinghua University), Xiao Yang (Tsinghua University), Shizhen Xu (RealAI), Hang Su (Tsinghua University), and Jun Zhu (Tsinghua University)</i>	
Searching by Generating: Flexible and Efficient One-Shot NAS With Architecture Generator	983
<i>Sian-Yao Huang (National Cheng Kung University) and Wei-Ta Chu (National Cheng Kung University)</i>	
Densely Connected Multi-Dilated Convolutional Networks for Dense Prediction Tasks	993
<i>Naoya Takahashi (Sony) and Yuki Mitsufuji (Sony Corporation)</i>	
DARCNN: Domain Adaptive Region-Based Convolutional Neural Network for Unsupervised Instance Segmentation in Biomedical Images	1003
<i>Joy Hsu (Stanford University), Wah Chiu (Stanford University), and Serena Yeung (Stanford University)</i>	
FedDG: Federated Domain Generalization on Medical Image Segmentation via Episodic Learning in Continuous Frequency Space	1013
<i>Quande Liu (The Chinese University of Hong Kong), Cheng Chen (The Chinese University of Hong Kong), Jing Qin (The Hong Kong Polytechnic University), Qi Dou (The Chinese University of Hong Kong), and Pheng-Ann Heng (The Chinese University of Hong Kong)</i>	
Distractor-Aware Fast Tracking via Dynamic Convolutions and MOT Philosophy	1024
<i>Zikai Zhang (Huaqiao University), Bineng Zhong (Guangxi Normal University), Shengping Zhang (Harbin Institute of Technology), Zhenjun Tang (Guangxi Normal University), Xin Liu (Beijing Seetatech Technology Co., Ltd), and Zhaoxiang Zhang (Chinese Academy of Sciences, China)</i>	

Mining Better Samples for Contrastive Learning of Temporal Correspondence	1034
<i>Sangryul Jeon (Yonsei university), Dongbo Min (Ewha Womans University), Seungryong Kim (Korea University), and Kwanghoon Sohn (Yonsei Univ.)</i>	
UPFlow: Upsampling Pyramid for Unsupervised Optical Flow Learning	1045
<i>Kunming Luo (Megvii), Chuan Wang (Face++, (Megvii)), Shuaicheng Liu (UESTC; Megvii), Haoqiang Fan (Megvii Incface++), Jue Wang (Tencent AI Lab), and Jian Sun (Megvii Technology)</i>	
KeepAugment: A Simple Information-Preserving Data Augmentation Approach	1055
<i>Chengyue Gong (UT Austin), Dilin Wang (Facebook), Meng Li (Facebook Inc), Vikas Chandra (Facebook), and Qiang Liu (UT Austin)</i>	
Keypoint-Graph-Driven Learning Framework for Object Pose Estimation	1065
<i>Shaobo Zhang (Northwest university of China), Wanqing Zhao (Northwest university of China), Ziyu Guan (Northwest University), Xianlin Peng (Northwest University), and Jinye Peng (Northwest University, Xi'an)</i>	
AdCo: Adversarial Contrast for Efficient Learning of Unsupervised Representations From Self-Trained Negative Adversaries	1074
<i>Qianjiang Hu (Peking University), Xiao Wang (Purdue University), Wei Hu (Peking University), and Guo-Jun Qi (Futurewei Technologies)</i>	
Generalized Domain Adaptation	1084
<i>Yu Mitsuzumi (NTT Corporation), Go Irie (NTT Corporation), Daiki Ikami (NTT Corporation), and Takashi Shibata (NTT Corporation)</i>	
FixBi: Bridging Domain Spaces for Unsupervised Domain Adaptation	1094
<i>Jaemin Na (Ajou University), Heechul Jung (Kyungpook National University), Hyung Jin Chang (University of Birmingham), and Wonjun Hwang (Ajou University)</i>	
Learning Invariant Representations and Risks for Semi-Supervised Domain Adaptation	1104
<i>Bo Li (UC Berkeley), Yezhen Wang (University of California San Diego), Shanghang Zhang (UC Berkeley), Dongsheng Li (Microsoft Research Asia), Kurt Keutzer (EECS, UC Berkeley), Trevor Darrell (UC Berkeley), and Han Zhao (University of Illinois at Urbana-Champaign)</i>	
Continual Semantic Segmentation via Repulsion-Attraction of Sparse and Disentangled Latent Representations	1114
<i>Umberto Michieli (University of Padova) and Pietro Zanuttigh (University of Padova)</i>	
Incremental Learning via Rate Reduction	1125
<i>Ziyang Wu (Cornell University), Christina Baek (UC Berkeley), Chong You (University of California, Berkeley), and Yi Ma (UC Berkeley)</i>	
Partially View-Aligned Representation Learning With Noise-Robust Contrastive Loss	1134
<i>Mouxing Yang (Sichuan University), Yunfan Li (Sichuan University), Zhenyu Huang (Sichuan University), Zitao Liu (TAL Education Group), Peng Hu (College of Computer Science, Sichuan University), and Xi Peng (College of Computer Science, Sichuan Univerisity)</i>	
Spatially Consistent Representation Learning	1144
<i>Byungseok Roh (KakaoBrain), Wuhyun Shin (Kakaobrain), Ildoo Kim (Kakao Brain), and Sungwoong Kim (Kakao Brain)</i>	

Mask Guided Matting via Progressive Refinement Network	1154
<i>Qihang Yu (Johns Hopkins University), Jianming Zhang (Adobe Research), He Zhang (Adobe), Yilin Wang (Adobe), Zhe Lin (Adobe Research), Ning Xu (Adobe Research), Yutong Bai (Johns Hopkins University), and Alan Yuille (Johns Hopkins University)</i>	
The Temporal Opportunist: Self-Supervised Multi-Frame Monocular Depth	1164
<i>Jamie Watson (Niantic), Oisin Mac Aodha (University of Edinburgh), Victor Prisacariu (University of Oxford), Gabriel Brostow (University College London), and Michael Firman (Niantic)</i>	
Exemplar-Based Open-Set Panoptic Segmentation Network	1175
<i>Jaedong Hwang (Seoul National University), Seoung Wug Oh (Adobe Research), Joon-Young Lee (Adobe Research), and Bohyung Han (Seoul National University)</i>	
Incremental Few-Shot Instance Segmentation	1185
<i>Dan Andrei Ganea (Cyclomedia Technology), Bas Boom (Cyclomedia), and Ronald Poppe (Utrecht University)</i>	
DoDNet: Learning To Segment Multi-Organ and Tumors From Multiple Partially Labeled Datasets	1195
<i>Jianpeng Zhang (Northwestern Polytechnical University), Yutong Xie (Northwestern Polytechnical University), Yong Xia (Northwestern Polytechnical University, Research & Development Institute of Northwestern Polytechnical University in Shenzhen), and Chunhua Shen (University of Adelaide)</i>	
Semi-Supervised Semantic Segmentation With Directional Context-Aware Consistency	1205
<i>Xin Lai (The Chinese University of Hong Kong), Zhuotao Tian (The Chinese University of Hong Kong), Li Jiang (The Chinese University of Hong Kong), Shu Liu (SmartMore), Hengshuang Zhao (University of Oxford), Liwei Wang (Tencent), and Jiaya Jia (Chinese University of Hong Kong)</i>	
Source-Free Domain Adaptation for Semantic Segmentation	1215
<i>Yuang Liu (East China Normal University), Wei Zhang (East China Normal University), and Jun Wang (ECNU)</i>	
Learning the Superpixel in a Non-Iterative and Lifelong Manner	1225
<i>Lei Zhu (Beijing University of Posts and Telecommunications), Qi She (Bytedance AI Lab), Bin Zhang (Bupt), Yanye Lu (Peking University), Zhilin Lu (Tsinghua), Duo Li (HKUST), and Jie Hu (Institute of Software Chinese Academy of Sciences)</i>	
ATSO: Asynchronous Teacher-Student Optimization for Semi-Supervised Image Segmentation ..	1235
<i>Xinyue Huo (University of Science and Technology of China), Lingxi Xie (Huawei Inc.), Jianzhong He (Peking University), Zijie Yang (Chinese Academy of Sciences), Wengang Zhou (University of Science and Technology of China), Houqiang Li (University of Science and Technology of China), and Qi Tian (Huawei Cloud & AI)</i>	
Can We Characterize Tasks Without Labels or Features?	1245
<i>Bram Wallace (Cornell University), Ziyang Wu (Cornell University), and Bharath Hariharan (Cornell University)</i>	

Reconsidering Representation Alignment for Multi-View Clustering	1255
<i>Daniel J. Trosten (UiT The Arctic University of Norway), Sigurd Løkse (UiT The Arctic University of Norway), Robert Jenssen (UiT - The Arctic University of Norway), and Michael Kampffmeyer (UiT The Arctic University of Norway)</i>	
Learning To Segment Rigid Motions From Two Frames	1266
<i>Gengshan Yang (Carnegie Mellon University) and Deva Ramanan (Carnegie Mellon University)</i>	
Self-Supervised Motion Learning From Static Images	1276
<i>Ziyuan Huang (National University of Singapore), Shiwei Zhang (DAMO Academy, Alibaba Group), Jianwen Jiang (Alibaba Group), Mingqian Tang (Alibaba Group), Rong Jin (alibaba group), and Marcelo H. Ang (National University of Singapore)</i>	
Efficient Regional Memory Network for Video Object Segmentation	1286
<i>Haozhe Xie (Harbin Institute of Technology), Hongxun Yao (Harbin Institute of Technology), Shangchen Zhou (Nanyang Technological University), Shengping Zhang (Harbin Institute of Technology), and Wenxiu Sun (SenseTime Research and Tetras.AI)</i>	
SwiftNet: Real-Time Video Object Segmentation	1296
<i>Haochen Wang (Alibaba Youku Cognitive and Intelligent Lab), Xiaolong Jiang (Alibaba Youku Cognitive and Intelligent Lab), Haibing Ren (Alibaba Group), Yao Hu (Alibaba Youku Cognitive and Intelligent Lab), and Song Bai (University of Oxford)</i>	
Improving OCR-Based Image Captioning by Incorporating Geometrical Relationship	1306
<i>Jing Wang (Nanjing University of Science and Technology), Jinhui Tang (Nanjing University of Science and Technology), Mingkun Yang (Huazhong University of Science and Technology), Xiang Bai (Huazhong University of Science and Technology), and Jiebo Luo (U. Rochester)</i>	
Improving Sign Language Translation With Monolingual Data by Sign Back-Translation	1316
<i>Hao Zhou (University of Science and Technology of China), Wengang Zhou (University of Science and Technology of China), Weizhen Qi (University of Science and Technology of China), Junfu Pu (University of Science and Technology of China), and Houqiang Li (University of Science and Technology of China)</i>	
Exploring Heterogeneous Clues for Weakly-Supervised Audio-Visual Video Parsing	1326
<i>Yu Wu (University of Technology Sydney) and Yi Yang (UTS)</i>	
Looking Into Your Speech: Learning Cross-Modal Affinity for Audio-Visual Speech Separation....	1336
<i>Jiyoung Lee (Yonsei University), Soo-Whan Chung (Yonsei University), Sunok Kim (Yonsei University), Hong-Goo Kang (Yonsei University), and Kwanghoon Sohn (Yonsei Univ.)</i>	
Co-Grounding Networks With Semantic Attention for Referring Expression Comprehension in Videos	1346
<i>Sijie Song (Peking University), Xudong Lin (Columbia University), Jiaying Liu (Peking University), Zongming Guo (Peking University), and Shih-Fu Chang (Columbia University)</i>	
Explicit Knowledge Incorporation for Visual Reasoning	1356
<i>Yifeng Zhang (University of Minnesota, Twin Cities), Ming Jiang (University of Minnesota), and Qi Zhao (University of Minnesota)</i>	

Deep Gradient Projection Networks for Pan-sharpening	1366
<i>Shuang Xu (Xi'an Jiaotong University), Jianshe Zhang (Xi'an Jiaotong University), Zixiang Zhao (Xi'an Jiaotong University), Kai Sun (Xi'an Jiaotong University), Junmin Liu (Xi'an Jiaotong University), and Chunxia Zhang (Xi'an Jiaotong University)</i>	
Capturing Omni-Range Context for Omnidirectional Segmentation	1376
<i>Kailun Yang (Karlsruhe Institute of Technology), Jiaming Zhang (Karlsruhe Institute of Technology), Simon Reiß (Karlsruhe Institute of Technology), Xinxin Hu (Huawei), and Rainer Stiefelhagen (Karlsruhe Institute of Technology)</i>	
Gradient-Based Algorithms for Machine Teaching	1387
<i>Pei Wang (UC San Diego), Kabir Nagrecha (UC San Diego), and Nuno Vasconcelos (UC San Diego)</i>	
Depth From Camera Motion and Object Detection	1397
<i>Brent A. Griffin (University of Michigan) and Jason J. Corso (University of Michigan)</i>	

Session 02

Deep RGB-D Saliency Detection With Depth-Sensitive Attention and Automatic Multi-Modal Fusion	1407
<i>Peng Sun (Zhejiang University), Wenhui Zhang (Zhejiang University), Huanyu Wang (Zhejiang University), Songyuan Li (Zhejiang University), and Xi Li (Zhejiang University)</i>	
SAIL-VOS 3D: A Synthetic Dataset and Baselines for Object Detection and 3D Mesh Reconstruction From Video Data	1418
<i>Yuan-Ting Hu (UIUC), Jiahong Wang (UIUC), Raymond A. Yeh (UIUC), and Alexander G. Schwing (UIUC)</i>	
Deep Implicit Templates for 3D Shape Representation	1429
<i>Zerong Zheng (Tsinghua University), Tao Yu (Tsinghua University), Qionghai Dai (Tsinghua University), and Yebin Liu (Tsinghua University)</i>	
Pulsar: Efficient Sphere-Based Neural Rendering	1440
<i>Christoph Lassner (Facebook Reality Labs) and Michael Zollhöfer (Facebook Reality Labs)</i>	
Neural Deformation Graphs for Globally-Consistent Non-Rigid Reconstruction	1450
<i>Aljaž Božič (Technical University Munich), Pablo Palafox (Technical University Munich), Michael Zollhöfer (Facebook Reality Labs), Justus Thies (Technical University of Munich), Angela Dai (Technical University of Munich), and Matthias Nießner (Technical University of Munich)</i>	
Modeling Multi-Label Action Dependencies for Temporal Action Localization	1460
<i>Praveen Tirupattur (University of Central Florida), Kevin Duarte (University of Central Florida), Yogesh S Rawat (University of Central Florida), and Mubarak Shah (University of Central Florida)</i>	

ContactOpt: Optimizing Contact To Improve Grasps	1471
<i>Patrick Grady (Georgia Institute of Technology), Chengcheng Tang (Facebook Reality Labs), Christopher D. Twigg (Facebook Reality Labs), Minh Vo (Facebook Reality Labs), Samarth Brahmhatt (Intel), and Charles C. Kemp (Georgia Institute of Technology)</i>	
From Synthetic to Real: Unsupervised Domain Adaptation for Animal Pose Estimation	1482
<i>Chen Li (National University of Singapore) and Gim Hee Lee (National University of Singapore)</i>	
Deep Homography for Efficient Stereo Image Compression	1492
<i>Xin Deng (Beihang university), Wenzhe Yang (Beihang University), Ren Yang (ETH Zurich), Mai Xu (BUAA), Enpeng Liu (Beihang University), Qianhan Feng (Beihang University), and Radu Timofte (ETH Zurich)</i>	
FVC: A New Framework Towards Deep Video Compression in Feature Space	1502
<i>Zhihao Hu (Beihang University), Guo Lu (Beijing Institute of Technology), and Dong Xu (University of Sydney)</i>	
Zero-Shot Adversarial Quantization	1512
<i>Yuang Liu (East China Normal University), Wei Zhang (East China Normal University), and Jun Wang (ECNU)</i>	
Farewell to Mutual Information: Variational Distillation for Cross-Modal Person Re-Identification	1522
<i>Xudong Tian (East China Normal University), Zhizhong Zhang (East China Normal University), Shaohui Lin (East China Normal University), Yanyun Qu (XMU), Yuan Xie (East China Normal University), and Lizhuang Ma (Shanghai Jiao Tong University)</i>	
Closed-Form Factorization of Latent Semantics in GANs	1532
<i>Yujun Shen (Dept. of IE, CUHK) and Bolei Zhou (CUHK)</i>	
High-Fidelity Neural Human Motion Transfer From Monocular Video	1541
<i>Moritz Kappel (Technische Universität Braunschweig), Vladislav Golyanik (MPI for Informatics), Mohamed Elgharib (Max Planck Institute for Informatics), Jann-Ole Henningson (Technische Universität Braunschweig), Hans-Peter Seidel (Max Planck Institute for Informatics), Susana Castillo (Technische Universität Braunschweig), Christian Theobalt (MPI Informatik), and Marcus Magnor (TU Braunschweig)</i>	
Correlated Input-Dependent Label Noise in Large-Scale Image Classification	1551
<i>Mark Collier (Google), Basil Mustafa (Google), Efi Kokiopoulou (Google AI), Rodolphe Jenatton (Google), and Jesse Berent (Google)</i>	
Bi-GCN: Binary Graph Convolutional Network	1561
<i>Junfu Wang (Beihang University, China), Yunhong Wang (State Key Laboratory of Virtual Reality Technology and System, Beihang University, China), Zhen Yang (Beihang University), Liang Yang (Hebei University of Technology), and Yuanfang Guo (Beihang University)</i>	
Transformer Meets Tracker: Exploiting Temporal Context for Robust Visual Tracking	1571
<i>Ning Wang (University of Science and Technology of China), Wengang Zhou (University of Science and Technology of China), Jie Wang (University of Science and Technology of China), and Houqiang Li (University of Science and Technology of China)</i>	

FS-Net: Fast Shape-Based Network for Category-Level 6D Object Pose Estimation With Decoupled Rotation Mechanism	1581
<i>Wei Chen (University of Birmingham), Xi Jia (University of Birmingham), Hyung Jin Chang (University of Birmingham), Jinming Duan (University of Birmingham), Linlin Shen (Shenzhen University), and Aleš Leonardis (University of Birmingham)</i>	
On Learning the Geodesic Path for Incremental Learning	1591
<i>Christian Simon (Australian National University), Piotr Koniusz (Data61/CSIRO, ANU), and Mehrtaash Harandi (Monash University)</i>	
UP-DETR: Unsupervised Pre-Training for Object Detection With Transformers	1601
<i>Zhigang Dai (South China University of Technology), Bolun Cai (Tencent Wechat AI), Yugeng Lin (Tencent Wechat AI), and Junying Chen (South China University of Technology)</i>	
Robust Consistent Video Depth Estimation	1611
<i>Johannes Kopf (Facebook), Xuejian Rong (Facebook), and Jia-Bin Huang (Virginia Tech)</i>	
Differentiable Multi-Granularity Human Representation Learning for Instance-Aware Human Semantic Parsing	1622
<i>Tianfei Zhou (ETH Zurich), Wenguan Wang (Eidgenössische Technische Hochschule Zürich), Si Liu (Beihang University), Yi Yang (UTS), and Luc Van Gool (ETH Zurich)</i>	
Global Transport for Fluid Reconstruction With Learned Self-Supervision	1632
<i>Erik Franz (Technical University of Munich), Barbara Solenthaler (ETH Zürich), and Nils Thuerey (Technical University of Munich)</i>	
VLN BERT: A Recurrent Vision-and-Language BERT for Navigation	1643
<i>Yicong Hong (Australian National University), Qi Wu (University of Adelaide), Yuankai Qi (The University of Adelaide), Cristian Rodriguez-Opazo (University of Adelaide), and Stephen Gould (Australian National University, Australia)</i>	
Single-View Robot Pose and Joint Angle Estimation via Render & Compare	1654
<i>Yann Labbé (Inria), Justin Carpentier (INRIA), Mathieu Aubry (École des ponts ParisTech), and Josef Sivic (Czech Technical University)</i>	
Learning Deep Classifiers Consistent With Fine-Grained Novelty Detection	1664
<i>Jiacheng Cheng (University of California, San Diego) and Nuno Vasconcelos (UC San Diego)</i>	
CRFace: Confidence Ranker for Model-Agnostic Face Detection Refinement	1674
<i>Noranart Vesdapunt (Microsoft Cloud&AI) and Baoyuan Wang (Xiaobing.AI)</i>	
Equalization Loss v2: A New Gradient Balance Approach for Long-Tailed Object Detection	1685
<i>Jingru Tan (Tongji University), Xin Lu (SenseTime Group Limited), Gang Zhang (Tsinghua University), Changqing Yin (Tongji University), and Quanquan Li (SenseTime Research)</i>	
Semantic-Aware Video Text Detection	1695
<i>Wei Feng (CAS), Fei Yin (Institute of Automation of Chinese Academy of Sciences), Xu-Yao Zhang (Institute of Automation of Chinese Academy of Sciences), and Cheng-Lin Liu (Institute of Automation of Chinese Academy of Sciences)</i>	

Improved Handling of Motion Blur in Online Object Detection	1706
<i>Mohamed Sayed (University College London) and Gabriel Brostow (University College London)</i>	
IQDet: Instance-Wise Quality Distribution Sampling for Object Detection	1717
<i>Yuchen Ma (Megvii), Songtao Liu (MEGVII), Zeming Li (Megvii(Face++ Inc), and Jian Sun (Megvii Technology)</i>	
One Thing One Click: A Self-Training Approach for Weakly Supervised 3D Semantic Segmentation	1726
<i>Zhengzhe Liu (The Chinese University of Hong Kong), Xiaojuan Qi (The University of Hong Kong), and Chi-Wing Fu (The Chinese University of Hong Kong)</i>	
Learning Monocular 3D Reconstruction of Articulated Categories From Motion	1737
<i>Filippos Kokkinos (University College London) and Iasonas Kokkinos (Snap / University College London)</i>	
SPSG: Self-Supervised Photometric Scene Generation From RGB-D Scans	1747
<i>Angela Dai (Technical University of Munich), Yawar Siddiqui (Technical University of Munich), Justus Thies (Technical University of Munich), Julien Valentin (Google), and Matthias Nießner (Technical University of Munich)</i>	
Semantic Segmentation for Real Point Cloud Scenes via Bilateral Augmentation and Adaptive Fusion	1757
<i>Shi Qiu (ANU), Saeed Anwar (The Australian National University), and Nick Barnes (ANU)</i>	
Unsupervised 3D Shape Completion Through GAN Inversion	1768
<i>Junzhe Zhang (Nanyang Technological University), Xinyi Chen (SenseTime International Pte Ltd), Zhongang Cai (SenseTime International Pte Ltd), Liang Pan (Nanyang Technological University), Haiyu Zhao (SenseTime International Pte Ltd), Shuai Yi (SenseTime Group Limited), Chai Kiat Yeo (Nanyang Technological University), Bo Dai (Nanyang Technological University), and Chen Change Loy (Nanyang Technological University)</i>	
3D AffordanceNet: A Benchmark for Visual Object Affordance Understanding	1778
<i>Shengheng Deng (South China University of Technology), Xun Xu (I2R, ASTAR), Chaozheng Wu (South China University of Technology), Ke Chen (South China University of Technology), and Kui Jia (South China University of Technology)</i>	
Deep Implicit Moving Least-Squares Functions for 3D Reconstruction	1788
<i>Shi-Lin Liu (University of Science and Technology of China), Hao-Xiang Guo (Tsinghua University), Hao Pan (Microsoft Research), Peng-Shuai Wang (Microsoft Research Asia), Xin Tong (Microsoft), and Yang Liu (MSR, Beijing)</i>	
Using Shape To Categorize: Low-Shot Learning With an Explicit Shape Bias	1798
<i>Stefan Stojanov (Georgia Institute of Technology), Anh Thai (Georgia Institute of Technology), and James M. Rehg (Georgia Institute of Technology)</i>	

Privacy Preserving Localization and Mapping From Uncalibrated Cameras	1809
<i>Marcel Geppert (ETH Zurich), Viktor Larsson (ETH Zurich), Pablo Speciale (Microsoft), Johannes L. Schönberger (Microsoft), and Marc Pollefeys (ETH Zurich / Microsoft)</i>	
HumanGPS: Geodesic PreServing Feature for Dense Human Correspondences	1820
<i>Feitong Tan (Simon Fraser University), Danhang Tang (Google), Mingsong Dou (Google Inc.), Kaiwen Guo (Google), Rohit Pandey (Google), Cem Keskin (Facebook), Ruofei Du (Google), Deqing Sun (Google), Sofien Bouaziz (Google), Sean Fanello (Google), Ping Tan (Simon Fraser University), and Yinda Zhang (Google)</i>	
Learning Camera Localization via Dense Scene Matching	1831
<i>Shitao Tang (Simon Fraser University), Chengzhou Tang (Simon Fraser University), Rui Huang (Alibaba A.I Labs), Siyu Zhu (Alibaba A.I. Labs), and Ping Tan (Simon Fraser University)</i>	
PlückerNet: Learn To Register 3D Line Reconstructions	1842
<i>Liu Liu (ANU, (Australian National University)), Hongdong Li (Australian National University, Australia), Haodong Yao (Australian National University, Australia), and Ruyi Zha (Australian National University, Australia)</i>	
MultiLink: Multi-Class Structure Recovery via Agglomerative Clustering and Model Selection.....	1853
<i>Luca Magri (Politecnico di Milano), Filippo Leveni (Politecnico di Milano), and Giacomo Boracchi (Politecnico di Milano)</i>	
3D-MAN: 3D Multi-Frame Attention Network for Object Detection	1863
<i>Zetong Yang (The Chinese University of Hong Kong), Yin Zhou (Waymo), Zhifeng Chen (Google), and Jiquan Ngiam (Google Brain)</i>	
Exploring intermediate representation for monocular vehicle pose estimation	1873
<i>Shichao Li (Hong Kong University of Science and Technology), Zengqiang Yan (Hong Kong University of Science and Technology), Hongyang Li (SenseTime), and Kwang-Ting Cheng (Hong Kong University of Science and Technology)</i>	
Towards Long-Form Video Understanding	1884
<i>Chao-Yuan Wu (UT Austin) and Philipp Krähenbühl (UT Austin)</i>	
TDN: Temporal Difference Networks for Efficient Action Recognition	1895
<i>Limin Wang (Nanjing University), Zhan Tong (Nanjing University), Bin Ji (Nanjing University), and Gangshan Wu (Nanjing University)</i>	
Self-Supervised Learning for Semi-Supervised Temporal Action Proposal	1905
<i>Xiang Wang (Huazhong University of Science and Technology), Shiwei Zhang (DAMO Academy, Alibaba Group), Zhiwu Qing (Huazhong University of Science and Technology), Yuanjie Shao (Huazhong University of Science and Technology), Changxin Gao (Huazhong University of Science and Technology), and Nong Sang (Huazhong University of Science and Technology)</i>	
WOAD: Weakly Supervised Online Action Detection in Untrimmed Videos	1915
<i>Mingfei Gao (Salesforce Research), Yingbo Zhou (Salesforce Research), Ran Xu (Salesforce Research), Richard Socher (Salesforce Research), and Caiming Xiong (Salesforce Research)</i>	

Enhancing the Transferability of Adversarial Attacks Through Variance Tuning	1924
<i>Xiaosen Wang (Huazhong University of Science and Technology) and Kun He (Huazhong University of Science and Technology)</i>	
You See What I Want You To See: Exploring Targeted Black-Box Transferability Attack for Hash-Based Image Retrieval Systems	1934
<i>Yanru Xiao (Old Dominion University) and Cong Wang (Old Dominion University)</i>	
Pose Recognition With Cascade Transformers	1944
<i>Ke Li (University of Chinese Academy of Sciences), Shijie Wang (Tsinghua University), Xiang Zhang (Tsinghua University), Yifan Xu (UC San Diego), Weijian Xu (University of California, San Diego), and Zhuowen Tu (UC San Diego)</i>	
End-to-End Human Pose and Mesh Reconstruction with Transformers	1954
<i>Kevin Lin (Microsoft), Lijuan Wang (Microsoft), and Zicheng Liu (Microsoft)</i>	
Beyond Static Features for Temporally Consistent 3D Human Pose and Shape From a Video	1964
<i>Hongsuk Choi (Seoul National University), Gyeongsik Moon (Seoul National University), Ju Yong Chang (Kwangwoon University), and Kyoung Mu Lee (Seoul National University)</i>	
A Generalized Loss Function for Crowd Counting and Localization	1974
<i>Jia Wan (City University of Hong Kong), Ziquan Liu (City University of Hong Kong), and Antoni B. Chan (City University of Hong Kong, Hong Kong)</i>	
LOHO: Latent Optimization of Hairstyles via Orthogonalization	1984
<i>Rohit Saha (University of Toronto), Brendan Duke (ModiFace Inc), Florian Shkurti (University of Toronto), Graham W. Taylor (University of Guelph), and Parham Aarabi (ModiFace Inc.)</i>	
Pseudo Facial Generation With Extreme Poses for Face Recognition	1994
<i>Guoli Wang (Tsinghua University), Jiaqi Ma (Wuhan University), Qian Zhang (Horizon Robotics), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)</i>	
Joint Generative and Contrastive Learning for Unsupervised Person Re-Identification	2004
<i>Hao Chen (INRIA), Yaohui Wang (INRIA), Benoit Lagadec (European System Integration), Antitza Dantcheva (INRIA), and Francois Bremond (Inria Sophia Antipolis, France)</i>	
BiCnet-TKS: Learning Efficient Spatial-Temporal Representation for Video Person Re-Identification	2014
<i>Ruibing Hou (Institute of Computing Technology, Chinese Academy of Sciences), Hong Chang (Chinese Academy of Sciences), Bingpeng Ma (University of Chinese Academy of Sciences), Rui Huang (The Chinese University of Hong Kong, Shenzhen), and Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences)</i>	
Learning To Reconstruct High Speed and High Dynamic Range Videos From Events	2024
<i>Yunhao Zou (Beijing Institute of Technology), Yinqiang Zheng (The University of Tokyo), Tsuyoshi Takatani (University of Tsukuba), and Ying Fu (Beijing Institute of Technology)</i>	

Iterative Filter Adaptive Network for Single Image Defocus Deblurring	2034
<i>Junyong Lee (POSTECH), Hyeongseok Son (POSTECH), Jaesung Rim (POSTECH), Sunghyun Cho (POSTECH), and Seungyong Lee (POSTECH)</i>	
Recorrputed-to-Recorrputed: Unsupervised Deep Learning for Image Denoising	2043
<i>Tongyao Pang (National University of Singapore), Huan Zheng (National University of Singapore), Yuhui Quan (South China University of Technology), and Hui Ji (National University of Singapore)</i>	
Closing the Loop: Joint Rain Generation and Removal via Disentangled Image Translation	2053
<i>Yuntong Ye (Huazhong University of Science and Technology), Yi Chang (Pengcheng Laboratory), Hanyu Zhou (Huazhong University of Science and Technology), and Luxin Yan (Huazhong University of Science and Technology)</i>	
Deep Denoising of Flash and No-Flash Pairs for Photography in Low-Light Environments	2063
<i>Zhihao Xia (Washington University in St. Louis), Michaël Gharbi (Adobe Research), Federico Perazzi (Facebook, Inc.), Kalyan Sunkavalli (Adobe Research), and Ayan Chakrabarti (Washington University in St. Louis)</i>	
Controllable Image Restoration for Under-Display Camera in Smartphones	2073
<i>Kinam Kwon (Samsung Advanced Institute of Technology, (SAIT)), Eunhee Kang (Samsung Electronics), Sangwon Lee (SAIT), Su-Jin Lee (Samsung), Hyong-Euk Lee (Samsung Advanced Institute of Technology), ByungIn Yoo (Samsung Advanced Institute of Technology), and Jae-Joon Han (Samsung)</i>	
MetaSCI: Scalable and Adaptive Reconstruction for Video Compressive Sensing	2083
<i>Zhengjue Wang (New Jersey Institute of Technology), Hao Zhang (Cornell University), Ziheng Cheng (Xidian University), Bo Chen (Xidian University), and Xin Yuan (Bell Labs)</i>	
Learning the Non-Differentiable Optimization for Blind Super-Resolution	2093
<i>Zheng Hui (Xidian University), Jie Li (Xidian University), Xiumei Wang (Xidian University), and Xinbo Gao (Xidian University)</i>	
Robust Reference-Based Super-Resolution via C2-Matching	2103
<i>Yuming Jiang (Nanyang Technological University), Kelvin C.K. Chan (Nanyang Technological University), Xintao Wang (Tencent), Chen Change Loy (Nanyang Technological University), and Ziwei Liu (Nanyang Technological University)</i>	
Space-Time Distillation for Video Super-Resolution	2113
<i>Zeyu Xiao (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), Jie Huang (University of Science and Technology of China), Zhen Cheng (University of Science and Technology of China), and Zhiwei Xiong (University of Science and Technology of China)</i>	
Person30K: A Dual-Meta Generalization Network for Person Re-Identification	2123
<i>Yan Bai (Peking University), Jile Jiao (Alibaba Group), Wang Ce (Peking University), Jun Liu (Singapore University of Technology and Design), Yihang Lou (Peking University), Xuetao Feng (Alibaba Group), and Ling-Yu Duan (Peking University)</i>	
Zillow Indoor Dataset: Annotated Floor Plans With 360° Panoramas and 3D Room Layouts	2133
<i>Steve Cruz (University of Colorado), Will Hutchcroft (Zillow Group), Yuguang Li (Zillow Group), Naji Khosravan (Zillow Group), Ivaylo Boyadzhiev (Zillow Group), and Sing Bing Kang (Zillow Group)</i>	

The Heterogeneity Hypothesis: Finding Layer-Wise Differentiated Network Architectures	2144
<i>Yawei Li (ETH Zurich), Wen Li (University of Electronic Science and Technology of China), Martin Danelljan (ETH Zurich), Kai Zhang (ETH Zurich), Shuhang Gu (ETH Zurich, Switzerland), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)</i>	
Distilling Object Detectors via Decoupled Features	2154
<i>Jianyuan Guo (Noah's Ark Lab, Huawei Technologies), Kai Han (Noah's Ark Lab, Huawei Technologies), Yunhe Wang (Huawei Technologies), Han Wu (Peking University), Xinghao Chen (Noah's Ark Lab, Huawei Technologies), Chunjing Xu (Huawei Noah's Ark Lab), and Chang Xu (University of Sydney)</i>	
S2-BNN: Bridging the Gap Between Self-Supervised Real and 1-Bit Neural Networks via Guided Distribution Calibration	2165
<i>Zhiqiang Shen (Carnegie Mellon University), Zechun Liu (Carnegie Mellon University), Jie Qin (Inception Institute of Artificial Intelligence), Lei Huang (Beihang University), Kwang-Ting Cheng (Hong Kong University of Science and Technology), and Marios Savvides (Carnegie Mellon University)</i>	
BCNet: Searching for Network Width With Bilaterally Coupled Network	2175
<i>Xiu Su (University of Sydney), Shan You (SenseTime), Fei Wang (SenseTime), Chen Qian (SenseTime), Changshui Zhang (Tsinghua University), and Chang Xu (University of Sydney)</i>	
Multi-Attentional Deepfake Detection	2185
<i>Hanqing Zhao (University of Science and Technology of China), Wenbo Zhou (University of Science and Technology of China), Dongdong Chen (Microsoft Cloud AI), Tianyi Wei (University of Science and Technology of China), Weiming Zhang (University of Science and Technology of China), and Nenghai Yu (University of Science and Technology of China)</i>	
A Peek Into the Reasoning of Neural Networks: Interpreting With Structural Visual Concepts.....	2195
<i>Yunhao Ge (University of Southern California), Yao Xiao (University of Southern California), Zhi Xu (University of Southern California), Meng Zheng (United Imaging Intelligence), Srikrishna Karanam (United Imaging Intelligence), Terrence Chen (United Imaging Intelligence), Laurent Itti (University of Southern California), and Ziyang Wu (United Imaging Intelligence)</i>	
Probabilistic Selective Encryption of Convolutional Neural Networks for Hierarchical Services	2205
<i>Jinyu Tian (University of Macau), Jiantao Zhou (University of Macau), and Jia Duan (JD)</i>	
Multi-Modal Relational Graph for Cross-Modal Video Moment Retrieval	2215
<i>Yawen Zeng (Hunan University), Da Cao (Hunan University), Xiaochi Wei (Baidu Inc.), Meng Liu (Shandong Jianzhu University), Zhou Zhao (Zhejiang University), and Zheng Qin (Hunan University)</i>	
PhD Learning: Learning With Pompeiu-Hausdorff Distances for Video-Based Vehicle Re-Identification	2225
<i>Jianan Zhao (Shanghai Em-Data Technology Co., Ltd.), Fengliang Qi (Shanghai Em-Data Technology Co., Ltd.), Guangyu Ren (Imperial College London), and Lin Xu (Shanghai Em-Data Technology Co., Ltd.)</i>	

Pareidolia Face Reenactment	2236
<i>Linsen Song (Institute of Automation, Chinese Academy of Sciences), Wayne Wu (SenseTime Research), Chaoyou Fu (Institute of Automation, Chinese Academy of Sciences), Chen Qian (SenseTime), Chen Change Loy (Nanyang Technological University), and Ran He (Institute of Automation, Chinese Academy of Sciences)</i>	
Hyper-LifelongGAN: Scalable Lifelong Learning for Image Conditioned Generation	2246
<i>Mengyao Zhai (Simon Fraser University), Lei Chen (Simon Fraser Univeristy), and Greg Mori (Simon Fraser University / Borealis AI)</i>	
TediGAN: Text-Guided Diverse Face Image Generation and Manipulation	2256
<i>Weihao Xia (Tsinghua University), Yujiu Yang (Tsinghua University), Jing-Hao Xue (University College London), and Baoyuan Wu (The Chinese University of Hong Kong, Shenzhen; Shenzhen Research Institute of Big Data)</i>	
TransFill: Reference-Guided Image Inpainting by Merging Multiple Color and Spatial Transformations	2266
<i>Yuqian Zhou (UIUC), Connelly Barnes (Adobe), Eli Shechtman (Adobe Research, US), and Sohrab Amirghodsi (Adobe Research)</i>	
ArtCoder: An End-to-End Method for Generating Scanning-Robust Stylized QR Codes	2277
<i>Hao Su (Beihang University), Jianwei Niu (beihang university), Xuefeng Liu (Beihang university), Qingfeng Li (Beihang University), Ji Wan (Beihang University), Mingliang Xu (Zhengzhou University), and Tao Ren (Beihang University)</i>	
Encoding in Style: A StyleGAN Encoder for Image-to-Image Translation	2287
<i>Elad Richardson (Penta-AI), Yuval Alaluf (Tel Aviv University), Or Patashnik (Tel Aviv University), Yotam Nitzan (Tel Aviv University), Yaniv Azar (Penta-AI), Stav Shapiro (Technion - Israel Institute of Technology), and Daniel Cohen-Or (Tel Aviv University)</i>	
Few-Shot Human Motion Transfer by Personalized Geometry and Texture Modeling	2297
<i>Zhichao Huang (Hong Kong University of Science and Technology), Xintong Han (Huya Inc), Jia Xu (HUYA Inc.), and Tong Zhang (The Hong Kong University of Science and Technology)</i>	
OCONet: Image Extrapolation by Object Completion	2307
<i>Richard Strong Bowen (Cornell Tech), Huiwen Chang (Google), Charles Herrmann (Cornell), Piotr Teterwak (Google), Ce Liu (Google), and Ramin Zabih (Cornell Tech/Google Research)</i>	
Greedy Hierarchical Variational Autoencoders for Large-Scale Video Prediction	2318
<i>Bohan Wu (Stanford University), Suraj Nair (Stanford University), Roberto Martín-Martín (Stanford University), Li Fei-Fei (Stanford University), and Chelsea Finn (Stanford)</i>	
Mutual CRF-GNN for Few-Shot Learning	2329
<i>Shixiang Tang (The University of Sydney), Dapeng Chen (The Chinese University of Hong Kong), Lei Bai (University of Sydney), Kaijian Liu (Zhejiang University), Yixiao Ge (The Chinese University of Hong Kong), and Wanli Ouyang (The University of Sydney)</i>	

Re-Labeling ImageNet: From Single to Multi-Labels, From Global to Localized Labels	2340
<i>Sangdoon Yun (NAVER AI LAB), Seong Joon Oh (NAVER AI Lab), Byeongho Heo (NAVER AI LAB), Dongyoon Han (NAVER AI Lab), Junsuk Choe (NAVER AI LAB), and Sanghyuk Chun (NAVER AI LAB)</i>	
Differentiable Patch Selection for Image Recognition	2351
<i>Jean-Baptiste Cordonnier (EPFL), Aravindh Mahendran (Google), Alexey Dosovitskiy (Google), Dirk Weissenborn (Google), Jakob Uszkoreit (Google Brain), and Thomas Unterthiner (Google Research, Brain Team)</i>	
Distribution Alignment: A Unified Framework for Long-Tail Visual Recognition	2361
<i>Songyang Zhang (ShanghaiTech University), Zeming Li (Megvii(Face++ Inc), Shipeng Yan (ShanghaiTech University), Xuming He (ShanghaiTech University), and Jian Sun (Megvii Technology)</i>	
Contrastive Embedding for Generalized Zero-Shot Learning	2371
<i>Zongyan Han (Nanjing University of Science and Technology), Zhenyong Fu (Nanjing University of Science and Technology), Shuo Chen (RIKEN), and Jian Yang (Nanjing University of Science and Technology)</i>	
Normal Integration via Inverse Plane Fitting With Minimum Point-to-Plane Distance	2382
<i>Xu Cao (Osaka University), Boxin Shi (Peking University), Fumio Okura (Osaka University), and Yasuyuki Matsushita (Osaka University)</i>	
Bayesian Nested Neural Networks for Uncertainty Calibration and Adaptive Compression	2392
<i>Yufei Cui (City University of Hong Kong), Ziquan Liu (City University of Hong Kong), Qiao Li (City University of Hong Kong), Antoni B. Chan (City University of Hong Kong, Hong, Kong), and Chun Jason Xue (City University of Hong Kong)</i>	
NetAdaptV2: Efficient Neural Architecture Search With Fast Super-Network Training and Architecture Optimization	2402
<i>Tien-Ju Yang (Massachusetts Institute of Technology), Yi-Lun Liao (Massachusetts Institute of Technology), and Vivienne Sze (Massachusetts Institute of Technology)</i>	
MIST: Multiple Instance Spatial Transformer	2412
<i>Baptiste Angles (University of Victoria), Yuhe Jin (University of British Columbia), Simon Kornblith (Google Brain), Andrea Tagliasacchi (Google Inc.), and Kwang Moo Yi (University of British Columbia)</i>	
Multi-Institutional Collaborations for Improving Deep Learning-Based Magnetic Resonance Image Reconstruction Using Federated Learning	2423
<i>Pengfei Guo (Johns Hopkins University), Puyang Wang (Johns Hopkins University), Jinyuan Zhou (Johns Hopkins University), Shanshan Jiang (Johns Hopkins School Of Medicine), and Vishal M. Patel (Johns Hopkins University)</i>	
A Self-Boosting Framework for Automated Radiographic Report Generation	2433
<i>Zhanyu Wang (University of Sydney), Luping Zhou (University of Sydney), Lei Wang (University of Wollongong, Australia), and Xiu Li (Tsinghua University)</i>	
Learning a Proposal Classifier for Multiple Object Tracking	2443
<i>Peng Dai (Tsinghua University), Renliang Weng (Aibee Inc), Wongun Choi (AiBee Inc), Changshui Zhang (Tsinghua University), Zhangping He (Aibee), and Wei Ding (CNU)</i>	

Improving Multiple Object Tracking With Single Object Tracking	2453
<i>Linyu Zheng (Institute of Automation, Chinese Academy of Sciences), Ming Tang (Institute of Automation, Chinese Academy of Sciences), Yingying Chen (CASIA), Guibo Zhu (Institute of Automation, Chinese Academy of Sciences), Jinqiao Wang (Institute of Automation, Chinese Academy of Sciences), and Hanqing Lu (NLPR, Institute of Automation, CAS)</i>	
Feature-Level Collaboration: Joint Unsupervised Learning of Optical Flow, Stereo Depth and Camera Motion	2463
<i>Cheng Chi (Huazhong University of Science and Technology), Qingjie Wang (Huazhong University of Science and Technology), Tianyu Hao (Huazhong University of Science and Technology), Peng Guo (Huazhong University of Science and Technology), and Xin Yang (Huazhong University of Science and Technology)</i>	
MaxUp: Lightweight Adversarial Training With Data Augmentation Improves Neural Network Training	2474
<i>Chengyue Gong (UT Austin), Tongzheng Ren (UT Austin), Mao Ye (Ut Austin), and Qiang Liu (UT Austin)</i>	
Unsupervised Human Pose Estimation Through Transforming Shape Templates	2484
<i>Luca Schmidtko (Imperial College London), Athanasios Vlontzos (Imperial College London), Simon Ellershaw (Imperial College London), Anna Lukens (Evelina London Children’s Hospital), Tomoki Arichi (King’s College London), and Bernhard Kainz (Imperial College London)</i>	
Understanding the Behaviour of Contrastive Loss	2495
<i>Feng Wang (Tsinghua University) and Huaping Liu (Tsinghua University)</i>	
Cross-Domain Adaptive Clustering for Semi-Supervised Domain Adaptation	2505
<i>Jichang Li (The University of Hong Kong), Guanbin Li (Sun Yat-sen University), Yemin Shi (Deepwise), and Yizhou Yu (The University of Hong Kong)</i>	
Divergence Optimization for Noisy Universal Domain Adaptation	2515
<i>Qing Yu (The University of Tokyo; OMRON SINIC X), Atsushi Hashimoto (OMRON SINIC X Corp.), and Yoshitaka Ushiku (OMRON SINIC X)</i>	
Limitations of Post-Hoc Feature Alignment for Robustness	2525
<i>Collin Burns (Columbia University) and Jacob Steinhardt (UC Berkeley)</i>	
Semantic-Aware Knowledge Distillation for Few-Shot Class-Incremental Learning	2534
<i>Ali Cheraghian (Australian National University, (ANU)), Shafin Rahman (North South University), Pengfei Fang (The Australian National University), Soumava Kumar Roy (Australian National University, (ANU)), Lars Petersson (Data61/CSIRO), and Mehrtash Harandi (Monash University)</i>	
Adaptive Aggregation Networks for Class-Incremental Learning	2544
<i>Yaoyao Liu (Max Planck Institute for Informatics), Bernt Schiele (MPI Informatics), and Qianru Sun (Singapore Management University)</i>	

Progressive Modality Reinforcement for Human Multimodal Emotion Recognition From Unaligned Multimodal Sequences	2554
<i>Fengmao Lv (Southwestern University of Finance and Economics), Xiang Chen (Tencent), Yanyong Huang (Southwestern University of Finance and Economics), Lixin Duan (University of Electronic Science and Technology of China), and Guosheng Lin (Nanyang Technological University)</i>	
Unsupervised Visual Representation Learning by Tracking Patches in Video	2563
<i>Guangting Wang (University of Science and Technology of China), Yizhou Zhou (University of Science and Technology of China), Chong Luo (Microsoft Research Asia), Wenxuan Xie (Microsoft Research Asia), Wenjun Zeng (Microsoft Research), and Zhiwei Xiong (University of Science and Technology of China)</i>	
HoHoNet: 360 Indoor Holistic Understanding With Latent Horizontal Features	2573
<i>Cheng Sun (National Tsing Hua University), Min Sun (NTHU), and Hwann-Tzong Chen (National Tsing Hua University)</i>	
Depth Completion With Twin Surface Extrapolation at Occlusion Boundaries	2583
<i>Saif Imran (Michigan State University), Xiaoming Liu (Michigan State University), and Daniel Morris (Michigan State University)</i>	
Zero-Shot Instance Segmentation	2593
<i>Ye Zheng (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Jiahong Wu (Peking University), Yongqiang Qin (AInnovation Technology Co., Ltd.), Faen Zhang (AInnovation Co. Ltd.), and Li Cui (Institute of computing technology of the Chinese Academy of Sciences)</i>	
Unsupervised Discovery of the Long-Tail in Instance Segmentation Using Hierarchical Self-Supervision	2603
<i>Zhenzhen Weng (Stanford University), Mehmet Giray Ogut (Stanford University), Shai Limonchik (Stanford University), and Serena Yeung (Stanford University)</i>	
Semi-Supervised Semantic Segmentation With Cross Pseudo Supervision	2613
<i>Xiaokang Chen (Peking University), Yuhui Yuan (Microsoft Research), Gang Zeng (Peking University), and Jingdong Wang (Microsoft)</i>	
Non-Salient Region Object Mining for Weakly Supervised Semantic Segmentation	2623
<i>Yazhou Yao (Nanjing University of Science and Technology), Tao Chen (Inception Institute of Artificial Intelligence), Guo-Sen Xie (Nanjing University of Science and Technology), Chuanyi Zhang (Nanjing University of Science and Technology), Fumin Shen (UESTC), Qi Wu (University of Adelaide), Zhenmin Tang (Nanjing University of Science and Technology), and Jian Zhang (UTS)</i>	
ABMDRNet: Adaptive-Weighted Bi-Directional Modality Difference Reduction Network for RGB-T Semantic Segmentation	2633
<i>Qiang Zhang (Xidian University), Shenlu Zhao (Xidian University), Yongjiang Luo (Xidian University), Dingwen Zhang (Xidian University), Nianchang Huang (Xidian University), and Jungong Han (Aberystwyth University)</i>	

BBAM: Bounding Box Attribution Map for Weakly Supervised Semantic and Instance Segmentation	2643
<i>Jungbeom Lee (Seoul National University), Jihun Yi (Seoul National University), Chaehun Shin (Seoul National University), and Sungroh Yoon (Seoul National University)</i>	
Positive-Unlabeled Data Purification in the Wild for Object Detection	2652
<i>Jianyuan Guo (Noah's Ark Lab, Huawei Technologies), Kai Han (Noah's Ark Lab, Huawei Technologies), Han Wu (Peking University), Chao Zhang (Noah's Ark Lab, Huawei Technologies), Xinghao Chen (Peking University), Chunjing Xu (Huawei Noah's Ark Lab), Chang Xu (University of Sydney), and Yunhe Wang (Huawei Technologies)</i>	
Ranking Neural Checkpoints	2662
<i>Yandong Li (University of Central Florida), Xuhui Jia (Google), Ruoxin Sang (Google), Yukun Zhu (Google Inc.), Bradley Green (Google Inc.), Liqiang Wang (University of Central Florida), and Boqing Gong (Google)</i>	
SelfAugment: Automatic Augmentation Policies for Self-Supervised Learning	2673
<i>Colorado J Reed (University of California, Berkeley), Sean Metzger (UC Berkeley), Aravind Srinivas (UC Berkeley), Trevor Darrell (EECS, UC Berkeley), and Kurt Keutzer (UC Berkeley)</i>	
Self-Supervised Multi-Frame Monocular Scene Flow	2683
<i>Junhwa Hur (TU Darmstadt) and Stefan Roth (TU Darmstadt)</i>	
Skip-Convolutions for Efficient Video Processing	2694
<i>Amirhossein Habibian (Qualcomm AI Research), Davide Abati (Qualcomm AI Research), Taco S. Cohen (Qualcomm AI Research), and Babak Ehteshami Bejnordi (Qualcomm AI Reseach)</i>	
Learning To Associate Every Segment for Video Panoptic Segmentation	2704
<i>Sanghyun Woo (KAIST), Dahun Kim (KAIST), Joon-Young Lee (Adobe Research), and In So Kweon (KAIST)</i>	
Triple-Cooperative Video Shadow Detection	2714
<i>Zhihao Chen (College of Intelligence and Computing, Tianjin University), Liang Wan (College of Intelligence and Computing, Tianjin University), Lei Zhu (The Chinese University of Hong Kong), Jia Shen (College of Intelligence and Computing, Tianjin University), Huazhu Fu (Inception Institute of Artificial Intelligence), Wennan Liu (Academy of Medical Engineering and Translational Medicine, Tianjin University), and Jing Qin (The Hong Kong Polytechnic University)</i>	
Image Change Captioning by Learning From an Auxiliary Task	2724
<i>Mehrdad Hosseinzadeh (University of Manitoba) and Yang Wang (University of Manitoba; Huawei Technologies Canada)</i>	
How2Sign: A Large-Scale Multimodal Dataset for Continuous American Sign Language	2734
<i>Amanda Duarte (Universitat Politecnica de Catalunya), Shruti Palaskar (Carnegie Mellon University), Lucas Ventura (UPC), Deepti Ghadiyaram (Facebook), Kenneth DeHaan (Gallaudet University), Florian Metze (Facebook), Jordi Torres (Barcelona Supercomputing Center), and Xavier Giro-i-Nieto (Universitat Politecnica de Catalunya)</i>	
Cyclic Co-Learning of Sounding Object Visual Grounding and Sound Separation	2744
<i>Yapeng Tian (University of Rochester), Di Hu (Renmin University of China), and Chenliang Xu (University of Rochester)</i>	

LipSync3D: Data-Efficient Learning of Personalized 3D Talking Faces From Video Using Pose and Lighting Normalization	2754
<i>Avisek Lahiri (Indian Institute of Technology Kharagpur), Vivek Kwatra (Google), Christian Frueh (Google Research), John Lewis (Google Research), and Chris Bregler (Google Research)</i>	
Interventional Video Grounding With Dual Contrastive Learning	2764
<i>Guoshun Nan (Singapore University of Technology and Design), Rui Qiao (Singapore University of Technology and Design), Yao Xiao (Shanghai Jiao Tong University), Jun Liu (Singapore University of Technology and Design), Sicong Leng (Singapore University of Technology and Design), Hao Zhang (Agency for Science, Technology and Research), and Wei Lu (SUTD)</i>	
Roses Are Red, Violets Are Blue... but Should VQA Expect Them To?	2775
<i>Corentin Kervadec (Orange Labs), Grigory Antipov (Orange Labs), Moez Baccouche (Orange Labs), and Christian Wolf (INSA Lyon, France)</i>	
ReDet: A Rotation-Equivariant Detector for Aerial Object Detection	2785
<i>Jiaming Han (Wuhan University), Jian Ding (Wuhan University), Nan Xue (Wuhan University), and Gui-Song Xia (Wuhan University)</i>	
Roof-GAN: Learning To Generate Roof Geometry and Relations for Residential Houses	2795
<i>Yiming Qian (Simon Fraser University), Hao Zhang (Simon Fraser University), and Yasutaka Furukawa (Simon Fraser University)</i>	
PANDA: Adapting Pretrained Features for Anomaly Detection and Segmentation	2805
<i>Tal Reiss (The Hebrew University of Jerusalem), Niv Cohen (The Hebrew University of Jerusalem), Liron Bergman (Hebrew University of Jerusalem), and Yedid Hoshen (The Hebrew University of Jerusalem)</i>	
Differentiable SLAM-Net: Learning Particle SLAM for Visual Navigation	2814
<i>Peter Karkus (National University of Singapore), Shaojun Cai (National University of Singapore), and David Hsu (NUS)</i>	

Session 03

DyStaB: Unsupervised Object Segmentation via Dynamic-Static Bootstrapping	2825
<i>Yanchao Yang (Stanford University), Brian Lai (UCLA), and Stefano Soatto (UCLA)</i>	
Diffusion Probabilistic Models for 3D Point Cloud Generation	2836
<i>Shitong Luo (Peking University) and Wei Hu (Peking University)</i>	
Learned Initializations for Optimizing Coordinate-Based Neural Representations	2845
<i>Matthew Tancik (UC Berkeley), Ben Mildenhall (UC Berkeley), Terrance Wang (UC Berkeley), Divi Schmidt (UC Berkeley), Pratul P. Srinivasan (Google Research), Jonathan T. Barron (Google Research), and Ren Ng (UC Berkeley)</i>	
Neural Scene Graphs for Dynamic Scenes	2855
<i>Julian Ost (Technical University of Munich / Algolux), Fahim Mannan (Algolux), Nils Thuerey (Technical University of Munich), Julian Knodt (Princeton), and Felix Heide (Princeton / Algolux)</i>	

Consensus Maximisation Using Influences of Monotone Boolean Functions	2865
<i>Ruwan Tennakoon (RMIT), David Suter (Edith Cowan University), Erchuan Zhang (Edith Cowan University), Tat-Jun Chin (University of Adelaide), and Alireza Bab-Hadiashar (RMIT)</i>	
Task Programming: Learning Data Efficient Behavior Representations	2875
<i>Jennifer J. Sun (Caltech), Ann Kennedy (Northwestern University), Eric Zhan (Caltech), David J. Anderson (Caltech), Yisong Yue (Caltech), and Pietro Perona (California Institute of Technology)</i>	
SCANimate: Weakly Supervised Learning of Skinned Clothed Avatar Networks	2885
<i>Shunsuke Saito (Facebook), Jinlong Yang (Max Planck Institute for Intelligent Systems), Qianli Ma (Max Planck Institute for Intelligent Systems), and Michael J. Black (Max Planck Institute for Intelligent Systems)</i>	
Diverse Part Discovery: Occluded Person Re-Identification With Part-Aware Transformer	2897
<i>Yulin Li (University of Science and Technology of China), Jianfeng He (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China), Xiang Liu (Dongguan University of Technology), Yongdong Zhang (University of Science and Technology of China), and Feng Wu (University of Science and Technology of China)</i>	
What's in the Image? Explorable Decoding of Compressed Images	2907
<i>Yuval Bahat (Technion Institute of Technology) and Tomer Michaeli (Technion)</i>	
Simple Copy-Paste Is a Strong Data Augmentation Method for Instance Segmentation	2917
<i>Golnaz Ghiasi (Google Brain), Yin Cui (Google), Aravind Srinivas (UC Berkeley), Rui Qian (Cornell University), Tsung-Yi Lin (Google Brain), Ekin D. Cubuk (Google Brain), Quoc V. Le (Google Brain), and Barret Zoph (Google)</i>	
Face Forgery Detection by 3D Decomposition	2928
<i>Xiangyu Zhu (Chinese Academy of Science), Hao Wang (CBSR & NLPR, Institute of Automation, Chinese Academy of Sciences), Hongyan Fei (University Of Science & Technology Beijing), Zhen Lei (Institute of Automation, Chinese Academy of Sciences), and Stan Z. Li (National Lab. of Pattern Recognition, China)</i>	
Convolutional Hough Matching Networks	2939
<i>Juhong Min (POSTECH) and Minsu Cho (POSTECH)</i>	
L2M-GAN: Learning To Manipulate Latent Space Semantics for Facial Attribute Editing	2950
<i>Guoxing Yang (Renmin University of China), Nanyi Fei (Renmin University of China), Mingyu Ding (The University of Hong Kong), Guangzhen Liu (Renmin University of China), Zhiwu Lu (Renmin University of China), and Tao Xiang (University of Surrey)</i>	
Patchwise Generative ConvNet: Training Energy-Based Models From a Single Natural Image for Internal Learning	2960
<i>Zilong Zheng (UCLA), Jianwen Xie (Baidu Research), and Ping Li (Baidu Research)</i>	

Generative Classifiers as a Basis for Trustworthy Image Classification	2970
<i>Radek Mackowiak (Heidelberg University), Lynton Ardizzone (Heidelberg University), Ullrich Köthe (University of Heidelberg), and Carsten Rother (University of Heidelberg)</i>	
HR-NAS: Searching Efficient High-Resolution Neural Architectures With Lightweight Transformers	2981
<i>Mingyu Ding (The University of Hong Kong), Xiaochen Lian (ByteDance), Linjie Yang (ByteDance AI Lab), Peng Wang (Bytedance USA LLC.), Xiaojie Jin (NUS), Zhiwu Lu (Renmin University of China), and Ping Luo (The University of Hong Kong)</i>	
Progressive Unsupervised Learning for Visual Object Tracking	2992
<i>Qiangqiang Wu (City University of Hong Kong), Jia Wan (City University of Hong Kong), and Antoni B. Chan (City University of Hong Kong, Hong Kong)</i>	
FFB6D: A Full Flow Bidirectional Fusion Network for 6D Pose Estimation	3002
<i>Yisheng He (Hong Kong University of Science and Technology), Haibin Huang (Kuaishou Technology), Haoqiang Fan (Megvii Incface++), Qifeng Chen (HKUST), and Jian Sun (Megvii Technology)</i>	
DER: Dynamically Expandable Representation for Class Incremental Learning	3013
<i>Shipeng Yan (ShanghaiTech University), Jiangwei Xie (ShanghaiTech University), and Xuming He (ShanghaiTech University)</i>	
Dense Contrastive Learning for Self-Supervised Visual Pre-Training	3023
<i>Xinlong Wang (University of Adelaide), Rufeng Zhang (Tongji University), Chunhua Shen (University of Adelaide), Tao Kong (Bytedance), and Lei Li (ByteDance AI Lab)</i>	
S2R-DepthNet: Learning a Generalizable Depth-Specific Structural Representation	3033
<i>Xiaotian Chen (University of Science and Technology of China), Yuwang Wang (Microsoft Research), Xuejin Chen (University of Science and Technology of China), and Wenjun Zeng (Microsoft Research)</i>	
Depth-Aware Mirror Segmentation	3043
<i>Haiyang Mei (Dalian University of Technology), Bo Dong (SRI International), Wen Dong (Dalian University of Technology), Pieter Peers (College of William & Mary), Xin Yang (Dalian University of Technology), Qiang Zhang (Dalian University of Technology), and Xiaopeng Wei (Dalian University of Technology)</i>	
Video Prediction Recalling Long-Term Motion Context via Memory Alignment Learning	3053
<i>Sangmin Lee (KAIST), Hak Gu Kim (EPFL), Dae Hwi Choi (KAIST), Hyung-Il Kim (ETRI), and Yong Man Ro (KAIST)</i>	
Room-and-Object Aware Knowledge Reasoning for Remote Embodied Referring Expression	3063
<i>Chen Gao (Beihang University), Jinyu Chen (Beihang University), Si Liu (Beihang University), Luting Wang (Beihang University), Qiong Zhang (Xiaomi Technology), and Qi Wu (University of Adelaide)</i>	
GATSBI: Generative Agent-Centric Spatio-Temporal Object Interaction	3073
<i>Cheol-Hui Min (Seoul National University), Jinseok Bae (Seoul National University), Junho Lee (Seoul National University), and Young Min Kim (Seoul National University)</i>	

Crossing Cuts Polygonal Puzzles: Models and Solvers	3083
<i>Peleg Harel (Ben-Gurion University) and Ohad Ben-Shahar (Ben Gurion University)</i>	
Transformation Invariant Few-Shot Object Detection	3093
<i>Aoxue Li (Noah's Ark Lab) and Zhenguo Li (Huawei Noah's Ark Lab)</i>	
Adaptive Class Suppression Loss for Long-Tail Object Detection	3102
<i>Tong Wang (CASIA), Yousong Zhu (Institute of Automation, Chinese Academy of Sciences), Chaoyang Zhao (National Laboratory of Pattern Recognition, CASIA), Wei Zeng (Peking University, China), Jinqiao Wang (Institute of Automation, Chinese Academy of Sciences), and Ming Tang (Institute of Automation, Chinese Academy of Sciences)</i>	
What if We Only Use Real Datasets for Scene Text Recognition? Toward Scene Text Recognition With Fewer Labels	3112
<i>Jeonghun Baek (The University of Tokyo), Yusuke Matsui (The University of Tokyo), and Kiyoharu Aizawa (The University of Tokyo)</i>	
Fourier Contour Embedding for Arbitrary-Shaped Text Detection	3122
<i>Yiqin Zhu (South China University of Technology), Jianyong Chen (South China University of Technology), Lingyu Liang (South China University of Technology), Zhanghui Kuang (SenseTime Ltd.), Lianwen Jin (South China University of Technology), and Wayne Zhang (SenseTime Research)</i>	
Humble Teachers Teach Better Students for Semi-Supervised Object Detection	3131
<i>Yihe Tang (Amazon Web Services), Weifeng Chen (University of Michigan, Ann Arbor), Yijun Luo (Amazon), and Yuting Zhang (Amazon)</i>	
Cross-Modal Center Loss for 3D Cross-Modal Retrieval	3141
<i>Longlong Jing (The City University of New York), Elahe Vahdani (The Graduate Center, The City University of New York), Jiaying Tan (City University of New York), and Yingli Tian (City University of New York)</i>	
Single-View 3D Object Reconstruction From Shape Priors in Memory	3151
<i>Shuo Yang (University of Technology Sydney), Min Xu (Univeristy of Technology Sydney), Haozhe Xie (Harbin Institute of Technology), Stuart Perry (University of Technology Sydney), and Jiahao Xia (School of Electrical and Data Engineering, FEIT, University of Technology Sydney)</i>	
NeuralFusion: Online Depth Fusion in Latent Space	3161
<i>Silvan Weder (ETH Zürich), Johannes L. Schönberger (Microsoft), Marc Pollefeys (ETH Zurich / Microsoft), and Martin R. Oswald (ETH Zurich)</i>	
PAConv: Position Adaptive Convolution With Dynamic Kernel Assembling on Point Clouds	3172
<i>Mutian Xu (The University of Hong Kong), Runyu Ding (The University of Hong Kong), Hengshuang Zhao (University of Oxford), and Xiaojuan Qi (The University of Hong Kong)</i>	
Self-Supervised Pillar Motion Learning for Autonomous Driving	3182
<i>Chenxu Luo (Johns Hopkins University), Xiaodong Yang (NVIDIA Research), and Alan Yuille (Johns Hopkins University)</i>	
Scan2Cap: Context-Aware Dense Captioning in RGB-D Scans	3192
<i>Zhenyu Chen (Technical University of Munich), Ali Gholami (Simon Fraser University), Matthias Nießner (Technical University of Munich), and Angel X. Chang (Simon Fraser University)</i>	

Neural Parts: Learning Expressive 3D Shape Abstractions With Invertible Neural Networks	3203
<i>Despoina Paschalidou (MPI-IS Tuebingen), Angelos Katharopoulos (Idiap), Andreas Geiger (MPI-IS and University of Tuebingen), and Sanja Fidler (University of Toronto, NVIDIA)</i>	
Universal Spectral Adversarial Attacks for Deformable Shapes	3215
<i>Arianna Rampini (Sapienza University of Rome), Franco Pestarini (Sapienza University of Rome), Luca Cosmo (Sapienza University of Rome), Simone Melzi (Sapienza University of Rome), and Emanuele Rodolà (Sapienza University of Rome)</i>	
Large-Scale Localization Datasets in Crowded Indoor Spaces	3226
<i>Donghwan Lee (NAVER LABS), Soohyun Ryu (NAVER LABS), Suyong Yeon (NAVER LABS), Yonghan Lee (NAVER LABS), Deokhwa Kim (Naverlabs), Cheolho Han (NAVER LABS), Yohann Cabon (Naver Labs Europe), Philippe Weinzaepfel (NAVER LABS Europe), Nicolas Guérin (Naver Labs Europe), Gabriela Csurka (Naver Labs Europe), and Martin Humenberger (Naver Labs Europe)</i>	
Learnable Motion Coherence for Correspondence Pruning	3236
<i>Yuan Liu (The University of Hong Kong), Lingjie Liu (Max Planck Institute for Informatics), Cheng Lin (The University of Hong Kong), Zhen Dong (Wuhan University), and Wenping Wang (The University of Hong Kong)</i>	
Back to the Feature: Learning Robust Camera Localization From Pixels To Pose	3246
<i>Paul-Edouard Sarlin (ETH Zurich), Ajaykumar Unagar (ETH Zurich), Måns Larsson (Chalmers and Eigenvision), Hugo Germain (Ecole des Ponts ParisTech), Carl Toft (Chalmers), Viktor Larsson (ETH Zurich), Marc Pollefeys (ETH Zurich / Microsoft), Vincent Lepetit (Ecole des Ponts ParisTech), Lars Hammarstrand (Chalmers university of technology), Fredrik Kahl (Chalmers), and Torsten Sattler (Czech Technical University in Prague)</i>	
Wide-Baseline Relative Camera Pose Estimation With Directional Learning	3257
<i>Kefan Chen (Google), Noah Snavely (Google), and Ameesh Makadia (Google Research)</i>	
Deep Optimized Priors for 3D Shape Modeling and Reconstruction	3268
<i>Mingyue Yang (South China University of Technology), Yuxin Wen (South China University of Technology), Weikai Chen (Tencent America), Yongwei Chen (South China University of Technology), and Kui Jia (South China University of Technology)</i>	
PVGNet: A Bottom-Up One-Stage 3D Object Detector With Integrated Multi-Level Features	3278
<i>Zhenwei Miao (DAMO Academy, Alibaba Group), Jikai Chen (DAMO Academy, Alibaba Group), Hongyu Pan (DAMO Academy, Alibaba Group), Ruiwen Zhang (Tsinghua University), Kaixuan Liu (DAMO Academy, Alibaba Group), Peihan Hao (DAMO Academy, Alibaba Group), Jun Zhu (DAMO Academy, Alibaba Group), Yang Wang (DAMO Academy, Alibaba Group), and Xin Zhan (DAMO Academy, Alibaba Group)</i>	
Objects Are Different: Flexible Monocular 3D Object Detection	3288
<i>Yunpeng Zhang (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)</i>	

A Large-Scale Study on Unsupervised Spatiotemporal Representation Learning	3298
<i>Christoph Feichtenhofer (Facebook AI Research), Haoqi Fan (Facebook AI Research), Bo Xiong (Facebook AI Research), Ross Girshick (FAIR), and Kaiming He (Facebook AI Research)</i>	
Representing Videos As Discriminative Sub-Graphs for Action Recognition	3309
<i>Dong Li (University of Science and Technology of China), Zhaofan Qiu (JD.com), Yingwei Pan (JD AI Research), Ting Yao (JD AI Research), Houqiang Li (University of Science and Technology of China), and Tao Mei (AI Research of JD.com)</i>	
Learning Salient Boundary Feature for Anchor-free Temporal Action Localization	3319
<i>Chuming Lin (Tencent), Chengming Xu (Fudan University), Donghao Luo (Tencent), Yabiao Wang (Tencent), Ying Tai (Tencent YouTu), Chengjie Wang (Tencent), Jilin Li (Tencent), Feiyue Huang (Tencent), and Yanwei Fu (Fudan University)</i>	
QAIR: Practical Query-Efficient Black-Box Attacks for Image Retrieval	3329
<i>Xiaodan Li (Alibaba Group), Jinfeng Li (Alibaba Group), Yuefeng Chen (Alibaba Group), Shaokai Ye (EPFL), Yuan He (Alibaba Group), Shuhui Wang (VIPL, ICT, Chinese academic of science), Hang Su (Tsinghua University), and Hui Xue (Alibaba)</i>	
Defending Multimodal Fusion Models Against Single-Source Adversaries	3339
<i>Karren Yang (MIT), Wan-Yi Lin (Bosch Center for Artificial Intelligence), Manash Barman (CMU), Filipe Condessa (Bosch Center for Artificial Intelligence), and Zico Kolter (Carnegie Mellon University)</i>	
Training Generative Adversarial Networks in One Stage	3349
<i>Chengchao Shen (Zhejiang University), Youtan Yin (Zhejiang University), Xinchao Wang (National University of Singapore), Xubin Li (Alibaba Group), Jie Song (Zhejiang University), and Mingli Song (Zhejiang University)</i>	
Learning Complete 3D Morphable Face Models From Images and Videos	3360
<i>Mallikarjun B R (Max Planck Institute for Informatics), Ayush Tewari (Max Planck Institute for Informatics), Hans-Peter Seidel (Max Planck Institute for Informatics), Mohamed Elgharib (Max Planck Institute for Informatics), and Christian Theobalt (MPI Informatik)</i>	
We Are More Than Our Joints: Predicting How 3D Bodies Move	3371
<i>Yan Zhang (ETH Zurich), Michael J. Black (Max Planck Institute for Intelligent Systems), and Siyu Tang (ETH Zurich)</i>	
HybrIK: A Hybrid Analytical-Neural Inverse Kinematics Solution for 3D Human Pose and Shape Estimation	3382
<i>Jiefeng Li (Shanghai Jiao Tong University), Chao Xu (Shanghai Jiao Tong University), Zhicun Chen (Shanghai Jiao Tong University), Siyuan Bian (Shanghai Jiao Tong University), Lixin Yang (Shanghai Jiao Tong University), and Cewu Lu (Shanghai Jiao Tong University)</i>	
Learning To Count Everything	3393
<i>Viresh Ranjan (Stony Brook University), Udbhav Sharma (Stony Brook University), Thu Nguyen (VinAI Research, Vietnam), and Minh Hoai (Stony Brook University)</i>	

Information Bottleneck Disentanglement for Identity Swapping	3403
<i>Gege Gao (Institute of Automation, Chinese Academy of Sciences), Huaibo Huang (Institute of Automation, Chinese Academy of Sciences), Chaoyou Fu (Institute of Automation, Chinese Academy of Sciences), Zhaoyang Li (Institute of Automation, Chinese Academy of Sciences), and Ran He (Institute of Automation, Chinese Academy of Sciences)</i>	
Mitigating Face Recognition Bias via Group Adaptive Classifier	3413
<i>Sixue Gong (Michigan State University), Xiaoming Liu (Michigan State University), and Anil K. Jain (Michigan State University)</i>	
Meta Batch-Instance Normalization for Generalizable Person Re-Identification	3424
<i>Seokeon Choi (KAIST), Taekyung Kim (KAIST), Minki Jeong (KAIST), Hyoungseob Park (KAIST), and Changick Kim (KAIST)</i>	
Refining Pseudo Labels With Clustering Consensus Over Generations for Unsupervised Object Re-Identification	3435
<i>Xiao Zhang (Chinese University of Hong Kong), Yixiao Ge (The Chinese University of Hong Kong), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), and Hongsheng Li (Chinese University of Hong Kong)</i>	
Back to Event Basics: Self-Supervised Learning of Image Reconstruction for Event Cameras via Photometric Constancy	3445
<i>Federico Paredes-Vallés (Delft University of Technology) and Guido C. H. E. de Croon (TU Delft)</i>	
DeFMO: Deblurring and Shape Recovery of Fast Moving Objects	3455
<i>Denys Rozumnyi (ETH Zurich / CTU in Prague), Martin R. Oswald (ETH Zurich), Vittorio Ferrari (Google Research), Jiří Matas (CMP CTU FEE), and Marc Pollefeys (ETH Zurich / Microsoft)</i>	
Efficient Multi-Stage Video Denoising With Recurrent Spatio-Temporal Fusion	3465
<i>Matteo Maggioni (Huawei Noah's Ark Lab), Yibin Huang (Huawei Noah's Ark Lab), Cheng Li (Huawei Noah's Ark Lab), Shuai Xiao (Huawei Noah's Ark Lab), Zhongqian Fu (Huawei Noah's Ark Lab), and Fenglong Song (Huawei Noah's Ark Lab)</i>	
ZeroScatter: Domain Transfer for Long Distance Imaging and Vision Through Scattering Media...	3475
<i>Zheng Shi (Princeton University), Ethan Tseng (Princeton University), Mario Bijelic (Mercedes-Benz AG), Werner Ritter (Mercedes-Benz AG), and Felix Heide (Princeton / Algolux)</i>	
Restoring Extremely Dark Images in Real Time	3486
<i>Mohit Lamba (Indian Institute of Technology Madras) and Kaushik Mitra (IIT Madras)</i>	
Practical Wide-Angle Portraits Correction With Deep Structured Models	3497
<i>Jing Tan (Megvii(face++) Research), Shan Zhao (Megvii), Pengfei Xiong (Tencent), Jiangyu Liu (megvii inc), Haoqiang Fan (Megvii Incface++), and Shuaicheng Liu (UESTC; Megvii)</i>	
End-to-End Learning for Joint Image Demosaicing, Denoising and Super-Resolution	3506
<i>Wenzhu Xing (Tampere University) and Karen Egiazarian (Tampere University)</i>	
Image Super-Resolution With Non-Local Sparse Attention	3516
<i>Yiqun Mei (University of Illinois), Yuchen Fan (University of Illinois at Urbana-Champaign), and Yuqian Zhou (UIUC)</i>	

Video Rescaling Networks With Joint Optimization Strategies for Downscaling and Upscaling	3526
<i>Yan-Cheng Huang (National Yang Ming Chiao Tung University), Yi-Hsin Chen (National Yang Ming Chiao Tung University), Cheng-You Lu (National Yang Ming Chiao Tung University), Hui-Po Wang (CISPA Helmholtz Center for Information Security), Wen-Hsiao Peng (National Chiao Tung University), and Ching-Chun Huang (National Yang Ming Chiao Tung University)</i>	
Restore From Restored: Video Restoration With Pseudo Clean Video	3536
<i>Seunghwan Lee (Hanyang Univ.), Donghyeon Cho (Chungnam National University), Jiwon Kim (SK T-Brain), and Tae Hyun Kim (Hanyang Univeristy)</i>	
Enriching ImageNet With Human Similarity Judgments and Psychological Embeddings	3546
<i>Brett D. Roads (University College London) and Bradley C. Love (University College London)</i>	
Conceptual 12M: Pushing Web-Scale Image-Text Pre-Training To Recognize Long-Tail Visual Concepts	3557
<i>Soravit Changpinyo (Google Research), Piyush Sharma (Google Research), Nan Ding (Google), and Radu Soricut (Google)</i>	
CondenseNet V2: Sparse Feature Reactivation for Deep Networks	3568
<i>Le Yang (Tsinghua University), Haojun Jiang (Tsinghua University), Ruojin Cai (Cornell University), Yulin Wang (Tsinghua University), Shiji Song (Tsinghua), Gao Huang (Tsinghua), and Qi Tian (Huawei Cloud & AI)</i>	
Revisiting Knowledge Distillation: An Inheritance and Exploration Framework	3578
<i>Zhen Huang (University of Science and Technology of China), Xu Shen (Alibaba Group), Jun Xing (miHoYo), Tongliang Liu (The University of Sydney), Xinmei Tian (USTC), Houqiang Li (University of Science and Technology of China), Bing Deng (Damo Academy, Alibaba Group), Jianqiang Huang (Alibaba Group), and Xian-Sheng Hua (Damo Academy, Alibaba Group)</i>	
Minimally Invasive Surgery for Sparse Neural Networks in Contrastive Manner	3588
<i>Chong Yu (NVIDIA)</i>	
Effective Sparsification of Neural Networks With Global Sparsity Constraint	3598
<i>Xiao Zhou (HKUST), Weizhong Zhang (HKUST), Hang Xu (Huawei Noah's Ark Lab), and Tong Zhang (The Hong Kong University of Science and Technology)</i>	
Improving the Efficiency and Robustness of Deepfakes Detection Through Precise Geometric Features	3608
<i>Zekun Sun (Shanghai Jiao Tong University), Yujie Han (Shanghai Jiaotong University), Zeyu Hua (Shanghai Jiao Tong University), Na Ruan (Shanghai Jiao Tong University), and Weijia Jia (Institute of AI and Future Networks, Beijing Normal University, (Zhuhai; BNU-HKBU United International College, Zhuhai, PR China Shanghai Jiao Tong University)</i>	
Right for the Right Concept: Revising Neuro-Symbolic Concepts by Interacting With Their Explanations	3618
<i>Wolfgang Stammer (TU Darmstadt), Patrick Schramowski (TU Darmstadt), and Kristian Kersting (TU Darmstadt)</i>	

Protecting Intellectual Property of Generative Adversarial Networks From Ambiguity Attacks....	3629
<i>Ding Sheng Ong (University of Malaya), Chee Seng Chan (University of Malaya), Kam Woh Ng (Webank AI Lab), Lixin Fan (WeBank), and Qiang Yang (Hong Kong UST)</i>	
VIGOR: Cross-View Image Geo-Localization Beyond One-to-One Retrieval	3639
<i>Sijie Zhu (University of North Carolina at Charlotte), Taojiannan Yang (University of North Carolina at Charlotte), and Chen Chen (University of North Carolina at Charlotte)</i>	
On Semantic Similarity in Video Retrieval	3649
<i>Michael Wray (University of Bristol), Hazel Doughty (University of Amsterdam), and Dima Damen (University of Bristol)</i>	
Flow-Guided One-Shot Talking Face Generation With a High-Resolution Audio-Visual Dataset ..	3660
<i>Zhimeng Zhang (Netease Fuxi AI Lab), Lincheng Li (Netease Fuxi AI Lab), Yu Ding (Netease Fuxi AI Lab), and Changjie Fan (NetEase Fuxi AI Lab)</i>	
Navigating the GAN Parameter Space for Semantic Image Editing	3670
<i>Anton Cherepkov (Yandex), Andrey Voynov (Yandex), and Artem Babenko (Yandex)</i>	
IMAGINE: Image Synthesis by Image-Guided Model Inversion	3680
<i>Pei Wang (UC San Diego), Yijun Li (Adobe Research), Krishna Kumar Singh (Adobe Research), Jingwan Lu (Adobe Research), and Nuno Vasconcelos (UC San Diego)</i>	
Human De-Occlusion: Invisible Perception and Recovery for Humans	3690
<i>Qiang Zhou (Huazhong Univ. of Science and Technology), Shiyin Wang (ByteDance), Yitong Wang (Bytedance), Zilong Huang (Tencent), and Xinggang Wang (Huazhong University of Science and Technology)</i>	
Learning To Warp for Style Transfer	3701
<i>Xiao-Chang Liu (University of Bath), Yong-Liang Yang (University of Bath), and Peter Hall (University of Bath)</i>	
StEP: Style-Based Encoder Pre-Training for Multi-Modal Image Synthesis	3711
<i>Moustafa Meshry (University of Maryland), Yixuan Ren (University of Maryland), Larry S. Davis (University of Maryland), and Abhinav Shrivastava (University of Maryland)</i>	
ANR: Articulated Neural Rendering for Virtual Avatars	3721
<i>Amit Raj (Georgia Institute of Technology), Julian Tanke (University of Bonn), James Hays (Georgia Institute of Technology, USA), Minh Vo (Facebook Reality Labs), Carsten Stoll (Facebook Reality Labs), and Christoph Lassner (Facebook Reality Labs)</i>	
LayoutTransformer: Scene Layout Generation With Conceptual and Spatial Diversity	3731
<i>Cheng-Fu Yang (National Taiwan University), Wan-Cyuan Fan (National Taiwan University), Fu-En Yang (National Taiwan University), and Yu-Chiang Frank Wang (National Taiwan University)</i>	
Stochastic Image-to-Video Synthesis Using cINNs	3741
<i>Michael Dorkenwald (Heidelberg University), Timo Milbich (Heidelberg University), Andreas Blattmann (Heidelberg University), Robin Rombach (Heidelberg University), Konstantinos G. Derpanis (Ryerson University), and Björn Ommer (Heidelberg University)</i>	

Prototype Completion With Primitive Knowledge for Few-Shot Learning	3753
<i>Baoquan Zhang (Harbin Institute of Technology, Shenzhen), Xutao Li (Harbin Institute of Technology Shenzhen Graduate School), Yunming Ye (Harbin Institute of Technology Shenzhen Graduate School), Zhichao Huang (Harbin Institute of Technology, Shenzhen), and Lisai Zhang (Harbin Institute of Technology, Shenzhen)</i>	
Dynamic Class Queue for Large Scale Face Recognition in the Wild	3762
<i>Bi Li (Huazhong University of Science and Technology), Teng Xi (Baidu Inc.), Gang Zhang (Baidu Inc.), Haocheng Feng (Baidu Inc.), Junyu Han (Baidu Inc.), Jingtuo Liu (baidu), Errui Ding (Baidu Inc.), and Wenyu Liu (Huazhong University of Science and Technology)</i>	
Truly Shift-Invariant Convolutional Neural Networks	3772
<i>Anadi Chaman (University of Illinois at Urbana-Champaign) and Ivan Dokmanić (University of Basel)</i>	
RSG: A Simple but Effective Module for Learning Imbalanced Datasets	3783
<i>Jianfeng Wang (University of Oxford), Thomas Lukasiewicz (University of Oxford), Xiaolin Hu (Tsinghua University), Jianfei Cai (Monash University), and Zhenghua Xu (Hebei University of Technology)</i>	
Goal-Oriented Gaze Estimation for Zero-Shot Learning	3793
<i>Yang Liu (Beihang University), Lei Zhou (Beihang University), Xiao Bai (Beihang University), Yifei Huang (The University of Tokyo), Lin Gu (RIKENAIP / The University of Tokyo), Jun Zhou (Griffith University), and Tatsuya Harada (The University of Tokyo / RIKEN)</i>	
Uncalibrated Neural Inverse Rendering for Photometric Stereo of General Surfaces	3803
<i>Berk Kaya (ETH Zurich), Suryansh Kumar (ETH Zurich), Carlos Oliveira (CVL Zurich), Vittorio Ferrari (Google Research), and Luc Van Gool (ETH Zurich)</i>	
Robust Bayesian Neural Networks by Spectral Expectation Bound Regularization	3814
<i>Jiaru Zhang (Shanghai Jiao Tong University), Yang Hua (Queen's University Belfast), Zhengui Xue (Shanghai Jiao Tong University), Tao Song (Shanghai Jiao Tong University), Chengyu Zheng (Shanghai Jiao Tong University), Ruhui Ma (Shanghai Jiao Tong University), and Haibing Guan (Shanghai Jiao Tong University)</i>	
MobileDets: Searching for Object Detection Architectures for Mobile Accelerators	3824
<i>Yunyang Xiong (University of Wisconsin-Madison), Hanxiao Liu (Google Brain), Suyog Gupta (Google), Berkin Akin (Google), Gabriel Bender (Google), Yongzhe Wang (Google), Pieter-Jan Kindermans (Google), Mingxing Tan (Google Brain), Vikas Singh (University of Wisconsin Madison), and Bo Chen (DJI)</i>	
Hilbert Sinkhorn Divergence for Optimal Transport	3834
<i>Qian Li (University of Technology Sydney), Zhichao Wang (School of Electrical Engineering and Telecommunications, University of New South Wales, Australia), Gang Li (Deakin Univeristy, Australia), Jun Pang (University of Luxembourg), and Guandong Xu (University of Technology Sydney, Australia)</i>	

Object Classification From Randomized EEG Trials	3844
<i>Hamad Ahmed (Purdue University), Ronnie B. Wilbur (Purdue University), Hari M. Bharadwaj (Purdue University), and Jeffrey Mark Siskind (Purdue University)</i>	
Leveraging Large-Scale Weakly Labeled Data for Semi-Supervised Mass Detection in Mammograms	3854
<i>Yuxing Tang (PAII Inc.), Zhenjie Cao (PAII Inc.), Yanbo Zhang (PAII Inc.), Zhicheng Yang (PAII Inc.), Zongcheng Ji (PAII Inc.), Yiwei Wang (PAII Inc.), Mei Han (PAII Inc.), Jie Ma (Department of Radiology, Shenzhen People's Hospital), Jing Xiao (Ping An Insurance, (Group) Company of China), and Peng Chang (PAII Inc.)</i>	
Tracking Pedestrian Heads in Dense Crowd	3864
<i>Ramana Sundararaman (INRIA Rennes - Bretagne Atlantique), Cédric De Almeida Braga (INRIA Rennes - Bretagne Atlantique), Eric Marchand (Universite de Rennes 1, IRISA), and Julien Pettré (INRIA Rennes - Bretagne Atlantique)</i>	
Multiple Object Tracking With Correlation Learning	3875
<i>Qiang Wang (Alibaba Group), Yun Zheng (Alibaba group), Pan Pan (Alibaba Group), and Yinghui Xu (Alibaba DAMO Academy)</i>	
SMURF: Self-Teaching Multi-Frame Unsupervised RAFT With Full-Image Warping	3886
<i>Austin Stone (Google), Daniel Maurer (Waymo), Alper Ayvaci (Waymo), Anelia Angelova (Google), and Rico Jonschkowski (Google)</i>	
Bilinear Parameterization for Non-Separable Singular Value Penalties	3896
<i>Marcus Valtonen Örn hag Örn hag (Lund University), José Pedro Iglesias (Chalmers), and Carl Olsson (Lund University, Sweden)</i>	
DSC-PoseNet: Learning 6DoF Object Pose Estimation via Dual-Scale Consistency	3906
<i>Zongxin Yang (UTS), Xin Yu (University of Technology Sydney), and Yi Yang (UTS)</i>	
Unsupervised Disentanglement of Linear-Encoded Facial Semantics	3916
<i>Yutong Zheng (Carnegie Mellon University), Yu-Kai Huang (Carnegie Mellon University), Ran Tao (Carnegie Mellon University), Zhiqiang Shen (Carnegie Mellon University), and Marios Savvides (Carnegie Mellon University)</i>	
MetaCorrection: Domain-Aware Meta Loss Correction for Unsupervised Domain Adaptation in Semantic Segmentation	3926
<i>Xiaoqing Guo (City University of Hong Kong), Chen Yang (City University of Hong Kong), Baopu Li (BAIDU USA LLC), and Yixuan Yuan (City University of Hong Kong)</i>	
Cross-Domain Gradient Discrepancy Minimization for Unsupervised Domain Adaptation	3936
<i>Zhekai Du (University of Electronic Science and Technology of China), Jingjing Li (University of Electronic Science and Technology of China), Hongzu Su (University of Electronic Science and Technology of China), Lei Zhu (Shandong Normal University), and Ke Lu (University of Electronic Science and Technology of China)</i>	

Generative Interventions for Causal Learning	3946
<i>Chengzhi Mao (Columbia University), Augustine Cha (Columbia University), Amogh Gupta (Columbia University), Hao Wang (MIT), Junfeng Yang (Columbia University), and Carl Vondrick (Columbia University)</i>	
Distilling Causal Effect of Data in Class-Incremental Learning	3956
<i>Xinting Hu (Nanyang Technological University), Kaihua Tang (Nanyang Technological University), Chunyan Miao (NTU), Xian-Sheng Hua (Damo Academy, Alibaba Group), and Hanwang Zhang (Nanyang Technological University)</i>	
Embedding Transfer With Label Relaxation for Improved Metric Learning	3966
<i>Sungyeon Kim (POSTECH), Dongwon Kim (POSTECH), Minsu Cho (POSTECH), and Suha Kwak (POSTECH)</i>	
M3P: Learning Universal Representations via Multitask Multilingual Multimodal Pre-Training....	3976
<i>Minheng Ni (Microsoft Research Asia; Research Center for Social Computing and Information Retrieval, Harbin Institute of Technology), Haoyang Huang (Microsoft Research Asia), Lin Su (STCA), Edward Cui (Microsoft), Taroon Bharti (Microsoft), Lijuan Wang (Microsoft), Dongdong Zhang (Microsoft Research Asia), and Nan Duan (Microsoft Research)</i>	
Instance Localization for Self-Supervised Detection Pretraining	3986
<i>Ceyuan Yang (Chinese University of Hong Kong), Zhirong Wu (Microsoft), Bolei Zhou (CUHK), and Stephen Lin (Microsoft Research)</i>	
VIP-DeepLab: Learning Visual Perception With Depth-Aware Video Panoptic Segmentation	3996
<i>Siyuan Qiao (Johns Hopkins University), Yukun Zhu (Google Inc.), Hartwig Adam (Google), Alan Yuille (Johns Hopkins University), and Liang-Chieh Chen (Google Inc.)</i>	
AdaBins: Depth Estimation Using Adaptive Bins	4008
<i>Shariq Farooq Bhat (KAUST), Ibraheem Alhashim (KAUST), and Peter Wonka (KAUST)</i>	
Deep Occlusion-Aware Instance Segmentation With Overlapping BiLayers	4018
<i>Lei Ke (HKUST), Yu-Wing Tai (Kuaishou Technology / HKUST), and Chi-Keung Tang (Hong Kong University of Science and Technology)</i>	
Information-Theoretic Segmentation by Inpainting Error Maximization	4028
<i>Pedro Savarese (Toyota Technological Institute at Chicago), Sunnie S. Y. Kim (Princeton University), Michael Maire (University of Chicago), Greg Shakhnarovich (Toyota Technological Institute at Chicago), and David McAllester (TTI Chicago)</i>	
PLOP: Learning Without Forgetting for Continual Semantic Segmentation	4039
<i>Arthur Douillard (Heuritech / Sorbonne University), Yifu Chen (Sorbonne University), Arnaud Dapogny (Pierre and Marie Curie University, UPMC), and Matthieu Cord (Sorbonne University)</i>	
Coarse-To-Fine Domain Adaptive Semantic Segmentation With Photometric Alignment and Category-Center Regularization	4050
<i>Haoyu Ma (the University of Hong Kong), Xiangru Lin (The University of Hong Kong), Zifeng Wu (DeepWise), and Yizhou Yu (The University of Hong Kong)</i>	

HyperSeg: Patch-Wise Hypernetwork for Real-Time Semantic Segmentation	4060
<i>Yuval Nirkin (Bar-Ilan University), Lior Wolf (Facebook AI), and Tal Hassner (Tel Aviv University, Israel)</i>	
Anti-Adversarially Manipulated Attributions for Weakly and Semi-Supervised Semantic Segmentation	4070
<i>Jungbeom Lee (Seoul National University), Eunji Kim (Seoul National University), and Sungroh Yoon (Seoul National University)</i>	
Instant-Teaching: An End-to-End Semi-Supervised Object Detection Framework	4079
<i>Qiang Zhou (Alibaba Group), Chaohui Yu (Alibaba Group), Zhibin Wang (Alibaba Group), Qi Qian (Alibaba Group), and Hao Li (Alibaba Group)</i>	
Unbiased Mean Teacher for Cross-Domain Object Detection	4089
<i>Jinhong Deng (University of Electronic Science and Technology of China), Wen Li (University of Electronic Science and Technology of China), Yuhua Chen (ETH Zurich), and Lixin Duan (University of Electronic Science and Technology of China)</i>	
MeanShift++: Extremely Fast Mode-Seeking With Applications to Segmentation and Object Tracking	4100
<i>Jennifer Jang (Waymo) and Heinrich Jiang (Google Research)</i>	
FlowStep3D: Model Unrolling for Self-Supervised Scene Flow Estimation	4112
<i>Yair Kittenplon (Tel-Aviv University), Yonina C. Eldar, and Dan Raviv (Tel Aviv University)</i>	
Recognizing Actions in Videos From Unseen Viewpoints	4122
<i>AJ Piergiovanni (Google) and Michael S. Ryoo (Google Brain; Stony Brook University)</i>	
VSPW: A Large-scale Dataset for Video Scene Parsing in the Wild	4131
<i>Jiaxu Miao (University of Technology Sydney), Yunchao Wei (UTS), Yu Wu (University of Technology Sydney), Chen Liang (Zhejiang University), Guangrui Li (University of Technology Sydney), and Yi Yang (UTS)</i>	
Learning Position and Target Consistency for Memory-Based Video Object Segmentation	4142
<i>Li Hu (Alibaba), Peng Zhang (Alibaba), Bang Zhang (Damo Academy Alibaba), Pan Pan (Alibaba Group), Yinghui Xu (Alibaba DAMO Academy), and Rong Jin (alibaba group)</i>	
UC2: Universal Cross-Lingual Cross-Modal Vision-and-Language Pre-Training	4153
<i>Mingyang Zhou (University of California, Davis), Luowei Zhou (Microsoft), Shuohang Wang (Microsoft), Yu Cheng (Microsoft Research), Linjie Li (Microsoft), Zhou Yu (Columbia), and Jingjing Liu (Microsoft)</i>	
Fingerspelling Detection in American Sign Language	4164
<i>Bowen Shi (Toyota Technological Institute at Chicago), Diane Brentari (University of Chicago), Greg Shakhnarovich (Toyota Technological Institute at Chicago), and Karen Livescu (TTI-Chicago)</i>	
Pose-Controllable Talking Face Generation by Implicitly Modularized Audio-Visual Representation	4174
<i>Hang Zhou (The Chinese University of Hong Kong), Yasheng Sun (Tokyo Institute of Technology), Wayne Wu (SenseTime Research), Chen Change Loy (Nanyang Technological University), Xiaogang Wang (Chinese University of Hong Kong, Hong Kong), and Ziwei Liu (Nanyang Technological University)</i>	

Collaborative Spatial-Temporal Modeling for Language-Queried Video Actor Segmentation	4185
<i>Tianrui Hui (Institute of Information Engineering, Chinese Academy of Sciences), Shaofei Huang (Institute of Information Engineering, Chinese Academy of Sciences), Si Liu (Beihang University), Zihan Ding (Beihang University), Guanbin Li (Sun Yat-sen University), Wenguan Wang (Eidgenössische Technische Hochschule Zürich), Jizhong Han (Institute of Information Engineering, Chinese Academy of Sciences), and Fei Wang (SenseTime)</i>	
Cascaded Prediction Network via Segment Tree for Temporal Video Grounding	4195
<i>Yang Zhao (Zhejiang University), Zhou Zhao (Zhejiang University), Zhu Zhang (Zhejiang University), and Zhijie Lin (Zhejiang University)</i>	
How Transferable Are Reasoning Patterns in VQA?	4205
<i>Corentin Kervadec (Orange Labs), Théo Jaunet (INSA-Lyon, LIRIS), Grigory Antipov (Orange Labs), Moez Baccouche (Orange Labs), Romain Vuillemot (Ecole Centrale de Lyon, LIRIS), and Christian Wolf (INSA Lyon, France)</i>	
PointFlow: Flowing Semantics Through Points for Aerial Image Segmentation	4215
<i>Xiangtai Li (Peking University), Hao He (NLPR CASIA), Xia Li (ETH Zurich), Duo Li (HKUST), Guangliang Cheng (Sensetime Group Limited), Jianping Shi (Sensetime Group Limited), Lubin Weng (CASIC), Yunhai Tong (Peking University), and Zhouchen Lin (Peking University)</i>	
HDMaGen: A Hierarchical Graph Generative Model of High Definition Maps	4225
<i>Lu Mi (MIT), Hang Zhao (Tsinghua University), Charlie Nash (DeepMind), Xiaohan Jin (Waymo), Jiyang Gao (Momenta.ai), Chen Sun (Google), Cordelia Schmid (Google), Nir Shavit (Massachusetts Institute of Technology), Yuning Chai (Alphabet), and Dragomir Anguelov (Waymo)</i>	
A Circular-Structured Representation for Visual Emotion Distribution Learning	4235
<i>Jingyuan Yang (Xidian University), Jie Li (Xidian University), Leida Li (Xidian University), Xiumei Wang (Xidian University), and Xinbo Gao (Xidian University)</i>	
More Photos Are All You Need: Semi-Supervised Learning for Fine-Grained Sketch Based Image Retrieval	4245
<i>Ayan Kumar Bhunia (University of Surrey), Pinaki Nath Chowdhury (University of Surrey), Aneeshan Sain (University of Surrey), Yongxin Yang (University of Surrey), Tao Xiang (University of Surrey), and Yi-Zhe Song (University of Surrey)</i>	

Session 04

Line Segment Detection Using Transformers Without Edges	4255
<i>Yifan Xu (UC San Diego), Weijian Xu (University of California, San Diego), David Cheung (University of California, San Diego), and Zhuowen Tu (UC San Diego)</i>	
Predator: Registration of 3D Point Clouds With Low Overlap	4265
<i>Shengyu Huang (ETH Zürich), Zan Gojcic (ETH Zürich), Mikhail Usoyatsov (ETH Zürich), Andreas Wieser (ETH Zürich), and Konrad Schindler (ETH Zurich)</i>	

Point2Skeleton: Learning Skeletal Representations from Point Clouds	4275
<i>Cheng Lin (The University of Hong Kong), Changjian Li (University College London), Yuan Liu (The University of Hong Kong), Nenglu Chen (The University of Hong Kong), Yi-King Choi (The University of Hong Kong), and Wenping Wang (The University of Hong Kong)</i>	
Neural Lumigraph Rendering	4285
<i>Petr Kellnhofer (Stanford University), Lars C. Jebe (Raxium), Andrew Jones (Raxium), Ryan Spicer (Raxium), Kari Pulli (University of Oulu), and Gordon Wetzstein (Stanford University)</i>	
Rotation Coordinate Descent for Fast Globally Optimal Rotation Averaging	4296
<i>Alvaro Parra (The University of Adelaide), Shin-Fang Chng (The University of Adelaide), Tat-Jun Chin (University of Adelaide), Anders Eriksson (University of Queensland), and Ian Reid (University of Adelaide, Australia)</i>	
Towards Evaluating and Training Verifiably Robust Neural Networks	4306
<i>Zhaoyang Lyu (The Chinese University of Hong Kong), Minghao Guo (The Chinese University of Hong Kong), Tong Wu (The Chinese University of Hong Kong), Guodong Xu (The Chinese University of Hong Kong), Kehuan Zhang (Chinese University of Hong Kong), and Dahua Lin (The Chinese University of Hong Kong)</i>	
Human POSEitioning System (HPS): 3D Human Pose Estimation and Self-Localization in Large Scenes From Body-Mounted Sensors	4316
<i>Vladimir Guzov (Max Planck Institute for Informatics), Aymen Mir (Max Planck Institute of Informatics), Torsten Sattler (Czech Technical University in Prague), and Gerard Pons-Moll (MPII, Germany)</i>	
Discover Cross-Modality Nuances for Visible-Infrared Person Re-Identification	4328
<i>Qiong Wu (Xiamen University), Pingyang Dai (Xiamen University), Jie Chen (Peking University), Chia-Wen Lin (National Tsing Hua University), Yongjian Wu (Tencent Technology, (Shanghai) Co., Ltd), Feiyue Huang (Tencent), Bineng Zhong (Guangxi Normal University), and Rongrong Ji (Xiamen University, China)</i>	
Dual Pixel Exploration: Simultaneous Depth Estimation and Image Restoration	4338
<i>Liyuan Pan (The Australian National University), Shah Chowdhury (Australian National University), Richard Hartley (Australian National University, Australia), Miaomiao Liu (The Australian National University), Hongguang Zhang (Australian National University), and Hongdong Li (Australian National University, Australia)</i>	
Towards Good Practices for Efficiently Annotating Large-Scale Image Classification Datasets	4348
<i>Yuan-Hong Liao (University of Toronto), Amlan Kar (University of Toronto, NVIDIA), and Sanja Fidler (University of Toronto, NVIDIA)</i>	
ForgeryNet: A Versatile Benchmark for Comprehensive Forgery Analysis	4358
<i>Yinan He (Beijing University of Posts and Telecommunications), Bei Gan (Sensetime), Siyu Chen (Peking University), Yichun Zhou (Beihang University), Guojun Yin (University of Science and Technology of China), Luchuan Song (University of Science and Technology of China), Lu Sheng (Beihang University), Jing Shao (Sensetime), and Ziwei Liu (Nanyang Technological University)</i>	

Spatial-Temporal Correlation and Topology Learning for Person Re-Identification in Videos	4368
<i>Jiawei Liu (University of Science and Technology of China), Zheng-Jun Zha (University of Science and Technology of China), Wei Wu (University of Science and Technology of China), Kecheng Zheng (University of Science and Technology of China), and Qibin Sun (University of Science and Technology of China)</i>	
SSN: Soft Shadow Network for Image Compositing	4378
<i>Yichen Sheng (Purdue University), Jianming Zhang (Adobe Research), and Bedrich Benes (Purdue University)</i>	
Soft-IntroVAE: Analyzing and Improving the Introspective Variational Autoencoder	4389
<i>Tal Daniel (Technion) and Aviv Tamar (Technion Israeli Institute of Technology)</i>	
Learning Placeholders for Open-Set Recognition	4399
<i>Da-Wei Zhou (Nanjing University), Han-Jia Ye (Nanjing University), and De-Chuan Zhan (Nanjing University)</i>	
ReNAS: Relativistic Evaluation of Neural Architecture Search	4409
<i>Yixing Xu (Huawei Technologies), Yunhe Wang (Huawei Technologies), Kai Han (Noah's Ark Lab, Huawei Technologies), Yehui Tang (Peking University), Shangling Jui (Huawei Kirin Solution), Chunjing Xu (Huawei Noah's Ark Lab), and Chang Xu (University of Sydney)</i>	
Learning To Filter: Siamese Relation Network for Robust Tracking	4419
<i>Siyuan Cheng (Huaqiao University), Bineng Zhong (Guangxi Normal University), Guorong Li (University of Chinese Academy of Sciences), Xin Liu (Beijing Seetatech Technology Co., Ltd), Zhenjun Tang (Guangxi Normal University), Xianxian Li (Guangxi Normal University), and Jing Wang (Huaqiao University)</i>	
Generative Hierarchical Features From Synthesizing Images	4430
<i>Yinghao Xu (Chinese University of Hong Kong), Yujun Shen (Dept. of IE, CUHK), Jiapeng Zhu (HKUST), Ceyuan Yang (Chinese University of Hong Kong), and Bolei Zhou (CUHK)</i>	
Continual Adaptation of Visual Representations via Domain Randomization and Meta-Learning	4441
<i>Riccardo Volpi (Naver Labs Europe), Diane Larlus (Naver Labs Europe), and Grégory Rogez (NAVER LABS Europe)</i>	
NewtonianVAE: Proportional Control and Goal Identification From Pixels via Physical Latent Spaces	4452
<i>Miguel Jaques (University of Edinburgh), Michael Burke (Monash University), and Timothy M. Hospedales (Edinburgh University)</i>	
3D-to-2D Distillation for Indoor Scene Parsing	4462
<i>Zhengzhe Liu (The Chinese University of Hong Kong), Xiaojuan Qi (The University of Hong Kong), and Chi-Wing Fu (The Chinese University of Hong Kong)</i>	
Repurposing GANs for One-Shot Semantic Part Segmentation	4473
<i>Nontawat Tritrong (Vidyasirimedhi Institute of Science and Technology), Pitchaporn Rewatbowornwong (Vidyasirimedhi Institute of Science and Technology), and Supasorn Suwajanakorn (Vidyasirimedhi Institute of Science and Technology)</i>	

Temporal Query Networks for Fine-Grained Video Understanding	4484
<i>Chuhan Zhang (University of Oxford), Ankush Gupta (DeepMind), and Andrew Zisserman (University of Oxford)</i>	
ManipulaTHOR: A Framework for Visual Object Manipulation	4495
<i>Kiana Ehsani (Allen Institute for Artificial Intelligence), Winson Han (Allen Institute for Artificial Intelligence), Alvaro Herrasti (alvaroh@allenai.org), Eli VanderBilt (Allen Institute for AI), Luca Weihs (Allen Institute for Artificial Intelligence), Eric Kolve (Allen AI), Aniruddha Kembhavi (Allen Institute for Artificial Intelligence), and Roozbeh Mottaghi (Allen Institute for AI)</i>	
Omnimatte: Associating Objects and Their Effects in Video	4505
<i>Erika Lu (University of Oxford), Forrester Cole (Google Research), Tali Dekel (Google), Andrew Zisserman (University of Oxford), William T. Freeman (Google), and Michael Rubinstein (Google)</i>	
MeGA-CDA: Memory Guided Attention for Category-Aware Unsupervised Domain Adaptive Object Detection	4514
<i>Vibashan VS (Johns Hopkins University), Vikram Gupta (MBRDI), Poojan Oza (Johns Hopkins University), Vishwanath A. Sindagi (Johns Hopkins University), and Vishal M. Patel (Johns Hopkins University)</i>	
Generalized Few-Shot Object Detection Without Forgetting	4525
<i>Zhibo Fan (Megvii Technology), Yuchen Ma (megvii), Zeming Li (Megvii(Face++) Inc), and Jian Sun (Megvii Technology)</i>	
DAP: Detection-Aware Pre-Training With Weak Supervision	4535
<i>Yuanyi Zhong (University of Illinois at Urbana-Champaign), Jianfeng Wang (Microsoft), Lijuan Wang (Microsoft), Jian Peng (University of Illinois at Urbana-Champaign), Yu-Xiong Wang (University of Illinois at Urbana-Champaign), and Lei Zhang (Microsoft)</i>	
A Multiplexed Network for End-to-End, Multilingual OCR	4545
<i>Jing Huang (Facebook), Guan Pang (Facebook), Rama Kovvuri (Facebook), Mandy Toh (Facebook), Kevin J Liang (Facebook), Praveen Krishnan (Facebook), Xi Yin (Facebook AI), and Tal Hassner (Facebook AI)</i>	
Scene Text Retrieval via Joint Text Detection and Similarity Learning	4556
<i>Hao Wang (Huazhong University of Science and Technology), Xiang Bai (Huazhong University of Science and Technology), Mingkun Yang (Huazhong University of Science and Technology), Shenggao Zhu (Huawei), Jing Wang (Huawei Cloud & AI), and Wenyu Liu (Huazhong University of Science and Technology)</i>	
Data-Uncertainty Guided Multi-Phase Learning for Semi-Supervised Object Detection	4566
<i>Zhenyu Wang (Tsinghua University), Yali Li (Tsinghua University), Ye Guo (Tsinghua University), Lu Fang (Tsinghua University), and Shengjin Wang (Tsinghua University)</i>	
pixelNeRF: Neural Radiance Fields From One or Few Images	4576
<i>Alex Yu (UC Berkeley), Vickie Ye (UC Berkeley), Matthew Tancik (UC Berkeley), and Angjoo Kanazawa (University of California Berkeley)</i>	
From Points to Multi-Object 3D Reconstruction	4586
<i>Francis Engelmann (RWTH Aachen University, Computer Vision Group), Konstantinos Rematas (Google), Bastian Leibe (RWTH Aachen University), and Vittorio Ferrari (Google Research)</i>	

4D Hyperspectral Photoacoustic Data Restoration With Reliability Analysis	4596
<i>Weihang Liao (National Institute of Informatics), Art Subpa-asa (National Institute of Informatics), Yinqiang Zheng (The University of Tokyo), and Imari Sato (National Institute of Informatics)</i>	
RfD-Net: Point Scene Understanding by Semantic Instance Reconstruction	4606
<i>Yinyu Nie (Bournemouth University), Ji Hou (Technical University of Munich), Xiaoguang Han (Shenzhen Research Institute of Big Data, the Chinese University of Hong Kong, (Shenzhen)), and Matthias Nießner (Technical University of Munich)</i>	
Style-Based Point Generator With Adversarial Rendering for Point Cloud Completion	4617
<i>Chulin Xie (University of Illinois at Urbana-Champaign), Chuxin Wang (University of Science and Technology of China), Bo Zhang (Microsoft Research Asia), Hao Yang (Microsoft Research Asia), Dong Chen (Microsoft Research Asia), and Fang Wen (Microsoft Research Asia)</i>	
Denoise and Contrast for Category Agnostic Shape Completion	4627
<i>Antonio Alliegro (Politecnico di Torino), Diego Valsesia (Politecnico di Torino), Giulia Fracastoro (Polito), Enrico Magli (POLITO), and Tatiana Tommasi (Politecnico di Torino)</i>	
Neural Surface Maps	4637
<i>Luca Morreale (University College London), Noam Aigerman (Adobe), Vladimir G. Kim (Adobe), and Niloy J. Mitra (University College London)</i>	
RGB-D Local Implicit Function for Depth Completion of Transparent Objects	4647
<i>Luyang Zhu (University of Washington), Arsalan Mousavian (NVIDIA), Yu Xiang (NVIDIA), Hammad Mazhar (NVIDIA), Jozef van Eenbergen (NVIDIA), Shoubhik Debnath (Nvidia), and Dieter Fox (NVIDIA)</i>	
Uncertainty-Aware Camera Pose Estimation From Points and Lines	4657
<i>Alexander Vakhitov (SLAMCore), Luis Ferraz (Universitat Pompeu Fabra), Antonio Agudo (Institut de Robotica i Informatica Industrial, CSIC-UPC), and Francesc Moreno-Noguer (IRI)</i>	
Patch2Pix: Epipolar-Guided Pixel-Level Correspondences	4667
<i>Qunjie Zhou (Technical University of Munich), Torsten Sattler (Czech Technical University in Prague), and Laura Leal-Taixé (TUM)</i>	
Deep Multi-Task Learning for Joint Localization, Perception, and Prediction	4677
<i>John Phillips (Uber ATG / University of Waterloo), Julieta Martinez (Uber ATG), Ioan Andrei Bârsan (Uber ATG, University of Toronto), Sergio Casas (Uber ATG / University of Toronto), Abbas Sadat (Uber ATG), and Raquel Urtasun (Uber ATG)</i>	
IBRNet: Learning Multi-View Image-Based Rendering	4688
<i>Qianqian Wang (Cornell), Zhicheng Wang (Google Research), Kyle Genova (Google Research), Pratul P. Srinivasan (Google Research), Howard Zhou (Google), Jonathan T. Barron (Google Research), Ricardo Martin-Brualla (Google), Noah Snavely (Cornell University and Google AI), and Thomas Funkhouser (Google Research)</i>	

Unsupervised Learning of 3D Object Categories From Videos in the Wild	4698
<i>Philipp Henzler (University College London), Jeremy Reizenstein (Facebook AI Research), Patrick Labatut (Facebook AI Research), Roman Shapovalov (Facebook AI Research), Tobias Ritschel (UCL), Andrea Vedaldi (University of Oxford / Facebook AI Research), and David Novotny (Facebook AI Research)</i>	
LiDAR-Aug: A General Rendering-Based Augmentation Framework for 3D Object Detection	4708
<i>Jin Fang (Baidu), Xinxin Zuo (University of Alberta), Dingfu Zhou (Baidu), Shengze Jin (ETH Zürich), Sen Wang (University of Alberta), and Liangjun Zhang (baidu)</i>	
Delving Into Localization Errors for Monocular 3D Object Detection	4719
<i>Xinzhu Ma (The University of Sydney), Yinmin Zhang (SenseTime Group Limited), Dan Xu (The Hong Kong University of Science and Technology), Dongzhan Zhou (The University of Sydney), Shuai Yi (SenseTime Group Limited), Haojie Li (Dalian University of Technology), and Wanli Ouyang (The University of Sydney)</i>	
3D CNNs With Adaptive Temporal Feature Resolutions	4729
<i>Mohsen Fayyaz (Microsoft), Emad Bahrami (University of Bonn), Ali Diba (KU Leuven), Mehdi Noroozi (Bosch Gmb), Ehsan Adeli (Stanford University), Luc Van Gool (ETH Zurich), and Jürgen Gall (University of Bonn)</i>	
3D Human Action Representation Learning via Cross-View Consistency Pursuit	4739
<i>Linguo Li (Shanghai Jiao Tong University), Minsi Wang (Shanghai Jiao Tong University), Bingbing Ni (Shanghai Jiao Tong University), Hang Wang (Shanghai Jiao Tong University), Jiancheng Yang (Shanghai Jiao Tong University), and Wenjun Zhang (Shanghai Jiao Tong University)</i>	
Three Birds with One Stone: Multi-Task Temporal Action Detection via Recycling Temporal Annotations	4749
<i>Zhihui Li (University of New South Wales) and Lina Yao (University of New South Wales)</i>	
Delving into Data: Effectively Substitute Training for Black-box Attack	4759
<i>Wenxuan Wang (Fudan University), Bangjie Yin (Tencent), Taiping Yao (Tencent YouTu), Li Zhang (University of Oxford), Yanwei Fu (Fudan University), Shouhong Ding (Tencent), Jilin Li (Tencent), Feiyue Huang (Tencent), and Xiangyang Xue (Fudan University)</i>	
Data-Free Model Extraction	4769
<i>Jean-Baptiste Truong (Worcester Polytechnic Institute), Pratyush Maini (IIT Delhi), Robert J. Walls (Worcester Polytechnic Institute), and Nicolas Papernot (University of Toronto and Vector Institute)</i>	
Adaptive Weighted Discriminator for Training Generative Adversarial Networks	4779
<i>Vasily Zadorozhnyy (University of Kentucky), Qiang Cheng (University of Kentucky), and Qiang Ye (University of Kentucky)</i>	

Monocular Reconstruction of Neural Face Reflectance Fields	4789
<i>Mallikarjun B R (Max Planck Institute for Informatics), Ayush Tewari (Max Planck Institute for Informatics), Tae-Hyun Oh (POSTECH), Tim Weyrich (University College London), Bernd Bickel (IST Austria), Hans-Peter Seidel (Max Planck Institute for Informatics), Hanspeter Pfister (Harvard University), Wojciech Matusik (MIT), Mohamed Elgharib (Max Planck Institute for Informatics), and Christian Theobalt (MPI Informatik)</i>	
Towards Accurate 3D Human Motion Prediction From Incomplete Observations	4799
<i>Qiongjie Cui (Nanjing University of Science and Technology) and Huaijiang Sun (Nanjing University of Science and Technology)</i>	
Monocular Real-Time Full Body Capture With Inter-Part Correlations	4809
<i>Yuxiao Zhou (Tsinghua University), Marc Habermann (Max Planck Institute for Informatics), Ikhsanul Habibie (Max Planck Institute for Informatics), Ayush Tewari (Max Planck Institute for Informatics), Christian Theobalt (MPI Informatik), and Feng Xu (Tsinghua University)</i>	
Cross-Modal Collaborative Representation Learning and a Large-Scale RGBT Benchmark for Crowd Counting	4821
<i>Lingbo Liu (Sun Yat-sen University), Jiaqi Chen (Sun Yat-sen University), Hefeng Wu (Sun Yat-sen University), Guanbin Li (Sun Yat-sen University), Chenglong Li (Anhui University), and Liang Lin (Sun Yat-sen University)</i>	
One Shot Face Swapping on Megapixels	4832
<i>Yuhao Zhu (Institute of Automation, Chinese Academy of Sciences), Qi Li (CASIA), Jian Wang (Center for Research on Intelligent Perception and Computing, (CRIPAC), Institute of Automation, Chinese Academy of Sciences, (CASIA), University of Chinese Academy of Sciences(UCAS)), Cheng-Zhong Xu (University of Macau), and Zhenan Sun (Chinese of Academy of Sciences)</i>	
Dynamic Probabilistic Graph Convolution for Facial Action Unit Intensity Estimation	4843
<i>Tengfei Song (Southeast University), Zijun Cui (Rensselaer Polytechnic Institute), Yuru Wang (NorthEast Normal University), Wenming Zheng (Southeast University), and Qiang Ji (Rensselaer Polytechnic Institute)</i>	
Joint Noise-Tolerant Learning and Meta Camera Shift Adaptation for Unsupervised Person Re-Identification	4853
<i>Fengxiang Yang (Xiamen University), Zhun Zhong (University of Trento), Zhiming Luo (Xiamen University), Yuanzheng Cai (Minjiang University), Yaojin Lin (Minnan Normal University), Shaozi Li (Xiamen University, China), and Nicu Sebe (University of Trento)</i>	
Prototype-Guided Saliency Feature Learning for Person Search	4863
<i>Hanjae Kim (Yonsei Univ.), Sunghun Joung (Yonsei Univ.), Ig-Jae Kim (KIST), and Kwanghoon Sohn (Yonsei Univ.)</i>	
Labeled From Unlabeled: Exploiting Unlabeled Data for Few-Shot Deep HDR Deghosting	4873
<i>K. Ram Prabhakar (Indian Institute of Science), Gowtham Senthil (Indian Institute of Technology Tirupati), Susmit Agrawal (Indian Institute of Science), R. Venkatesh Babu (Indian Institute of Science), and Rama Krishna Sai S S Gorthi (IIT Tirupati)</i>	

Learning Spatially-Variant MAP Models for Non-Blind Image Deblurring	4884
<i>Jiangxin Dong (Max Planck Institute for Informatics), Stefan Roth (TU Darmstadt), and Bernt Schiele (MPI Informatics)</i>	
NBNet: Noise Basis Learning for Image Denoising With Subspace Projection	4894
<i>Shen Cheng (Megvii), Yuzhi Wang (Tsinghua University), Haibin Huang (Kuaishou Technology), Donghao Liu (Megvii), Haoqiang Fan (Megvii Incface++), and Shuaicheng Liu (UESTC; Megvii)</i>	
Image De-Raining via Continual Learning	4905
<i>Man Zhou (University of Science and Technology of China), Jie Xiao (University of Science and Technology of China), Yifan Chang (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), Aiping Liu (University of Science and Technology of China), Jinshan Pan (Nanjing University of Science and Technology), and Zheng-Jun Zha (University of Science and Technology of China)</i>	
Exploring Sparsity in Image Super-Resolution for Efficient Inference	4915
<i>Longguang Wang (National University of Defense Technology), Xiaoyu Dong (The University of Tokyo / RIKEN AIP), Yingqian Wang (National University of Defense Technology), Xinyi Ying (National University of Defense Technology), Zaiping Lin (National University of Defense Technology), Wei An (National University of Defense Technology), and Yulan Guo (National University of Defense Technology)</i>	
From Shadow Generation To Shadow Removal	4925
<i>Zhihao Liu (Beijing Jiaotong University), Hui Yin (Beijing Jiaotong University), Xinyi Wu (University of South Carolina), Zhenyao Wu (University of South Carolina), Yang Mi (University of South Carolina), and Song Wang (University of South Carolina)</i>	
Spatiotemporal Registration for Event-Based Visual Odometry	4935
<i>Daqi Liu (University of Adelaide), Alvaro Parra (The University of Adelaide), and Tat-Jun Chin (University of Adelaide)</i>	
BasicVSR: The Search for Essential Components in Video Super-Resolution and Beyond	4945
<i>Kelvin C.K. Chan (Nanyang Technological University), Xintao Wang (Tencent), Ke Yu (The Chinese University of Hong Kong), Chao Dong (SIAT), and Chen Change Loy (Nanyang Technological University)</i>	
Fast Bayesian Uncertainty Estimation and Reduction of Batch Normalized Single Image Super-Resolution Network	4955
<i>Aupendu Kar (Indian Institute of Technology Kharagpur) and Prabir Kumar Biswas (IIT Khargpur)</i>	
Learning Temporal Consistency for Low Light Video Enhancement From Single Images	4965
<i>Fan Zhang (Beijing Institute of Technology), Yu Li (Tencent), Shaodi You (University of Amsterdam), and Ying Fu (Beijing Institute of Technology)</i>	
Towards Semantic Segmentation of Urban-Scale 3D Point Clouds: A Dataset, Benchmarks and Challenges	4975
<i>Qingyong Hu (University of Oxford), Bo Yang (The Hong Kong Polytechnic University), Sheikh Khalid (Sensat), Wen Xiao (Newcastle University), Niki Trigoni (University of Oxford), and Andrew Markham (University of Oxford)</i>	

Neural Side-by-Side: Predicting Human Preferences for No-Reference Super-Resolution Evaluation	4986
<i>Valentin Khrulkov (Yandex) and Artem Babenko (Yandex)</i>	
Slimmable Compressive Autoencoders for Practical Neural Image Compression	4996
<i>Fei Yang (CVC (computer vision center) in UAB), Luis Herranz (Computer Vision Center), Yongmei Cheng (NWPUI), and Mikhail G. Mozerov (Computer Vision Center, UAB, Barcelona)</i>	
Distilling Knowledge via Knowledge Review	5006
<i>Pengguang Chen (Chinese University of Hong Kong), Shu Liu (SmartMore), Hengshuang Zhao (University of Oxford), and Jiaya Jia (Chinese University of Hong Kong)</i>	
Manifold Regularized Dynamic Network Pruning	5016
<i>Yehui Tang (Peking University), Yunhe Wang (Huawei Technologies), Yixing Xu (Huawei Technologies), Yiping Deng (Huawei), Chao Xu (Peking University), Dacheng Tao (The University of Sydney), and Chang Xu (University of Sydney)</i>	
Learnable Companding Quantization for Accurate Low-Bit Neural Networks	5027
<i>Kohei Yamamoto (Oki Electric Industry Co., Ltd.)</i>	
Lips Don't Lie: A Generalisable and Robust Approach To Face Forgery Detection	5037
<i>Alexandros Haliassos (Imperial College London), Konstantinos Vougioukas (Imperial College London), Stavros Petridis (Imperial College London), and Maja Pantic (Samsung AI Centre Cambridge/Imperial College London)</i>	
Guided Integrated Gradients: An Adaptive Path Method for Removing Noise	5048
<i>Andrei Kapishnikov (Google), Subhashini Venugopalan (Google), Besim Avci (Google), Ben Wedin (Google), Michael Terry (Google), and Tolga Bolukbasi (Google)</i>	
Scalable Differential Privacy With Sparse Network Finetuning	5057
<i>Zelun Luo (Stanford University), Daniel J. Wu (Stanford University), Ehsan Adeli (Stanford University), and Li Fei-Fei (Stanford University)</i>	
Deep Graph Matching Under Quadratic Constraint	5067
<i>Quankai Gao (Wuhan University), Fudong Wang (Wuhan University), Nan Xue (Wuhan University), Jin-Gang Yu (South China University of Technology), and Gui-Song Xia (Wuhan University)</i>	
T2VLAD: Global-Local Sequence Alignment for Text-Video Retrieval	5075
<i>Xiaohan Wang (University of Technology, Sydney), Linchao Zhu (University of Technology, Sydney), and Yi Yang (UTS)</i>	
FaceInpainter: High Fidelity Face Adaptation to Heterogeneous Domains	5085
<i>Jia Li (Institute of Automation, Chinese Academy of Sciences), Zhaoyang Li (Institute of Automation, Chinese Academy of Sciences), Jie Cao (Institute of Automation, Chinese Academy of Sciences), Xingguang Song (Institute of Automation, Chinese Academy of Sciences), and Ran He (Institute of Automation, Chinese Academy of Sciences)</i>	
Partition-Guided GANs	5095
<i>Mohammadreza Armandpour (Texas A&M University), Ali Sadeghian (University of Florida), Chunyuan Li (Microsoft Research, Redmond), and Mingyuan Zhou (University of Texas at Austin)</i>	

Repopulating Street Scenes	5106
<i>Yifan Wang (University of Washington), Andrew Liu (Google), Richard Tucker (Google), Jiajun Wu (Stanford University), Brian L. Curless (University of Washington), Steven M. Seitz (University of Washington), and Noah Snavely (Cornell University and Google AI)</i>	
Image Inpainting With External-Internal Learning and Monochromic Bottleneck	5116
<i>Tengfei Wang (HKUST), Hao Ouyang (HKUST), and Qifeng Chen (HKUST)</i>	
DG-Font: Deformable Generative Networks for Unsupervised Font Generation	5126
<i>Yangchen Xie (East China Normal University), Xinyuan Chen (East China Normal University), Li Sun (East China Normal University), and Yue Lu (East China Normal University)</i>	
Drafting and Revision: Laplacian Pyramid Network for Fast High-Quality Artistic Style Transfer	5137
<i>Tianwei Lin (Baidu Inc), Zhuoqi Ma (Xidian University), Fu Li (Baidu), Dongliang He (Baidu), Xin Li (Baidu), Errui Ding (Baidu Inc.), Nannan Wang (Xidian University), Jie Li (Xidian University), and Xinbo Gao (Chongqing University of Posts and Telecommunications)</i>	
StylePeople: A Generative Model of Fullbody Human Avatars	5147
<i>Artur Grigorev (Samsung), Karim Isakov (Samsung), Anastasia Ianina (Samsung), Renat Bashirov (Samsung), Ilya Zakharkin (Skoltech), Alexander Vakhitov (SLAMCore), and Victor Lempitsky (Samsung)</i>	
Synthesize-It-Classifier: Learning a Generative Classifier Through Recurrent Self-Analysis.....	5157
<i>Arghya Pal (Indian Institute of Technology Hyderabad), Raphaël C.-W. Phan (Monash University), and KokSheik Wong (Monash University Malaysia)</i>	
Understanding Object Dynamics for Interactive Image-to-Video Synthesis	5167
<i>Andreas Blattmann (Heidelberg University), Timo Milbich (Heidelberg University), Michael Dorkenwald (Heidelberg University), and Björn Ommer (Heidelberg University)</i>	
Learning Dynamic Alignment via Meta-Filter for Few-Shot Learning	5178
<i>Chengming Xu (Fudan University), Yanwei Fu (Fudan University), Chen Liu (Fudan University), Chengjie Wang (Tencent), Jilin Li (Tencent), Feiyue Huang (Tencent), Li Zhang (University of Oxford), and Xiangyang Xue (Fudan University)</i>	
Jo-SRC: A Contrastive Approach for Combating Noisy Labels	5188
<i>Yazhou Yao (Nanjing University of Science and Technology), Zeren Sun (Nanjing University of Science and Technology), Chuanyi Zhang (Nanjing University of Science and Technology), Fumin Shen (UESTC), Qi Wu (University of Adelaide), Jian Zhang (UTS), and Zhenmin Tang (Nanjing University of Science and Technology)</i>	
On Focal Loss for Class-Posterior Probability Estimation: A Theoretical Perspective	5198
<i>Nontawat Charoenphakdee (The University of Tokyo / RIKEN), Jayakorn Vongkulbhisal (IBM Research), Nuttapon Chairatanakul (Tokyo Institute of Technology), and Masashi Sugiyama (RIKEN/The University of Tokyo)</i>	

MetaSAug: Meta Semantic Augmentation for Long-Tailed Visual Recognition	5208
<i>Shuang Li (Beijing Institute of Technology), Kaixiong Gong (Beijing Institute of Technology), Chi Harold Liu (Beijing Institute of Technology), Yulin Wang (Tsinghua University), Feng Qiao (Inceptio Tech.), and Xinjing Cheng (Inceptio Tech.)</i>	
Open World Compositional Zero-Shot Learning	5218
<i>Massimiliano Mancini (University of Tübingen), Muhammad Ferjad Naeem (ETH Zürich), Yongqin Xian (Max Planck Institute Informatics), and Zeynep Akata (University of Tübingen)</i>	
Deep Texture Recognition via Exploiting Cross-Layer Statistical Self-Similarity	5227
<i>Zhile Chen (South China University of Technology), Feng Li (South China University of Technology), Yuhui Quan (South China University of Technology), Yong Xu (South China University of Technology), and Hui Ji (National University of Singapore)</i>	
Combinatorial Learning of Graph Edit Distance via Dynamic Embedding	5237
<i>Runzhong Wang (Shanghai Jiao Tong University), Tianqi Zhang (Shanghai Jiao Tong University), Tianshu Yu (Arizona State University), Junchi Yan (Shanghai Jiao Tong University), and Xiaokang Yang (Shanghai Jiao Tong University of China)</i>	
TransNAS-Bench-101: Improving Transferability and Generalizability of Cross-Task Neural Architecture Search	5247
<i>Yawen Duan (The University of Hong Kong), Xin Chen (The University of Hong Kong), Hang Xu (Huawei Noah's Ark Lab), Zewei Chen (Huawei Noah's Ark Lab), Xiaodan Liang (Sun Yat-sen University), Tong Zhang (The Hong Kong University of Science and Technology), and Zhenguo Li (Huawei Noah's Ark Lab)</i>	
An Alternative Probabilistic Interpretation of the Huber Loss	5257
<i>Gregory P. Meyer (Uber)</i>	
Joint Deep Model-Based MR Image and Coil Sensitivity Reconstruction Network (Joint-ICNet) for Fast MRI	5266
<i>Yohan Jun (Yonsei University), Hyungseob Shin (Yonsei university), Taejoon Eo (Yonsei University), and Dosik Hwang (Yonsei University)</i>	
Automatic Vertebra Localization and Identification in CT by Spine Rectification and Anatomically-Constrained Optimization	5276
<i>Fakai Wang (University of Maryland, College Park), Kang Zheng (PAII Inc.), Le Lu (PAII Inc.), Jing Xiao (Ping An Insurance, (Group) Company of China), Min Wu (University of Maryland), and Shun Miao (PAII)</i>	
Alpha-Refine: Boosting Tracking Performance by Precise Bounding Box Estimation	5285
<i>Bin Yan (Dalian University of Technology), Xinyu Zhang (Dalian University of Technology), Dong Wang (Dalian University of Technology), Huchuan Lu (Dalian University of Technology), and Xiaoyun Yang (Remark Holdings)</i>	
Learnable Graph Matching: Incorporating Graph Partitioning With Deep Feature Learning for Multiple Object Tracking	5295
<i>Jiawei He (Institute of Automation, Chinese Academy of Sciences), Zehao Huang (TuSimple), Naiyan Wang (TuSimple), and Zhaoxiang Zhang (Chinese Academy of Sciences, China)</i>	

Group-aware Label Transfer for Domain Adaptive Person Re-identification	5306
<i>Kecheng Zheng (University of Science and Technology of China), Wu Liu (AI Research of JD.com), Lingxiao He (AI Research of JD.com), Tao Mei (AI Research of JD.com), Jiebo Luo (U. Rochester), and Zheng-Jun Zha (University of Science and Technology of China)</i>	
Double Low-Rank Representation With Projection Distance Penalty for Clustering	5316
<i>Zhiqiang Fu (Beijing Jiaotong University), Yao Zhao (Beijing Jiaotong University), Dongxia Chang (Beijing Jiaotong University), Xingxing Zhang (Tsinghua University), and Yiming Wang (Beijing Jiaotong University)</i>	
Multiple Instance Active Learning for Object Detection	5326
<i>Tianning Yuan (University of Chinese Academy of Sciences), Fang Wan (University of Chinese Academy of Sciences), Mengying Fu (University of Chinese Academy of Sciences), Jianzhuang Liu (Huawei Noah's Ark Lab), Songcen Xu (Noah's Ark Lab, Huawei Technologies Co., Ltd.), Xiangyang Ji (Tsinghua University), and Qixiang Ye (University of Chinese Academy of Sciences, China)</i>	
Learning Compositional Representation for 4D Captures With Neural ODE	5336
<i>Boyan Jiang (Fudan University), Yinda Zhang (Google), Xingkui Wei (Fudan University), Xiangyang Xue (Fudan University), and Yanwei Fu (Fudan University)</i>	
Curriculum Graph Co-Teaching for Multi-Target Domain Adaptation	5347
<i>Subhankar Roy (University of Trento), Evgeny Krivosheev (University of Trento), Zhun Zhong (University of Trento), Nicu Sebe (University of Trento), and Elisa Ricci (University of Trento)</i>	
Instance Level Affinity-Based Transfer for Unsupervised Domain Adaptation	5357
<i>Astuti Sharma (UCSD), Tarun Kalluri (UCSD), and Manmohan Chandraker (UC San Diego)</i>	
Deep Stable Learning for Out-of-Distribution Generalization	5368
<i>Xingxuan Zhang (Tsinghua University), Peng Cui (Tsinghua University), Renzhe Xu (Tsinghua University), Linjun Zhou (Tsinghua University), Yue He (Tsinghua University), and Zheyang Shen (Tsinghua University)</i>	
ORDisCo: Effective and Efficient Usage of Incremental Unlabeled Data for Semi-Supervised Continual Learning	5379
<i>Liyuan Wang (Tsinghua University), Kuo Yang (Huawei Noah's Ark Lab), Chongxuan Li (Tsinghua University), Lanqing Hong (Huawei Noah's Ark Lab), Zhenguo Li (Huawei Noah's Ark Lab), and Jun Zhu (Tsinghua University)</i>	
Dynamic Metric Learning: Towards a Scalable Metric Space To Accommodate Multiple Semantic Scales	5389
<i>Yifan Sun (Megvii), Yuke Zhu (Megvii Research Shanghai), Yuhan Zhang (Beihang University), Pengkun Zheng (Megvii), Xi Qiu (Megvii Inc), Chi Zhang (Megvii Inc.), and Yichen Wei (Megvii Research Shanghai)</i>	
Learning Cross-Modal Retrieval With Noisy Labels	5399
<i>Peng Hu (Sichuan University), Xi Peng (Sichuan University), Hongyuan Zhu (Astar), Liangli Zhen (Institute of High Performance Computing, A*STAR), and Jie Lin (Institute for Infocomm Research, I2R, Singapore)</i>	

How Well Do Self-Supervised Models Transfer?	5410
<i>Linus Ericsson (University of Edinburgh), Henry Gouk (University of Edinburgh), and Timothy M. Hospedales (Edinburgh University)</i>	
Generic Perceptual Loss for Modeling Structured Output Dependencies	5420
<i>Yifan Liu (University of Adelaide), Hao Chen (The University of Adelaide), Yu Chen (Motovis Inc.), Wei Yin (University of Adelaide), and Chunhua Shen (University of Adelaide)</i>	
EDNet: Efficient Disparity Estimation With Cost Volume Combination and Attention-Based Spatial Residual	5429
<i>Songyan Zhang (Tongji University), Zhicheng Wang (Tongji University), Qiang Wang (Hong Kong Baptist University), Jinshuo Zhang (Tongji University), Gang Wei (Tongji University), and Xiaowen Chu (Hong Kong Baptist University)</i>	
BoxInst: High-Performance Instance Segmentation With Box Annotations	5439
<i>Zhi Tian (The University of Adelaide), Chunhua Shen (University of Adelaide), Xinlong Wang (University of Adelaide), and Hao Chen (The University of Adelaide)</i>	
PhySG: Inverse Rendering With Spherical Gaussians for Physics-Based Material Editing and Relighting	5449
<i>Kai Zhang (Cornell University), Fujun Luan (Cornell University), Qianqian Wang (Cornell), Kavita Bala (Cornell University), and Noah Snavely (Cornell University and Google AI)</i>	
MaX-DeepLab: End-to-End Panoptic Segmentation With Mask Transformers	5459
<i>Huiyu Wang (Johns Hopkins University), Yukun Zhu (Google Inc.), Hartwig Adam (Google), Alan Yuille (Johns Hopkins University), and Liang-Chieh Chen (Google Inc.)</i>	
Scale-Aware Graph Neural Network for Few-Shot Semantic Segmentation	5471
<i>Guo-Sen Xie (Inception Institute of Artificial Intelligence), Jie Liu (Northeastern University), Huan Xiong (Mohamed bin Zayed University of Artificial Intelligence, (MBZUAI)), and Ling Shao (Inception Institute of Artificial Intelligence)</i>	
Part-Aware Panoptic Segmentation	5481
<i>Daan de Geus (Eindhoven University of Technology), Panagiotis Meletis (Eindhoven University of Technology), Chenyang Lu (Eindhoven University of Technology), Xiaoxiao Wen (University of Amsterdam), and Gijs Dubbelman (Eindhoven University of Technology)</i>	
Railroad Is Not a Train: Saliency As Pseudo-Pixel Supervision for Weakly Supervised Semantic Segmentation	5491
<i>Seungho Lee (Yonsei University), Minhyun Lee (Yonsei University), Jongwuk Lee (Sungkyunkwan University), and Hyunjung Shim (Yonsei University)</i>	

Mask-Embedded Discriminator With Region-Based Semantic Regularization for Semi-Supervised Class-Conditional Image Synthesis	5502
<i>Yi Liu (South China University of Technology), Xiaoyang Huo (south china university of technology), Tianyi Chen (South China University of Technology), Xiangping Zeng (South China University of Technology), Si Wu (South China University of Technology), Zhiwen Yu (South China University of Technology), and Hau-San Wong (City University of Hong Kong)</i>	
Unsupervised Hyperbolic Representation Learning via Message Passing Auto-Encoders	5512
<i>Jiwoong Park (Seoul National University), Junho Cho (Seoul National University), Hyung Jin Chang (University of Birmingham), and Jin Young Choi (Seoul National University)</i>	
4D Panoptic LiDAR Segmentation	5523
<i>Mehmet Aygün (TU Munich), Aljoša Ošep (TUM Munich), Mark Weber (Technical University Munich), Maxim Maximov (TUM), Cyrill Stachniss (University of Bonn), Jens Behley (University of Bonn), and Laura Leal-Taixé (TUM)</i>	
EffiScene: Efficient Per-Pixel Rigidity Inference for Unsupervised Joint Learning of Optical Flow, Depth, Camera Pose and Motion Segmentation	5534
<i>Yang Jiao (Xidian University), Trac D. Tran (Johns Hopkins University), and Guangming Shi (Xidian University)</i>	
Learning by Aligning Videos in Time	5544
<i>Sanjay Haresh (Retrocausal, Inc), Sateesh Kumar (Retrocausal), Huseyin Coskun (Technical University of Munich), Shahram N. Syed (Retrocausal Inc.), Andrey Konin (Retrocausal), Zeeshan Zia (Retrocausal, Inc.), and Quoc-Huy Tran (Retrocausal, Inc.)</i>	
Modular Interactive Video Object Segmentation: Interaction-to-Mask, Propagation and Difference-Aware Fusion	5555
<i>Ho Kei Cheng (HKUST), Yu-Wing Tai (Kuaishou Technology / HKUST), and Chi-Keung Tang (Hong Kong University of Science and Technology)</i>	
Polygonal Point Set Tracking	5565
<i>Gunhee Nam (Yonsei Univ.), Miran Heo (Yonsei University), Seoung Wug Oh (Adobe Research), Joon-Young Lee (Adobe Research), and Seon Joo Kim (Yonsei University)</i>	
VinVL: Revisiting Visual Representations in Vision-Language Models	5575
<i>Pengchuan Zhang (Microsoft Research AI), Xiujun Li (Microsoft Research), Xiaowei Hu (Microsoft), Jianwei Yang (Microsoft Research), Lei Zhang (Microsoft), Lijuan Wang (Microsoft), Yejin Choi (University of Washington), and Jianfeng Gao (Microsoft Research)</i>	
Visual Semantic Role Labeling for Video Understanding	5585
<i>Arka Sadhu (University of Southern California), Tanmay Gupta (Allen Institute for Artificial Intelligence), Mark Yatskar (University of Washington), Ram Nevatia (U of Southern California), and Aniruddha Kembhavi (Allen Institute for Artificial Intelligence)</i>	
Can Audio-Visual Integration Strengthen Robustness Under Multimodal Attacks?	5597
<i>Yapeng Tian (University of Rochester) and Chenliang Xu (University of Rochester)</i>	

Relation-aware Instance Refinement for Weakly Supervised Visual Grounding	5608
<i>Yongfei Liu (ShanghaiTech), Bo Wan (ShanghaiTech University), Lin Ma (Meituan), and Xuming He (ShanghaiTech University)</i>	
Learning Better Visual Dialog Agents With Pretrained Visual-Linguistic Representation	5618
<i>Tao Tu (National Taiwan University), Qing Ping (Amazon Alexa AI), Govindarajan Thattai (Amazon), Gokhan Tur (Amazon Alexa AI), and Prem Natarajan (Amazon.com Inc.)</i>	
Separating Skills and Concepts for Novel Visual Question Answering	5628
<i>Spencer Whitehead (University of Illinois at Urbana-Champaign), Hui Wu (MIT-IBM Watson AI Lab, IBM Research), Heng Ji (University of Illinois at Urbana-Champaign), Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research), and Kate Saenko (Boston University)</i>	
Generating Manga From Illustrations via Mimicking Manga Creation Workflow	5638
<i>Lumin Zhang (Soochow University / Style2Paints Research), Xinrui Wang (The University of Tokyo/Tencent), Qingnan Fan (Stanford University), Yi Ji (Soochow University), and Chunping Liu (School of Computer Science and Technology, Soochow University)</i>	
SelfDoc: Self-Supervised Document Representation Learning	5648
<i>Peizhao Li (Brandeis University), Jiuxiang Gu (Adobe Research), Jason Kuen (Adobe Research), Vlad I. Morariu (Adobe Research), Handong Zhao (Adobe Research), Rajiv Jain (Adobe Research), Varun Manjunatha (Adobe Research), and Hongfu Liu (Brandeis University)</i>	
Affect2MM: Affective Analysis of Multimedia Content Using Emotion Causality	5657
<i>Trisha Mittal (University of Maryland), Puneet Mathur (University of Maryland), Aniket Bera (University of Maryland, College Park), and Dinesh Manocha (University of Maryland at College Park)</i>	
Vectorization and Rasterization: Self-Supervised Learning for Sketch and Handwriting	5668
<i>Ayan Kumar Bhunia (University of Surrey), Pinaki Nath Chowdhury (University of Surrey), Yongxin Yang (University of Surrey), Timothy M. Hospedales (Edinburgh University), Tao Xiang (University of Surrey), and Yi-Zhe Song (University of Surrey)</i>	

Session 05

Layer-Wise Searching for 1-Bit Detectors	5678
<i>Sheng Xu (Beihang University), Junhe Zhao (Beihang University), Jinhua Lü (Beihang University, Beijing, China), Baochang Zhang (Beihang University), Shumin Han (Baidu Inc.), and David Doermann (University at Buffalo)</i>	
Weakly Supervised Learning of Rigid 3D Scene Flow	5688
<i>Zan Gojic (ETH Zürich), Or Litany (NVIDIA), Andreas Wieser (ETH Zürich), Leonidas J. Guibas (Stanford University), and Tolga Birdal (TU Munich)</i>	
Learning Compositional Radiance Fields of Dynamic Human Heads	5700
<i>Ziyang Wang (Carnegie Mellon University), Timur Bagautdinov (Facebook), Stephen Lombardi (Facebook), Tomas Simon (Facebook Reality Labs), Jason Saragih (Facebook), Jessica Hodgins (Carnegie Mellon University), and Michael Zollhöfer (Facebook Reality Labs)</i>	

Learning Accurate Dense Correspondences and When To Trust Them	5710
<i>Prune Truong (ETH Zurich), Martin Danelljan (ETH Zurich), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)</i>	
RSN: Range Sparse Net for Efficient, Accurate LiDAR 3D Object Detection	5721
<i>Pei Sun (Waymo), Weiyue Wang (Waymo), Yuning Chai (Alphabet), Gamaleldin Elsayed (Google Research, Brain Team), Alex Bewley (Google), Xiao Zhang (Google), Cristian Sminchisescu (Google), and Dragomir Anguelov (Waymo)</i>	
LAFEAT: Piercing Through Adversarial Defenses With Latent Features	5731
<i>Yunrui Yu (University of Macau), Xitong Gao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), and Cheng-Zhong Xu (University of Macau)</i>	
Function4D: Real-Time Human Volumetric Capture From Very Sparse Consumer RGBD Sensors	5742
<i>Tao Yu (Tsinghua University), Zerong Zheng (Tsinghua University), Kaiwen Guo (Google), Pengpeng Liu (Institute of Automation, Chinese Academy of Sciences), Qionghai Dai (Tsinghua University), and Yebin Liu (Tsinghua University)</i>	
Polka Lines: Learning Structured Illumination and Reconstruction for Active Stereo	5753
<i>Seung-Hwan Baek (Princeton University) and Felix Heide (Princeton / Algolux)</i>	
FBI-Denoiser: Fast Blind Image Denoiser for Poisson-Gaussian Noise	5764
<i>Jaeseok Byun (Sungkyunkwan university), Sungmin Cha (Sungkyunkwan University), and Taesup Moon (Sungkyunkwan University)</i>	
Face Forensics in the Wild	5774
<i>Tianfei Zhou (ETH Zurich), Wenguan Wang (Eidgenössische Technische Hochschule Zürich), Zhiyuan Liang (Beijing Institute of technology), and Jianbing Shen (Inception Institute of Artificial Intelligence)</i>	
Exploring Adversarial Fake Images on Face Manifold	5785
<i>Dongze Li (Institute of Automation, Chinese Academy of Science), Wei Wang (Center for Research on Intelligent Perception and Computing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences), Hongxing Fan (Institute of Automation Chinese Academy of Sciences), and Jing Dong (Chinese Academy of Sciences)</i>	
Pi-GAN: Periodic Implicit Generative Adversarial Networks for 3D-Aware Image Synthesis	5795
<i>Eric R. Chan (Stanford University), Marco Monteiro (Stanford University), Petr Kellnhofer (Stanford University), Jiajun Wu (Stanford University), and Gordon Wetzstein (Stanford University)</i>	
Animating Pictures With Eulerian Motion Fields	5806
<i>Aleksander Holynski (University of Washington), Brian L. Curless (University of Washington), Steven M. Seitz (University of Washington), and Richard Szeliski (FaceBook)</i>	
DriveGAN: Towards a Controllable High-Quality Neural Simulation	5816
<i>Seung Wook Kim (University of Toronto, NVIDIA), Jonah Philion (University of Toronto, NVIDIA), Antonio Torralba (MIT), and Sanja Fidler (University of Toronto, NVIDIA)</i>	

Towards Open World Object Detection	5826
<i>K J Joseph (Indian Institute of Technology, Hyderabad), Salman Khan (Australian National University, (ANU)), Fahad Shahbaz Khan (MBZUAI), and Vineeth N Balasubramanian (Indian Institute of Technology, Hyderabad)</i>	
DiNTS: Differentiable Neural Network Topology Search for 3D Medical Image Segmentation	5837
<i>Yufan He (Johns Hopkins University), Dong Yang (NVIDIA Corporation), Holger Roth (NVIDIA), Can Zhao (Nvidia), and Daguang Xu (NVIDIA Corporation)</i>	
Siamese Natural Language Tracker: Tracking by Natural Language Descriptions With Siamese Trackers	5847
<i>Qi Feng (Boston University), Vitaly Ablavsky (University of Washington), Qinxun Bai (Horizon Robotics), and Stan Sclaroff (Boston University)</i>	
Where and What? Examining Interpretable Disentangled Representations	5857
<i>Xinqi Zhu (University of Sydney), Chang Xu (University of Sydney), and Dacheng Tao (The University of Sydney)</i>	
Prototype Augmentation and Self-Supervision for Incremental Learning	5867
<i>Fei Zhu (Institute of Automation of Chinese Academy of Science), Xu-Yao Zhang (Institute of Automation of Chinese Academy of Sciences), Chuang Wang (Institute of Automation of Chinese Academy of Sciences), Fei Yin (Institute of Automation of Chinese Academy of Sciences), and Cheng-Lin Liu (Institute of Automation of Chinese Academy of Sciences)</i>	
Brain Image Synthesis With Unsupervised Multivariate Canonical CSCI4Net	5877
<i>Yawen Huang (Malong LLC), Feng Zheng (SUSTech), Danyang Wang (Malong LLC), Weilin Huang (Malong LLC), Matthew R. Scott (Malong LLC), and Ling Shao (Inception Institute of Artificial Intelligence)</i>	
Polygonal Building Extraction by Frame Field Learning	5887
<i>Nicolas Girard (Inria Sophia-Antipolis), Dmitriy Smirnov (MIT), Justin Solomon (MIT), and Yuliya Tarabalka (Inria Sophia-Antipolis)</i>	
InverseForm: A Loss Function for Structured Boundary-Aware Segmentation	5897
<i>Shubhankar Borse (Qualcomm AI Research), Ying Wang (Qualcomm AI research), Yizhe Zhang (Qualcomm AI Research), and Fatih Porikli (Qualcomm AI Research)</i>	
SSTVOS: Sparse Spatiotemporal Transformers for Video Object Segmentation	5908
<i>Brendan Duke (ModiFace Inc), Abdalla Ahmed (ModiFace Inc.), Christian Wolf (ModiFace Inc.), Parham Aarabi (INSA Lyon, France), and Graham W. Taylor (University of Guelph)</i>	
Visual Room Rearrangement	5918
<i>Luca Weihs (Allen Institute for Artificial Intelligence), Matt Deitke (Allen Institute for AI), Aniruddha Kembhavi (Allen Institute for Artificial Intelligence), and Roozbeh Mottaghi (Allen Institute for AI)</i>	
A Deep Emulator for Secondary Motion of 3D Characters	5928
<i>Mianlun Zheng (University of Southern California), Yi Zhou (University of Southern California), Duygu Ceylan (Adobe Research), and Jernej Barbič (University of Southern California)</i>	

Interactive Self-Training With Mean Teachers for Semi-Supervised Object Detection	5937
<i>Qize Yang (DAMO Academy, Alibaba Group), Xihan Wei (Alibaba), Biao Wang (Alibaba), Xian-Sheng Hua (Hong Kong Polytechnic University, Hong Kong, China), and Lei Zhang (Alibaba Group)</i>	
UniT: Unified Knowledge Transfer for Any-Shot Object Detection and Segmentation	5947
<i>Siddhesh Khandelwal (University of British Columbia), Raghav Goyal (University of British Columbia), and Leonid Sigal (University of British Columbia)</i>	
Unsupervised Object Detection With LIDAR Clues	5958
<i>Hao Tian (SenseTime), Yuntao Chen (CASIA), Jifeng Dai (SenseTime), Zhaoxiang Zhang (Chinese Academy of Sciences, China), and Xizhou Zhu (SenseTime)</i>	
Implicit Feature Alignment: Learn To Convert Text Recognizer to Text Spotter	5969
<i>Tianwei Wang (South China University of Technology), Yuanzhi Zhu (South China University of Technology), Lianwen Jin (South China University of Technology), Dezhi Peng (South China University of Technology), Zhe Li (South China University of Technology), Mengchao He (DAMO Academy, Alibaba Group), Yongpan Wang (Alibaba Group), and Canjie Luo (South China University of Technology)</i>	
Self-Attention Based Text Knowledge Mining for Text Detection	5979
<i>Qi Wan (Shenzhen University), Haoqin Ji (Shenzhen University), and Linlin Shen (Shenzhen University)</i>	
Shallow Feature Matters for Weakly Supervised Object Localization	5989
<i>Jun Wei (The Chinese University of Hong Kong, Shenzhen), Qin Wang (The Chinese University of Hongkong, Shenzhen), Zhen Li (Chinese University of Hong Kong, Shenzhen), Sheng Wang (Tencent), S. Kevin Zhou (CAS), and Shuguang Cui (The Chinese University of Hong Kong, Shenzhen)</i>	
Self-Supervised 3D Mesh Reconstruction From Single Images	5998
<i>Tao Hu (The Chinese University of Hong Kong), Liwei Wang (Tencent), Xiaogang Xu (The Chinese University of Hong Kong), Shu Liu (SmartMore), and Jiaya Jia (Chinese University of Hong Kong)</i>	
Sketch2Model: View-Aware 3D Modeling From Single Free-Hand Sketches	6008
<i>Song-Hai Zhang (Tsinghua University), Yuan-Chen Guo (Tsinghua University), and Qing-Wen Gu (Tsinghua University)</i>	
Learning Parallel Dense Correspondence From Spatio-Temporal Descriptors for Efficient and Robust 4D Reconstruction	6018
<i>Jiapeng Tang (South China University of Technology), Dan Xu (The Hong Kong University of Science and Technology), Kui Jia (South China University of Technology), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)</i>	
Refer-It-in-RGBD: A Bottom-Up Approach for 3D Visual Grounding in RGBD Images	6028
<i>Haolin Liu (The Chinese University of Hong Kong, Shenzhen), Anran Lin (The Chinese University of Hong Kong, Shenzhen), Xiaoguang Han (Shenzhen Research Institute of Big Data, the Chinese University of Hong Kong, (Shenzhen)), Lei Yang (University of Hong Kong), Yizhou Yu (The University of Hong Kong), and Shuguang Cui (The Chinese University of Hong Kong, Shenzhen)</i>	

VoxelContext-Net: An Octree Based Framework for Point Cloud Compression	6038
<i>Zizheng Que (Beihang University), Guo Lu (Beijing Institute of Technology), and Dong Xu (University of Sydney)</i>	
CorrNet3D: Unsupervised End-to-End Learning of Dense Correspondence for 3D Point Clouds ..	6048
<i>Yiming Zeng (City University of Hong Kong), Yue Qian (City University of Hong Kong, Hong Kong), Zhiyu Zhu (City University of Hong Kong), Junhui Hou (City University of Hong Kong, Hong Kong), Hui Yuan (School of Control Science and Engineering, Shandong University), and Ying He (Nanyang Technological University)</i>	
Inferring CAD Modeling Sequences Using Zone Graphs	6058
<i>Xianghao Xu (Brown University), Wenzhe Peng (MIT), Chin-Yi Cheng (Autodesk Research), Karl D.D. Willis (Autodesk Research), and Daniel Ritchie (Brown University)</i>	
Seeing Behind Objects for 3D Multi-Object Tracking in RGB-D Sequences	6067
<i>Norman Müller (Technical University of Munich), Yu-Shiang Wong (University College London), Niloy J. Mitra (University College London), Angela Dai (Technical University of Munich), and Matthias Nießner (Technical University of Munich)</i>	
View Generalization for Single Image Textured 3D Models	6077
<i>Anand Bhattad (UIUC), Aysegul Dundar (Bilkent University), Guilin Liu (NVIDIA), Andrew Tao (NVIDIA), and Bryan Catanzaro (NVIDIA)</i>	
A Decomposition Model for Stereo Matching	6087
<i>Chengtang Yao (Beijing Institute of Technology), Yunde Jia (Beijing Institute of Technology), Huijun Di (Beijing Institute of Technology), Pengxiang Li (Beijing Institute of Technology), and Yuwei Wu (Beijing Institute of Technology)</i>	
VS-Net: Voting With Segmentation for Visual Localization	6097
<i>Zhaoyang Huang (Zhejiang University), Han Zhou (Zhejiang University), Yijin Li (Zhejiang University), Bangbang Yang (Zhejiang University), Yan Xu (Chinese University of Hong Kong), Xiaowei Zhou (Zhejiang Univ., China), Hujun Bao (Zhejiang University), Guofeng Zhang (Zhejiang University), and Hongsheng Li (Chinese University of Hong Kong)</i>	
MonoRec: Semi-Supervised Dense Reconstruction in Dynamic Environments From a Single Moving Camera	6108
<i>Felix Wimbauer (Technical University of Munich), Nan Yang (Technical University of Munich), Lukas von Stumberg (TUM), Niclas Zeller (Artisense), and Daniel Cremers (TU Munich)</i>	
Shape and Material Capture at Home	6119
<i>Daniel Lichy (University of Maryland), Jiaye Wu (University of Maryland), Soumyadip Sengupta (University of Washington), and David W. Jacobs (University of Maryland, USA)</i>	
Offboard 3D Object Detection From Point Cloud Sequences	6130
<i>Charles R. Qi (Waymo), Yin Zhou (Waymo), Mahyar Najibi (Waymo LLC), Pei Sun (Waymo), Khoa Vo (Waymo LLC), Boyang Deng (Waymo LLC), and Dragomir Anguelov (Waymo)</i>	

M3DSSD: Monocular 3D Single Stage Object Detector	6141
<i>Shujie Luo (Zhejiang University), Hang Dai (Mohamed bin Zayed University of Artificial Intelligence), Ling Shao (Inception Institute of Artificial Intelligence), and Yong Ding (Zhejiang University)</i>	
2D or not 2D? Adaptive 3D Convolution Selection for Efficient Video Recognition	6151
<i>Hengduo Li (University of Maryland, College Park), Zuxuan Wu (UMD), Abhinav Shrivastava (University of Maryland), and Larry S. Davis (University of Maryland)</i>	
Deep Analysis of CNN-Based Spatio-Temporal Representations for Action Recognition	6161
<i>Chun-Fu Richard Richard Chen (MIT-IBM Watson AI Lab, IBM Research AI), Rameswar Panda (MIT-IBM Watson AI Lab, IBM Research), Kandan Ramakrishnan (Baylor college of medicine), Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research), John Cohn (IBM), Aude Oliva (MIT), and Quanfu Fan (IBM Research)</i>	
The Blessings of Unlabeled Background in Untrimmed Videos	6172
<i>Yuan Liu (Damo Academy, Alibaba Group), Jingyuan Chen (Damo Academy, Alibaba Group), Zhenfang Chen (The university of Hong Kong), Bing Deng (Damo Academy, Alibaba Group), Jianqiang Huang (Damo Academy, Alibaba Group), and Hanwang Zhang (Nanyang Technological University)</i>	
PointGuard: Provably Robust 3D Point Cloud Classification	6182
<i>Hongbin Liu (Duke University), Jinyuan Jia (Duke University), and Neil Zhenqiang Gong (Duke University)</i>	
DSRNA: Differentiable Search of Robust Neural Architectures	6192
<i>Ramtin Hosseini (University of California San Diego), Xingyi Yang (University of California San Diego), and Pengtao Xie (UC San Diego)</i>	
Backdoor Attacks Against Deep Learning Systems in the Physical World	6202
<i>Emily Wenger (University of Chicago), Josephine Passananti (University of Chicago), Arjun Nitin Bhagoji (University of Chicago), Yuanshun Yao (University of Chicago), Haitao Zheng (University of Chicago), and Ben Y. Zhao (University of Chicago)</i>	
Riggable 3D Face Reconstruction via In-Network Optimization	6212
<i>Ziqian Bai (Simon Fraser University), Zhaopeng Cui (Zhejiang University), Xiaoming Liu (Michigan State University), and Ping Tan (Simon Fraser University)</i>	
NeuralHumanFVV: Real-Time Neural Volumetric Human Performance Rendering Using RGB Cameras	6222
<i>Xin Suo (ShanghaiTech university), Yuheng Jiang (ShanghaiTech University), Pei Lin (ShanghaiTech University), Yingliang Zhang (DGene), Minye Wu (ShanghaiTech University), Kaiwen Guo (Google), and Lan Xu (HKUST)</i>	
Context Modeling in 3D Human Pose Estimation: A Unified Perspective	6234
<i>Xiaoxuan Ma (Peking University), Jiajun Su (Peking University), Chunyu Wang (Microsoft Research asia), Hai Ci (Peking University), and Yizhou Wang (PKU)</i>	

Dive Into Ambiguity: Latent Distribution Mining and Pairwise Uncertainty Estimation for Facial Expression Recognition	6244
<i>Jiahui She (Nanjing University), Yibo Hu (JD AI Research), Hailin Shi (JD AI Research), Jun Wang (JD AI Research), Qiu Shen (Nanjing University), and Tao Mei (AI Research of JD.com)</i>	
Lifting 2D StyleGAN for 3D-Aware Face Generation	6254
<i>Yichun Shi (Michigan State University), Divyansh Aggarwal (Michigan State University), and Anil K. Jain (Michigan State University)</i>	
Hybrid Message Passing With Performance-Driven Structures for Facial Action Unit Detection....	6263
<i>Tengfei Song (Southeast University), Zijun Cui (Rensselaer Polytechnic Institute), Wenming Zheng (Southeast University), and Qiang Ji (Rensselaer Polytechnic Institute)</i>	
Learning to Generalize Unseen Domains via Memory-based Multi-Source Meta-Learning for Person Re-Identification	6273
<i>Yuyang Zhao (Xiamen University), Zhun Zhong (University of Trento), Fengxiang Yang (Xiamen University), Zhiming Luo (Xiamen University), Yaojin Lin (Minnan Normal University), Shaozi Li (Xiamen University, China), and Nicu Sebe (University of Trento)</i>	
Invertible Image Signal Processing	6283
<i>Yazhou Xing (HKUST), Zian Qian (HKUST), and Qifeng Chen (HKUST)</i>	
End-to-End High Dynamic Range Camera Pipeline Optimization	6293
<i>Nicolas Robidoux (Algolux), Luis E. García Capel (Algolux), Dong-eun Seo (Algolux), Avinash Sharma (Algolux), Federico Ariza (Algolux), and Felix Heide (Princeton / Algolux)</i>	
Blind Deblurring for Saturated Images	6304
<i>Liang Chen (East China Normal University), Jiawei Zhang (SenseTime Research), Songnan Lin (Beijing Institute of Technology), Faming Fang (East China Normal University), and Jimmy S. Ren (SenseTime Research; Qing Yuan Research Institute, Shanghai Jiao Tong University)</i>	
Extreme Low-Light Environment-Driven Image Denoising Over Permanently Shadowed Lunar Regions With a Physical Noise Model	6313
<i>Ben Moseley (University of Oxford), Valentin Bickel (Max Planck Institute for Solar System Research), Ignacio G. López-Francos (NASA), and Loveneesh Rana (University of Luxembourg)</i>	
Controlling the Rain: From Removal to Rendering	6324
<i>Siqi Ni (Nanjing University), Xueyun Cao (Nanjing University), Tao Yue (Nanjing University), and Xuemei Hu (Nanjing University)</i>	
De-Rendering the World’s Revolutionary Artefacts	6334
<i>Shangzhe Wu (University of Oxford), Ameesh Makadia (Google Research), Jiajun Wu (Stanford University), Noah Snavely (Cornell University and Google AI), Richard Tucker (Google), and Angjoo Kanazawa (University of California Berkeley)</i>	
Progressively Complementary Network for Fisheye Image Rectification Using Appearance Flow	6344
<i>Shangrong Yang (Beijing Jiaotong University), Chunyu Lin (Beijing Jiaotong University), Kang Liao (Beijing Jiaotong University), Chunjie Zhang (Beijing Jiaotong University), and Yao Zhao (Beijing Jiaotong University)</i>	

High-Speed Image Reconstruction Through Short-Term Plasticity for Spiking Cameras	6354
<i>Yajing Zheng (Peking University), Lingxiao Zheng (Peking University), Zhaofei Yu (Peking University), Boxin Shi (Peking University), Yonghong Tian (Peking University), and Tiejun Huang (Peking University)</i>	
MASA-SR: Matching Acceleration and Spatial Adaptation for Reference-Based Image Super-Resolution	6364
<i>Liyang Lu (The Chinese University of Hong Kong), Wenbo Li (The Chinese University of Hong Kong), Xin Tao (Kuaishou), Jiangbo Lu (SmartMore Corporation), and Jiaya Jia (Chinese University of Hong Kong)</i>	
Single Pair Cross-Modality Super Resolution	6374
<i>Guy Shacht (Tel-Aviv University), Dov Danon (Amazon), Sharon Fogel (Tel Aviv University), and Daniel Cohen-Or (Tel Aviv University)</i>	
Temporal Modulation Network for Controllable Space-Time Video Super-Resolution	6384
<i>Gang Xu (Nankai University), Jun Xu (Nankai University), Zhen Li (Nankai University), Liang Wang (NLPR, China), Xing Sun (Tencent), and Ming-Ming Cheng (Nankai University)</i>	
The Multi-Temporal Urban Development SpaceNet Dataset	6394
<i>Adam Van Etten (In-Q-Tel), Daniel Hogan (CosmiQ Works, In-Q-Tel), Jesus Martinez Manso (Planet), Jacob Shermeyer (CosmiQ Works, In-Q-Tel), Nicholas Weir (Amazon), and Ryan Lewis (Amazon)</i>	
Euro-PVI: Pedestrian Vehicle Interactions in Dense Urban Centers	6404
<i>Apratim Bhattacharyya (Max Planck Institute for Informatics), Daniel Olmeda Reino (Toyota Motor Europe), Mario Fritz (CISPA Helmholtz Center for Information Security), and Bernt Schiele (MPI Informatics)</i>	
AttentiveNAS: Improving Neural Architecture Search via Attentive Sampling	6414
<i>Dilin Wang (Facebook), Meng Li (Facebook Inc), Chengyue Gong (UT Austin), and Vikas Chandra (Facebook)</i>	
Learning Student Networks in the Wild	6424
<i>Hanting Chen (Peking University), Tianyu Guo (Peking University), Chang Xu (University of Sydney), Wenshuo Li (Huawei Technologies), Chunjing Xu (Huawei Noah's Ark Lab), Chao Xu (Peking University), and Yunhe Wang (Huawei Technologies)</i>	
Towards Compact CNNs via Collaborative Compression	6434
<i>Yuchao Li (Alibaba Group), Shaohui Lin (East China Normal University), Jianzhuang Liu (Huawei Noah's Ark Lab), Qixiang Ye (University of Chinese Academy of Sciences, China), Mengdi Wang (Alibaba Group), Fei Chao (School of Informatics), Fan Yang (Huawei), Jincheng Ma (HiSilicon Technologies), Qi Tian (Huawei Cloud & AI), and Rongrong Ji (Xiamen University, China)</i>	
Network Quantization With Element-Wise Gradient Scaling	6444
<i>Junghyup Lee (Yonsei University), Dohyung Kim (Yonsei University), and Bumsub Ham (Yonsei University)</i>	

Frequency-Aware Discriminative Feature Learning Supervised by Single-Center Loss for Face Forgery Detection	6454
<i>Jiaming Li (University of Science and Technology of China), Hongtao Xie (University of Science and Technology of China), Jiahong Li (Kuaishou), Zhongyuan Wang (Kuaishou), and Yongdong Zhang (University of Science and Technology of China)</i>	
Building Reliable Explanations of Unreliable Neural Networks: Locally Smoothing Perspective of Model Interpretation	6464
<i>Dohun Lim (Jeonbuk National University), Hyeonseok Lee (Jeonbuk National University), and Sungchan Kim (Jeonbuk National University)</i>	
Perceptual Indistinguishability-Net (PI-Net): Facial Image Obfuscation With Manipulable Semantics	6474
<i>Jia-Wei Chen (Academia Sinica), Li-Ju Chen (Academia Sinica), Chia-Mu Yu (National Chiao Tung University), and Chun-Shien Lu (Academia Sinica)</i>	
Coming Down to Earth: Satellite-to-Street View Synthesis for Geo-Localization	6484
<i>Aysim Toker (TUM), Qunjie Zhou (Technical University of Munich), Maxim Maximov (TUM), and Laura Leal-Taixé (TUM)</i>	
Neural Scene Flow Fields for Space-Time View Synthesis of Dynamic Scenes	6494
<i>Zhengqi Li (Cornell University), Simon Niklaus (Adobe Research), Noah Snavely (Cornell University and Google AI), and Oliver Wang (Adobe Systems Inc)</i>	
Not Just Compete, but Collaborate: Local Image-to-Image Translation via Cooperative Mask Prediction	6505
<i>Daejin Kim (Korea Advanced Institute of Science and Technology), Mohammad Azam Khan (Dhaka Power Distribution Company Limited), and Jaegul Choo (Korea Advanced Institute of Science and Technology)</i>	
Posterior Promoted GAN With Distribution Discriminator for Unsupervised Image Synthesis	6515
<i>Xianchao Zhang (Dalian University of Technology), Ziyang Cheng (Dalian University Of Technology), Xiaotong Zhang (Dalian University of Technology), and Han Liu (Dalian University of Technology)</i>	
Surrogate Gradient Field for Latent Space Manipulation	6525
<i>Minjun Li (Preferred Networks), Yanghua Jin (Preferred Networks), and Huachun Zhu (Preferred Networks)</i>	
Image Inpainting Guided by Coherence Priors of Semantics and Textures	6535
<i>Liang Liao (National Institute of Informatics), Jing Xiao (Wuhan University), Zheng Wang (The University of Tokyo), Chia-Wen Lin (National Tsing Hua University), and Shin'ichi Satoh (National Institute of Informatics)</i>	
Spatially-Invariant Style-Codes Controlled Makeup Transfer	6545
<i>Han Deng (South China University of Technology), Chu Han (Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences), Hongmin Cai (South China University of Technology), Guoqiang Han (South China University of Technology), and Shengfeng He (South China University of Technology)</i>	
Memory-Guided Unsupervised Image-to-Image Translation	6554
<i>Somi Jeong (Yonsei University), Youngjung Kim (ADD), Eungbean Lee (Yonsei University), and Kwanghoon Sohn (Yonsei Univ.)</i>	

Hierarchical Motion Understanding via Motion Programs	6564
<i>Sumith Kulal (Stanford University), Jiayuan Mao (MIT), Alex Aiken (Stanford University), and Jiajun Wu (Stanford University)</i>	
Adaptive Rank Estimate in Robust Principal Component Analysis	6573
<i>Zhengqin Xu (Wuhan University of Science and Technology), Rui He (Wuhan University of Science and Technology), Shoulie Xie (Institute for Infocomm Research, Singapore), and Shiqian Wu (School of Machinery and Automation, Wuhan University of Science and Technology)</i>	
Deep Animation Video Interpolation in the Wild	6583
<i>Li Siyao (SenseTime), Shiyu Zhao (SenseTime), Weijiang Yu (SUN YAT-SEN UNIVERSITY), Wenxiu Sun (SenseTime Research and Tetras.AI), Dimitris Metaxas (Rutgers), Chen Change Loy (Nanyang Technological University), and Ziwei Liu (Nanyang Technological University)</i>	
ECKPN: Explicit Class Knowledge Propagation Network for Transductive Few-Shot Learning	6592
<i>Chaofan Chen (University of Science and Technology of China), Xiaoshan Yang (CASIA), Changsheng Xu (CASIA), Xuhui Huang (X Lab, The Second Academy of CASIC), and Zhe Ma (X Lab, The Second Academy of CASIC)</i>	
Multi-Objective Interpolation Training for Robustness To Label Noise	6602
<i>Diego Ortego (Insight Centre for Data Analytics, (DCU)), Eric Arazo (Insight Centre for Data Analytics, (DCU)), Paul Albert (Insight Centre for Data Analytics, (DCU)), Noel E. O’Connor (Dublin City University, (DCU)), and Kevin McGuinness (Insight Centre for Data Analytics)</i>	
T-vMF Similarity for Regularizing Intra-Class Feature Distribution	6612
<i>Takumi Kobayashi (National Institute of Advanced Industrial Science and Technology)</i>	
Disentangling Label Distribution for Long-Tailed Visual Recognition	6622
<i>Youngkyu Hong (KAIST), Seungju Han (Hyperconnect), Kwanghee Choi (Hyperconnect), Seokjun Seo (Hyperconnect), Beomsu Kim (Hyperconnect), and Buru Chang (Hyperconnect)</i>	
Leveraging the Availability of Two Cameras for Illuminant Estimation	6633
<i>Abdelrahman Abdelhamed (Samsung), Abhijith Punnappurath (Samsung AI Center Toronto), and Michael S. Brown (York University)</i>	
Decoupled Dynamic Filter Networks	6643
<i>Jingkai Zhou (South China University of Technology), Varun Jampani (Google), Zhixiong Pi (Huazhong University of Science and Technology), Qiong Liu (South China University of Technology), and Ming-Hsuan Yang (University of California at Merced)</i>	
Rethinking Graph Neural Architecture Search From Message-Passing	6653
<i>Shaofei Cai (Institute of Computing Technology, Chinese Academy of Sciences), Liang Li (Institute of Computing Technology, Chinese Academy of Sciences), Jincan Deng (Institute of Computing Technology, Chinese Academy of Sciences), Beichen Zhang (University of Chinese Academy of Sciences), Zheng-Jun Zha (University of Science and Technology of China), Li Su (University of Chinese Academy of Sciences), and Qingming Huang (University of Chinese Academy of Sciences)</i>	

Towards Improving the Consistency, Efficiency, and Flexibility of Differentiable Neural Architecture Search	6663
<i>Yibo Yang (Peking University), Shan You (SenseTime), Hongyang Li (Peking University), Fei Wang (SenseTime), Chen Qian (SenseTime), and Zhouchen Lin (Peking University)</i>	
Unsupervised Visual Attention and Invariance for Reinforcement Learning	6673
<i>Xudong Wang (UC Berkeley / ICSI), Long Lian (UC Berkeley), and Stella X. Yu (UC Berkeley / ICSI)</i>	
Mol2Image: Improved Conditional Flow Models for Molecule to Image Synthesis	6684
<i>Karren Yang (MIT), Samuel Goldman (MIT), Wengong Jin (MIT), Alex X. Lu (University of Toronto), Regina Barzilay (Massachusetts institute of technology), Tommi Jaakkola (MIT), and Caroline Uhler (MIT)</i>	
TSGCNet: Discriminative Geometric Feature Learning With Two-Stream Graph Convolutional Network for 3D Dental Model Segmentation	6695
<i>Lingming Zhang (Chongqing University of Posts and Telecommunications), Yue Zhao (Chongqing University of Posts and telecommunications), Deyu Meng (Chongqing University of Posts and Telecommunications), Zhiming Cui (HKU), Chenqiang Gao (Xi'an Jiaotong University), Xinbo Gao (Chongqing University of Posts and Telecommunications), Chunfeng Lian (Xi'an Jiaotong University), and Dinggang Shen (United Imaging Intelligence)</i>	
IoU Attack: Towards Temporally Coherent Black-Box Adversarial Attack for Visual Object Tracking	6705
<i>Shuai Jia (Shanghai Jiao Tong University), Yibing Song (Tencent), Chao Ma (Shanghai Jiao Tong University), and Xiaokang Yang (Shanghai Jiao Tong University of China)</i>	
GMOT-40: A Benchmark for Generic Multiple Object Tracking	6715
<i>Hexin Bai (Temple University), Wensheng Cheng (Stony Brook University), Peng Chu (Microsoft), Juehuan Liu (Temple University), Kai Zhang (Temple University), and Haibin Ling (Stony Brook University)</i>	
Combined Depth Space Based Architecture Search for Person Re-Identification	6725
<i>Hanjun Li (Sun Yat-sen University), Gaojie Wu (Sun Yat-sen University), and Wei-Shi Zheng (Sun Yat-sen University, China)</i>	
Learning an Explicit Weighting Scheme for Adapting Complex HSI Noise	6735
<i>Xiangyu Rui (Xi'an Jiaotong University), Xiangyong Cao (Xi'an Jiaotong University), Qi Xie (Xi'an Jiaotong University), Zongsheng Yue (Xi'an Jiaotong University), Qian Zhao (Xi'an Jiaotong University), and Deyu Meng (Xi'an Jiaotong University)</i>	
VaB-AL: Incorporating Class Imbalance and Difficulty With Variational Bayes for Active Learning	6745
<i>Jongwon Choi (Chung-Ang University), Kwang Moo Yi (University of British Columbia), Jihoon Kim (Samsung SDS), Jinho Choo (Samsung SDS), Byoungjip Kim (Samsung SDS), Jinyeop Chang (Samsung SDS), Youngjune Gwon (Samsung SDS), and Hyung Jin Chang (University of Birmingham)</i>	

Learning a Facial Expression Embedding Disentangled From Identity	6755
<i>Wei Zhang (Netease Fuxi AI Lab), Xianpeng Ji (Netease Fuxi AI Lab), Keyu Chen (University of Science and Technology of China), Yu Ding (Netease Fuxi AI Lab), and Changjie Fan (NetEase Fuxi AI Lab)</i>	
SRDAN: Scale-Aware and Range-Aware Domain Adaptation Network for Cross-Dataset 3D Object Detection	6765
<i>Weichen Zhang (University of Sydney), Wen Li (University of Electronic Science and Technology of China), and Dong Xu (University of Sydney)</i>	
Regressive Domain Adaptation for Unsupervised Keypoint Detection	6776
<i>Junguang Jiang (Tsinghua University), Yifei Ji (Tsinghua University), Ximei Wang (Tsinghua University), Yufeng Liu (Kuaishou Technology), Jianmin Wang (Tsinghua University, China), and Mingsheng Long (Tsinghua University)</i>	
Uncertainty-Guided Model Generalization to Unseen Domains	6786
<i>Fengchun Qiao (University of Delaware) and Xi Peng (University of Delaware)</i>	
Self-Promoted Prototype Refinement for Few-Shot Class-Incremental Learning	6797
<i>Kai Zhu (USTC), Yang Cao (University of Science and Technology of China), Wei Zhai (University of Science and Technology of China), Jie Cheng (Huawei Technologies Co. Ltd.), and Zheng-Jun Zha (University of Science and Technology of China)</i>	
Noise-Resistant Deep Metric Learning With Ranking-Based Instance Selection	6807
<i>Chang Liu (Nanyang Technological University), Han Yu (Nanyang Technological University, (NTU)), Boyang Li (Nanyang Technological University), Zhiqi Shen (NTU), Zhanning Gao (Alibaba Group), Peiran Ren (Alibaba Group), Xuansong Xie (Alibaba), Lizhen Cui (ShanDong University), and Chunyan Miao (NTU)</i>	
DAT: Training Deep Networks Robust To Label-Noise by Matching the Feature Distributions	6817
<i>Yuntao Qu (beihang University), Shasha Mo (beihang university), and Jianwei Niu (beihang university)</i>	
OBoW: Online Bag-of-Visual-Words Generation for Self-Supervised Learning	6826
<i>Spyros Gidaris (valeo.ai), Andrei Bursuc (valeo.ai), Gilles Puy (Valeo), Nikos Komodakis (ENPC, France), Matthieu Cord (Sorbonne University), and Patrick Pérez (Valeo.ai)</i>	
Learning Affinity-Aware Upsampling for Deep Image Matting	6837
<i>Yutong Dai (The University of Adelaide), Hao Lu (Huazhong University of Science and Technology), and Chunhua Shen (University of Adelaide)</i>	
PLADE-Net: Towards Pixel-Level Accuracy for Self-Supervised Single-View Depth Estimation With Neural Positional Encoding and Distilled Matting Loss	6847
<i>Juan Luis Gonzalez (KAIST-VICLab) and Munchurl Kim (Korea Advanced Institute of Science and Technology)</i>	
RefineMask: Towards High-Quality Instance Segmentation With Fine-Grained Features	6857
<i>Gang Zhang (Tsinghua University), Xin Lu (SenseTime Group Limited), Jingru Tan (Tongji University), Jianmin Li (Tsinghua University), Zhaoxiang Zhang (Chinese Academy of Sciences, China), Quanquan Li (SenseTime Research), and Xiaolin Hu (Tsinghua University)</i>	

CompositeTasking: Understanding Images by Spatial Composition of Tasks	6866
<i>Nikola Popović (ETH Zürich), Danda Pani Paudel (ETH Zürich), Thomas Probst (ETH Zurich), Guolei Sun (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
Rethinking Semantic Segmentation From a Sequence-to-Sequence Perspective With Transformers.....	6877
<i>Sixiao Zheng (Fudan University), Jiachen Lu (Shanghai Jiaotong University), Hengshuang Zhao (University of Oxford), Xiatian Zhu (University of Surrey), Zekun Luo (Tencent), Yabiao Wang (Tencent), Yanwei Fu (Fudan University), Jianfeng Feng (Fudan University), Tao Xiang (University of Surrey), Philip H.S. Torr (University of Oxford), and Li Zhang (University of Oxford)</i>	
FSDR: Frequency Space Domain Randomization for Domain Generalization	6887
<i>Jiaxing Huang (Nanyang Technological University), Dayan Guan (Nanyang Technological University), Aoran Xiao (Nanyang Technological University), and Shijian Lu (Nanyang Technological University)</i>	
Transformation Driven Visual Reasoning	6899
<i>Xin Hong (CAS Key Laboratory of Network Data Science and Technology, Institute of Computing Technology, Chinese Academy of Sciences), Yanyan Lan (Institute of Computing Technology), Liang Pang (Institute of Computing Technology, Chinese Academy of Sciences), Jiafeng Guo (Institute of Computing Technology, Chinese Academy of Sciences), and Xueqi Cheng (Institute of Computing Technology, Chinese Academy of Sciences)</i>	
Background-Aware Pooling and Noise-Aware Loss for Weakly-Supervised Semantic Segmentation	6909
<i>Youngmin Oh (Yonsei University), Beomjun Kim (Yonsei University), and Bumsub Ham (Yonsei University)</i>	
Adaptive Consistency Regularization for Semi-Supervised Transfer Learning	6919
<i>Abulikemu Abuduweili (Peking University), Xingjian Li (Baidu Research), Humphrey Shi (U of Oregon; UIUC), Cheng-Zhong Xu (University of Macau), and Dejing Dou (Baidu)</i>	
Self-Generated Defocus Blur Detection via Dual Adversarial Discriminators	6929
<i>Wenda Zhao (Dalian University of Technology), Cai Shang (Dalian University of Technology), and Huchuan Lu (Dalian University of Technology)</i>	
Ego-Exo: Transferring Visual Representations From Third-Person to First-Person Videos	6939
<i>Yanghao Li (Facebook AI Research), Tushar Nagarajan (UT Austin), Bo Xiong (Facebook AI Research), and Kristen Grauman (Facebook AI Research & UT Austin)</i>	
PV-RAFT: Point-Voxel Correlation Fields for Scene Flow Estimation of Point Clouds	6950
<i>Yi Wei (Tsinghua University), Ziyi Wang (Tsinghua University), Yongming Rao (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)</i>	

Spatiotemporal Contrastive Video Representation Learning	6960
<i>Rui Qian (Cornell University), Tianjian Meng (Google Brain), Boqing Gong (Google), Ming-Hsuan Yang (University of California at Merced), Huisheng Wang (Google), Serge Belongie (Cornell University), and Yin Cui (Google)</i>	
Deep Video Matting via Spatio-Temporal Alignment and Aggregation	6971
<i>Yanan Sun (HKUST), Guanzhi Wang (Stanford University), Qiao Gu (Carnegie Mellon University), Chi-Keung Tang (Hong Kong University of Science and Technology), and Yu-Wing Tai (Kuaishou Technology / HKUST)</i>	
Target-Aware Object Discovery and Association for Unsupervised Video Multi-Object Segmentation	6981
<i>Tianfei Zhou (ETH Zurich), Jianwu Li (Beijing Institute of Technology), Xuwei Li (Beijing Institute of Technology), and Ling Shao (Inception Institute of Artificial Intelligence)</i>	
Multimodal Contrastive Training for Visual Representation Learning	6991
<i>Xin Yuan (University of Chicago), Zhe Lin (Adobe Research), Jason Kuen (Adobe Research), Jianming Zhang (Adobe Research), Yilin Wang (Adobe), Michael Maire (University of Chicago), Ajinkya Kale (Adobe Research), and Baldo Faieta (Adobe)</i>	
Vx2Text: End-to-End Learning of Video-Based Text Generation From Multimodal Inputs	7001
<i>Xudong Lin (Columbia University), Gedas Bertasius (Facebook AI), Jue Wang (Facebook AI), Shih-Fu Chang (Columbia University), Devi Parikh (Georgia Tech & Facebook AI Research), and Lorenzo Torresani (Facebook AI)</i>	
Distilling Audio-Visual Knowledge by Compositional Contrastive Learning	7012
<i>Yanbei Chen (University of Tübingen), Yongqin Xian (Max Planck Institute Informatics), A. Sophia Koepke (University of Tübingen), Ying Shan (Tencent), and Zeynep Akata (University of Tübingen)</i>	
Structured Multi-Level Interaction Network for Video Moment Localization via Language Query	7022
<i>Hao Wang (University of Science and Technology of China), Zheng-Jun Zha (University of Science and Technology of China), Liang Li (Institute of Computing Technology, Chinese Academy of Sciences), Dong Liu (University of Science and Technology of China), and Jiebo Luo (U. Rochester)</i>	
Scene-Intuitive Agent for Remote Embodied Visual Grounding	7032
<i>Xiangru Lin (The University of Hong Kong), Guanbin Li (Sun Yat-sen University), and Yizhou Yu (The University of Hong Kong)</i>	
Domain-Robust VQA With Diverse Datasets and Methods but No Target Labels	7042
<i>Mingda Zhang (University of Pittsburgh), Tristan Maidment (University of Pittsburgh), Ahmad Diab (University of Pittsburgh), Adriana Kovashka (University of Pittsburgh), and Rebecca Hwa (University of Pittsburgh)</i>	

Composing Photos Like a Photographer	7053
<i>Chaoyi Hong (Huazhong Univ. of Sci.&Tech.), Shuaiyuan Du (Huazhong Univ. of Sci.&Tech.), Ke Xian (Huazhong University of Science and Technology), Hao Lu (Huazhong University of Science and Technology), Zhiguo Cao (Huazhong Univ. of Sci.&Tech.), and Weicai Zhong (Huawei CBG Consumer Cloud Service Big Data Platform Dept.)</i>	
Dogfight: Detecting Drones From Drones Videos	7063
<i>Muhammad Waseem Ashraf (ITU), Waqas Sultani (Information Technology University), and Mubarak Shah (University of Central Florida)</i>	
Multi-Modal Fusion Transformer for End-to-End Autonomous Driving	7073
<i>Aditya Prakash (Max Planck Institute for Intelligent Systems, Tuebingen), Kashyap Chitta (MPI-IS and University of Tuebingen), and Andreas Geiger (MPI-IS and University of Tuebingen)</i>	
Cloud2Curve: Generation and Vectorization of Parametric Sketches	7084
<i>Ayan Das (University of Surrey), Yongxin Yang (University of Surrey), Timothy M. Hospedales (Edinburgh University), Tao Xiang (University of Surrey), and Yi-Zhe Song (University of Surrey)</i>	

Session 06

Read Like Humans: Autonomous, Bidirectional and Iterative Language Modeling for Scene Text Recognition	7094
<i>Shancheng Fang (University of Science and Technology of China), Hongtao Xie (University of Science and Technology of China), Yuxin Wang (University of Science and Technology of China), Zhendong Mao (University of Science and Technology of China), and Yongdong Zhang (University of Science and Technology of China)</i>	
MultiBodySync: Multi-Body Segmentation and Motion Estimation via 3D Scan Synchronization ..	7104
<i>Jiahui Huang (Tsinghua University), He Wang (Stanford University), Tolga Birdal (TU Munich), Minhyuk Sung (KAIST), Federica Arrigoni (University of Trento), Shi-Min Hu (Tsinghua University), and Leonidas J. Guibas (Stanford University)</i>	
NeuTex: Neural Texture Mapping for Volumetric Neural Rendering	7115
<i>Fanbo Xiang (UCSD), Zexiang Xu (Adobe Research), Miloš Hašan (Adobe Research), Yannick Hold-Geoffroy (Adobe Research), Kalyan Sunkavalli (Adobe Research), and Hao Su (UCSD)</i>	
UnsupervisedR&R: Unsupervised Point Cloud Registration via Differentiable Rendering	7125
<i>Mohamed El Banani (University of Michigan), Luya Gao (University of Michigan), and Justin Johnson (University of Michigan)</i>	
RangeIoUDet: Range Image Based Real-Time 3D Object Detector Optimized by Intersection Over Union	7136
<i>Zhidong Liang (Hikvision), Zehan Zhang (Shanghai Jiao Tong University & Hangzhou Hikvision Digital Technology Co. Ltd), Ming Zhang (hikvision), Xian Zhao (Hikvision), and Shiliang Pu (Hikvision Research Institute)</i>	
Architectural Adversarial Robustness: The Case for Deep Pursuit	7146
<i>George Cazenavette (Carnegie Mellon University), Calvin Murdock (Carnegie Mellon University), and Simon Lucey (CMU)</i>	

SimPoE: Simulated Character Control for 3D Human Pose Estimation	7155
<i>Ye Yuan (Carnegie Mellon University), Shih-En Wei (Facebook), Tomas Simon (Facebook Reality Labs), Kris Kitani (Carnegie Mellon University), and Jason Saragih (Facebook)</i>	
CodedStereo: Learned Phase Masks for Large Depth-of-Field Stereo	7166
<i>Shiyu Tan (Rice University), Yicheng Wu (Rice University), Shoou-I Yu (Facebook Reality Labs Research Pittsburgh), and Ashok Veeraraghavan (Rice University)</i>	
PSD: Principled Synthetic-to-Real Dehazing Guided by Physical Priors	7176
<i>Zeyuan Chen (University of Science and Technology of China), Yangchao Wang (University of Electronic Science and Technology of China), Yang Yang (University of Electronic Science and Technology of China), and Dong Liu (University of Science and Technology of China)</i>	
OpenRooms: An Open Framework for Photorealistic Indoor Scene Datasets	7186
<i>Zhengqin Li (UC San Diego), Ting-Wei Yu (UC San Diego), Shen Sang (University of California, San Diego), Sarah Wang (UC San Diego), Meng Song (University of California, San Diego), Yuhan Liu (University of California, San Diego), Yu-Ying Yeh (UC San Diego), Rui Zhu (University of California San Diego), Nitesh Gundavarapu (University of California San Diego), Jia Shi (UC San Diego), Sai Bi (UC San Diego), Hong-Xing Yu (Stanford University), Zexiang Xu (Adobe Research), Kalyan Sunkavalli (Adobe Research), Miloš Hašan (Adobe Research), Ravi Ramamoorthi (University of California San Diego), and Manmohan Chandraker (UC San Diego)</i>	
A Closer Look at Fourier Spectrum Discrepancies for CNN-Generated Images Detection	7196
<i>Keshigeyan Chandrasegaran (Singapore University of Technology and Design), Ngoc-Trung Tran (Singapore University of Technology and Design), and Ngai-Man Cheung (Singapore University of Technology and Design)</i>	
NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections	7206
<i>Ricardo Martin-Brualla (Google), Noha Radwan (Google), Mehdi S. M. Sajjadi (Google), Jonathan T. Barron (Google Research), Alexey Dosovitskiy (Google), and Daniel Duckworth (Google)</i>	
ID-Unet: Iterative Soft and Hard Deformation for View Synthesis	7216
<i>Mingyu Yin (East China Normal University), Li Sun (East China Normal University), and Qingli Li (East China Normal University)</i>	
GeoSim: Realistic Video Simulation via Geometry-Aware Composition for Self-Driving	7226
<i>Yun Chen (Uber ATG), Frieda Rong (Uber ATG), Shivam Duggal (Delhi Technological University), Shenlong Wang (Uber ATG, University of Toronto), Xinchun Yan (Uber ATG), Sivabalan Manivasagam (University of Toronto), Shangjie Xue (MIT), Ersin Yumer (Uber ATG), and Raquel Urtasun (Uber ATG)</i>	
All Labels Are Not Created Equal: Enhancing Semi-Supervision via Label Grouping and Co-Training	7237
<i>Islam Nassar (Monash University), Samitha Herath (Monash University), Ehsan Abbasnejad (The University of Adelaide), Wray Buntine (Monash University), and Gholamreza Haffari (Monash University, Australia)</i>	

Orthogonal Over-Parameterized Training	7247
<i>Weiyang Liu (University of Cambridge), Rongmei Lin (Emory University), Zhen Liu (Mila, University of Montreal), James M. Rehg (Georgia Institute of Technology), Liam Paull (Université de Montréal), Li Xiong (Emory University), Le Song (Georgia Institute of Technology & Ant Financial), and Adrian Weller (University of Cambridge)</i>	
DeepTag: An Unsupervised Deep Learning Method for Motion Tracking on Cardiac Tagging Magnetic Resonance Images	7257
<i>Meng Ye (Rutgers University), Mikael Kanski (New York University School of Medicine), Dong Yang (NVIDIA Corporation), Qi Chang (Rutgers), Zhennan Yan (SenseBrain Technology), Qiaoying Huang (Rutgers University), Leon Axel (NYU), and Dimitris Metaxas (Rutgers)</i>	
Transferable Query Selection for Active Domain Adaptation	7268
<i>Bo Fu (Tsinghua University), Zhangjie Cao (Tsinghua University), Jianmin Wang (Tsinghua University, China), and Mingsheng Long (Tsinghua University)</i>	
When Age-Invariant Face Recognition Meets Face Age Synthesis: A Multi-Task Learning Framework	7278
<i>Zhizhong Huang (Fudan University), Junping Zhang (Fudan University), and Hongming Shan (Fudan University)</i>	
Simpler Certified Radius Maximization by Propagating Covariances	7288
<i>Xingjian Zhen (University of Wisconsin-Madison), Rudrasis Chakraborty (Amazon), and Vikas Singh (University of Wisconsin Madison)</i>	
Improving Panoptic Segmentation at All Scales	7298
<i>Lorenzo Porzi (Facebook), Samuel Rota Bulò (Facebook), and Peter Kontschieder (Facebook)</i>	
Learning Triadic Belief Dynamics in Nonverbal Communication From Videos	7308
<i>Lifeng Fan (University of California, Los Angeles), Shuwen Qiu (University of California, Los Angeles), Zilong Zheng (UCLA), Tao Gao (UCLA), Song-Chun Zhu (UCLA), and Yixin Zhu (UCLA)</i>	
Guided Interactive Video Object Segmentation Using Reliability-Based Attention Maps	7318
<i>Yuk Heo (Korea University), Yeong Jun Koh (Chungnam National University), and Chang-Su Kim (Korea university)</i>	
Less Is More: ClipBERT for Video-and-Language Learning via Sparse Sampling	7327
<i>Jie Lei (UNC Chapel Hill), Linjie Li (Microsoft), Luowei Zhou (Microsoft), Zhe Gan (Microsoft), Tamara L. Berg (UNC Chapel Hill, USA), Mohit Bansal (University of North Carolina at Chapel Hill), and Jingjing Liu (Microsoft)</i>	
Im2Vec: Synthesizing Vector Graphics Without Vector Supervision	7338
<i>Pradyumna Reddy (University College London), Michaël Gharbi (Adobe Research), Michal Lukáč (Adobe Research), and Niloy J. Mitra (University College London)</i>	
FSCE: Few-Shot Object Detection via Contrastive Proposal Encoding	7348
<i>Bo Sun (University of Southern California), Banghuai Li (Megvii), Shengcai Cai (Megvii), Ye Yuan (Megvii), and Chi Zhang (Megvii Inc.)</i>	

Beyond Max-Margin: Class Margin Equilibrium for Few-Shot Object Detection	7359
<i>Bohao Li (University of Chinese Academy of Sciences), Boyu Yang (University of Chinese Academy of Sciences), Chang Liu (University of Chinese Academy of Sciences), Feng Liu (University of Chinese Academy of Sciences), Rongrong Ji (Xiamen University, China), and Qixiang Ye (University of Chinese Academy of Sciences, China)</i>	
Dynamic Head: Unifying Object Detection Heads With Attentions	7369
<i>Xiyang Dai (Microsoft), Yinpeng Chen (Microsoft), Bin Xiao (Microsoft), Dongdong Chen (Microsoft Cloud AI), Mengchen Liu (Microsoft), Lu Yuan (Microsoft), and Lei Zhang (Microsoft)</i>	
Dictionary-Guided Scene Text Recognition	7379
<i>Nguyen Nguyen (VinAI Research, Vietnam), Thu Nguyen (VinAI Research, Vietnam), Vinh Tran (Stony Brook University), Minh-Triet Tran (Ho Chi Minh University of Science, VNU), Thanh Duc Ngo (University of Information Technology, VNU-HCM), Thien Huu Nguyen (University of Oregon), and Minh Hoai (Stony Brook University)</i>	
Progressive Contour Regression for Arbitrary-Shape Scene Text Detection	7389
<i>Pengwen Dai (Institute of Information Engineering, Chinese Academy of Sciences), Sanyi Zhang (Tianjin University), Hua Zhang (Institute of Information Engineering, Chinese Academy of Sciences), and Xiaochun Cao (Chinese Academy of Sciences)</i>	
Strengthen Learning Tolerance for Weakly Supervised Object Localization	7399
<i>Guangyu Guo (Northwestern Polytechnical University), Junwei Han (NWPU, China), Fang Wan (University of Chinese Academy of Sciences), and Dingwen Zhang (Xidian University)</i>	
StruMonoNet: Structure-Aware Monocular 3D Prediction	7409
<i>Zhenpei Yang (The University of Texas at Austin), Li Erran Li (Amazon / Columbia University), and Qixing Huang (The University of Texas at Austin)</i>	
Fully Understanding Generic Objects: Modeling, Segmentation, and Reconstruction	7419
<i>Feng Liu (Michigan State University), Luan Tran (Facebook), and Xiaoming Liu (Michigan State University)</i>	
Exploiting & Refining Depth Distributions With Triangulation Light Curtains	7430
<i>Yaadhav Raaj (CMU), Siddharth Ancha (Carnegie Mellon University), Robert Tamburo (Carnegie Mellon University), David Held (CMU), and Srinivasa G. Narasimhan (Carnegie Mellon University, USA)</i>	
PMP-Net: Point Cloud Completion by Learning Multi-Step Point Moving Paths	7439
<i>Xin Wen (Tsinghua University), Peng Xiang (Tsinghua University), Zhizhong Han (Wayne State University), Yan-Pei Cao (Y-tech, Kuaishou Technology), Pengfei Wan (Kuaishou Technology), Wen Zheng (Kuaishou Technology), and Yu-Shen Liu (Tsinghua University)</i>	
TearingNet: Point Cloud Autoencoder To Learn Topology-Friendly Representations	7449
<i>Jiahao Pang (InterDigital Communications, Inc.), Duanshun Li (University of Alberta), and Dong Tian (InterDigital)</i>	
3D Object Detection With Pointformer	7459
<i>Xuran Pan (Tsinghua University), Zhuofan Xia (Tsinghua University), Shiji Song (Tsinghua), Li Erran Li (Amazon / Columbia University), and Gao Huang (Tsinghua)</i>	

NeuroMorph: Unsupervised Shape Interpolation and Correspondence in One Go	7469
<i>Marvin Eisenberger (TU Munich), David Novotny (Facebook AI Research), Gael Kerchenbaum (Facebook AI Research), Patrick Labatut (Facebook AI Research), Natalia Neverova (Facebook AI Research), Daniel Cremers (TU Munich), and Andrea Vedaldi (University of Oxford / Facebook AI Research)</i>	
Towards Part-Based Understanding of RGB-D Scans	7480
<i>Alexey Bokhovkin (Skoltech), Vladislav Ishimtsev (SkolTech), Emil Bogomolov (Skoltech), Denis Zorin (New York University), Alexey Artemov (Skoltech), Evgeny Burnaev (Skoltech), and Angela Dai (Technical University of Munich)</i>	
NeRV: Neural Reflectance and Visibility Fields for Relighting and View Synthesis	7491
<i>Pratul P. Srinivasan (Google Research), Boyang Deng (Google), Xiuming Zhang (MIT), Matthew Tancik (UC Berkeley), Ben Mildenhall (UC Berkeley), and Jonathan T. Barron (Google Research)</i>	
Probabilistic Model Distillation for Semantic Correspondence	7501
<i>Xin Li (Group 42), Deng-Ping Fan (Inception Institute of Artificial Intelligence), Fan Yang (Group 42), Ao Luo (Megvii), Hong Cheng (UESTC), and Zicheng Liu (Microsoft)</i>	
SceneGraphFusion: Incremental 3D Scene Graph Prediction From RGB-D Sequences	7511
<i>Shun-Cheng Wu (Technical University of Munich), Johanna Wald (Technical University of Munich), Keisuke Tateno (Technical University Munich), Nassir Navab (TU Munich, Germany), and Federico Tombari (Google, TU Munich)</i>	
Self-Supervised Learning of Depth Inference for Multi-View Stereo	7522
<i>Jiayu Yang (The Australian National University), Jose M. Alvarez (NVIDIA), and Miaomiao Liu (The Australian National University)</i>	
Mesoscopic Photogrammetry With an Unstabilized Phone Camera	7531
<i>Kevin C. Zhou (Duke University), Colin Cooke (Duke University), Jaehee Park (Duke University), Ruobing Qian (Duke University), Roarke Horstmeyer (Duke University), Joseph A. Izatt (Duke University), and Sina Farsiu (Duke University)</i>	
LiDAR R-CNN: An Efficient and Universal 3D Object Detector	7542
<i>Zhichao Li (TuSimple), Feng Wang (TuSimple), and Naiyan Wang (TuSimple)</i>	
Monocular 3D Object Detection: An Extrinsic Parameter Free Approach	7552
<i>Yunsong Zhou (Shanghai Jiao Tong University), Yuan He (SenseTime Research), Hongzi Zhu (Shanghai Jiao Tong University), Cheng Wang (SenseTime), Hongyang Li (SenseTime), and Qinhong Jiang (SenseTime Research; Shanghai AI Laboratory)</i>	
Beyond Short Clips: End-to-End Video-Level Learning With Collaborative Memories	7563
<i>Xitong Yang (University of Maryland), Haoqi Fan (Facebook AI Research), Lorenzo Torresani (Facebook AI), Larry S. Davis (University of Maryland), and Heng Wang (Facebook AI)</i>	
Multimodal Motion Prediction With Stacked Transformers	7573
<i>Yicheng Liu (CUHK), Jinghuai Zhang (SenseTime Research), Liangji Fang (SenseTime Research), Qinhong Jiang (SenseTime Research; Shanghai AI Laboratory), and Bolei Zhou (CUHK)</i>	

Weakly Supervised Action Selection Learning in Video	7583
<i>Junwei Ma (Layer6 AI), Satya Krishna Gorti (Layer6 AI), Maksims Volkovs (Layer6 AI), and Guangwei Yu (Layer6 AI)</i>	
BASAR:Black-Box Attack on Skeletal Action Recognition	7593
<i>Yunfeng Diao (Southwest Jiaotong University), Tianjia Shao (Zhejiang University), Yong-Liang Yang (University of Bath), Kun Zhou (Zhejiang University), and He Wang (Leeds University)</i>	
Adversarial Robustness Across Representation Spaces	7604
<i>Pranjal Awasthi (Rutgers University/Google), George Yu (Google), Chun-Sung Ferng (Google), Andrew Tomkins (Google), and Da-Cheng Juan (Google)</i>	
img2pose: Face Alignment and Detection via 6DoF, Face Pose Estimation	7613
<i>Vitor Albiero (University of Notre Dame), Xingyu Chen (Facebook AI), Xi Yin (Facebook AI), Guan Pang (Facebook), and Tal Hassner (Facebook AI)</i>	
OSTeC: One-Shot Texture Completion	7624
<i>Baris Gecer (Imperial College London), Jiankang Deng (Imperial College London), and Stefanos Zafeiriou (Imperial College London)</i>	
Locally Aware Piecewise Transformation Fields for 3D Human Mesh Registration	7635
<i>Shaofei Wang (ETH Zurich), Andreas Geiger (MPI-IS and University of Tuebingen), and Siyu Tang (ETH Zurich)</i>	
Monocular 3D Multi-Person Pose Estimation by Integrating Top-Down and Bottom-Up Networks	7645
<i>Yu Cheng (National University of Singapore), Bo Wang (Tencent America), Bo Yang (Tencent America), and Robby T. Tan (National University of Singapore)</i>	
Feature Decomposition and Reconstruction Learning for Effective Facial Expression Recognition	7656
<i>Delian Ruan (Xiamen University), Yan Yan (Xiamen University), Shenqi Lai (Meituan), Zhenhua Chai (Meituan), Chunhua Shen (University of Adelaide), and Hanzi Wang (Xiamen University)</i>	
SDD-FIQA: Unsupervised Face Image Quality Assessment With Similarity Distribution Distance.	7666
<i>Fu-Zhao Ou (Guangzhou University), Xingyu Chen (Youtu Lab), Ruixin Zhang (Tencent), Yuge Huang (Tencent YouTu), Shaoxin Li (Tencent), Jilin Li (Tencent), Yong Li (Nanjing University of Science and Technology), Liujuan Cao (Xiamen University), and Yuan-Gen Wang (Guangzhou University)</i>	
Facial Action Unit Detection With Transformers	7676
<i>Geethu Miriam Jacob (Rakuten Institute of Technology) and Björn Stenger (Rakuten Institute of Technology)</i>	

Anchor-Free Person Search	7686
<i>Yichao Yan (inception institute of artificial intelligence), Jinpeng Li (Inception Institute of Artificial Intelligence), Jie Qin (Inception Institute of Artificial Intelligence), Song Bai (University of Oxford), Shengcai Liao (Inception Institute of Artificial Intelligence), Li Liu (the inception institute of artificial intelligence), Fan Zhu (Inception Institute of Artificial Intelligence), and Ling Shao (Inception Institute of Artificial Intelligence)</i>	
Neural Camera Simulators	7696
<i>Hao Ouyang (HKUST), Zifan Shi (HKUST), Chenyang Lei (HKUST), Ka Lung Law (SenseTime), and Qifeng Chen (HKUST)</i>	
Neural Auto-Exposure for High-Dynamic Range Object Detection	7706
<i>Emmanuel Onzon (Algolux), Fahim Mannan (Algolux), and Felix Heide (Princeton / Algolux)</i>	
ARVo: Learning All-Range Volumetric Correspondence for Video Deblurring	7717
<i>Dongxu Li (THE AUSTRALIAN NATIONAL UNIVERSITY), Chenchen Xu (The Australian National University), Kaihao Zhang (Australian National University), Xin Yu (Australian National University), Yiran Zhong (Australian National University), Wenqi Ren (Institute of Information Engineering, Chinese Academy of Sciences), Hanna Suominen (The Australian National University, Data61/CSIRO, and University of Turku, (Finland)), and Hongdong Li (Australian National University, Australia)</i>	
Memory Oriented Transfer Learning for Semi-Supervised Image Deraining	7728
<i>Huaibo Huang (Institute of Automation, Chinese Academy of Sciences), Aijing Yu (Institute of Automation, Chinese Academy of Sciences), and Ran He (Institute of Automation, Chinese Academy of Sciences)</i>	
Robust Representation Learning With Feedback for Single Image Deraining	7738
<i>Chenghao Chen (Shanghai Jiao Tong University) and Hao Li (Shanghai Jiao Tong University)</i>	
A Multi-Task Network for Joint Specular Highlight Detection and Removal	7748
<i>Gang Fu (Wuhan University), Qing Zhang (Sun Yat-sen University), Lei Zhu (The Chinese University of Hong Kong), Ping Li (The Hong Kong Polytechnic University), and Chunxia Xiao (Wuhan University)</i>	
Panoramic Image Reflection Removal	7758
<i>Yuchen Hong (Peking University), Qian Zheng (Nanyang Technological University), Lingran Zhao (Peking University), Xudong Jiang (Nanyang Technological University), Alex C. Kot (Nanyang Technological University), and Boxin Shi (Peking University)</i>	
Turning Frequency to Resolution: Video Super-Resolution via Event Cameras	7768
<i>Yongcheng Jing (The University of Sydney), Yiding Yang (Stevens Institute of Technology), Xinchao Wang (National University of Singapore), Mingli Song (Zhejiang University), and Dacheng Tao (The University of Sydney)</i>	
SRWarp: Generalized Image Super-Resolution under Arbitrary Transformation	7778
<i>Sanghyun Son (Seoul National University) and Kyoung Mu Lee (Seoul National University)</i>	

Learning Scene Structure Guidance via Cross-Task Knowledge Transfer for Single Depth Super-Resolution	7788
<i>Baoli Sun (Dalian University of Technology), Xinchen Ye (Dalian University of Technology), Baopu Li (BAIDU USA LLC), Haojie Li (Dalian University of Technology), Zhihui Wang (Dalian University of Technology), and Rui Xu (Dalian University of Techonology)</i>	
Gated Spatio-Temporal Attention-Guided Video Deblurring	7798
<i>Maitreya Suin (Indian Institute of Technology Madras) and A. N. Rajagopalan (Indian Institute of Technology Madras)</i>	
Detection, Tracking, and Counting Meets Drones in Crowds: A Benchmark	7808
<i>Longyin Wen (Bytedance, Inc.), Dawei Du (Kitware, Inc.), Pengfei Zhu (tianjin university), Qinghua Hu (Tianjin University), Qilong Wang (Tianjin University), Liefeng Bo (JD Finance), and Siwei Lyu (University at Buffalo)</i>	
Objectron: A Large Scale Dataset of Object-Centric Videos in the Wild With Pose Annotations	7818
<i>Adel Ahmadyan (Google), Liangkai Zhang (Google AI), Artsiom Ablavatski (Google), Jianing Wei (Google AI), and Matthias Grundmann (Google Research)</i>	
Dynamic Domain Adaptation for Efficient Inference	7828
<i>Shuang Li (Beijing Institute of Technology), JinMing Zhang (Beijing Institute of Technology), Wenxuan Ma (Beijing Institute of Technology), Chi Harold Liu (Beijing Institute of Technology), and Wei Li (Peking University)</i>	
General Instance Distillation for Object Detection	7838
<i>Xing Dai (Megvii), Zeren Jiang (Megvii), Zhao Wu (Megvii), Yiping Bao (Megvii(Face++) Inc), Zhicheng Wang (Megvii), Si Liu (Beihang University), and Erjin Zhou (Megvii Research)</i>	
Data-Free Knowledge Distillation for Image Super-Resolution	7848
<i>Yiman Zhang (Huawei Noah's Ark Lab), Hanting Chen (Peking University), Xinghao Chen (Noah's Ark Lab, Huawei Technologies), Yiping Deng (Huawei), Chunjing Xu (Huawei Noah's Ark Lab), and Yunhe Wang (Huawei Technologies)</i>	
Improving Accuracy of Binary Neural Networks Using Unbalanced Activation Distribution	7858
<i>Hyungjun Kim (POSTECH), Jihoon Park (POSTECH), Changhun Lee (POSTECH), and Jae-Joon Kim (POSTECH)</i>	
Hijack-GAN: Unintended-Use of Pretrained, Black-Box GANs	7868
<i>Hui-Po Wang (CISPA Helmholtz Center for Information Security), Ning Yu (University of Maryland and Max Planck Institute for Informatics), and Mario Fritz (CISPA Helmholtz Center for Information Security)</i>	
Cross Modal Focal Loss for RGBD Face Anti-Spoofing	7878
<i>Anjith George (Idiap Research Institute) and Sébastien Marcel (IDIAP)</i>	
On the Difficulty of Membership Inference Attacks	7888
<i>Shahbaz Rezaei (University of California at Davis) and Xin Liu (University of California)</i>	

Lifelong Person Re-Identification via Adaptive Knowledge Accumulation	7897
<i>Nan Pu (Leiden University), Wei Chen (Leiden University), Yu Liu (Dalian University of Technology), Erwin M. Bakker (Leiden University), and Michael S. Lew (Leiden University)</i>	
Stereo Radiance Fields (SRF): Learning View Synthesis for Sparse Views of Novel Scenes	7907
<i>Julian Chibane (Max Planck Institute for Informatics, University of Wuerzburg), Aayush Bansal (Carnegie Mellon University), Verica Lazova (Max Planck Institute for Informatics), and Gerard Pons-Moll (MPII, Germany)</i>	
Regularizing Generative Adversarial Networks Under Limited Data	7917
<i>Hung-Yu Tseng (University of California, Merced), Lu Jiang (Google Research), Ce Liu (Google), Ming-Hsuan Yang (University of California at Merced), and Weilong Yang (Google Inc.)</i>	
Automatic Correction of Internal Units in Generative Neural Networks	7928
<i>Ali Tousi (Korea Advanced Institute of Science and Technology), Haedong Jeong (Ulsan National Institute of Science and Technology), Jiyeon Han (Korea Advanced Institute of Science and Technology), Hwanil Choi (Korea Advanced Institute of Science and Technology), and Jaesik Choi (KAIST)</i>	
HistoGAN: Controlling Colors of GAN-Generated and Real Images via Color Histograms	7937
<i>Mahmoud Afifi (York University), Marcus A. Brubaker (York University), and Michael S. Brown (York University)</i>	
Prior Based Human Completion	7947
<i>Zibo Zhao (ShanghaiTech University), Wen Liu (ShanghaiTech University), Yanyu Xu (Institute of High Performance Computing, A*Star), Xianing Chen (ShanghaiTech), Weixin Luo (Meituan), Lei Jin (ShanghaiTech University), Bohui Zhu (Alibaba Group), Tong Liu (Taobao), Binqiang Zhao (Alibaba), and Shenghua Gao (Shanghaitech University)</i>	
Diverse Semantic Image Synthesis via Probability Distribution Modeling	7958
<i>Zhentao Tan (University of Science and Technology of China), Menglei Chai (Snap Inc.), Dongdong Chen (Microsoft Cloud AI), Jing Liao (City University of Hong Kong), Qi Chu (University of Science and Technology of China), Bin Liu (University of Science and Technology of China), Gang Hua (Wormpex AI Research), and Nenghai Yu (University of Science and Technology of China)</i>	
Adaptive Convolutions for Structure-Aware Style Transfer	7968
<i>Prashanth Chandran (ETH, Zurich), Gaspard Zoss (ETH, Zurich), Paulo Gotardo (Disney Research; Studios), Markus Gross (ETH Zurich), and Derek Bradley (Disney Research; Studios)</i>	
PISE: Person Image Synthesis and Editing With Decoupled GAN	7978
<i>Jinsong Zhang (Tianjin University), Kun Li (Tianjin University), Yu-Kun Lai (Cardiff University), and Jingyu Yang (Tianjin University)</i>	
Semi-Supervised Synthesis of High-Resolution Editable Textures for 3D Humans	7987
<i>Bindita Chaudhuri (University of Washington), Nikolaos Sarafianos (Facebook Reality Labs), Linda Shapiro (University of Washington), and Tony Tung (Facebook Reality Labs)</i>	

CDFI: Compression-Driven Network Design for Frame Interpolation	7997
<i>Tianyu Ding (Johns Hopkins University), Luming Liang (Microsoft), Zhihui Zhu (University of Denver), and Ilya Zharkov (Microsoft)</i>	
Few-Shot Classification With Feature Map Reconstruction Networks	8008
<i>Davis Wertheimer (Cornell), Luming Tang (Cornell University), and Bharath Hariharan (Cornell University)</i>	
Augmentation Strategies for Learning With Noisy Labels	8018
<i>Kento Nishi (University of California, Santa Barbara), Yi Ding (UCSB), Alex Rich (University of California, Santa Barbara), and Tobias Höllner (UCSB)</i>	
Activate or Not: Learning Customized Activation	8028
<i>Ningning Ma (hkust), Xiangyu Zhang (Megvii Technology), Ming Liu (HKUST), and Jian Sun (Megvii Technology)</i>	
Background Splitting: Finding Rare Classes in a Sea of Background	8039
<i>Ravi Teja Mullapudi (CMU), Fait Poms (Stanford), William R. Mark (Google), Deva Ramanan (Carnegie Mellon University), and Kayvon Fatahalian (Stanford)</i>	
CLCC: Contrastive Learning for Color Constancy	8049
<i>Yi-Chen Lo (MediaTek Inc.), Chia-Che Chang (MediaTek Inc.), Hsuan-Chao Chiu (Mediatek), Yu-Hao Huang (Mediatek Inc.), Chia-Ping Chen (MediaTek Inc.), Yu-Lin Chang (MediaTek Inc.), and Kevin Jou (MediaTek Inc.)</i>	
Dynamic Region-Aware Convolution	8060
<i>Jin Chen (Fudan University), Xijun Wang (MEGVII), Zichao Guo (Megvii Inc), Xiangyu Zhang (Megvii Technology), and Jian Sun (Megvii Technology)</i>	
Learning Dynamics via Graph Neural Networks for Human Pose Estimation and Tracking	8070
<i>Yiding Yang (Stevens Institute of Technology), Zhou Ren (Wormpex AI Research), Haoxiang Li (Wormpex AI Research), Chunluan Zhou (Wormpex AI Research), Xinchao Wang (National University of Singapore), and Gang Hua (Wormpex AI Research)</i>	
Searching for Fast Model Families on Datacenter Accelerators	8081
<i>Sheng Li (Google), Mingxing Tan (Google Brain), Ruoming Pang (Google Brain), Andrew Li (Google), Liqun Cheng (Google), Quoc V. Le (Google Brain), and Norman P. Jouppi (Google)</i>	
Discrete-Continuous Action Space Policy Gradient-Based Attention for Image-Text Matching	8092
<i>Shiyang Yan (Nanjing University of Information Science & Technology), Li Yu (Nanjing University of Information Science and Technology), and Yuan Xie (East China Normal University)</i>	
Quantifying Explainers of Graph Neural Networks in Computational Pathology	8102
<i>Guillaume Jaume (EPFL/IBM), Pushpak Pati (IBM Research Zurich), Behzad Bozorgtabar (EPFL), Antonio Foncubierto (IBM Research), Anna Maria Anniciello (National Cancer Institute - IRCCS), Florinda Feroce (National Cancer Institute - IRCCS), Tilman Rau (Institute of Pathology Bern), Jean-Philippe Thiran (École Polytechnique Fédérale de Lausanne), Maria Gabrani (IBM Research Zurich), and Orcun Goksel (Uppsala University)</i>	

Forecasting Irreversible Disease via Progression Learning	8113
<i>Botong Wu (Peking University), Sijie Ren (Beijing Stars Universal Technology Co., Ltd), Jing Li (Peking University), Xinwei Sun (MSRA), Shi-Ming Li (Beijing Tongren Hospital, Capital Medical University), and Yizhou Wang (PKU)</i>	
Transformer Tracking	8122
<i>Xin Chen (Dalian University of Technology), Bin Yan (Dalian University of Technology), Jiawen Zhu (Dalian University of Technology), Dong Wang (Dalian University of Technology), Xiaoyun Yang (Remark Holdings), and Huchuan Lu (Dalian University of Technology)</i>	
Online Multiple Object Tracking With Cross-Task Synergy	8132
<i>Song Guo (University of Sydney), Jingya Wang (ShanghaiTech University), Xinchao Wang (National University of Singapore), and Dacheng Tao (The University of Sydney)</i>	
Learning 3D Shape Feature for Texture-Insensitive Person Re-Identification	8142
<i>Jiaxing Chen (School of Data and Computer Science, Sun Yat-sen University), Xinyang Jiang (Tencent), Fudong Wang (Tencent), Jun Zhang (Tencent), Feng Zheng (SUSTech), Xing Sun (Tencent), and Wei-Shi Zheng (Sun Yat-sen University, China)</i>	
Regularizing Neural Networks via Adversarial Model Perturbation	8152
<i>Yaowei Zheng (Beihang University), Richong Zhang (Beihang University, China), and Yongyi Mao (University of Ottawa)</i>	
Task-Aware Variational Adversarial Active Learning	8162
<i>Kwanyoung Kim (Korea Advanced Institute of Science and Technology), Dongwon Park (Ulsan National Institute of Science and Technology), Kwang In Kim (UNIST), and Se Young Chun (Seoul National University)</i>	
VDSM: Unsupervised Video Disentanglement With State-Space Modeling and Deep Mixtures of Experts	8172
<i>Matthew J. Vowels (University of Surrey), Necati Cihan Camgoz (University of Surrey), and Richard Bowden (University of Surrey)</i>	
Multi-Target Domain Adaptation With Collaborative Consistency Learning	8183
<i>Takashi Isobe (Tsinghua University), Xu Jia (Dalian University of Technology), Shuaijun Chen (Huawei Noah's Ark Lab), Jianzhong He (Huawei), Yongjie Shi (Peking University), Jianzhuang Liu (Huawei Noah's Ark Lab), Huchuan Lu (Dalian University of Technology), and Shengjin Wang (Tsinghua University)</i>	
Learning To Relate Depth and Semantics for Unsupervised Domain Adaptation	8193
<i>Suman Saha (ETH Zurich), Anton Obukhov (ETH Zurich), Danda Pani Paudel (ETH Zürich), Menelaos Kanakis (ETH Zurich), Yuhua Chen (ETH Zurich), Stamatios Georgoulis (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
Adversarially Adaptive Normalization for Single Domain Generalization	8204
<i>Xinjie Fan (UT Austin), Qifei Wang (Google), Junjie Ke (Google), Feng Yang (Google Research), Boqing Gong (Google), and Mingyuan Zhou (University of Texas at Austin)</i>	
Rainbow Memory: Continual Learning With a Memory of Diverse Samples	8214
<i>Jihwan Bang (Search Solutions, Inc), Heesu Kim (Naver corp.), YoungJoon Yoo (Clova AI Research, NAVER Corp.), Jung-Woo Ha (NAVER AI, NAVER Corp.), and Jonghyun Choi (GIST)</i>	

Asymmetric Metric Learning for Knowledge Transfer	8224
<i>Mateusz Budnik (Inria) and Yannis Avrithis (Inria)</i>	
Scalability vs. Utility: Do We Have To Sacrifice One for the Other in Data Importance Quantification?	8235
<i>Ruoxi Jia (Virginia Tech), Fan Wu (UIUC), Xuehui Sun (SJTU), Jiachen Xu (UC Irvine), David Dao (ETH Zurich), Bhavya Kailkhura (Lawrence Livermore National Laboratory), Ce Zhang (ETH), Bo Li (UIUC), and Dawn Song (UC Berkeley)</i>	
Self-Supervised Learning on 3D Point Clouds by Learning Discrete Generative Models	8244
<i>Benjamin Eckart (NVIDIA), Wentao Yuan (University of Washington), Chao Liu (Nvidia Research), and Jan Kautz (NVIDIA)</i>	
Multi-view Depth Estimation using Epipolar Spatio-Temporal Networks	8254
<i>Xiaoxiao Long (The University of Hong Kong), Lingjie Liu (Max Planck Institute for Informatics), Wei Li (Peking University), Christian Theobalt (MPI Informatik), and Wenping Wang (The University of Hong Kong)</i>	
Beyond Image to Depth: Improving Depth Prediction Using Echoes	8264
<i>Kranti Kumar Parida (IIT Kanpur), Siddharth Srivastava (IIT Delhi), and Gaurav Sharma (IIT Kanpur)</i>	
Deeply Shape-Guided Cascade for Instance Segmentation	8274
<i>Hao Ding (Johns Hopkins University), Siyuan Qiao (Johns Hopkins University), Alan Yuille (Johns Hopkins University), and Wei Shen (Shanghai Jiao Tong University)</i>	
Linguistic Structures As Weak Supervision for Visual Scene Graph Generation	8285
<i>Keren Ye (University of Pittsburgh) and Adriana Kovashka (University of Pittsburgh)</i>	
Semantic Segmentation With Generative Models: Semi-Supervised Learning and Strong Out-of-Domain Generalization	8296
<i>Daiqing Li (NVIDIA), Junlin Yang (Yale University), Karsten Kreis (NVIDIA), Antonio Torralba (MIT), and Sanja Fidler (University of Toronto, NVIDIA)</i>	
Self-Guided and Cross-Guided Learning for Few-Shot Segmentation	8308
<i>Bingfeng Zhang (Xi'an Jiaotong-Liverpool University), Jimin Xiao (Xi'an Jiaotong-Liverpool University), and Terry Qin (Suzhou Dinnar Automation Technology Co., Ltd)</i>	
Scene Essence	8318
<i>Jiayan Qiu (University of Sydney), Yiding Yang (Stevens Institute of Technology), Xinchao Wang (National University of Singapore), and Dacheng Tao (The University of Sydney)</i>	
Adaptive Prototype Learning and Allocation for Few-Shot Segmentation	8330
<i>Gen Li (Sungkyunkwan University), Varun Jampani (Google), Laura Sevilla-Lara (Facebook), Deqing Sun (Google), Jonghyun Kim (Sungkyunkwan University), and Joongkyu Kim (Sungkyunkwan University)</i>	

Cluster, Split, Fuse, and Update: Meta-Learning for Open Compound Domain Adaptive Semantic Segmentation	8340
<i>Rui Gong (ETH Zurich), Yuhua Chen (ETH Zurich), Danda Pani Paudel (ETH Zürich), Yawei Li (ETH Zurich), Ajad Chhatkuli (ETH Zurich), Wen Li (University of Electronic Science and Technology of China), Dengxin Dai (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
Unsupervised Part Segmentation Through Disentangling Appearance and Shape	8351
<i>Shilong Liu (Tsinghua University), Lei Zhang (Microsoft), Xiao Yang (Tsinghua University), Hang Su (Tsinghua University), and Jun Zhu (Tsinghua University)</i>	
Temporal Action Segmentation From Timestamp Supervision	8361
<i>Zhe Li (University of Bonn), Yazan Abu Farha (University of Bonn), and Jürgen Gall (University of Bonn)</i>	
RAFT-3D: Scene Flow Using Rigid-Motion Embeddings	8371
<i>Zachary Teed (Princeton University) and Jia Deng (Princeton University)</i>	
Coarse-Fine Networks for Temporal Activity Detection in Videos	8381
<i>Kumara Kahatapitiya (Stony Brook University) and Michael S. Ryoo (Stony Brook/Google)</i>	
Learning Discriminative Prototypes With Dynamic Time Warping	8391
<i>Xiaobin Chang (Simon Fraser University), Frederick Tung (Borealis AI), and Greg Mori (Simon Fraser University / Borealis AI)</i>	
Learning Dynamic Network Using a Reuse Gate Function in Semi-Supervised Video Object Segmentation	8401
<i>Hyojin Park (Seoul National University), Jayeon Yoo (Seoul National University), Seohyeong Jeong (Seoul National University), Ganesh Venkatesh (Facebook), and Nojun Kwak (Seoul National University)</i>	
Probabilistic Embeddings for Cross-Modal Retrieval	8411
<i>Sanghyuk Chun (NAVER AI LAB), Seong Joon Oh (NAVER AI Lab), Rafael Sampaio de Rezende (Naver Labs Europe), Yannis Kalantidis (NAVER Labs Europe), and Diane Larlus (Naver Labs Europe)</i>	
Towards Bridging Event Captioner and Sentence Localizer for Weakly Supervised Dense Event Captioning	8421
<i>Shaoxiang Chen (Fudan University) and Yu-Gang Jiang (Fudan University)</i>	
Positive Sample Propagation Along the Audio-Visual Event Line	8432
<i>Jinxing Zhou (Hefei University of Technology), Liang Zheng (Australian National University), Yiran Zhong (Australian National University), Shijie Hao (Hefei University of Technology), and Meng Wang (Hefei University of Technology)</i>	
Embracing Uncertainty: Decoupling and De-Bias for Robust Temporal Grounding	8441
<i>Hao Zhou (Shanghai Jiao Tong University), Chongyang Zhang (Shanghai Jiao Tong University), Yan Luo (Shanghai Jiao Tong University), Yanjun Chen (Shanghai Jiao Tong University), and Chuanping Hu (The Third Research Institute of Ministry of Public Security)</i>	

Structured Scene Memory for Vision-Language Navigation	8451
<i>Hanqing Wang (Beijing Institute of Technology), Wenguan Wang (Eidgenössische Technische Hochschule Zürich), Wei Liang (Beijing Institute of Technology), Caiming Xiong (Salesforce Research), and Jianbing Shen (Inception Institute of Artificial Intelligence)</i>	
Found a Reason for me? Weakly-supervised Grounded Visual Question Answering using Capsules....	8461
<i>Aisha Urooj (University of Central Florida), Hilde Kuehne (IBM), Kevin Duarte (University of Central Florida), Chuang Gan (MIT-IBM Watson AI Lab), Niels Lobo (University of Central Florida), and Mubarak Shah (University of Central Florida)</i>	
Hierarchical Layout-Aware Graph Convolutional Network for Unified Aesthetics Assessment	8471
<i>Dongyu She (Tsinghua University), Yu-Kun Lai (Cardiff University), Gaoxiong Yi (Tencent), and Kun Xu (Tsinghua University)</i>	
Parser-Free Virtual Try-On via Distilling Appearance Flows	8481
<i>Yuying Ge (The University of Hong Kong), Yibing Song (Tencent), Ruimao Zhang (The Chinese University of Hong Kong, Shenzhen), Chongjian Ge (The University of Hong Kong, (HKU)), Wei Liu (Tencent), and Ping Luo (The University of Hong Kong)</i>	
Self-Supervised Simultaneous Multi-Step Prediction of Road Dynamics and Cost Map	8490
<i>Elmira Amirloo (Huawei Technologies Canada Co. Ltd.), Mohsen Rohani (Huawei Noah's Ark), Ershad Banijamali (University of Waterloo), Jun Luo (Huawei Technologies Canada Co. Ltd.), and Pascal Poupart (University of Waterloo)</i>	
StyleMeUp: Towards Style-Agnostic Sketch-Based Image Retrieval	8500
<i>Aneeshan Sain (University of Surrey), Ayan Kumar Bhunia (University of Surrey), Yongxin Yang (University of Surrey), Tao Xiang (University of Surrey), and Yi-Zhe Song (University of Surrey)</i>	

Session 07

VarifocalNet: An IoU-Aware Dense Object Detector	8510
<i>Haoyang Zhang (Queensland University of Technology), Ying Wang (University of Queensland), Feras Dayoub (Queensland University of Technology), and Niko Sünderhauf (Queensland University of Technology)</i>	
Variational Relational Point Completion Network	8520
<i>Liang Pan (Nanyang Technological University), Xinyi Chen (SenseTime International Pte Ltd), Zhongang Cai (SenseTime International Pte Ltd), Junzhe Zhang (Nanyang Technological University), Haiyu Zhao (SenseTime International Pte Ltd), Shuai Yi (SenseTime Group Limited), and Ziwei Liu (Nanyang Technological University)</i>	
NeX: Real-Time View Synthesis With Neural Basis Expansion	8530
<i>Suttisak Wizadwongsa (Vidyasirimedhi Institute of Science and Technology), Pakkapon Phongthawee (Vidyasirimedhi Institute of Science and Technology), Jiraphon Yenphraphai (Vidyasirimedhi Institute of Science and Technology), and Supasorn Suwajanakorn (Vidyasirimedhi Institute of Science and Technology)</i>	

Robust Neural Routing Through Space Partitions for Camera Relocalization in Dynamic Indoor Environments	8540
<i>Siyang Dong (Shandong University), Qingnan Fan (Stanford University), He Wang (Stanford University), Ji Shi (Peking University), Li Yi (Google Research), Thomas Funkhouser (Google Research), Baoquan Chen (Peking University), and Leonidas J. Guibas (Stanford University)</i>	
Categorical Depth Distribution Network for Monocular 3D Object Detection	8551
<i>Cody Reading (University of Toronto), Ali Harakeh (University of Toronto), Julia Chae (University of Toronto), and Steven L. Waslander (University of Toronto)</i>	
Dual Attention Suppression Attack: Generate Adversarial Camouflage in Physical World	8561
<i>Jiakai Wang (Beihang University), Aishan Liu (Beihang University), Zixin Yin (Beihang University), Shunchang Liu (Beihang University), Shiyu Tang (Beihang University), and Xianglong Liu (BUAA)</i>	
PoseAug: A Differentiable Pose Augmentation Framework for 3D Human Pose Estimation	8571
<i>Kehong Gong (National University of Singapore), Jianfeng Zhang (NUS), and Jiashi Feng (NUS)</i>	
Passive Inter-Photon Imaging	8581
<i>Atul Ingle (University of Wisconsin-Madison), Trevor Seets (University of Wisconsin-Madison), Mauro Buttavava (Politecnico di Milano), Shantanu Gupta (University of Wisconsin-Madison), Alberto Tosi (Politecnico di Milano), Mohit Gupta (University of Wisconsin-Madison, USA), and Andreas Velten (University of Wisconsin - Madison)</i>	
Adaptive Consistency Prior Based Deep Network for Image Denoising	8592
<i>Chao Ren (Sichuan University), Xiaohai He (Sichuan University), Chuncheng Wang (Sichuan University), and Zhibo Zhao (Sichuan University)</i>	
Dynamic Slimmable Network	8603
<i>Changlin Li (Monash University), Guangrun Wang (University of Oxford), Bing Wang (Alibaba group), Xiaodan Liang (Sun Yat-sen University), Zhihui Li (University of New South Wales), and Xiaojun Chang (Monash University)</i>	
The Neural Tangent Link Between CNN Denoisers and Non-Local Filters	8614
<i>Julián Tachella (University of Edinburgh), Junqi Tang (University of Edinburgh), and Mike Davies (University of Edinburgh)</i>	
Learning Continuous Image Representation With Local Implicit Image Function	8624
<i>Yinbo Chen (UC San Diego), Sifei Liu (NVIDIA), and Xiaolong Wang (UCSD)</i>	
Image-to-Image Translation via Hierarchical Style Disentanglement	8635
<i>Xinyang Li (Xiamen University), Shengchuan Zhang (Xiamen University), Jie Hu (Xiamen University), Liujuan Cao (Xiamen University), Xiaopeng Hong (Xi'an Jiaotong University), Xudong Mao (Xiamen University), Feiyue Huang (Tencent), Yongjian Wu (Tencent Technology, (Shanghai) Co., Ltd), and Rongrong Ji (Xiamen University, China)</i>	
Dynamic Neural Radiance Fields for Monocular 4D Facial Avatar Reconstruction	8645
<i>Guy Gafni (Technical University of Munich), Justus Thies (Technical University of Munich), Michael Zollhöfer (Facebook Reality Labs), and Matthias Nießner (Technical University of Munich)</i>	

Adversarial Robustness Under Long-Tailed Distribution	8655
<i>Tong Wu (The Chinese University of Hong Kong), Ziwei Liu (Nanyang Technological University), Qingqiu Huang (The Chinese University of Hong Kong), Yu Wang (Tsinghua University), and Dahua Lin (The Chinese University of Hong Kong)</i>	
Representative Batch Normalization With Feature Calibration	8665
<i>Shang-Hua Gao (Nankai University), Qi Han (Nankai University), Duo Li (HKUST), Ming-Ming Cheng (Nankai University), and Pai Peng (Tencent Youtu Lab)</i>	
Learning to Track Instances without Video Annotations	8676
<i>Yang Fu (UIUC), Sifei Liu (NVIDIA), Umar Iqbal (NVIDIA Research), Shalini De Mello (NVIDIA Research), Humphrey Shi (U of Oregon; UIUC), and Jan Kautz (NVIDIA)</i>	
Reducing Domain Gap by Reducing Style Bias	8686
<i>Hyeonseob Nam (Lunit Inc.), HyunJae Lee (Lunit Inc.), Jongchan Park (Lunit), Wonjun Yoon (Lunit Inc.), and Donggeun Yoo (Lunit)</i>	
Taskology: Utilizing Task Relations at Scale	8696
<i>Yao Lu (Google Research), Sören Pirk (Google), Jan Dlabal (Google Research), Anthony Brohan (Google Research), Ankita Pasad (Toyota Technological Institute at Chicago), Zhao Chen (Waymo LLC), Vincent Casser (Waymo), Anelia Angelova (Google), and Ariel Gordon (Google Research)</i>	
MOS: Towards Scaling Out-of-Distribution Detection for Large Semantic Space	8706
<i>Rui Huang (University of Wisconsin-Madison) and Yixuan Li (University of Wisconsin-Madison)</i>	
DCT-Mask: Discrete Cosine Transform Mask Representation for Instance Segmentation	8716
<i>Xing Shen (School of Mathematical Sciences, Zhejiang University), Jirui Yang (DAMO Academy, Alibaba Group), Chunbo Wei (Damo Academy, Alibaba Group), Bing Deng (Damo Academy, Alibaba Group), Jianqiang Huang (Alibaba Group), Xian-Sheng Hua (Alibaba Group), Xiaoliang Cheng (Zhejiang University), and Kewei Liang (Zhejiang University)</i>	
Fine-Grained Angular Contrastive Learning With Coarse Labels	8726
<i>Guy Bukchin (Penta AI, Tel Aviv University), Eli Schwartz (IBM-Research), Kate Saenko (Boston University), Ori Shahar (Penta AI), Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research), Raja Giryes (Tel Aviv University), and Leonid Karlinsky (IBM-Research)</i>	
End-to-End Video Instance Segmentation With Transformers	8737
<i>Yuqing Wang (Meituan), Zhaoliang Xu (Meituan), Xinlong Wang (University of Adelaide), Chunhua Shen (University of Adelaide), Baoshan Cheng (Meituan), Hao Shen (Meituan), and Huaxia Xia (Meituan)</i>	
TAP: Text-Aware Pre-Training for Text-VQA and Text-Caption	8747
<i>Zhengyuan Yang (University of Rochester), Yijuan Lu (Microsoft), Jianfeng Wang (Microsoft), Xi Yin (Microsoft), Dinei Florencio (Microsoft Research), Lijuan Wang (Microsoft), Cha Zhang (Microsoft Research), Lei Zhang (Microsoft), and Jiebo Luo (U. Rochester)</i>	

Real-Time High-Resolution Background Matting	8758
<i>Shanchuan Lin (University of Washington), Andrey Ryabtsev (University of Washington), Soumyadip Sengupta (University of Washington), Brian L. Curless (University of Washington), Steven M. Seitz (University of Washington), and Ira Kemelmacher-Shlizerman (University of Washington)</i>	
Camouflaged Object Segmentation With Distraction Mining	8768
<i>Haiyang Mei (Dalian University of Technology), Ge-Peng Ji (Wuhan University), Ziqi Wei (Tsinghua University), Xin Yang (Dalian University of Technology), Xiaopeng Wei (Dalian University of Technology), and Deng-Ping Fan (Inception Institute of Artificial Intelligence)</i>	
Semantic Relation Reasoning for Shot-Stable Few-Shot Object Detection	8778
<i>Chenchen Zhu (Carnegie Mellon University), Fangyi Chen (Carnegie Mellon University), Uzair Ahmed (Carnegie Mellon University), Zhiqiang Shen (Carnegie Mellon University), and Marios Savvides (Carnegie Mellon University)</i>	
Beyond Bounding-Box: Convex-Hull Feature Adaptation for Oriented and Densely Packed Object Detection	8788
<i>Zonghao Guo (University of Chinese Academy of Sciences), Chang Liu (University of Chinese Academy of Sciences), Xiaosong Zhang (University of Chinese Academy of Sciences), Jianbin Jiao (University of Chinese Academy of Sciences), Xiangyang Ji (Tsinghua University), and Qixiang Ye (University of Chinese Academy of Sciences, China)</i>	
TextOCR: Towards Large-Scale End-to-End Reasoning for Arbitrary-Shaped Scene Text	8798
<i>Amanpreet Singh (Facebook), Guan Pang (Facebook), Mandy Toh (Facebook), Jing Huang (Facebook), Wojciech Galuba (Facebook), and Tal Hassner (Facebook AI)</i>	
MOST: A Multi-Oriented Scene Text Detector With Localization Refinement	8809
<i>Minghang He (Huazhong University of Science and Technology), Minghui Liao (Huazhong University of Science and Technology), Zhibo Yang (Alibaba Group), Humen Zhong (Nanjing University), Jun Tang (Alibaba Group), Wenqing Cheng (Huazhong University of Science and Technology), Cong Yao (Alibaba DAMO Academy), Yongpan Wang (Alibaba Group), and Xiang Bai (Huazhong University of Science and Technology)</i>	
Points As Queries: Weakly Semi-Supervised Object Detection by Points	8819
<i>Liangyu Chen (Megvii Technology), Tong Yang (Megvii Technology), Xiangyu Zhang (Megvii Technology), Wei Zhang (Fudan University), and Jian Sun (Megvii Technology)</i>	
Holistic 3D Scene Understanding From a Single Image With Implicit Representation	8829
<i>Cheng Zhang (UESTC), Zhaopeng Cui (Zhejiang University), Yinda Zhang (Google), Bing Zeng (University of Electronic Science and Technology of China), Marc Pollefeys (ETH Zurich / Microsoft), and Shuaicheng Liu (UESTC; Megvii)</i>	
Shelf-Supervised Mesh Prediction in the Wild	8839
<i>Yufei Ye (Carnegie Mellon University), Shubham Tulsiani (Facebook AI Research), and Abhinav Gupta (CMU/FAIR)</i>	

Mesh Saliency: An Independent Perceptual Measure or a Derivative of Image Saliency?	8849
<i>Ran Song (School of Control Science and Engineering, Shandong University), Wei Zhang (Shandong University), Yitian Zhao (Cixi Institute of Biomedical Engineering, Ningbo Institute of Industrial Technology, Chinese Academy of Sciences), Yonghuai Liu (Edge Hill University), and Paul L. Rosin (Cardiff University)</i>	
MetaSets: Meta-Learning on Point Sets for Generalizable Representations	8859
<i>Chao Huang (Tsinghua University), Zhangjie Cao (Tsinghua University), Yunbo Wang (Shanghai Jiao Tong University), Jianmin Wang (Tsinghua University, China), and Mingsheng Long (Tsinghua University)</i>	
Few-Shot 3D Point Cloud Semantic Segmentation	8869
<i>Na Zhao (NUS), Tat-Seng Chua (National university of Singapore), and Gim Hee Lee (National University of Singapore)</i>	
Point Cloud Instance Segmentation Using Probabilistic Embeddings	8879
<i>Biao Zhang (KAUST) and Peter Wonka (KAUST)</i>	
Robust Point Cloud Registration Framework Based on Deep Graph Matching	8889
<i>Kexue Fu (Digital Medical Research Center, School of Basic Medical Science, Fudan University), Shaolei Liu (Digital Medical Research Center, School of Basic Medical Science, Fudan University), Xiaoyuan Luo (Digital Medical Research Center, School of Basic Medical Science, Fudan University), and Manning Wang (Fudan University)</i>	
Nutrition5k: Towards Automatic Nutritional Understanding of Generic Food	8899
<i>Quin Thames (Google), Arjun Karpur (Google), Wade Norris (Google), Fangting Xia (Google), Liviu Panait (Google), Tobias Weyand (Google), and Jack Sim (Google LLC)</i>	
Differentiable Diffusion for Dense Depth Estimation From Multi-View Images	8908
<i>Numair Khan (Brown University), Min H. Kim (KAIST), and James Tompkin (Brown University)</i>	
LoFTR: Detector-Free Local Feature Matching With Transformers	8918
<i>Jiaming Sun (SenseTime), Zehong Shen (Zhejiang University), Yuang Wang (Zhejiang University), Hujun Bao (Zhejiang University), and Xiaowei Zhou (Zhejiang University)</i>	
DI-Fusion: Online Implicit 3D Reconstruction With Deep Priors	8928
<i>Jiahui Huang (Tsinghua University), Shi-Sheng Huang (Tsinghua University), Haoxuan Song (Tsinghua University), and Shi-Min Hu (Tsinghua University)</i>	
SMD-Nets: Stereo Mixture Density Networks	8938
<i>Fabio Tosi (University of Bologna), Yiyi Liao (MPI-IS and University of Tübingen), Carolin Schmitt (MPI-IS and University of Tübingen), and Andreas Geiger (MPI-IS and University of Tuebingen)</i>	
Deep Two-View Structure-From-Motion Revisited	8949
<i>Jianyuan Wang (Australian National University), Yiran Zhong (Australian National University), Yuchao Dai (Northwestern Polytechnical University), Stan Birchfield (NVIDIA), Kaihao Zhang (Australian National University), Nikolai Smolyanskiy (NVIDIA), and Hongdong Li (Australian National University, Australia)</i>	

Back-Tracing Representative Points for Voting-Based 3D Object Detection in Point Clouds	8959
<i>Bowen Cheng (Beihang University), Lu Sheng (Beihang University), Shaoshuai Shi (The Chinese University of Hong Kong), Ming Yang (Beihang University), and Dong Xu (University of Sydney)</i>	
GrooMeD-NMS: Grouped Mathematically Differentiable NMS for Monocular 3D Object Detection	8969
<i>Abhinav Kumar (Michigan State University), Garrick Brazil (Michigan State University), and Xiaoming Liu (Michigan State University)</i>	
Graph-Based High-Order Relation Modeling for Long-Term Action Recognition	8980
<i>Jiaming Zhou (Sun Yat-sen University), Kun-Yu Lin (Sun Yat-sen University), Haoxin Li (Sun Yat-sen University), and Wei-Shi Zheng (Sun Yat-sen University, China)</i>	
SGCN: Sparse Graph Convolution Network for Pedestrian Trajectory Prediction	8990
<i>Liushuai Shi (Xi'an Jiaotong University), Le Wang (Xi'an Jiaotong University), Chengjiang Long (JD Finance America Corporation), Sanping Zhou (Xi'an Jiaotong University), Mo Zhou (Xi'an Jiaotong University), Zhenxing Niu (Alibaba Group), and Gang Hua (Wormpex AI Research)</i>	
Reformulating HOI Detection As Adaptive Set Prediction	9000
<i>Mingfei Chen (Huazhong University of Science and Technology), Yue Liao (Beihang University), Si Liu (Beihang University), Zhiyuan Chen (Westlake University), Fei Wang (SenseTime), and Chen Qian (SenseTime)</i>	
MagDR: Mask-Guided Detection and Reconstruction for Defending Deepfakes	9010
<i>Zhikai Chen (Xi'an Jiaotong University), Lingxi Xie (Huawei Inc.), Shanmin Pang (Xi'an Jiaotong University), Yong He (Xi'an jiaotong university), and Bo Zhang (Tencent)</i>	
Improving the Transferability of Adversarial Samples With Adversarial Transformations	9020
<i>Weibin Wu (The Chinese University of Hong Kong), Yuxin Su (The Chinese University of Hong Kong), Michael R. Lyu (The Chinese University of Hong Kong), and Irwin King (The Chinese University of Hong Kong)</i>	
FCPose: Fully Convolutional Multi-Person Pose Estimation With Dynamic Instance-Aware Convolutions	9030
<i>Weian Mao (the university of adelaide), Zhi Tian (The University of Adelaide), Xinlong Wang (University of Adelaide), and Chunhua Shen (University of Adelaide)</i>	
DexYCB: A Benchmark for Capturing Hand Grasping of Objects	9040
<i>Yu-Wei Chao (NVIDIA), Wei Yang (NVIDIA), Yu Xiang (NVIDIA), Pavlo Molchanov (NVIDIA), Ankur Handa (NVIDIA), Jonathan Tremblay (NVIDIA), Yashraj S. Narang (NVIDIA), Karl Van Wyk (NVIDIA), Umar Iqbal (NVIDIA Research), Stan Birchfield (NVIDIA), Jan Kautz (NVIDIA), and Dieter Fox (NVIDIA)</i>	
Neural Body: Implicit Neural Representations With Structured Latent Codes for Novel View Synthesis of Dynamic Humans	9050
<i>Sida Peng (Zhejiang University), Yuanqing Zhang (Zhejiang University), Yinghao Xu (Chinese University of Hong Kong), Qianqian Wang (Cornell), Qing Shuai (Zhejiang University), Hujun Bao (Zhejiang University), and Xiaowei Zhou (Zhejiang University)</i>	

PCLs: Geometry-Aware Neural Reconstruction of 3D Pose With Perspective Crop Layers	9060
<i>Frank Yu (University of British Columbia), Mathieu Salzmann (EPFL), Pascal Fua (EPFL, Switzerland), and Helge Rhodin (UBC)</i>	
Affective Processes: Stochastic Modelling of Temporal Context for Emotion and Facial Expression Recognition	9070
<i>Enrique Sanchez (Samsung AI Centre), Mani Kumar Tellamekala (University of Nottingham), Michel Valstar (University of Nottingham), and Georgios Tzimiropoulos (Queen Mary University of London)</i>	
Structure-Aware Face Clustering on a Large-Scale Graph With 107 Nodes	9081
<i>Shuai Shen (Tsinghua University), Wanhua Li (Tsinghua University), Zheng Zhu (Tsinghua University), Guan Huang (Institute of Automation, Chinese Academy of Sciences), Dalong Du (XForwardAI Technology Co., Ltd), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)</i>	
Cross-View Gait Recognition With Deep Universal Linear Embeddings	9091
<i>Shaoxiong Zhang (Beihang University), Yunhong Wang (State Key Laboratory of Virtual Reality Technology and System, Beihang University, China), and Annan Li (Beijing University of Aeronautics and Astronautics)</i>	
Partial Person Re-Identification With Part-Part Correspondence Learning	9101
<i>Tianyu He (Alibaba Group), Xu Shen (Alibaba Group), Jianqiang Huang (Alibaba Group), Zhibo Chen (University of Science and Technology of China), and Xian-Sheng Hua (Damo Academy, Alibaba Group)</i>	
Mask-ToF: Learning Microlens Masks for Flying Pixel Correction in Time-of-Flight Imaging	9112
<i>Ilya Chugunov (Princeton University), Seung-Hwan Baek (Princeton University), Qiang Fu (KAUST), Wolfgang Heidrich (KAUST), and Felix Heide (Princeton / Algolux)</i>	
Effective Snapshot Compressive-Spectral Imaging via Deep Denoising and Total Variation Priors	9123
<i>Haiquan Qiu (Xi'an Jiaotong University), Yao Wang (Xi'an Jiaotong University), and Deyu Meng (Xi'an Jiaotong University)</i>	
Test-Time Fast Adaptation for Dynamic Scene Deblurring via Meta-Auxiliary Learning	9133
<i>Zhixiang Chi (Huawei Noah's Ark Laboratory), Yang Wang (University of Manitoba; Huawei Technologies Canada), Yuanhao Yu (Huawei Noah's Ark Laboratory), and Jin Tang (Huawei Noah's Ark Laboratory)</i>	
Removing Raindrops and Rain Streaks in One Go	9143
<i>Ruijie Quan (University of Technology Sydney), Xin Yu (University of Technology Sydney), Yuanzhi Liang (Baidu), and Yi Yang (UTS)</i>	
Learning Multi-Scale Photo Exposure Correction	9153
<i>Mahmoud Afifi (York University), Konstantinos G. Derpanis (Ryerson University), Björn Ommer (Heidelberg University), and Michael S. Brown (York University)</i>	
Towards Real-World Blind Face Restoration With Generative Facial Prior	9164
<i>Xintao Wang (Tencent), Yu Li (Tencent), Honglun Zhang (Applied Research Center, Tencent PCG), and Ying Shan (Tencent)</i>	

Image Restoration for Under-Display Camera	9175
<i>Yuqian Zhou (UIUC), David Ren (UC Berkeley), Neil Emerton (Microsoft Corporation), Sehoon Lim (Microsoft Corporation), and Timothy Large (Microsoft Corporation)</i>	
LAU-Net: Latitude Adaptive Upscaling Network for Omnidirectional Image Super-Resolution ...	9185
<i>Xin Deng (Beihang university), Hao Wang (beihang university), Mai Xu (BUAA), Yichen Guo (Beihang university), Yuhang Song (University of Oxford), and Li Yang (Beihang university)</i>	
Interpreting Super-Resolution Networks With Local Attribution Maps	9195
<i>Jinjin Gu (The University of Sydney) and Chao Dong (SIAT)</i>	
Deep Burst Super-Resolution	9205
<i>Goutam Bhat (ETH Zurich), Martin Danelljan (ETH Zurich), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)</i>	
Towards Rolling Shutter Correction and Deblurring in Dynamic Scenes	9215
<i>Zhihang Zhong (The University of Tokyo), Yinqiang Zheng (The University of Tokyo), and Imari Sato (National Institute of Informatics)</i>	
Towards Fast and Accurate Real-World Depth Super-Resolution: Benchmark Dataset and Baseline	9225
<i>Lingzhi He (Beijing JiaoTong University), Hongguang Zhu (Beijing Jiaotong university), Feng Li (Beijing Jiaotong University), Huihui Bai (Beijing Jiaotong University), Runmin Cong (Beijing Jiaotong University), Chunjie Zhang (Beijing Jiaotong University), Chunyu Lin (Beijing Jiaotong University), Meiqin Liu (Beijing Jiaotong University), and Yao Zhao (Beijing Jiaotong University)</i>	
Learning To Restore Hazy Video: A New Real-World Dataset and a New Method	9235
<i>Xinyi Zhang (Tencent Youtu), Hang Dong (Xi'an Jiaotong University), Jinshan Pan (Nanjing University of Science and Technology), Chao Zhu (Xi'an Jiaotong University), Ying Tai (Tencent YouTu), Chengjie Wang (Tencent), Jilin Li (Tencent), Feiyue Huang (Tencent), and Fei Wang (Xi'an Jiaotong University)</i>	
Gradient Forward-Propagation for Large-Scale Temporal Video Modelling	9245
<i>Mateusz Malinowski (DeepMind), Dimitrios Vytiniotis (DeepMind), Grzegorz Świrszcz (DeepMind), Viorica Pătrăucean (DeepMind), and João Carreira (DeepMind)</i>	
Complementary Relation Contrastive Distillation	9256
<i>Jinguo Zhu (Xi'an Jiaotong University), Shixiang Tang (The University of Sydney), Dapeng Chen (The Chinese University of Hong Kong), Shijie Yu (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Yakun Liu (SenseTime), Mingzhe Rong (Xi'an Jiaotong University), Aijun Yang (Xi'an Jiaotong University), and Xiaohua Wang (Xi'an Jiaotong University)</i>	
Network Pruning via Performance Maximization	9266
<i>Shangqian Gao (University of Pittsburgh), Feihu Huang (University of Pittsburgh), Weidong Cai (University of Sydney), and Heng Huang (University of Pittsburgh & JD Tech)</i>	

Distribution-Aware Adaptive Multi-Bit Quantization	9277
<i>Sijie Zhao (Nanjing University), Tao Yue (Nanjing University), and Xuemei Hu (Nanjing University)</i>	
The Affective Growth of Computer Vision	9287
<i>Norman Makoto Su (Indiana University Bloomington) and David J. Crandall (Indiana University)</i>	
Fair Attribute Classification Through Latent Space De-Biasing	9297
<i>Vikram V. Ramaswamy (Princeton University), Sunnie S. Y. Kim (Princeton University), and Olga Russakovsky (Princeton University)</i>	
Soteria: Provable Defense Against Privacy Leakage in Federated Learning From Representation Perspective	9307
<i>Jingwei Sun (Duke University), Ang Li (Duke University), Binghui Wang (Duke University), Huanrui Yang (Duke University), Hai Li (Duke University), and Yiran Chen (Duke University)</i>	
Deep Compositional Metric Learning	9316
<i>Wenzhao Zheng (Tsinghua University), Chengkun Wang (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)</i>	
Physically-Aware Generative Network for 3D Shape Modeling	9326
<i>Mariem Mezghanni (École Polytechnique), Malika Boulkenafed (None), André Lieutier (none), and Maks Ovsjanikov (Ecole polytechnique)</i>	
Semantic Palette: Guiding Scene Generation With Class Proportions	9338
<i>Guillaume Le Moing (Inria), Tuan-Hung Vu (Valeo.ai), Himalaya Jain (Valeo.ai), Patrick Pérez (Valeo.ai), and Matthieu Cord (Sorbonne University)</i>	
Linear Semantics in Generative Adversarial Networks	9347
<i>Jianjin Xu (Columbia University) and Changxi Zheng (Columbia University)</i>	
Region-Aware Adaptive Instance Normalization for Image Harmonization	9357
<i>Jun Ling (Shanghai Jiao Tong University), Han Xue (Shanghai Jiao Tong University), Li Song (Shanghai Jiao Tong University), Rong Xie (Shanghai Jiao Tong University), and Xiao Gu (Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University)</i>	
PD-GAN: Probabilistic Diverse GAN for Image Inpainting	9367
<i>Hongyu Liu (Huya Inc), Ziyu Wan (City University of Hong Kong), Wei Huang (Hunan University), Yibing Song (Tencent), Xintong Han (Huya Inc), and Jing Liao (City University of Hong Kong)</i>	
In the Light of Feature Distributions: Moment Matching for Neural Style Transfer	9377
<i>Nikolai Kalischek (ETH Zürich), Jan D. Wegner (ETH Zurich), and Konrad Schindler (ETH Zurich)</i>	
High-Resolution Photorealistic Image Translation in Real-Time: A Laplacian Pyramid Translation Network	9387
<i>Jie Liang (The Hong Kong Polytechnic University), Hui Zeng (The Hong Kong Polytechnic University), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)</i>	

Synthesizing Long-Term 3D Human Motion and Interaction in 3D Scenes	9396
<i>Jiashun Wang (UCSD), Huazhe Xu (UC Berkeley), Jingwei Xu (Shanghai Jiao Tong University), Sifei Liu (NVIDIA), and Xiaolong Wang (UCSD)</i>	
A Sliced Wasserstein Loss for Neural Texture Synthesis	9407
<i>Eric Heitz (Unity Technologies), Kenneth Vanhoey (Unity Technologies), Thomas Chambon (Unity Technologies), and Laurent Belcour (Unity Technologies)</i>	
Space-Time Neural Irradiance Fields for Free-Viewpoint Video	9416
<i>Wenqi Xian (Cornell University), Jia-Bin Huang (Virginia Tech), Johannes Kopf (Facebook), and Changil Kim (Facebook)</i>	
Rethinking Class Relations: Absolute-Relative Supervised and Unsupervised Few-Shot Learning	9427
<i>Hongguang Zhang (Australian National University), Piotr Koniusz (Data61/CSIRO, ANU), Songlei Jian (NUDT), Hongdong Li (Australian National University, Australia), and Philip H. S. Torr (University of Oxford)</i>	
Joint Negative and Positive Learning for Noisy Labels	9437
<i>Youngdong Kim (KAIST), Juseung Yun (KAIST), Hyounguk Shon (KAIST), and Junmo Kim (KAIST)</i>	
Out-of-Distribution Detection Using Union of 1-Dimensional Subspaces	9447
<i>Alireza Zaeemzadeh (University of Central Florida), Niccolò Bisagno (Università di Trento), Zeno Sarnigaro (University of Trento), Nicola Conci (UNITN), Nazanin Rahnavard (University of Central Florida), and Mubarak Shah (University of Central Florida)</i>	
OpenMix: Reviving Known Knowledge for Discovering Novel Visual Categories in an Open World....	9457
<i>Zhun Zhong (University of Trento), Linchao Zhu (University of Technology, Sydney), Zhiming Luo (Xiamen University), Shaozi Li (Xiamen University, China), Yi Yang (UTS), and Nicu Sebe (University of Trento)</i>	
Calibrated RGB-D Salient Object Detection	9466
<i>Wei Ji (University of Alberta), Jingjing Li (University of Alberta), Shuang Yu (Tencent), Miao Zhang (Dalian University of Technology), Yongri Piao (Dalian University of Technology), Shunyu Yao (Dalian University of Technology), Qi Bi (University of Amsterdam), Kai Ma (Tencent), Yefeng Zheng (Tencent), Huchuan Lu (Dalian University of Technology), and Li Cheng (ECE dept., University of Alberta)</i>	
Permuted AdaIN: Reducing the Bias Towards Global Statistics in Image Classification	9477
<i>Oren Nuriel (Tel Aviv University), Sagie Benaim (Tel Aviv University), and Lior Wolf (Tel Aviv University, Israel)</i>	
Binary Graph Neural Networks	9487
<i>Mehdi Bahri (Imperial College London), Gaétan Bahl (INRIA), and Stefanos Zafeiriou (Imperial College London)</i>	

Contrastive Neural Architecture Search With Neural Architecture Comparators	9497
<i>Yaofu Chen (South China University of Technology), Yong Guo (South China University of Technology), Qi Chen (University of Adelaide), Minli Li (School of Software Engineering, South China University of Technology), Wei Zeng (Peking University, China), Yaowei Wang (PengCheng Laboratory), and Mingkui Tan (South China University of Technology)</i>	
Group Whitening: Balancing Learning Efficiency and Representational Capacity	9507
<i>Lei Huang (Beihang University), Yi Zhou (Southeast University), Li Liu (the inception institute of artificial intelligence), Fan Zhu (Inception Institute of Artificial Intelligence), and Ling Shao (Inception Institute of Artificial Intelligence)</i>	
Towards Unified Surgical Skill Assessment	9517
<i>Daochang Liu (Peking University), Qiyue Li (Peking University), Tingting Jiang (Peking University), Yizhou Wang (PKU), Rulin Miao (Peking University Cancer Hospital), Fei Shan (Peking University Cancer Hospital & Institute), and Ziyu Li (Peking University Cancer Hospital)</i>	
Every Annotation Counts: Multi-Label Deep Supervision for Medical Image Segmentation	9527
<i>Simon Reiß (Karlsruhe Institute of Technology), Constantin Seibold (Karlsruhe Institute of Technology), Alexander Freytag (Carl Zeiss AG, Jena, Germany), Erik Rodner (University of Applied Sciences Berlin), and Rainer Stiefelhagen (Karlsruhe Institute of Technology)</i>	
Graph Attention Tracking	9538
<i>Dongyan Guo (Zhejiang University of Technology), Yanyan Shao (Zhejiang University of Technology), Ying Cui (Zhejiang University of Technology), Zhenhua Wang (Zhejiang University of Technology), Liyan Zhang (Nanjing University of Aeronautics and Astronautics), and Chunhua Shen (University of Adelaide)</i>	
Discriminative Appearance Modeling With Multi-Track Pooling for Real-Time Multi-Object Tracking	9548
<i>Chanho Kim (Georgia Tech), Li Fuxin (Oregon State University), Mazen Alotaibi (Oregon State University), and James M. Rehg (Georgia Institute of Technology)</i>	
Scale-Aware Automatic Augmentation for Object Detection	9558
<i>Yukang Chen (The Chinese University of Hong Kong), Yanwei Li (The Chinese University of Hong Kong), Tao Kong (Bytedance), Lu Qi (The Chinese University of Hong Kong), Ruihang Chu (The Chinese University of Hong Kong), Lei Li (ByteDance AI Lab), and Jiaya Jia (Chinese University of Hong Kong)</i>	
Confluent Vessel Trees With Accurate Bifurcations	9568
<i>Zhongwen Zhang (University of Waterloo), Dmitrii Marin (University of Waterloo), Maria Drangova (Robarts Research Institute), and Yuri Boykov (University of Waterloo)</i>	
Sequential Graph Convolutional Network for Active Learning	9578
<i>Razvan Caramalau (Imperial College), Binod Bhattarai (Imperial College London), and Tae-Kyun Kim (Imperial College London)</i>	

CausalVAE: Disentangled Representation Learning via Neural Structural Causal Models	9588
<i>Mengyue Yang (Huawei Noah's Ark Lab), Furuo Liu (Huawei Noah's Ark Lab), Zhitang Chen (Huawei Noah's Ark Lab), Xinwei Shen (HKUST), Jianye Hao (Huawei Noah's Ark Lab), and Jun Wang (UCL)</i>	
Domain-Specific Suppression for Adaptive Object Detection	9598
<i>Yu Wang (Institute of Computing Technology, Chinese Academy of Sciences), Rui Zhang (Institute of Computing Technology, Chinese Academy of Sciences), Shuo Zhang (Cambricon Tech.Ltd), Miao Li (Institute of Computing Technology, Chinese Academy of Sciences), Yangyang Xia (Cambricon), Xishan Zhang (Cambricon Tech.Ltd), and Shaoli Liu (Cambricon Tech.Ltd)</i>	
Uncertainty Reduction for Model Adaptation in Semantic Segmentation	9608
<i>Prabhu Teja S (Idiap Research Institute) and François Fleuret (University of Geneva)</i>	
Open Domain Generalization with Domain-Augmented Meta-Learning	9619
<i>Yang Shu (Tsinghua University), Zhangjie Cao (Tsinghua University), Chenyu Wang (Tsinghua University), Jianmin Wang (Tsinghua University, China), and Mingsheng Long (Tsinghua University)</i>	
Layerwise Optimization by Gradient Decomposition for Continual Learning	9629
<i>Shixiang Tang (The University of Sydney), Dapeng Chen (The Chinese University of Hong Kong), Jinguo Zhu (Xi'an Jiaotong University), Shijie Yu (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), and Wanli Ouyang (The University of Sydney)</i>	
SLADE: A Self-Training Framework for Distance Metric Learning	9639
<i>Jiali Duan (University of Southern California), Yen-Liang Lin (Amazon), Son Tran (Amazon A9), Larry S. Davis (AMAZON COM SERVICES INC), and C.-C. Jay Kuo (USC)</i>	
DualGraph: A Graph-Based Method for Reasoning About Label Noise	9649
<i>HaiYang Zhang (Beijing University of Posts and Telecommunications), XiMing Xing (Beijing University of Posts and Telecommunications), and Liang Liu (Beijing University of Posts and Telecommunications)</i>	
CutPaste: Self-Supervised Learning for Anomaly Detection and Localization	9659
<i>Chun-Liang Li (Google), Kihyuk Sohn (Google), Jinsung Yoon (Google), and Tomas Pfister (Google)</i>	
Self-Supervised Visibility Learning for Novel View Synthesis	9670
<i>Yujiao Shi (ANU), Hongdong Li (Australian National University, Australia), and Xin Yu (University of Technology Sydney)</i>	
Boosting Monocular Depth Estimation Models to High-Resolution via Content-Adaptive Multi-Resolution Merging	9680
<i>S. Mahdi H. Miangoleh (Simon Fraser University), Sebastian Dille (Simon Fraser University), Long Mai (Adobe Research), Sylvain Paris (Adobe Research), and Yağiz Aksoy (Simon Fraser University)</i>	

Seesaw Loss for Long-Tailed Instance Segmentation	9690
<i>Jiaqi Wang (The Chinese University of Hong Kong), Wenwei Zhang (NTU), Yuhang Zang (Nanyang Technological University), Yuhang Cao (CUHK), Jiangmiao Pang (Zhejiang University), Tao Gong (University of Science and Technology of China), Kai Chen (SenseTime), Ziwei Liu (Nanyang Technological University), Chen Change Loy (Nanyang Technological University), and Dahua Lin (The Chinese University of Hong Kong)</i>	
Exploiting Edge-Oriented Reasoning for 3D Point-Based Scene Graph Analysis	9700
<i>Chaoyi Zhang (University of Sydney), Jianhui Yu (University of Sydney), Yang Song (University of New South Wales), and Weidong Cai (University of Sydney)</i>	
Rethinking BiSeNet for Real-Time Semantic Segmentation	9711
<i>Mingyuan Fan (Meituan), Shenqi Lai (Meituan), Junshi Huang (Meituan), Xiaoming Wei (Meituan), Zhenhua Chai (Meituan), Junfeng Luo (Meituan), and Xiaolin Wei (Meituan)</i>	
Exploit Visual Dependency Relations for Semantic Segmentation	9721
<i>Mingyuan Liu (University of Illinois at Chicago), Dan Schonfeld (University of Illinois at Chicago), and Wei Tang (University of Illinois at Chicago)</i>	
Abstract Spatial-Temporal Reasoning via Probabilistic Abduction and Execution	9731
<i>Chi Zhang (University of California, Los Angeles), Baoxiong Jia (UCLA), Song-Chun Zhu (UCLA), and Yixin Zhu (UCLA)</i>	
Anti-Aliasing Semantic Reconstruction for Few-Shot Semantic Segmentation	9742
<i>Binghao Liu (University of Chinese Academy of Sciences), Yao Ding (University of Chinese Academy of Sciences), Jianbin Jiao (University of Chinese Academy of Sciences), Xiangyang Ji (Tsinghua University), and Qixiang Ye (University of Chinese Academy of Sciences, China)</i>	
Domain Consensus Clustering for Universal Domain Adaptation	9752
<i>Guangrui Li (University of Technology Sydney), Guoliang Kang (Carnegie Mellon University), Yi Zhu (Amazon), Yunchao Wei (UTS), and Yi Yang (UTS)</i>	
Progressive Stage-Wise Learning for Unsupervised Feature Representation Enhancement	9762
<i>Zefan Li (Shanghai Jiao Tong University), Chenxi Liu (Waymo), Alan Yuille (Johns Hopkins University), Bingbing Ni (Shanghai Jiao Tong University), Wenjun Zhang (Shanghai Jiao Tong University), and Wen Gao (PKU)</i>	
NExT-QA: Next Phase of Question-Answering to Explaining Temporal Actions	9772
<i>Junbin Xiao (National University of Singapore), Xindi Shang (National University of Singapore), Angela Yao (National University of Singapore), and Tat-Seng Chua (National Univ. of Singapore)</i>	
Spatio-temporal Contrastive Domain Adaptation for Action Recognition	9782
<i>Xiaolin Song (Tianjin University), Sicheng Zhao (University of California Berkeley), Jingyu Yang (Tianjin University), Huanjing Yue (Tianjin University), Pengfei Xu (Didi Chuxing), Runbo Hu (Didi Chuxing), and Hua Chai (Didi Chuxing)</i>	

Shot Contrastive Self-Supervised Learning for Scene Boundary Detection	9791
<i>Shixing Chen (Amazon), Xiaohan Nie (Amazon), David Fan (Amazon), Dongqing Zhang (Amazon), Vimal Bhat (Amazon), and Raffay Hamid (Amazon)</i>	
Anchor-Constrained Viterbi for Set-Supervised Action Segmentation	9801
<i>Jun Li (Oregon State University) and Sinisa Todorovic (Oregon State U)</i>	
SG-Net: Spatial Granularity Network for One-Stage Video Instance Segmentation	9811
<i>Dongfang Liu (Purdue University), Yiming Cui (University of Florida), Wenbo Tan (HangZhou DianZi University), and Yingjie Chen (Purdue University)</i>	
Thinking Fast and Slow: Efficient Text-to-Visual Retrieval With Transformers	9821
<i>Antoine Miech (DeepMind), Jean-Baptiste Alayrac (DeepMind), Ivan Laptev (INRIA Paris), Josef Sivic (Czech Technical University), and Andrew Zisserman (University of Oxford)</i>	
Open-Book Video Captioning With Retrieve-Copy-Generate Network	9832
<i>Ziqi Zhang (CASIA), Zhongang Qi (Tencent), Chunfeng Yuan (NLPR), Ying Shan (Tencent), Bing Li (National Laboratory of Pattern Recognition, (NLPR), Institute of Automation, Chinese Academy of Sciences), Ying Deng (School of Aeronautical Manufacturing Engineering, Nanchang Hangkong University), and Weiming Hu (Institute of Automation Chinese Academy of Sciences)</i>	
Causal Attention for Vision-Language Tasks	9842
<i>Xu Yang (Nanyang Technological University), Hanwang Zhang (Nanyang Technological University), Guojun Qi (Futurewei Technologies), and Jianfei Cai (Monash University)</i>	
Locate Then Segment: A Strong Pipeline for Referring Image Segmentation	9853
<i>Ya Jing (Institute of Automation, Chinese Academy of Sciences), Tao Kong (Bytedance), Wei Wang (Institute of Automation Chinese Academy of Sciences), Liang Wang (NLPR, China), Lei Li (ByteDance AI Lab), and Tieniu Tan (NLPR, China)</i>	
Pushing It Out of the Way: Interactive Visual Navigation	9863
<i>Kuo-Hao Zeng (University of Washington), Luca Weihs (Allen Institute for Artificial Intelligence), Ali Farhadi (University of Washington, Allen Institue for AI), and Roozbeh Mottaghi (Allen Institute for AI)</i>	
SUTD-TrafficQA: A Question Answering Benchmark and an Efficient Network for Video Reasoning Over Traffic Events	9873
<i>Li Xu (Singapore University of Technology and Design), He Huang (Singapore University of Technology and Design), and Jun Liu (Singapore University of Technology and Design)</i>	
User-Guided Line Art Flat Filling With Split Filling Mechanism	9884
<i>Lvmin Zhang (Soochow University / Style2Paints Research), Chengze Li (The Chinese University of Hong Kong; Style2Paints Research; Caritas Institute of Higher Education), Edgar Simo-Serra (Waseda University), Yi Ji (Soochow University), Tien-Tsin Wong (The Chinese University of Hong Kong), and Chunping Liu (School of Computer Science and Technology, Soochow University)</i>	

CT-Net: Complementary Transferring Network for Garment Transfer With Arbitrary Geometric Changes	9894
<i>Fan Yang (Nanjing University) and Guosheng Lin (Nanyang Technological University)</i>	
AdvSim: Generating Safety-Critical Scenarios for Self-Driving Vehicles	9904
<i>Jingkang Wang (University of Toronto), Ava Pun (University of Waterloo), James Tu (University of Toronto), Sivabalan Manivasagam (University of Toronto), Abbas Sadat (Uber ATG), Sergio Casas (Uber ATG / University of Toronto), Mengye Ren (University of Toronto), and Raquel Urtasun (Uber ATG)</i>	
AIFit: Automatic 3D Human-Interpretable Feedback Models for Fitness Training	9914
<i>Mihai Fieraru (IMAR), Mihai Zanfir (IMAR), Silviu Cristian Pirlea (IMAR), Vlad Olaru (IMAR), and Cristian Sminchisescu (Lund University)</i>	

Session 08

Informative and Consistent Correspondence Mining for Cross-Domain Weakly Supervised Object Detection	9924
<i>Luwei Hou (Beihang University), Yu Zhang (Beihang University), Kui Fu (Beihang University), and Jia Li (Beihang University)</i>	
Cylindrical and Asymmetrical 3D Convolution Networks for LiDAR Segmentation	9934
<i>Xinge Zhu (The Chinese University of Hong Kong), Hui Zhou (Sensetime Group Limited.), Tai Wang (The Chinese University of Hong Kong), Fangzhou Hong (Nanyang Technological University), Yuexin Ma (ShanghaiTech University), Wei Li (Peking University), Hongsheng Li (Chinese University of Hong Kong), and Dahua Lin (The Chinese University of Hong Kong)</i>	
Neural Splines: Fitting 3D Surfaces With Infinitely-Wide Neural Networks	9944
<i>Francis Williams (New York University), Matthew Trager (NYU), Joan Bruna (Courant Institute of Mathematical Sciences, NYU, USA), and Denis Zorin (New York University)</i>	
Learning Neural Representation of Camera Pose with Matrix Representation of Pose Shift via View Synthesis	9954
<i>Yaxuan Zhu (UCLA), Ruiqi Gao (UCLA), Siyuan Huang (UCLA), Song-Chun Zhu (UCLA), and Ying Nian Wu (University of California, Los Angeles)</i>	
Action Unit Memory Network for Weakly Supervised Temporal Action Localization	9964
<i>Wang Luo (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China), Wenfei Yang (University of Science and Technology of China), Jingen Liu (JD AI Research), Tao Mei (AI Research of JD.com), Feng Wu (University of Science and Technology of China), and Yongdong Zhang (University of Science and Technology of China)</i>	
Weakly-Supervised Physically Unconstrained Gaze Estimation	9975
<i>Rakshit Kothari (Imaging Science), Shalini De Mello (NVIDIA Research), Umar Iqbal (NVIDIA Research), Wonmin Byeon (NVIDIA Research), Seonwook Park (Lunit Inc.), and Jan Kautz (NVIDIA)</i>	

On Self-Contact and Human Pose	9985
<i>Lea Müller (Max Planck Institute for Intelligent Systems), Ahmed A. A. Osman (Max Planck Institute for Intelligent Systems), Siyu Tang (ETH Zurich), Chun-Hao P. Huang (Max Planck Institute for Intelligent Systems), and Michael J. Black (Max Planck Institute for Intelligent Systems)</i>	
Seeing in Extra Darkness Using a Deep-Red Flash	9995
<i>Jinhui Xiong (KAUST), Jian Wang (Snap), Wolfgang Heidrich (KAUST), and Shree Nayar (Snap)</i>	
Light Field Super-Resolution With Zero-Shot Learning	10005
<i>Zhen Cheng (University of Science and Technology of China), Zhiwei Xiong (University of Science and Technology of China), Chang Chen (University of Science and Technology of China), Dong Liu (University of Science and Technology of China), and Zheng-Jun Zha (University of Science and Technology of China)</i>	
Neural Cellular Automata Manifold	10015
<i>Alejandro Hernandez (Institut de Robotica i Informatica Industrial), Armand Vilalta (Barcelona Supercomputing Center, (BSC)), and Francesc Moreno-Noguer (IRI)</i>	
Convolutional Dynamic Alignment Networks for Interpretable Classifications	10024
<i>Moritz Böhle (Max-Planck-Institute for Informatics), Mario Fritz (CISPA Helmholtz Center for Information Security), and Bernt Schiele (MPI Informatics)</i>	
One-Shot Free-View Neural Talking-Head Synthesis for Video Conferencing	10034
<i>Ting-Chun Wang (NVIDIA), Arun Mallya (NVIDIA), and Ming-Yu Liu (NVIDIA)</i>	
Stay Positive: Non-Negative Image Synthesis for Augmented Reality	10045
<i>Katie Luo (Cornell University), Guandao Yang (Cornell University), Wenqi Xian (Cornell University), Harald Haraldsson (Cornell Tech), Bharath Hariharan (Cornell University), and Serge Belongie (Cornell University)</i>	
Playable Video Generation	10056
<i>Willi Menapace (University of Trento), Stéphane Lathuilière (Telecom-Paris), Sergey Tulyakov (Snap Inc), Aliaksandr Siarohin (University of Trento), and Elisa Ricci (University of Trento)</i>	
Uncertainty-Aware Joint Salient Object and Camouflaged Object Detection	10066
<i>Aixuan Li (Northwestern Polytechnical University), Jing Zhang (Australian National University), Yunqiu Lv (Northwestern Polytechnical University), Bowen Liu (Northwestern Polytechnical University), Tong Zhang (EPFL), and Yuchao Dai (Northwestern Polytechnical University)</i>	
Discovering Hidden Physics Behind Transport Dynamics	10077
<i>Peirong Liu (UNC Chapel Hill), Lin Tian (Department of Computer Science University of North Carolina at Chapel Hill), Yubo Zhang (The University of North Carolina at Chapel Hill), Stephen Aylward (Kitware), Yueh Lee (UNC Chapel Hill), and Marc Niethammer (UNC)</i>	

AutoFlow: Learning a Better Training Set for Optical Flow	10088
<i>Deqing Sun (Google), Daniel Vlasic (Google), Charles Herrmann (Cornell), Varun Jampani (Google), Michael Krainin (Google), Huiwen Chang (Google), Ramin Zabih (Google), William T. Freeman (Google), and Ce Liu (Google)</i>	
Unsupervised Multi-Source Domain Adaptation Without Access to Source Data	10098
<i>Sk Miraj Ahmed (University of California Riverside), Dripta S. Raychaudhuri (University of California, Riverside), Sujoy Paul (Google Research), Samet Oymak (University of California, Riverside), and Amit K. Roy-Chowdhury (University of California, Riverside)</i>	
A Second-Order Approach to Learning With Instance-Dependent Label Noise	10108
<i>Zhaowei Zhu (University of California, Santa Cruz), Tongliang Liu (The University of Sydney), and Yang Liu (UC Santa Cruz)</i>	
Post-Hoc Uncertainty Calibration for Domain Drift Scenarios	10119
<i>Christian Tomani (TUM), Sebastian Gruber (Siemens AG), Muhammed Ebrar Erdem (Siemens AG), Daniel Cremers (TU Munich), and Florian Buettner (Siemens AG)</i>	
Cross-View Regularization for Domain Adaptive Panoptic Segmentation	10128
<i>Jiaxing Huang (Nanyang Technological University), Dayan Guan (Nanyang Technological University), Aoran Xiao (Nanyang Technological University), and Shijian Lu (Nanyang Technological University)</i>	
DatasetGAN: Efficient Labeled Data Factory With Minimal Human Effort	10140
<i>Yuxuan Zhang (NVIDIA, University of Waterloo), Huan Ling (University of Toronto, NVIDIA), Jun Gao (University of Toronto, Nvidia), Kangxue Yin (NVIDIA), Jean-Francois Lafleche (Nvidia), Adela Barriuso (MIT), Antonio Torralba (MIT), and Sanja Fidler (University of Toronto, NVIDIA)</i>	
Learning To Segment Actions From Visual and Language Instructions via Differentiable Weak Sequence Alignment	10151
<i>Yuhan Shen (Northeastern University), Lu Wang (UMich), and Ehsan Elhamifar (Northeastern University)</i>	
SIPSA-Net: Shift-Invariant Pan Sharpening With Moving Object Alignment for Satellite Imagery	10161
<i>Jaehyup Lee (KAIST), Soomin Seo (KAIST), and Munchurl Kim (Korea Advanced Institute of Science and Technology)</i>	
Joint-DetNAS: Upgrade Your Detector With NAS, Pruning and Dynamic Distillation	10170
<i>Lewei Yao (Huawei Noah's Ark Lab), Renjie Pi (Huawei Noah's Ark Lab), Hang Xu (Huawei Noah's Ark Lab), Wei Zhang (Noah's Ark Lab, Huawei Technologies), Zhenguo Li (Huawei Noah's Ark Lab), and Tong Zhang (Hong Kong University of Science and Technology)</i>	
Dense Relation Distillation With Context-Aware Aggregation for Few-Shot Object Detection	10180
<i>Hanzhe Hu (Peking University), Shuai Bai (Beijing University of Posts and Telecommunications), Aoxue Li (Noah's Ark Lab), Jinshi Cui (Peking University, China), and Liwei Wang (Peking University)</i>	
OPANAS: One-Shot Path Aggregation Network Architecture Search for Object Detection	10190
<i>Tingting Liang (Peking University), Yongtao Wang (Peking University), Zhi Tang (Peking University), Guosheng Hu (AnyVision), and Haibin Ling (Stony Brook University)</i>	

Pedestrian and Ego-Vehicle Trajectory Prediction From Monocular Camera	10199
<i>Lukáš Neumann (Czech Technical University) and Andrea Vedaldi (Oxford University)</i>	
DetectoRS: Detecting Objects With Recursive Feature Pyramid and Switchable Atrous Convolution	10208
<i>Siyuan Qiao (Johns Hopkins University), Liang-Chieh Chen (Google Inc.), and Alan Yuille (Johns Hopkins University)</i>	
Weakly-Supervised Instance Segmentation via Class-Agnostic Learning With Salient Images	10220
<i>Xinggong Wang (Huazhong University of Science and Technology), Jiapei Feng (Huazhong University of Science and Technology), Bin Hu (Huazhong University of Science and Technology), Qi Ding (VIVO), Longjin Ran (VIVO), Xiaoxin Chen (VIVO), and Wenyu Liu (Huazhong University of Science and Technology)</i>	
3DCaricShop: A Dataset and a Baseline Method for Single-View 3D Caricature Face Reconstruction	10231
<i>Yuda Qiu (CUHK, SZ), Xiaojie Xu (University of Science and Technology of China), Lingteng Qiu (Shenzhen Research Institute of Big Data, the Chinese University of Hong Kong, (Shenzhen)), Yan Pan (The Chinese University of Hong Kong, Shenzhen), Yushuang Wu (CUHKSZ), Weikai Chen (Tencent America), and Xiaoguang Han (Shenzhen Research Institute of Big Data, the Chinese University of Hong Kong, (Shenzhen))</i>	
D2IM-Net: Learning Detail Disentangled Implicit Fields From Single Images	10241
<i>Manyi Li (Simon Fraser University) and Hao Zhang (Simon Fraser University)</i>	
Sign-Agnostic Implicit Learning of Surface Self-Similarities for Shape Modeling and Reconstruction From Raw Point Clouds	10251
<i>Wenbin Zhao (South China University of Technology), Jiabao Lei (South China University of Technology), Yuxin Wen (South China University of Technology), Jianguo Zhang (Southern University of Science and Technology), and Kui Jia (South China University of Technology)</i>	
Learning Progressive Point Embeddings for 3D Point Cloud Generation	10261
<i>Cheng Wen (The University of Sydney), Baosheng Yu (The University of Sydney), and Dacheng Tao (The University of Sydney)</i>	
Learning Fine-Grained Segmentation of 3D Shapes Without Part Labels	10271
<i>Xiaogang Wang (Southwest University), Xun Sun (State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University), Xinyu Cao (State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University), Kai Xu (National University of Defense Technology), and Bin Zhou (State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University)</i>	
Deformed Implicit Field: Modeling 3D Shapes With Learned Dense Correspondence	10281
<i>Yu Deng (Tsinghua University), Jiaolong Yang (Microsoft Research), and Xin Tong (Microsoft)</i>	

Recurrent Multi-View Alignment Network for Unsupervised Surface Registration	10292
<i>Wanquan Feng (University of Science and Technology of China), Juyong Zhang (University of Science and Technology of China), Hongrui Cai (University of Science and Technology of China), Haofei Xu (University of Science and Technology of China), Junhui Hou (City University of Hong Kong, Hong Kong), and Hujun Bao (Zhejiang University)</i>	
DeepLM: Large-Scale Nonlinear Least Squares on Deep Learning Frameworks Using Stochastic Domain Decomposition	10303
<i>Jingwei Huang (Huawei), Shan Huang (Huawei Technologies Co., Ltd.), and Mingwei Sun (Wuhan University)</i>	
D-NeRF: Neural Radiance Fields for Dynamic Scenes	10313
<i>Albert Pumarola (IRI), Enric Corona (IRI), Gerard Pons-Moll (MPII, Germany), and Francesc Moreno-Noguer (IRI)</i>	
AdaStereo: A Simple and Efficient Approach for Adaptive Stereo Matching	10323
<i>Xiao Song (SenseTime Group Limited), Guorun Yang (SenseTime Research), Xinge Zhu (The Chinese University of Hong Kong), Hui Zhou (SenseTime Group Limited.), Zhe Wang (SenseTime Research), and Jianping Shi (SenseTime Group Limited)</i>	
Tangent Space Backpropagation for 3D Transformation Groups	10333
<i>Zachary Teed (Princeton University) and Jia Deng (Princeton University)</i>	
Unsupervised Learning for Robust Fitting: A Reinforcement Learning Approach	10343
<i>Giang Truong (Edith Cowan University), Huu Le (Chalmers University of Technology), David Suter (Edith Cowan University), Erchuan Zhang (Edith Cowan University), and Syed Zulqarnain Gilani (Edith Cowan University)</i>	
Hybrid Rotation Averaging: A Fast and Robust Rotation Averaging Approach	10353
<i>Yu Chen (Peking University), Ji Zhao (TuSimple), and Laurent Kneip (ShanghaiTech University)</i>	
ST3D: Self-Training for Unsupervised Domain Adaptation on 3D Object Detection	10363
<i>Jihan Yang (The University of Hong Kong), Shaoshuai Shi (The Chinese University of Hong Kong), Zhe Wang (SenseTime Research), Hongsheng Li (Chinese University of Hong Kong), and Xiaojuan Qi (The University of Hong Kong)</i>	
MonoRUn: Monocular 3D Object Detection by Reconstruction and Uncertainty Propagation	10374
<i>Hansheng Chen (Tongji University), Yuyao Huang (Tongji University), Wei Tian (Tongji University), Zhong Gao (Tongji University), and Lu Xiong (Tongji University)</i>	
Semi-Supervised Action Recognition With Temporal Contrastive Learning	10384
<i>Ankit Singh (IIT Madras), Omprakash Chakraborty (Indian Institute of Technology Kharagpur), Ashutosh Varshney (IIT Kharagpur), Rameswar Panda (MIT-IBM Watson AI Lab, IBM Research), Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research), Kate Saenko (Boston University), and Abir Das (IIT Kharagpur)</i>	
TrafficSim: Learning To Simulate Realistic Multi-Agent Behaviors	10395
<i>Simon Suo (Uber ATG, University of Toronto), Sebastian Regalado (University of Waterloo), Sergio Casas (Uber ATG / University of Toronto), and Raquel Urtasun (Uber ATG)</i>	

QPIC: Query-Based Pairwise Human-Object Interaction Detection With Image-Wide Contextual Information	10405
<i>Masato Tamura (Hitachi, Ltd.), Hiroki Ohashi (Hitachi Ltd), and Tomoaki Yoshinaga (Hitachi, Ltd.)</i>	
Class-Aware Robust Adversarial Training for Object Detection	10415
<i>Pin-Chun Chen (Academia Sinica), Bo-Han Kung (Academia Sinica), and Jun-Cheng Chen (Academia Sinica)</i>	
SurFree: A Fast Surrogate-Free Black-Box Attack	10425
<i>Thibault Maho (Inria), Teddy Furon (Inria), and Erwan Le Merrer (Inria)</i>	
Lite-HRNet: A Lightweight High-Resolution Network	10435
<i>Changqian Yu (Huazhong University of Science and Technology), Bin Xiao (Microsoft), Changxin Gao (Huazhong University of Science and Technology), Lu Yuan (Microsoft), Lei Zhang (Microsoft), Nong Sang (Huazhong University of Science and Technology), and Jingdong Wang (Microsoft)</i>	
Model-Based 3D Hand Reconstruction via Self-Supervised Learning	10446
<i>Yujin Chen (Wuhan University), Zhigang Tu (Wuhan University), Di Kang (Tencent), Linchao Bao (Tencent AI Lab), Ying Zhang (Tencent), Xuefei Zhe (Tencent AI lab), Ruizhi Chen (Wuhan University), and Junsong Yuan (State University of New York at Buffalo, USA)</i>	
LEAP: Learning Articulated Occupancy of People	10456
<i>Marko Mihajlovic (ETH Zurich), Yan Zhang (ETH Zurich), Michael J. Black (Max Planck Institute for Intelligent Systems), and Siyu Tang (ETH Zurich)</i>	
Bilevel Online Adaptation for Out-of-Domain Human Mesh Reconstruction	10467
<i>Shanyan Guan (Shanghai Jiao Tong University), Jingwei Xu (Shanghai Jiao Tong University), Yunbo Wang (Shanghai Jiao Tong University), Bingbing Ni (Shanghai Jiao Tong University), and Xiaokang Yang (Shanghai Jiao Tong University of China)</i>	
Exploiting Semantic Embedding and Visual Feature for Facial Action Unit Detection	10477
<i>Huiyuan Yang (Binghamton University-SUNY), Lijun Yin (State University of New York at Binghamton), Yi Zhou (IBM), and Jiuxiang Gu (Adobe Research)</i>	
WebFace260M: A Benchmark Unveiling the Power of Million-Scale Deep Face Recognition	10487
<i>Zheng Zhu (Tsinghua University), Guan Huang (Institute of Automation, Chinese Academy of Sciences), Jiankang Deng (Imperial College London), Yun Ye (XForwardAI Technology Co., Ltd, Beijing, China), Junjie Huang (Xforwardai), Xinze Chen (XFORWARDAI), Jiagang Zhu (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Tian Yang (XForwardAI), Jiwen Lu (Tsinghua University), Dalong Du (XForwardAI Technology Co., Ltd), and Jie Zhou (Tsinghua University)</i>	
PML: Progressive Margin Loss for Long-Tailed Age Classification	10498
<i>Zongyong Deng (Ningxia University), Hao Liu (Ningxia University), Yaoxing Wang (Ningxia University), Chenyang Wang (Ningxia University), Zekuan Yu (Fudan University), and Xuehong Sun (Ningxia University)</i>	

Fine-Grained Shape-Appearance Mutual Learning for Cloth-Changing Person Re-Identification	10508
<i>Peixian Hong (Sun Yat-sen university), Tao Wu (Sun Yat-sen university), Ancong Wu (Sun Yat-sen University), Xintong Han (Huya Inc), and Wei-Shi Zheng (Sun Yat-sen University, China)</i>	
Physics-Based Iterative Projection Complex Neural Network for Phase Retrieval in Lensless Microscopy Imaging	10518
<i>Feilong Zhang (Harbin Institute of Technology), Xianming Liu (Harbin Institute of Technology), Cheng Guo (Harbin Institute of Technology), Shiyi Lin (Harbin Institute of Technology), Junjun Jiang (Harbin Institute of Technology), and Xiangyang Ji (Tsinghua University)</i>	
Asymmetric Gained Deep Image Compression With Continuous Rate Adaptation	10527
<i>Ze Cui (Huawei Technologies), Jing Wang (Huawei), Shangyin Gao (Huawei), Tiansheng Guo (Huawei Technologies), Yihui Feng (HUAWEI Technology Co., Ltd), and Bo Bai (Huawei Technologies)</i>	
Learning a Non-Blind Deblurring Network for Night Blurry Images	10537
<i>Liang Chen (East China Normal University), Jiawei Zhang (SenseTime Research), Jinshan Pan (Nanjing University of Science and Technology), Songnan Lin (Beijing Institute of Technology), Faming Fang (East China Normal University), and Jimmy S. Ren (SenseTime Research; Qing Yuan Research Institute, Shanghai Jiao Tong University)</i>	
Contrastive Learning for Compact Single Image Dehazing	10546
<i>Haiyan Wu (East China Normal University), Yanyun Qu (XMU), Shaohui Lin (Tencent Youtu Lab), Jian Zhou (Tencent Youtu Lab), Ruizhi Qiao (East China Normal University), Zhizhong Zhang (East China Normal University), Yuan Xie (East China Normal University), and Lizhuang Ma (Shanghai Jiao Tong University)</i>	
Retinex-Inspired Unrolling With Cooperative Prior Architecture Search for Low-Light Image Enhancement	10556
<i>Risheng Liu (Dalian University of Technology), Long Ma (Dalian University of Technology), Jiaao Zhang (Dalian University of Technology), Xin Fan (Dalian University of Technology), and Zhongxuan Luo (Dalian University of Technology)</i>	
Auto-Exposure Fusion for Single-Image Shadow Removal	10566
<i>Lan Fu (University of South Carolina), Changqing Zhou (Nanyang Technological University), Qing Guo (Nanyang Technological University), Felix Juefei-Xu (Alibaba Group, USA), Hongkai Yu (Cleveland State University), Wei Feng (College of Intelligence and Computing, Tianjin University, China), Yang Liu (Nanyang Technology University, Singapore), and Song Wang (University of South Carolina)</i>	
Unsupervised Degradation Representation Learning for Blind Super-Resolution	10576
<i>Longguang Wang (National University of Defense Technology), Yingqian Wang (National University of Defense Technology), Xiaoyu Dong (The University of Tokyo / RIKEN AIP), Qingyu Xu (National University of Defense Technology), Jungang Yang (National University of Defense Technology), Wei An (National University of Defense Technology), and Yulan Guo (National University of Defense Technology)</i>	

Lighting, Reflectance and Geometry Estimation From 360° Panoramic Stereo	10586
<i>Junxuan Li (The Australian National University), Hongdong Li (Australian National University, Australia), and Yasuyuki Matsushita (Osaka University)</i>	
Flow-Based Kernel Prior With Application to Blind Super-Resolution	10596
<i>Jingyun Liang (ETH Zurich), Kai Zhang (ETH, Zurich), Shuhang Gu (ETH Zurich, Switzerland), Luc Van Gool (ETH Zurich), and Radu Timofte (ETH Zurich)</i>	
KOALANet: Blind Super-Resolution Using Kernel-Oriented Adaptive Local Adjustment	10606
<i>Soo Ye Kim (KAIST), Hyeonjun Sim (KAIST), and Munchurl Kim (Korea Advanced Institute of Science and Technology)</i>	
3D Video Stabilization With Depth Estimation by CNN-Based Optimization	10616
<i>Yao-Chih Lee (Academia Sinica), Kuan-Wei Tseng (National Taiwan University), Yu-Ta Chen (National Taiwan University), Chien-Cheng Chen (National Taiwan University), Chu-Song Chen (Academia Sinica), and Yi-Ping Hung (National Taiwan University)</i>	
iMiGUE: An Identity-Free Video Dataset for Micro-Gesture Understanding and Emotion Analysis	10626
<i>Xin Liu (Tianjin University), Henglin Shi (University of Oulu), Haoyu Chen (University of Oulu), Zitong Yu (CMVS, University of Oulu), Xiaobai Li (University of Oulu), and Guoying Zhao (University of Oulu)</i>	
ACRE: Abstract Causal REasoning Beyond Covariation	10638
<i>Chi Zhang (University of California, Los Angeles), Baoxiong Jia (UCLA), Mark Edmonds (UCLA), Song-Chun Zhu (UCLA), and Yixin Zhu (UCLA)</i>	
Enhance Curvature Information by Structured Stochastic Quasi-Newton Methods	10649
<i>Minghan Yang (Peking University), Dong Xu (Peking University), Hongyu Chen (Peking University), Zaiwen Wen (Peking University), and Mengyun Chen (Huawei Technologies Co. Ltd)</i>	
Refine Myself by Teaching Myself: Feature Refinement via Self-Knowledge Distillation	10659
<i>Mingi Ji (KAIST), Seungjae Shin (KAIST), Seunghyun Hwang (Looko Inc.), Gibeom Park (KAIST), and Il-Chul Moon (KAIST)</i>	
Towards Efficient Tensor Decomposition-Based DNN Model Compression With Optimization Framework	10669
<i>Miao Yin (Rutgers University), Yang Sui (Rutgers University), Siyu Liao (Rutgers University), and Bo Yuan (rutgers university)</i>	
QPP: Real-Time Quantization Parameter Prediction for Deep Neural Networks	10679
<i>Vladimir Kryzhanovskiy (Huawei Noah's Ark Lab), Gleb Balitskiy (Skolkovo Institute of Science and Technology), Nikolay Kozyrskiy (Skolkovo Institute of Science and Technology), and Aleksandr Zuruev (Novosibirsk State University)</i>	
Explaining Classifiers Using Adversarial Perturbations on the Perceptual Ball	10688
<i>Andrew Elliott (Uni. of Glasgow/The Alan Turing Insitute), Stephen Law (University College London/ Alan Turing Institute), and Chris Russell (Amazon)</i>	

Verifiability and Predictability: Interpreting Utilities of Network Architectures for Point Cloud Processing	10698
<i>Wen Shen (Tongji University), Zhihua Wei (Tongji University), Shikun Huang (Tongji University), Binbin Zhang (Tongji University), Panyue Chen (Tongji University), Ping Zhao (Tongji university), and Quanshi Zhang (Shanghai Jiao Tong University)</i>	
Model-Contrastive Federated Learning	10708
<i>Qinbin Li (National University of Singapore), Bingsheng He (National University of Singapore), and Dawn Song (UC Berkeley)</i>	
Compatibility-Aware Heterogeneous Visual Search	10718
<i>Rahul Duggal (Georgia Tech), Hao Zhou (Amazon), Shuo Yang (Amazon), Yuanjun Xiong (Amazon), Wei Xia (Amazon), Zhuowen Tu (UC San Diego), and Stefano Soatto (AWS Amazon ML)</i>	
Plan2Scene: Converting Floorplans to 3D Scenes	10728
<i>Madhawa Vidanapathirana (Simon Fraser University), Qirui Wu (Simon Fraser University), Yasutaka Furukawa (Simon Fraser University), Angel X. Chang (Simon Fraser University), and Manolis Savva (Simon Fraser University)</i>	
Few-Shot Image Generation via Cross-Domain Correspondence	10738
<i>Utkarsh Ojha (University of California, Davis), Yijun Li (Adobe Research), Jingwan Lu (Adobe Research), Alexei A. Efros (UC Berkeley), Yong Jae Lee (University of California, Davis), Eli Shechtman (Adobe Research, US), and Richard Zhang (Adobe)</i>	
Adversarial Generation of Continuous Images	10748
<i>Ivan Skorokhodov (KAUST), Savva Ignatyev (Skolkovo Institute of Science and Technology), and Mohamed Elhoseiny (KAUST)</i>	
DeFLOCNet: Deep Image Editing via Flexible Low-Level Controls	10760
<i>Hongyu Liu (Huya Inc), Ziyu Wan (City University of Hong Kong), Wei Huang (Hunan University), Yibing Song (Tencent), Xintong Han (Huya Inc), Jing Liao (City University of Hong Kong), Bin Jiang (Hunan University), and Wei Liu (Tencent)</i>	
Generating Diverse Structure for Image Inpainting With Hierarchical VQ-VAE	10770
<i>Jialun Peng (University of Science and Technology of China), Dong Liu (University of Science and Technology of China), Songcen Xu (Noah's Ark Lab, Huawei Technologies Co., Ltd.), and Houqiang Li (University of Science and Technology of China)</i>	
Smoothing the Disentangled Latent Style Space for Unsupervised Image-to-Image Translation ..	10780
<i>Yahui Liu (University of Trento), Enver Sangineto (University of Trento), Yajing Chen (Shanghai Jiao Tong University), Linchao Bao (Tencent AI Lab), Haoxian Zhang (Tencent), Nicu Sebe (University of Trento), Bruno Lepri (FBK, Trento, Italy), Wei Wang (EPFL), and Marco De Nadai (Fondazione Bruno Kessler)</i>	
Flow Guided Transformable Bottleneck Networks for Motion Retargeting	10790
<i>Jian Ren (Snap Inc.), Menglei Chai (Snap Inc.), Oliver J. Woodford (Snap Inc), Kyle Olszewski (University of Southern California), and Sergey Tulyakov (Snap Inc)</i>	

Learning Semantic Person Image Generation by Region-Adaptive Normalization	10801
<i>Zhengyao Lv (Harbin Institute of Technology), Xiaoming Li (Harbin Institute of Technology), Xin Li (Baidu), Fu Li (Baidu), Tianwei Lin (Baidu Inc), Dongliang He (Baidu), and Wangmeng Zuo (Harbin Institute of Technology, China)</i>	
Large-Capacity Image Steganography Based on Invertible Neural Networks	10811
<i>Shao-Ping Lu (Nankai University), Rong Wang (Nankai University), Tao Zhong (Nankai University), and Paul L. Rosin (Cardiff University)</i>	
Self-Supervised Video GANs: Learning for Appearance Consistency and Motion Coherency	10821
<i>Sangeek Hyun (Sungkyunkwan University), Jihwan Kim (Sungkyunkwan University), and Jae-Pil Heo (Sungkyunkwan University)</i>	
Exploring Complementary Strengths of Invariant and Equivariant Representations for Few-Shot Learning	10831
<i>Mamshad Nayeem Rizve (University of Central Florida), Salman Khan (Australian National University, (ANU)), Fahad Shahbaz Khan (Inception Institute of Artificial Intelligence), and Mubarak Shah (University of Central Florida)</i>	
Clusformer: A Transformer Based Clustering Approach to Unsupervised Large-Scale Face and Visual Landmark Recognition	10842
<i>Xuan-Bac Nguyen (VinAI Research Vietnam), Duc Toan Bui (VinAI Research, Vietnam), Chi Nhan Duong (Concordia University), Tien D. Bui (Concordia University), and Khoa Luu (University of Arkansas)</i>	
CReST: A Class-Rebalancing Self-Training Framework for Imbalanced Semi-Supervised Learning.....	10852
<i>Chen Wei (Johns Hopkins University), Kihyuk Sohn (Google), Clayton Mellina (Google Cloud), Alan Yuille (Johns Hopkins University), and Fan Yang (Google)</i>	
Neighborhood Contrastive Learning for Novel Class Discovery	10862
<i>Zhun Zhong (University of Trento), Enrico Fini (University of Trento), Subhankar Roy (University of Trento), Zhiming Luo (Xiamen University), Elisa Ricci (University of Trento), and Nicu Sebe (University of Trento)</i>	
Predicting Human Scanpaths in Visual Question Answering	10871
<i>Xianyu Chen (University of Minnesota, Twin Cities), Ming Jiang (University of Minnesota), and Qi Zhao (University of Minnesota)</i>	
Diverse Branch Block: Building a Convolution as an Inception-Like Unit	10881
<i>Xiaohan Ding (Tsinghua University), Xiangyu Zhang (Megvii Technology), Jungong Han (Tsinghua University, China), and Guiguang Ding (Aberystwyth University)</i>	
HourNAS: Extremely Fast Neural Architecture Search Through an Hourglass Lens	10891
<i>Zhaohui Yang (Peking University), Yunhe Wang (Huawei Technologies), Xinghao Chen (Noah's Ark Lab, Huawei Technologies), Jianyuan Guo (Noah's Ark Lab, Huawei Technologies), Wei Zhang (Noah's Ark Lab, Huawei Technologies), Chao Xu (Peking University), Chunjing Xu (Huawei Noah's Ark Lab), Dacheng Tao (The University of Sydney), and Chang Xu (University of Sydney)</i>	

Neural Architecture Search With Random Labels	10902
<i>Xuanyang Zhang (Megvii Technology), Pengfei Hou (Alibaba), Xiangyu Zhang (Megvii Technology), and Jian Sun (Megvii Technology)</i>	
Metadata Normalization	10912
<i>Mandy Lu (Stanford University), Qingyu Zhao (Stanford University), Jiequan Zhang (Stanford University), Kilian M. Pohl (Stanford University), Li Fei-Fei (Stanford University), Juan Carlos Niebles (Stanford University), and Ehsan Adeli (Stanford University)</i>	
Learning-Based Image Registration With Meta-Regularization	10923
<i>Ebrahim Al Safadi (Oregon Health and Science University) and Xubo Song (Oregon Health and Science University)</i>	
Lesion-Aware Transformers for Diabetic Retinopathy Grading	10933
<i>Rui Sun (University of Science and Technology of China), Yihao Li (University of Science and Technology of China), Tianzhu Zhang (University of Science and Technology of China), Zhendong Mao (University of Science and Technology of China), Feng Wu (University of Science and Technology of China), and Yongdong Zhang (University of Science and Technology of China)</i>	
CapsuleRRT: Relationships-Aware Regression Tracking via Capsules	10943
<i>Ding Ma (Harbin Institute of Technology) and Xiangqian Wu (Harbin Institute of Technology, China)</i>	
Improving Multiple Pedestrian Tracking by Track Management and Occlusion Handling	10953
<i>Daniel Stadler (Karlsruhe Institute of Technology) and Jürgen Beyerer (Fraunhofer IOSB)</i>	
Prioritized Architecture Sampling With Monto-Carlo Tree Search	10963
<i>Xiu Su (University of Sydney), Tao Huang (SenseTime), Yanxi Li (University of Sydney), Shan You (SenseTime), Fei Wang (SenseTime), Chen Qian (SenseTime), Changshui Zhang (Tsinghua University), and Chang Xu (University of Sydney)</i>	
Stochastic Whitening Batch Normalization	10973
<i>Shengdong Zhang (LG Electronics Canada), Ehsan Nezhadarya (LG Electronics), Homa Fashandi (LG Electronics), Jiayi Liu (Kwai Inc.), Darin Graham (LG Electronics Canada), and Mohak Shah (LG Electronics)</i>	
Revisiting Superpixels for Active Learning in Semantic Segmentation With Realistic Annotation Costs	10983
<i>Lile Cai (Institute for Infocomm Research), Xun Xu (I2R, ASTAR), Jun Hao Liew (NUS), and Chuan Sheng Foo (Institute for Infocomm Research, A*STAR)</i>	
Dynamic Transfer for Multi-Source Domain Adaptation	10993
<i>Yunsheng Li (UCSD), Lu Yuan (Microsoft), Yinpeng Chen (Microsoft), Pei Wang (UC San Diego), and Nuno Vasconcelos (UC San Diego)</i>	
Multi-Source Domain Adaptation With Collaborative Learning for Semantic Segmentation	11003
<i>Jianzhong He (Peking University), Xu Jia (Dalian University of Technology), Shuaijun Chen (Huawei Noah's Ark Lab), and Jianzhuang Liu (Huawei Noah's Ark Lab)</i>	

Semi-Supervised Domain Adaptation Based on Dual-Level Domain Mixing for Semantic Segmentation	11013
<i>Shuaijun Chen (Huawei Noah's Ark Lab), Xu Jia (Dalian University of Technology), Jianzhong He (Peking University), Yongjie Shi (Peking University), and Jianzhuang Liu (Huawei Noah's Ark Lab)</i>	
What Can Style Transfer and Paintings Do for Model Robustness?	11023
<i>Hubert Lin (Cornell University), Mitchell van Zuijlen (Delft university of technology), Sylvia C. Pont (TU Delft), Maarten W.A. Wijntjes (TU Delft), and Kavita Bala (Cornell University)</i>	
IIRC: Incremental Implicitly-Refined Classification	11033
<i>Mohamed Abdelsalam (Mila, Université de Montréal), Mojtaba Faramarzi (Mila, Université de Montréal), Shagun Sodhani (Facebook AI), and Sarath Chandar (Mila)</i>	
LayoutGMN: Neural Graph Matching for Structural Layout Similarity	11043
<i>Akshay Gadi Patil (Simon Fraser University), Manyi Li (Simon Fraser University), Matthew Fisher (Adobe Research), Manolis Savva (Simon Fraser University), and Hao Zhang (Simon Fraser University)</i>	
CASTing Your Model: Learning To Localize Improves Self-Supervised Representations	11053
<i>Ramprasaath R. Selvaraju (Georgia Institute of Technology), Karan Desai (University of Michigan), Justin Johnson (University of Michigan), and Nikhil Naik (MIT)</i>	
Representation Learning via Global Temporal Alignment and Cycle-Consistency	11063
<i>Isma Hadji (Samsung AI Center - Toronto), Konstantinos G. Derpanis (SAIC Toronto), and Allan D. Jepson (Samsung Toronto AIC)</i>	
Sparse Auxiliary Networks for Unified Monocular Depth Prediction and Completion	11073
<i>Vitor Guizilini (Toyota Research Institute), Rareş Ambruş Rareş Ambruş (Toyota Research Institute), Wolfram Burgard (Toyota Research Institute), and Adrien Gaidon (Toyota Research Institute)</i>	
Single Image Depth Prediction With Wavelet Decomposition	11084
<i>Michaël Ramamonjisoa (Ecole des Ponts), Michael Firman (Niantic), Jamie Watson (Niantic), Vincent Lepetit (Ecole des Ponts ParisTech), and Daniyar Turmukhambetov (Niantic)</i>	
FAPIS: A Few-Shot Anchor-Free Part-Based Instance Segmenter	11094
<i>Khoi Nguyen (Oregon State University) and Sinisa Todorovic (Oregon State U)</i>	
Bipartite Graph Network With Adaptive Message Passing for Unbiased Scene Graph Generation	11104
<i>Rongjie Li (ShanghaiTech University), Songyang Zhang (ShanghaiTech University), Bo Wan (ShanghaiTech University), and Xuming He (ShanghaiTech University)</i>	
Semantic Image Matting	11115
<i>Yanan Sun (HKUST), Chi-Keung Tang (Hong Kong University of Science and Technology), and Yu-Wing Tai (Kuaishou Technology / HKUST)</i>	
Three Ways To Improve Semantic Segmentation With Self-Supervised Depth Estimation	11125
<i>Lukas Hoyer (ETH Zurich), Dengxin Dai (ETH Zurich), Yuhua Chen (ETH Zurich), Adrian Köring (University of Bonn), Suman Saha (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	

Robust Instance Segmentation Through Reasoning About Multi-Object Occlusion	11136
<i>Xiaoding Yuan (Tongji University), Adam Kortylewski (Johns Hopkins University), Yihong Sun (Johns Hopkins University), and Alan Yuille (Johns Hopkins University)</i>	
Learning Graphs for Knowledge Transfer With Limited Labels	11146
<i>Pallabi Ghosh (Amazon), Nirat Saini (University of Maryland), Larry S. Davis (University of Maryland), and Abhinav Shrivastava (University of Maryland)</i>	
VirTex: Learning Visual Representations From Textual Annotations	11157
<i>Karan Desai (University of Michigan) and Justin Johnson (University of Michigan)</i>	
COMPLETER: Incomplete Multi-View Clustering via Contrastive Prediction	11169
<i>Yijie Lin (Sichuan University), Yuanbiao Gou (College of Computer Science, Sichuan University), Zitao Liu (TAL Education Group), Boyun Li (College of Computer Science, Sichuan University), Jiancheng Lv (Sichuan University), and Xi Peng (College of Computer Science, Sichuan University)</i>	
Home Action Genome: Cooperative Compositional Action Understanding	11179
<i>Nishant Rai (Stanford University), Haofeng Chen (Stanford University), Jingwei Ji (Stanford University), Rishi Desai (Stanford University), Kazuki Kozuka (Panasonic Corporation), Shun Ishizaka (Panasonic Corporation), Ehsan Adeli (Stanford University), and Juan Carlos Niebles (Stanford University)</i>	
Learning Goals From Failure	11189
<i>Dave Epstein (UC Berkeley) and Carl Vondrick (Columbia University)</i>	
VideoMoCo: Contrastive Video Representation Learning With Temporally Adversarial Examples	11200
<i>Tian Pan (Tencent), Yibing Song (Tencent), Tianyu Yang (Tencent AI Lab), Wenhao Jiang (Tencent), and Wei Liu (Tencent)</i>	
Spatial Feature Calibration and Temporal Fusion for Effective One-Stage Video Instance Segmentation	11210
<i>Minghan Li (The Hong Kong Polytechnic University), Shuai Li (The Hong Kong Polytechnic University), Lida Li (The Hong Kong Polytechnic University), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)</i>	
Temporally-Weighted Hierarchical Clustering for Unsupervised Action Segmentation	11220
<i>Saqib Sarfraz (Karlsruhe Institute of Technology), Naila Murray (Naver Labs), Vivek Sharma (Harvard, MIT), Ali Diba (KU Leuven), Luc Van Gool (ETH Zurich), and Rainer Stiefelhagen (Karlsruhe Institute of Technology)</i>	
Context-Aware Biaffine Localizing Network for Temporal Sentence Grounding	11230
<i>Daizong Liu (Huazhong University of Science and Technology), Xiaoye Qu (Huawei Cloud), Jianfeng Dong (Zhejiang Gongshang University), Pan Zhou (Huazhong University of Science and Technology), Yu Cheng (Microsoft Research), Wei Wei (Huazhong University of Science and Technology), Zichuan Xu (Dalian University of Technology), and Yulai Xie (huazhong university of science and technology)</i>	

Towards Diverse Paragraph Captioning for Untrimmed Videos	11240
<i>Yuqing Song (Renmin University of China), Shizhe Chen (INRIA), and Qin Jin (Renmin University of China)</i>	
Intelligent Carpet: Inferring 3D Human Pose From Tactile Signals	11250
<i>Yiyue Luo (Massachusetts Institute of Technology), Yunzhu Li (MIT), Michael Foshey (MIT), Wan Shou (Massachusetts Institute of Technology), Pratyusha Sharma (MIT), Tomás Palacios (MIT), Antonio Torralba (MIT), and Wojciech Matusik (MIT)</i>	
Bottom-Up Shift and Reasoning for Referring Image Segmentation	11261
<i>Sibei Yang (ShanghaiTech University), Meng Xia (Sun Yat-sen University), Guanbin Li (Sun Yat-sen University), Hong-Yu Zhou (The University of Hong Kong), and Yizhou Yu (The University of Hong Kong)</i>	
Topological Planning With Transformers for Vision-and-Language Navigation	11271
<i>Kevin Chen (Stanford University), Junshen K. Chen (Stanford University), Jo Chuang (Stanford University), Marynel Vázquez (Yale University), and Silvio Savarese (Stanford University)</i>	
AGQA: A Benchmark for Compositional Spatio-Temporal Reasoning	11282
<i>Madeleine Grunde-McLaughlin (University of Pennsylvania), Ranjay Krishna (Stanford University), and Maneesh Agrawala (Stanford University)</i>	
HDR Environment Map Estimation for Real-Time Augmented Reality	11293
<i>Gowri Somanath (Apple) and Daniel Kurz (Apple)</i>	
Fashion IQ: A New Dataset Towards Retrieving Images by Natural Language Feedback	11302
<i>Hui Wu (MIT-IBM Watson AI Lab, IBM Research), Yupeng Gao (IBM Research AI), Xiaoxiao Guo (IBM Research), Ziad Al-Halah (UT Austin), Steven Rennie (Fusemachines), Kristen Grauman (Facebook AI Research & UT Austin), and Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research)</i>	
TPCN: Temporal Point Cloud Networks for Motion Forecasting	11313
<i>Maosheng Ye (HKUST), Tongyi Cao (Deeproute.ai), and Qifeng Chen (HKUST)</i>	
Generalizable Pedestrian Detection: The Elephant in the Room	11323
<i>Irtiza Hasan (Inception Institute of Artificial Intelligence), Shengcai Liao (Inception Institute of Artificial Intelligence), Jinpeng Li (Inception Institute of Artificial Intelligence), Saad Ullah Akram (Aalto University), and Ling Shao (Inception Institute of Artificial Intelligence)</i>	

Session 09

Indoor Panorama Planar 3D Reconstruction via Divide and Conquer	11333
<i>Cheng Sun (National Tsing Hua University), Chi-Wei Hsiao (National Tsing Hua University), Ning-Hsu Wang (National Tsing Hua University; MediaTek), Min Sun (NTHU), and Hwann-Tzong Chen (National Tsing Hua University)</i>	

SOE-Net: A Self-Attention and Orientation Encoding Network for Point Cloud Based Place Recognition	11343
<i>Yan Xia (Technical University of Munich), Yusheng Xu (Technical University of Munich), Shuang Li (Beijing Institute of Technology), Rui Wang (Technical University of Munich), Juan Du (Technical University of Munich), Daniel Cremers (TU Munich), and Uwe Stilla (Technical University of Munich)</i>	
Neural Geometric Level of Detail: Real-Time Rendering With Implicit 3D Shapes	11353
<i>Towaki Takikawa (University of Toronto), Joey Litalien (McGill University), Kangxue Yin (NVIDIA), Karsten Kreis (NVIDIA), Charles Loop (NVIDIA Research), Derek Nowrouzezahrai (University of Toronto), Alec Jacobson (McGill University), Morgan McGuire (NVIDIA), and Sanja Fidler (University of Toronto, NVIDIA)</i>	
SOLD2: Self-Supervised Occlusion-Aware Line Description and Detection	11363
<i>Rémi Pautrat (ETH Zurich), Juan-Ting Lin (ETH Zurich), Viktor Larsson (ETH Zurich), Martin R. Oswald (ETH Zurich), and Marc Pollefeys (ETH Zurich / Microsoft)</i>	
PGT: A Progressive Method for Training Models on Long Videos	11374
<i>Bo Pang (Shanghai Jiao Tong University), Gao Peng (Shanghai Jiao Tong University), Yizhuo Li (Shanghai Jiao Tong University), and Cewu Lu (Shanghai Jiao Tong University)</i>	
Dual Attention Guided Gaze Target Detection in the Wild	11385
<i>Yi Fang (Shanghai Jiao Tong University), Jiapeng Tang (Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University), Wang Shen (Shanghai Jiao Tong University), Wei Shen (Shanghai Jiao Tong University), Xiao Gu (Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University), Li Song (Shanghai Jiao Tong University), and Guangtao Zhai (Shanghai Jiao Tong University)</i>	
ChallenCap: Monocular 3D Capture of Challenging Human Performances Using Multi-Modal References	11395
<i>Yannan He (ShanghaiTech University), Anqi Pang (ShanghaiTech University), Xin Chen (ShanghaiTech University), Han Liang (ShanghaiTech University), Minye Wu (Shanghaitech University), Yuexin Ma (ShanghaiTech University), and Lan Xu (HKUST)</i>	
Blocks-World Cameras	11407
<i>Jongho Lee (University of Wisconsin-Madison) and Mohit Gupta (University of Wisconsin-Madison, USA)</i>	
Real-Time Sphere Sweeping Stereo From Multiview Fisheye Images	11418
<i>Andreas Meuleman (KAIST), Hyeonjoong Jang (KAIST), Daniel S. Jeon (KAIST), and Min H. Kim (KAIST)</i>	
Optimal Gradient Checkpoint Search for Arbitrary Computation Graphs	11428
<i>Jianwei Feng (Carnegie Mellon University) and Dong Huang (Carnegie Mellon University)</i>	

Black-Box Explanation of Object Detectors via Saliency Maps	11438
<i>Vitali Petsiuk (Boston University), Rajiv Jain (Adobe Research), Varun Manjunatha (Adobe Research), Vlad I. Morariu (Adobe Research), Ashutosh Mehra (Adobe Inc), Vicente Ordonez (University of Virginia), and Kate Saenko (Boston University)</i>	
GIRAFFE: Representing Scenes As Compositional Generative Neural Feature Fields	11448
<i>Michael Niemeyer (Max Planck Institute for Intelligent Systems, Tübingen and University of Tübingen) and Andreas Geiger (MPI-IS and University of Tuebingen)</i>	
CoCosNet v2: Full-Resolution Correspondence Learning for Image Translation	11460
<i>Xingran Zhou (Zhejiang University), Bo Zhang (Microsoft Research Asia), Ting Zhang (MSRA), Pan Zhang (USTC), Jianmin Bao (Microsoft Research Asia), Dong Chen (Microsoft Research Asia), Zhongfei Zhang (Binghamton University), and Fang Wen (Microsoft Research Asia)</i>	
Your “Flamingo” is My “Bird”: Fine-Grained, or Not	11471
<i>Dongliang Chang (Beijing University of Posts and Telecommunications), Kaiyue Pang (Queen Mary University of London), Yixiao Zheng (Beijing University of Posts and Telecommunications), Zhanyu Ma (Beijing University of Posts and Telecommunications), Yi-Zhe Song (University of Surrey), and Jun Guo (Beijing University of Posts and Telecommunications)</i>	
Inception Convolution With Efficient Dilation Search	11481
<i>Jie Liu (Beihang University), Chuming Li (sensenet), Feng Liang (Sensenet Research), Chen Lin (University of Oxford), Ming Sun (SenseTime Group Limited), Junjie Yan (Sensenet Group Limited), Wanli Ouyang (The University of Sydney), and Dong Xu (University of Sydney)</i>	
Geo-FARM: Geodesic Factor Regression Model for Misaligned Pre-Shape Responses in Statistical Shape Analysis	11491
<i>Chao Huang (Florida State University), Anuj Srivastava (Florida State University), and Rongjie Liu (Florida State University)</i>	
UnrealPerson: An Adaptive Pipeline Towards Costless Person Re-Identification	11501
<i>Tianyu Zhang (Beihang University), Lingxi Xie (Huawei Inc.), Longhui Wei (University of Science and Technology of China), Zijie Zhuang (Tsinghua University), Yongfei Zhang (Beihang University), Bo Li (Beihang University), and Qi Tian (Xidian University)</i>	
Transferable Semantic Augmentation for Domain Adaptation	11511
<i>Shuang Li (Beijing Institute of Technology), Mixue Xie (Beijing Institute of Technology), Kaixiong Gong (Beijing Institute of Technology), Chi Harold Liu (Beijing Institute of Technology), Yulin Wang (Tsinghua University), and Wei Li (Peking University)</i>	
Jigsaw Clustering for Unsupervised Visual Representation Learning	11521
<i>Pengguang Chen (Chinese University of Hong Kong), Shu Liu (SmartMore), and Jiaya Jia (Chinese University of Hong Kong)</i>	
SliceNet: Deep Dense Depth Estimation From a Single Indoor Panorama Using a Slice-Based Representation	11531
<i>Giovanni Pintore (CRS4), Marco Agus (HBKU), Eva Almansa (CRS4), Jens Schneider (HBKU), and Enrico Gobbetti (CRS4)</i>	

Fully Convolutional Scene Graph Generation	11541
<i>Hengyue Liu (UC Riverside), Ning Yan (Futurewei Inc.), Masood Mortazaoui (Futurewei), and Bir Bhanu (University of California Riverside)</i>	
Meta Pseudo Labels	11552
<i>Hieu Pham (Google AI), Zihang Dai (Google AI), Qizhe Xie (Google AI), and Quoc V. Le (Google AI)</i>	
ArtEmis: Affective Language for Visual Art	11564
<i>Panos Achlioptas (Stanford University), Maks Ovsjanikov (Ecole polytechnique), Kilichbek Haydarov (KAUST), Mohamed Elhoseiny (KAUST), and Leonidas J. Guibas (Stanford University)</i>	
RobustNet: Improving Domain Generalization in Urban-Scene Segmentation via Instance Selective Whitening	11575
<i>Sungha Choi (Korea University), Sanghun Jung (KAIST), Huiwon Yun (KakaoEnterprise), Joanne T. Kim (Korea University), Seungryong Kim (Korea University), and Jaegul Choo (Korea Advanced Institute of Science and Technology)</i>	
Simultaneously Localize, Segment and Rank the Camouflaged Objects	11586
<i>Yunqiu Lv (Northwestern Polytechnical University), Jing Zhang (Australian National University), Yuchao Dai (Northwestern Polytechnical University), Aixuan Li (Northwestern Polytechnical University), Bowen Liu (Northwestern Polytechnical University), Nick Barnes (ANU), and Deng-Ping Fan (Inception Institute of Artificial Intelligence)</i>	
Interpolation-Based Semi-Supervised Learning for Object Detection	11597
<i>Jisoo Jeong (Seoul National University), Vikas Verma (Aalto University), Minsung Hyun (Seoul National University), Juho Kannala (Aalto University, Finland), and Nojun Kwak (Seoul National University)</i>	
There Is More Than Meets the Eye: Self-Supervised Multi-Object Detection and Tracking With Sound by Distilling Multimodal Knowledge	11607
<i>Francisco Rivera Valverde (University of Freiburg), Juana Valeria Hurtado (University of Freiburg), and Abhinav Valada (University of Freiburg)</i>	
Variational Pedestrian Detection	11617
<i>Yuang Zhang (Shanghai Jiao Tong University), Huanyu He (Shanghai Jiaotong University), Jianguo Li (Ant Group), Yuxi Li (Tencent), John See (Heriot-Watt University Malaysia), and Weiyao Lin (Shanghai Jiao Tong university)</i>	
Generalized Focal Loss V2: Learning Reliable Localization Quality Estimation for Dense Object Detection	11627
<i>Xiang Li (Nanjing University of Science and Technology), Wenhai Wang (Nanjing University), Xiaolin Hu (Tsinghua University), Jun Li (Nanjing University of Science and Technology), Jinhui Tang (Nanjing University of Science and Technology), and Jian Yang (Nanjing University of Science and Technology)</i>	

Unveiling the Potential of Structure Preserving for Weakly Supervised Object Localization	11637
<i>Xingjia Pan (Tencent), Yingguo Gao (Peking University), Zhiwen Lin (Zhejiang University), Fan Tang (Jilin University), Weiming Dong (NLPR, Institute of Automation, Chinese Academy of Sciences), Haolei Yuan (Tencent Youtu Lab), Feiyue Huang (Tencent), and Changsheng Xu (CASIA)</i>	
Deep Active Surface Models	11647
<i>Udaranga Wickramasinghe (EPFL, Switzerland), Pascal Fua (EPFL, Switzerland), and Graham Knott (EPFL, Switzerland)</i>	
Normalized Avatar Synthesis Using StyleGAN and Perceptual Refinement	11657
<i>Huiwen Luo (Pinscreen), Koki Nagano (NVIDIA), Han-Wei Kung (Pinscreen), Qingguo Xu (Pinscreen), Zejian Wang (Pinscreen), Lingyu Wei (Pinscreen), Liwen Hu (Pinscreen), and Hao Li (Pinscreen)</i>	
Omni-Supervised Point Cloud Segmentation via Gradual Receptive Field Component Reasoning	11668
<i>Jingyu Gong (Shanghai Jiao Tong University), Jiachen Xu (Shanghai Jiao Tong University), Xin Tan (Shanghai Jiao Tong University), Haichuan Song (East China Normal University), Yanyun Qu (XMU), Yuan Xie (East China Normal University), and Lizhuang Ma (Shanghai Jiao Tong University)</i>	
PU-GCN: Point Cloud Upsampling Using Graph Convolutional Networks	11678
<i>Guocheng Qian (KAUST), Abdullellah Abualshour (KAUST), Guohao Li (King Abdullah University of Science and Technology, (KAUST)), Ali Thabet (KAUST), and Bernard Ghanem (KAUST)</i>	
CGA-Net: Category Guided Aggregation for Point Cloud Semantic Segmentation	11688
<i>Tao Lu (Nanjing University), Limin Wang (Nanjing University), and Gangshan Wu (Nanjing University)</i>	
UV-Net: Learning From Boundary Representations	11698
<i>Pradeep Kumar Jayaraman (Autodesk Research), Aditya Sanghi (Autodesk; Autodesk AI Lab), Joseph G. Lambourne (Autodesk AI Lab), Karl D.D. Willis (Autodesk Research), Thomas Davies (University of Toronto), Hooman Shayani (Autodesk AI Lab), and Nigel Morris (Autodesk)</i>	
Joint Learning of 3D Shape Retrieval and Deformation	11708
<i>Mikaela Angelina Uy (Stanford University), Vladimir G. Kim (Adobe), Minhyuk Sung (KAIST), Noam Aigerman (Adobe), Siddhartha Chaudhuri (Adobe Research), and Leonidas J. Guibas (Stanford University)</i>	
Square Root Bundle Adjustment for Large-Scale Reconstruction	11718
<i>Nikolaus Demmel (TU Munich), Christiane Sommer (Technical University of Munich), Daniel Cremers (TU Munich), and Vladyslav Usenko (TU Munich)</i>	
Pixel-Aligned Volumetric Avatars	11728
<i>Amit Raj (Georgia Institute of Technology), Michael Zollhöfer (Facebook Reality Labs), Tomas Simon (Facebook Reality Labs), Jason Saragih (Facebook), Shunsuke Saito (Facebook), James Hays (Georgia Institute of Technology, USA), and Stephen Lombardi (Facebook)</i>	

Learning To Identify Correct 2D-2D Line Correspondences on Sphere	11738
<i>Haoang Li (The Chinese University of Hong Kong), Kai Chen (The Chinese University of Hong Kong), Ji Zhao (TuSimple), Jiangliu Wang (CUHK), Pyojin Kim (Sookmyung Women's University), Zhe Liu (University of Cambridge), and Yun-Hui Liu (CUHK)</i>	
SpinNet: Learning a General Surface Descriptor for 3D Point Cloud Registration	11748
<i>Sheng Ao (Sun Yat-sen University), Qingyong Hu (University of Oxford), Bo Yang (The Hong Kong Polytechnic University), Andrew Markham (University of Oxford), and Yulan Guo (National University of Defense Technology)</i>	
Self-Supervised Collision Handling via Generative 3D Garment Models for Virtual Try-On	11758
<i>Igor Santesteban (Universidad Rey Juan Carlos), Nils Thuerey (Technical University of Munich), Miguel A. Otaduy (Universidad Rey Juan Carlos), and Dan Casas (Universidad Rey Juan Carlos)</i>	
End-to-End Rotation Averaging With Multi-Source Propagation	11769
<i>Luwei Yang (Simon Fraser University), Heng Li (Simon Fraser University), Jamal Ahmed Rahim (Simon Fraser University), Zhaopeng Cui (Zhejiang University), and Ping Tan (Simon Fraser University)</i>	
Center-Based 3D Object Detection and Tracking	11779
<i>Tianwei Yin (UT Austin), Xingyi Zhou (The University of Texas at Austin), and Philipp Krähenbühl (UT Austin)</i>	
PointAugmenting: Cross-Modal Augmentation for 3D Object Detection	11789
<i>Chunwei Wang (Shanghai Jiao Tong University), Chao Ma (Shanghai Jiao Tong University), Ming Zhu (Shanghai Jiao Tong University), and Xiaokang Yang (Shanghai Jiao Tong University of China)</i>	
Removing the Background by Adding the Background: Towards Background Robust Self-Supervised Video Representation Learning	11799
<i>Jinpeng Wang (Sun Yat-sen University), Yuting Gao (tencent), Ke Li (Tencent), Yiqi Lin (School of Data and Computer Science, Sun Yat-sen University), Andy J. Ma (Sun Yat-sen University), Hao Cheng (Tencent), Pai Peng (Tencent Youtu Lab), Feiyue Huang (Tencent), Rongrong Ji (Xiamen University, China), and Xing Sun (Tencent)</i>	
Trajectory Prediction With Latent Belief Energy-Based Model	11809
<i>Bo Pang (University of California Los Angeles), Tianyang Zhao (UCLA), Xu Xie (University of California, Los Angeles), and Ying Nian Wu (University of California, Los Angeles)</i>	
End-to-End Human Object Interaction Detection With HOI Transformer	11820
<i>Cheng Zou (MEGVII), Bohan Wang (Megvii technology), Yue Hu (Megvii), Junqi Liu (Megvii), Qian Wu (Megvii), Yu Zhao (Megvii), Boxun Li (Megvii Inc.), Chenguang Zhang (Megvii), Chi Zhang (Megvii Inc.), Yichen Wei (Megvii Research Shanghai), and Jian Sun (Megvii Technology)</i>	
Simulating Unknown Target Models for Query-Efficient Black-Box Attacks	11830
<i>Chen Ma (Tsinghua University), Li Chen (Tsinghua University), and Jun-Hai Yong (Tsinghua University)</i>	

Improving Transferability of Adversarial Patches on Face Recognition With Generative Models	11840
<i>Zihao Xiao (RealAI, Inc.), Xianfeng Gao (REALAI), Chilin Fu (antgroup), Yinpeng Dong (Tsinghua University), Wei Gao (Nanyang Technological University), Xiaolu Zhang (Ant Financial Services Group), Jun Zhou (Ant Financial), and Jun Zhu (Tsinghua University)</i>	
When Human Pose Estimation Meets Robustness: Adversarial Algorithms and Benchmarks	11850
<i>Jiahang Wang (Huazhong University of Science and Technology), Sheng Jin (SenseTime Research), Wentao Liu (Sensetime), Weizhong Liu (Huazhong University of Science and Technology), Chen Qian (SenseTime), and Ping Luo (The University of Hong Kong)</i>	
Body2Hands: Learning To Infer 3D Hands From Conversational Gesture Body Dynamics	11860
<i>Evonne Ng (UC Berkeley), Shiry Ginosar (UC Berkeley), Trevor Darrell (UC Berkeley), and Hanbyul Joo (Facebook AI Research)</i>	
SMPLicit: Topology-Aware Generative Model for Clothed People	11870
<i>Enric Corona (IRI), Albert Pumarola (IRI), Guillem Alenyà (IRI), Gerard Pons-Moll (MPII, Germany), and Francesc Moreno-Noguer (IRI)</i>	
Multi-View Multi-Person 3D Pose Estimation With Plane Sweep Stereo	11881
<i>Jiahao Lin (National University of Singapore) and Gim Hee Lee (National University of Singapore)</i>	
Progressive Semantic-Aware Style Transformation for Blind Face Restoration	11891
<i>Chaofeng Chen (The University of Hong Kong), Xiaoming Li (Harbin Institute of Technology), Lingbo Yang (Peking University), Xianhui Lin (Alibaba Group), Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China), and Kwan-Yee K. Wong (The University of Hong Kong)</i>	
Variational Prototype Learning for Deep Face Recognition	11901
<i>Jiankang Deng (Imperial College London), Jia Guo (InsightFace.ai), Jing Yang (University of Nottingham), Alexandros Lattas (Imperial College London), and Stefanos Zafeiriou (Imperial College London)</i>	
Learning Spatial-Semantic Relationship for Facial Attribute Recognition With Limited Labeled Data	11911
<i>Ying Shu (Xiamen University), Yan Yan (Xiamen University), Si Chen (Xiamen University of Technology), Jing-Hao Xue (University College London), Chunhua Shen (University of Adelaide), and Hanzi Wang (Xiamen University)</i>	
Intra-Inter Camera Similarity for Unsupervised Person Re-Identification	11921
<i>Shiyu Xuan (Peking University) and Shiliang Zhang (Peking University)</i>	
Digital Gimbal: End-to-End Deep Image Stabilization With Learnable Exposure Times	11931
<i>Omer Dahary (Technion), Matan Jacoby (Technion), and Alex M. Bronstein (Tel Aviv University, Israel)</i>	
Learning Scalable \mathbb{F} -Constrained Near-Lossless Image Compression via Joint Lossy Image and Residual Compression	11941
<i>Yuanchao Bai (Pengcheng Laboratory), Xianming Liu (Harbin Institute of Technology), Wangmeng Zuo (Harbin Institute of Technology, China), Yaowei Wang (PengCheng Laboratory), and Xiangyang Ji (Tsinghua University)</i>	

Explore Image Deblurring via Encoded Blur Kernel Space	11951
<i>Phong Tran (VinAI Research), Anh Tuan Tran (VinAI Research), Quynh Phung (Vinai Research), and Minh Hoai (Stony Brook University)</i>	
Self-Aligned Video Deraining With Transmission-Depth Consistency	11961
<i>Wending Yan (National University of Singapore), Robby T. Tan (National University of Singapore), Wenhan Yang (City University of Hong Kong), and Dengxin Dai (ETH Zurich)</i>	
Nighttime Visibility Enhancement by Increasing the Dynamic Range and Suppression of Light Effects	11972
<i>Aashish Sharma (National University of Singapore) and Robby T. Tan (Yale-NUS College)</i>	
High-Quality Stereo Image Restoration From Double Refraction	11982
<i>Hakyeon Kim (KAIST), Andreas Meuleman (KAIST), Daniel S. Jeon (KAIST), and Min H. Kim (KAIST)</i>	
Spk2ImgNet: Learning To Reconstruct Dynamic Scene From Continuous Spike Stream	11991
<i>Jing Zhao (Peking University), Ruiqin Xiong (Peking University), Hangfan Liu (University of Pennsylvania), Jian Zhang (Peking University Shenzhen Graduate School), and Tiejun Huang (Peking University)</i>	
Learning Tensor Low-Rank Prior for Hyperspectral Image Reconstruction	12001
<i>Shipeng Zhang (Xi'an Jiaotong University), Lizhi Wang (Beijing Institute of Technology), Lei Zhang (Beijing Institute of Technology), and Hua Huang (Beijing Normal University)</i>	
ClassSR: A General Framework to Accelerate Super-Resolution Networks by Data Characteristic	12011
<i>Xiangtao Kong (SIAT), Hengyuan Zhao (Baidu), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), and Chao Dong (SIAT)</i>	
Scene Text Telescope: Text-Focused Scene Image Super-Resolution	12021
<i>Jingye Chen (Fudan University), Bin Li (Fudan University), and Xiangyang Xue (Fudan University)</i>	
Real-Time Selfie Video Stabilization	12031
<i>Jiyang Yu (JD AI Research), Ravi Ramamoorthi (University of California San Diego), Keli Cheng (Qualcomm Technologies Inc.), Michel Sarkis (Qualcomm Technologies Inc.), and Ning Bi (Qualcomm)</i>	
Rethinking Text Segmentation: A Novel Dataset and a Text-Specific Refinement Approach	12040
<i>Xingqian Xu (UIUC), Zhifei Zhang (Adobe Research), Zhaowen Wang (Adobe Research), Brian Price (Adobe), Zhonghao Wang (UIUC), and Humphrey Shi (U of Oregon; UIUC)</i>	
PQA: Perceptual Question Answering	12051
<i>Yonggang Qi (Beijing University of Posts and Telecommunications), Kai Zhang (Beijing University of Posts and Telecommunications), Aneeshan Sain (University of Surrey), and Yi-Zhe Song (University of Surrey)</i>	
Communication Efficient SGD via Gradient Sampling With Bayes Prior	12060
<i>LiuYihan Song (Alibaba Group), Kang Zhao (Alibaba), Pan Pan (Alibaba Group), Yu Liu (Alibaba Group), Yingya Zhang (Alibaba Group), Yinghui Xu (Alibaba DAMO Academy), and Rong Jin (alibaba group)</i>	

Student-Teacher Learning From Clean Inputs to Noisy Inputs	12070
<i>Guanzhe Hong (Purdue University), Zhiyuan Mao (Purdue University), Xiaojun Lin (Purdue University), and Stanley H. Chan (Purdue University, USA)</i>	
Towards Extremely Compact RNNs for Video Recognition With Fully Decomposed Hierarchical Tucker Structure	12080
<i>Miao Yin (Rutgers University), Siyu Liao (Rutgers University), Xiao-Yang Liu (Columbia University), Xiaodong Wang (Columbia University), and Bo Yuan (rutgers university)</i>	
Optimal Quantization Using Scaled Codebook	12090
<i>Yerlan Idelbayev (University of California, Merced), Pavlo Molchanov (NVIDIA), Maying Shen (NVIDIA), Hongxu Yin (NVIDIA), Miguel Á. Carreira-Perpiñán (UC Merced), and Jose M. Alvarez (NVIDIA)</i>	
Causal Hidden Markov Model for Time Series Disease Forecasting	12100
<i>Jing Li (Peking University), Botong Wu (Peking University), Xinwei Sun (MSRA), and Yizhou Wang (PKU)</i>	
Fair Feature Distillation for Visual Recognition	12110
<i>Sangwon Jung (SKKU), Donggyu Lee (Sungkyunkwan University), Taeon Park (Sungkyunkwan University), and Taesup Moon (Sungkyunkwan University)</i>	
DISCO: Dynamic and Invariant Sensitive Channel Obfuscation for Deep Neural Networks	12120
<i>Abhishek Singh (MIT), Ayush Chopra (MIT), Ethan Garza (MIT), Emily Zhang (MIT), Praneeth Vepakomma (MIT), Vivek Sharma (Harvard, MIT), and Ramesh Raskar (Massachusetts Institute of Technology)</i>	
Person Re-Identification Using Heterogeneous Local Graph Attention Networks	12131
<i>Zhong Zhang (Tianjin Normal University), Haijia Zhang (Tianjin Normal University), and Shuang Liu (Tianjin Normal University)</i>	
Hierarchical Video Prediction Using Relational Layouts for Human-Object Interactions	12141
<i>Navaneeth Bodla (University of Maryland), Gaurav Shrivastava (University of Maryland), Rama Chellappa (Johns Hopkins University), and Abhinav Shrivastava (University of Maryland)</i>	
Content-Aware GAN Compression	12151
<i>Yuchen Liu (Princeton University), Zhixin Shu (Adobe Research), Yijun Li (Adobe Research), Zhe Lin (Adobe Research), Federico Perazzi (Facebook, Inc.), and Sun-Yuan Kung (Princeton University)</i>	
Efficient Conditional GAN Transfer With Knowledge Propagation Across Classes	12162
<i>Mohamad Shahbazi (ETH Zurich), Zhiwu Huang (ETH Zurich), Danda Pani Paudel (ETH Zürich), Ajad Chhatkuli (ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
Discovering Interpretable Latent Space Directions of GANs Beyond Binary Attributes	12172
<i>Huiting Yang (South China University of Technology), Liangyu Chai (South China University of Technology), Qiang Wen (South China University of Technology), Shuang Zhao (Tencent Inc.), Zixun Sun (Tencent), and Shengfeng He (South China University of Technology)</i>	

Leveraging Line-Point Consistence To Preserve Structures for Wide Parallax Image Stitching.....	12181
<i>Qi Jia (Dalian University of Technology), ZhengJun Li (dalian university of technology), Xin Fan (Dalian University of Technology), Haotian Zhao (Dalian University of Technology), Shiyu Teng (Dalian University of Technology), Xinchen Ye (Dalian University of Technology), and Longin Jan Latecki (Temple University)</i>	
Rethinking Style Transfer: From Pixels to Parameterized Brushstrokes	12191
<i>Dmytro Kotovenko (Heidelberg University), Matthias Wright (Heidelberg University), Arthur Heimbrecht (Heidelberg University), and Björn Ommer (Heidelberg University)</i>	
Scene-Aware Generative Network for Human Motion Synthesis	12201
<i>Jingbo Wang (The Chinese University of HongKong), Sijie Yan (Chinese University of Hong Kong), Bo Dai (Nanyang Technological University), and Dahua Lin (The Chinese University of Hong Kong)</i>	
Stable View Synthesis	12211
<i>Gernot Riegler (Intel Labs) and Vladlen Koltun (Intel Labs)</i>	
Understanding and Simplifying Perceptual Distances	12221
<i>Dan Amir (Hebrew University of Jerusalem) and Yair Weiss (Hebrew University)</i>	
Behavior-Driven Synthesis of Human Dynamics	12231
<i>Andreas Blattmann (Heidelberg University), Timo Milbich (Heidelberg University), Michael Dorkenwald (Heidelberg University), and Björn Ommer (Heidelberg University)</i>	
Adaptive Image Transformer for One-Shot Object Detection	12242
<i>Ding-Jie Chen (Academia Sinica), He-Yen Hsieh (Academia Sinica), and Tyng-Luh Liu (Academia Sinica)</i>	
Quality-Agnostic Image Recognition via Invertible Decoder	12252
<i>Insoo Kim (Samsung Advanced Institute of Technology), Seungju Han (Samsung Advanced Institute of Technology), Ji-won Baek (Samsung Advanced Institute of Technology), Seong-Jin Park (Samsung Advanced Institute of Technology), Jae-Joon Han (Samsung), and Jinwoo Shin (KAIST)</i>	
Self-Supervised Wasserstein Pseudo-Labeling for Semi-Supervised Image Classification	12262
<i>Fariborz Taherkhani (West Virginia University), Ali Dabouei (West Virginia university), Sobhan Soleymani (West Virginia University), Jeremy Dawson (West Virginia University), and Nasser M. Nasrabadi (West Virginia University)</i>	
Improving Unsupervised Image Clustering With Robust Learning	12273
<i>Sungwon Park (KAIST), Sungwon Han (KAIST), Sundong Kim (Institute for Basic Science), Danu Kim (KAIST), Sungkyu Park (KAIST), Seunghoon Hong (KAIST), and Meeyoung Cha (KAIST & IBS)</i>	
Group Collaborative Learning for Co-Salient Object Detection	12283
<i>Qi Fan (Hong Kong University of Science and Technology), Deng-Ping Fan (Inception Institute of Artificial Intelligence), Huazhu Fu (Inception Institute of Artificial Intelligence), Chi-Keung Tang (Hong Kong University of Science and Technology), Ling Shao (Inception Institute of Artificial Intelligence), and Yu-Wing Tai (Kuaishou Technology / HKUST)</i>	

Pre-Trained Image Processing Transformer	12294
<i>Hanting Chen (Peking University), Yunhe Wang (Huawei Technologies), Tianyu Guo (Peking University), Chang Xu (University of Sydney), Yiping Deng (Huawei), Zhenhua Liu (Peking University), Siwei Ma (Peking University, China), Chunjing Xu (Huawei Noah's Ark Lab), Chao Xu (Peking University), and Wen Gao (PKU)</i>	
DOTS: Decoupling Operation and Topology in Differentiable Architecture Search	12306
<i>Yu-Chao Gu (Nankai University), Li-Juan Wang (Nankai University), Yun Liu (ETH Zurich), Yi Yang (UTS), Yu-Huan Wu (Nankai University), Shao-Ping Lu (Nankai University), and Ming-Ming Cheng (Nankai University)</i>	
Involution: Inverting the Inherence of Convolution for Visual Recognition	12316
<i>Duo Li (HKUST), Jie Hu (Institute of Software Chinese Academy of Sciences), Changhu Wang (ByteDance.Inc), Xiangtai Li (Peking University), Qi She (Bytedance AI Lab), Lei Zhu (Beijing University of Posts and Telecommunications), Tong Zhang (Hong Kong University of Science and Technology), and Qifeng Chen (HKUST)</i>	
Cross-Iteration Batch Normalization	12326
<i>Zhuliang Yao (Tsinghua University), Yue Cao (Microsoft Research), Shuxin Zheng (Microsoft Research), Gao Huang (Tsinghua), and Stephen Lin (Microsoft Research)</i>	
Learning Calibrated Medical Image Segmentation via Multi-Rater Agreement Modeling	12336
<i>Wei Ji (University of Alberta), Shuang Yu (Tencent), Junde Wu (Harbin Institute of Technology), Kai Ma (Tencent), Cheng Bian (Tencent), Qi Bi (University of Amsterdam), Jingjing Li (University of Alberta), Hanruo Liu (Beijing Tongren Hospital, Capital Medical University; School of Information and Electronics, Beijing Institute of Technology), Li Cheng (University of Alberta), and Yefeng Zheng (Tencent)</i>	
Track To Detect and Segment: An Online Multi-Object Tracker	12347
<i>Jialian Wu (State University of New York at Buffalo), Jiale Cao (Tianjin University), Liangchen Song (University at Buffalo), Yu Wang (Horizon Robotics), Ming Yang (Horizon Robotics), and Junsong Yuan (State University of New York at Buffalo, USA)</i>	
Rotation Equivariant Siamese Networks for Tracking	12357
<i>Deepak K. Gupta (University of Amsterdam), Devanshu Arya (University of Amsterdam), and Efstratios Gavves (University of Amsterdam)</i>	
SiamMOT: Siamese Multi-Object Tracking	12367
<i>Bing Shuai (Amazon), Andrew Berneshawi (Amazon), Xinyu Li (Amazon), Davide Modolo (Amazon), and Joseph Tighe (Amazon)</i>	
On Feature Normalization and Data Augmentation	12378
<i>Boyi Li (Cornell University), Felix Wu (Cornell University), Ser-Nam Lim (Facebook AI), Serge Belongie (Cornell University), and Kilian Q. Weinberger (Cornell University)</i>	

Learning a Self-Expressive Network for Subspace Clustering	12388
<i>Shangzhi Zhang (Beijing University of Posts and Telecommunications), Chong You (University of California, Berkeley), René Vidal (Johns Hopkins University, USA), and Chun-Guang Li (Beijing University of Posts & Telecommunications)</i>	
Dual-GAN: Joint BVP and Noise Modeling for Remote Physiological Measurement	12399
<i>Hao Lu (Institute of Computing Technology, Chinese Academy of Sciences), Hu Han (Chinese Academy of Sciences), and S. Kevin Zhou (CAS)</i>	
Prototypical Pseudo Label Denoising and Target Structure Learning for Domain Adaptive Semantic Segmentation	12409
<i>Pan Zhang (USTC), Bo Zhang (Microsoft Research Asia), Ting Zhang (MSRA), Dong Chen (Microsoft Research Asia), Yong Wang (University of Science and Technology of China), and Fang Wen (Microsoft Research Asia)</i>	
RPN Prototype Alignment for Domain Adaptive Object Detector	12420
<i>Yixin Zhang (University of Science and Technology of China), Zilei Wang (University of Science and Technology of China), and Yushi Mao (University of Science and Technology of China)</i>	
PixMatch: Unsupervised Domain Adaptation via Pixelwise Consistency Training	12430
<i>Luke Melas-Kyriazi (Harvard University) and Arjun K. Manrai (Harvard Medical School)</i>	
Adversarial Invariant Learning	12441
<i>Nanyang Ye (Shanghai Jiao Tong University), Jingxuan Tang (Peking University), Huayu Deng (Shanghai Jiao Tong University), Xiao-Yun Zhou (PAII INC.), Qianxiao Li (National University of Singapore), Zhenguo Li (Huawei Noah's Ark Lab), Guang-Zhong Yang (Hamlyn Centre for Robotic Surgery), and Zhanxing Zhu (Peking University)</i>	
Few-Shot Incremental Learning With Continually Evolved Classifiers	12450
<i>Chi Zhang (Nanyang Technological University), Nan Song (Nanyang Technological University), Guosheng Lin (Nanyang Technological University), Yun Zheng (Alibaba group), Pan Pan (Alibaba Group), and Yinghui Xu (Alibaba DAMO Academy)</i>	
Unsupervised Hyperbolic Metric Learning	12460
<i>Jiexi Yan (Xidian University), Lei Luo (University of Pittsburgh), Cheng Deng (Xidian University), and Heng Huang (University of Pittsburgh & JD Tech)</i>	
Audio-Visual Instance Discrimination with Cross-Modal Agreement	12470
<i>Pedro Morgado (University of California, San Diego), Nuno Vasconcelos (UCSD, USA), and Ishan Misra (Facebook AI Research)</i>	
CoCoNets: Continuous Contrastive 3D Scene Representations	12482
<i>Shamit Lal (CMU), Mihir Prabhudesai (Carnegie Mellon University), Ishita Mediratta (BITS Pilani K.K. Birla Goa Campus), Adam W. Harley (Carnegie Mellon University), and Katerina Fragkiadaki (Carnegie Mellon University)</i>	

Bilateral Grid Learning for Stereo Matching Networks	12492
<i>Bin Xu (Orbbec), Yuhua Xu (Orbbec Research), Xiaoli Yang (Orbbec), Wei Jia (Heifei University of Technology), and Yulan Guo (National University of Defense Technology)</i>	
Radar-Camera Pixel Depth Association for Depth Completion	12502
<i>Yunfei Long (Michigan State University), Daniel Morris (Michigan State University), Xiaoming Liu (Michigan State University), Marcos Castro (Ford Motor Company), Punarjay Chakravarty (Ford Motor Company), and Praveen Narayanan (Ford Motor Company)</i>	
Panoptic Segmentation Forecasting	12512
<i>Colin Graber (UIUC), Grace Tsai (Niantic), Michael Firman (Niantic), Gabriel Brostow (University College London), and Alexander G. Schwing (UIUC)</i>	
Probabilistic Modeling of Semantic Ambiguity for Scene Graph Generation	12522
<i>Gengcong Yang (Tsinghua University), Jingyi Zhang (University of Electronic Science and Technology of China), Yong Zhang (Tencent AI Lab), Baoyuan Wu (The Chinese University of Hong Kong, Shenzhen; Shenzhen Research Institute of Big Data), and Yujiu Yang (Tsinghua University)</i>	
Learning Statistical Texture for Semantic Segmentation	12532
<i>Lanyun Zhu (Beihang University), Deyi Ji (SenseTime Group Limited), Shiping Zhu (Beihang University), Weihao Gan (SenseTime Group Limited), Wei Wu (SenseTime Group Limited), and Junjie Yan (Sensetime Group Limited)</i>	
(AF)2-S3Net: Attentive Feature Fusion With Adaptive Feature Selection for Sparse Semantic Segmentation Network	12542
<i>Ran Cheng (Huawei), Ryan Razani (Huawei Technology Canada), Ehsan Taghavi (Huawei Technologies Canada), Enxu Li (Huawei), and Bingbing Liu (Huawei Noah's Ark Lab, Canada)</i>	
Scale-Localized Abstract Reasoning	12552
<i>Yaniv Benny (Tel Aviv University), Niv Pekar (Tel Aviv University), and Lior Wolf (Tel Aviv University, Israel)</i>	
Few-Shot Open-Set Recognition by Transformation Consistency	12561
<i>Minki Jeong (KAIST), Seokeon Choi (KAIST), and Changick Kim (KAIST)</i>	
I3Net: Implicit Instance-Invariant Network for Adapting One-Stage Object Detectors	12571
<i>Chaoqi Chen (The University of Hong Kong), Zebiao Zheng (Xiamen University), Yue Huang (Xiamen University), Xinghao Ding (Xiamen University), and Yizhou Yu (The University of Hong Kong)</i>	
Unsupervised Feature Learning by Cross-Level Instance-Group Discrimination	12581
<i>Xudong Wang (UC Berkeley / ICSI), Ziwei Liu (Nanyang Technological University), and Stella X. Yu (UC Berkeley / ICSI)</i>	
Multi-Shot Temporal Event Localization: A Benchmark	12591
<i>Xiaolong Liu (Huazhong University of Science and Technology), Yao Hu (Alibaba Youku Cognitive and Intelligent Lab), Song Bai (University of Oxford), Fei Ding (Alibaba Youku Cognitive and Intelligent Lab), Xiang Bai (Huazhong University of Science and Technology), and Philip H. S. Torr (University of Oxford)</i>	

Learning the Predictability of the Future	12602
<i>Dídac Surís (Columbia University), Ruoshi Liu (Columbia University), and Carl Vondrick (Columbia University)</i>	
SSAN: Separable Self-Attention Network for Video Representation Learning	12613
<i>Xudong Guo (Tsinghua University), Xun Guo (Microsoft Research Asia), and Yan Lu (Microsoft Research Asia)</i>	
Action Shuffle Alternating Learning for Unsupervised Action Segmentation	12623
<i>Jun Li (Oregon State University) and Sinisa Todorovic (Oregon State U)</i>	
Towards Accurate Text-Based Image Captioning With Content Diversity Exploration	12632
<i>Guanghui Xu (South China University of Technology), Shuaicheng Niu (South China University of Technology), Mingkui Tan (South China University of Technology), Yucheng Luo (South China University of Technology), Qing Du (South China University of Technology), and Qi Wu (University of Adelaide)</i>	
Kaleido-BERT: Vision-Language Pre-Training on Fashion Domain	12642
<i>Mingchen Zhuge (China University of Geosciences), Dehong Gao (Alibaba Group), Deng-Ping Fan (Inception Institute of Artificial Intelligence), Linbo Jin (Alibaba Group), Ben Chen (Huazhong University of Science and Technology; Alibaba Group), Haoming Zhou (Alibaba Group), Minghui Qiu (Alibaba Group), and Ling Shao (Inception Institute of Artificial Intelligence)</i>	
Transitional Adaptation of Pretrained Models for Visual Storytelling	12653
<i>Youngjae Yu (Seoul National University Vision and Learning Lab), Jiwan Chung (Seoul National University), Heeseung Yun (Seoul National University), Jongseok Kim (Violet), and Gunhee Kim (Seoul National University)</i>	
Multi-Stage Aggregated Transformer Network for Temporal Language Localization in Videos ..	12664
<i>Mingxing Zhang (University of Electronic Science and Technology of China), Yang Yang (University of Electronic Science and Technology of China), Xinghan Chen (UESTC), Yanli Ji (UESTC), Xing Xu (University of Electronic Science and Technology of China), Jingjing Li (University of Electronic Science and Technology of China), and Heng Tao Shen (University of Electronic Science and Technology of China, (UESTC))</i>	
Connecting What To Say With Where To Look by Modeling Human Attention Traces	12674
<i>Zihang Meng (University of Wisconsin Madison), Licheng Yu (Facebook), Ning Zhang (UC Berkeley), Tamara L. Berg (UNC Chapel Hill, USA), Babak Damavandi (Facebook), Vikas Singh (University of Wisconsin Madison), and Amy Bearman (Facebook)</i>	
SOON: Scenario Oriented Object Navigation With Graph-Based Exploration	12684
<i>Fengda Zhu (Monash), Xiwen Liang (Sun Yat-sen University), Yi Zhu (University of Chinese Academy of Sciences), Qizhi Yu (Monash University), Xiaojun Chang (Zhejiang Laboratory), and Xiaodan Liang (Sun Yat-sen University)</i>	

Counterfactual VQA: A Cause-Effect Look at Language Bias	12695
<i>Yulei Niu (Nanyang Technological University), Kaihua Tang (Nanyang Technological University), Hanwang Zhang (Nanyang Technological University), Zhiwu Lu (Renmin University of China), Xian-Sheng Hua (Damo Academy, Alibaba Group), and Ji-Rong Wen (Renmin University of China)</i>	
Learning by Watching	12706
<i>Jimuyang Zhang (Boston University) and Eshed Ohn-Bar (Boston University)</i>	
Personalized Outfit Recommendation With Learnable Anchors	12717
<i>Zhi Lu (University of Electronic Science and Technology of China), Yang Hu (University of Science and Technology of China), Yan Chen (University of Science and Technology of China), and Bing Zeng (University of Electronic Science and Technology of China)</i>	
Safe Local Motion Planning With Self-Supervised Freespace Forecasting	12727
<i>Peiyun Hu (Carnegie Mellon University), Aaron Huang (Carnegie Mellon University), John Dolan (Carnegie Mellon University), David Held (CMU), and Deva Ramanan (Carnegie Mellon University)</i>	
Anomaly Detection in Video via Self-Supervised and Multi-Task Learning	12737
<i>Mariana-Iuliana Georgescu (University of Bucharest), Antonio Bărbălău (University of Bucharest), Radu Tudor Ionescu (University of Bucharest), Fahad Shahbaz Khan (MBZUAI), Marius Popescu (University of Bucharest), and Mubarak Shah (University of Central Florida)</i>	

Session 10

Learning High Fidelity Depths of Dressed Humans by Watching Social Media Dance Videos	12748
<i>Yasamin Jafarian (University of Minnesota) and Hyun Soo Park (The University of Minnesota)</i>	
PointNetLK Revisited	12758
<i>Xueqian Li (Carnegie Mellon University), Jhony Kaesemodel Pontes (Argo AI), and Simon Lucey (CMU)</i>	
BRepNet: A Topological Message Passing System for Solid Models	12768
<i>Joseph G. Lambourne (Autodesk AI Lab), Karl D.D. Willis (Autodesk Research), Pradeep Kumar Jayaraman (Autodesk Research), Aditya Sanghi (Autodesk; Autodesk AI Lab), Peter Meltzer (University College London), and Hooman Shayani (Autodesk AI Lab)</i>	
KeypointDeformer: Unsupervised 3D Keypoint Discovery for Shape Control	12778
<i>Tomas Jakab (University of Oxford), Richard Tucker (Google), Ameesh Makadia (Google Research), Jiajun Wu (Stanford University), Noah Snavely (Cornell University and Google AI), and Angjoo Kanazawa (University of California Berkeley)</i>	

Learning View-Disentangled Human Pose Representation by Contrastive Cross-View Mutual Information Maximization	12788
<i>Long Zhao (Rutgers University), Yuxiao Wang (Google Research), Jiaping Zhao (Google Inc.), Liangzhe Yuan (Google Research), Jennifer J. Sun (Caltech), Florian Schroff (Google Inc.), Hartwig Adam (Google), Xi Peng (University of Delaware), Dimitris Metaxas (Rutgers), and Ting Liu (Google Research)</i>	
i3DMM: Deep Implicit 3D Morphable Model of Human Heads	12798
<i>Tarun Yenamandra (TU Munich), Ayush Tewari (Max Planck Institute for Informatics), Florian Bernard (TUM), Hans-Peter Seidel (Max Planck Institute for Informatics), Mohamed Elgharib (Max Planck Institute for Informatics), Daniel Cremers (TU Munich), and Christian Theobalt (MPI Informatik)</i>	
Reconstructing 3D Human Pose by Watching Humans in the Mirror	12809
<i>Qi Fang (Zhejiang University), Qing Shuai (Zhejiang University), Junting Dong (Zhejiang University), Hujun Bao (Zhejiang University), and Xiaowei Zhou (Zhejiang University)</i>	
EventZoom: Learning To Denoise and Super Resolve Neuromorphic Events	12819
<i>Peiqi Duan (Peking University), Zihao W. Wang (Northwestern University), Xinyu Zhou (Peking University), Yi Ma (Peking University), and Boxin Shi (Peking University)</i>	
Spatially-Varying Outdoor Lighting Estimation From Intrinsic	12829
<i>Yongjie Zhu (Beijing University of Posts and Telecommunications), Yinda Zhang (Google), Si Li (Beijing University of Posts and Telecommunications), and Boxin Shi (Peking University)</i>	
Knowledge Evolution in Neural Networks	12838
<i>Ahmed Taha (University of Maryland), Abhinav Shrivastava (University of Maryland), and Larry S. Davis (University of Maryland)</i>	
Understanding Failures of Deep Networks via Robust Feature Extraction	12848
<i>Sahil Singla (University of Maryland), Besmira Nushi (Microsoft Research), Shital Shah (Microsoft), Ece Kamar (Microsoft Research), and Eric Horvitz (MSR)</i>	
StyleSpace Analysis: Disentangled Controls for StyleGAN Image Generation	12858
<i>Zongze Wu (Hebrew University of Jerusalem), Dani Lischinski (The Hebrew University of Jerusalem), and Eli Shechtman (Adobe Research, US)</i>	
Taming Transformers for High-Resolution Image Synthesis	12868
<i>Patrick Esser (Heidelberg University), Robin Rombach (Heidelberg University), and Björn Ommer (Heidelberg University)</i>	
Benchmarking Representation Learning for Natural World Image Collections	12879
<i>Grant Van Horn (Cornell University), Elijah Cole (Caltech), Sara Beery (Caltech), Kimberly Wilber (Cornell Tech, Google), Serge Belongie (Cornell University), and Oisin Mac Aodha (University of Edinburgh)</i>	
Scaling Local Self-Attention for Parameter Efficient Visual Backbones	12889
<i>Ashish Vaswani (Google Brain), Prajit Ramachandran (Google), Aravind Srinivas (UC Berkeley), Niki Parmar (Google), Blake Hechtman (Google), and Jonathon Shlens (Google)</i>	

IMODAL: Creating Learnable User-Defined Deformation Models	12900
<i>Leander Lacroix (Sorbonne Université), Benjamin Charlier (University of Montpellier), Alain Trouvé (Ecole Normale Supérieure de Cachan), and Barbara Gris (Sorbonne Université)</i>	
Unsupervised Multi-Source Domain Adaptation for Person Re-Identification	12909
<i>Zechen Bai (University of Chinese Academy of Sciences), Zhigang Wang (Baidu Inc.), Jian Wang (Baidu), Di Hu (Renmin University of China), and Errui Ding (Baidu Inc.)</i>	
Generalization on Unseen Domains via Inference-Time Label-Preserving Target Projections	12919
<i>Prashant Pandey (Indian Institute of Technology Delhi), Mrigank Raman (Indian Institute of Technology Delhi), Sumanth Varambally (Indian Institute of Technology Delhi), and Prathosh AP (Indian Institute of Technology Delhi)</i>	
Robust Audio-Visual Instance Discrimination	12929
<i>Pedro Morgado (University of California, San Diego), Ishan Misra (Facebook AI Research), and Nuno Vasconcelos (UCSD, USA)</i>	
Binary TTC: A Temporal Geofence for Autonomous Navigation	12941
<i>Abhishek Badki (University of California, Santa Barbara), Orazio Gallo (NVIDIA Research), Jan Kautz (NVIDIA), and Pradeep Sen (UC Santa Barbara)</i>	
LED2-Net: Monocular 360° Layout Estimation via Differentiable Depth Rendering	12951
<i>Fu-En Wang (National Tsing Hua University), Yu-Hsuan Yeh (National Chiao Tung University), Min Sun (NTHU), Wei-Chen Chiu (National Chiao Tung University), and Yi-Hsuan Tsai (NEC Labs America)</i>	
A Realistic Evaluation of Semi-Supervised Learning for Fine-Grained Classification	12961
<i>Jong-Chyi Su (University of Massachusetts, Amherst), Zezhou Cheng (University of Massachusetts, Amherst), and Subhransu Maji (University of Massachusetts, Amherst)</i>	
Seeing Out of the Box: End-to-End Pre-Training for Vision-Language Representation Learning ...	12971
<i>Zhicheng Huang (University of Science and Technology Beijing), Zhaoyang Zeng (Sun Yat-sen University), Yupan Huang (Sun Yat-sen University), Bei Liu (Microsoft Research), Dongmei Fu (University of Science and Technology Beijing), and Jianlong Fu (Microsoft Research)</i>	
Intentonomy: A Dataset and Study Towards Human Intent Understanding	12981
<i>Menglin Jia (Cornell University), Zuxuan Wu (UMD), Austin Reiter (Facebook AI), Claire Cardie (Cornell University), Serge Belongie (Cornell University), and Ser-Nam Lim (Facebook AI)</i>	
Mutual Graph Learning for Camouflaged Object Detection	12992
<i>Qiang Zhai (UESTC), Xin Li (Group 42), Fan Yang (Group 42), Chenglizhao Chen (Qingdao University), Hong Cheng (UESTC), and Deng-Ping Fan (Inception Institute of Artificial Intelligence)</i>	
Hallucination Improves Few-Shot Object Detection	13003
<i>Weilin Zhang (University of Illinois Urbana-Champaign) and Yu-Xiong Wang (University of Illinois at Urbana-Champaign)</i>	

Learning To Predict Visual Attributes in the Wild	13013
<i>Khoi Pham (University of Maryland, College Park), Kushal Kafle (Adobe Research), Zhe Lin (Adobe Research), Zhihong Ding (Adobe Research), Scott Cohen (Adobe Research), Quan Tran (Adobe Research), and Abhinav Shrivastava (University of Maryland)</i>	
Scaled-YOLOv4: Scaling Cross Stage Partial Network	13024
<i>Chien-Yao Wang (Institute of Information Science, Academia Sinica), Alexey Bochkovskiy (Intel ISL), and Hong-Yuan Mark Mark Liao (Institute of Information Science, Academia Sinica, Taiwan)</i>	
You Only Look One-Level Feature	13034
<i>Qiang Chen (CASIA), Yingming Wang (Megvii Technology), Tong Yang (Megvii Technology), Xiangyu Zhang (Megvii Technology), Jian Cheng (Chinese Academy of Sciences, China), and Jian Sun (Megvii Technology)</i>	
Neighborhood Normalization for Robust Geometric Feature Learning	13044
<i>Xingtong Liu (Johns Hopkins University), Benjamin D. Killeen (Johns Hopkins University), Ayushi Sinha (Philips Research North America), Masaru Ishii (Johns Hopkins Medical Institutions), Gregory D. Hager (The Johns Hopkins University), Russell H. Taylor (Johns Hopkins University), and Mathias Unberath (Johns Hopkins University)</i>	
High-Fidelity Face Tracking for AR/VR via Deep Lighting Adaptation	13054
<i>Lele Chen (Facebook), Chen Cao (Facebook Reality Labs), Fernando De la Torre (Facebook), Jason Saragih (Facebook), Chenliang Xu (University of Rochester), and Yaser Sheikh (Facebook Reality Labs)</i>	
Cuboids Revisited: Learning Robust 3D Shape Fitting to Single RGB Images	13065
<i>Florian Kluger (Leibniz Universität Hannover), Hanno Ackermann (Leibniz University Hannover), Eric Brachmann (Niantic), Michael Ying Yang (University of Twente), and Bodo Rosenhahn (Leibniz University Hannover)</i>	
Cycle4Completion: Unpaired Point Cloud Completion Using Cycle Transformation With Missing Region Coding	13075
<i>Xin Wen (Tsinghua University), Zhizhong Han (Wayne State University), Yan-Pei Cao (Y-tech, Kuaishou Technology), Pengfei Wan (Kuaishou Technology), Wen Zheng (Kuaishou Technology), and Yu-Shen Liu (Tsinghua University)</i>	
LiDAR-Based Panoptic Segmentation via Dynamic Shifting Network	13085
<i>Fangzhou Hong (Nanyang Technological University), Hui Zhou (SenseTime Group Limited.), Xinge Zhu (The Chinese University of Hong Kong), Hongsheng Li (Chinese University of Hong Kong), and Ziwei Liu (Nanyang Technological University)</i>	
RPSRNet: End-to-End Trainable Rigid Point Set Registration Network Using Barnes-Hut 2D-Tree Representation	13095
<i>Sk Aziz Ali (Technische Universität Kaiserslautern), Kerem Kahraman (Technische Universität Kaiserslautern), Gerd Reis (German Research Center for Artificial Intelligence), and Didier Stricker (DFKI)</i>	
Online Learning of a Probabilistic and Adaptive Scene Representation	13106
<i>Zike Yan (Peking University), Xin Wang (Peking University), and Hongbin Zha (Peking University, China)</i>	

Quantum Permutation Synchronization	13117
<i>Tolga Birdal (TU Munich), Vladislav Golyanik (MPI for Informatics), Christian Theobalt (MPI Informatik), and Leonidas J. Guibas (Stanford University)</i>	
Wide-Baseline Multi-Camera Calibration Using Person Re-Identification	13129
<i>Yan Xu (Carnegie Mellon University), Yu-Jhe Li (Carnegie Mellon University), Xinshuo Weng (Carnegie Mellon University), and Kris Kitani (Carnegie Mellon University)</i>	
StaR: Self-Supervised Tracking and Reconstruction of Rigid Objects in Motion With Neural Rendering	13139
<i>Wentao Yuan (University of Washington), Zhaoyang Lv (Facebook), Tanner Schmidt (Facebook Reality Labs), and Steven Lovegrove (Facebook Reality Labs)</i>	
PatchMatch-Based Neighborhood Consensus for Semantic Correspondence	13148
<i>Jae Yong Lee (University of Illinois), Joseph DeGol (Microsoft), Victor Fragoso (Microsoft), and Sudipta N. Sinha (Microsoft)</i>	
Learning Feature Aggregation for Deep 3D Morphable Models	13159
<i>Zhixiang Chen (Imperial College London) and Tae-Kyun Kim (Imperial College London)</i>	
A Functional Approach to Rotation Equivariant Non-Linearities for Tensor Field Networks.	13169
<i>Adrien Poulenard (Stanford) and Leonidas J. Guibas (Stanford University)</i>	
Generalizing to the Open World: Deep Visual Odometry With Online Adaptation	13179
<i>Shunkai Li (Peking University), Xin Wu (Peking University), Yingdian Cao (Peking University), and Hongbin Zha (Peking University, China)</i>	
Panoptic-PolarNet: Proposal-Free LiDAR Point Cloud Panoptic Segmentation	13189
<i>Zixiang Zhou (University of Central Florida), Yang Zhang (University of Central Florida), and Hassan Foroosh (University of Central Florida)</i>	
3D Spatial Recognition Without Spatially Labeled 3D	13199
<i>Zhongzheng Ren (UIUC), Ishan Misra (Facebook AI Research), Alexander G. Schwing (UIUC), and Rohit Girdhar (Facebook AI Research)</i>	
ACTION-Net: Multipath Excitation for Action Recognition	13209
<i>Zhengwei Wang (Trinity Collge Dublin), Qi She (Bytedance AI Lab), and Aljosa Smolic (Trinity College Dublin)</i>	
Anticipating Human Actions by Correlating Past With the Future With Jaccard Similarity Measures	13219
<i>Basura Fernando (Agency for Science, Technology and Research, A*STAR, Singapore) and Samitha Herath (Monash University)</i>	
Glance and Gaze: Inferring Action-Aware Points for One-Stage Human-Object Interaction Detection	13229
<i>Xubin Zhong (South China University of Technology), Xian Qu (South China University of Technology), Changxing Ding (South China University of Technology), and Dacheng Tao (JD.com)</i>	

How Robust Are Randomized Smoothing Based Defenses to Data Poisoning?	13239
<i>Akshay Mehra (Tulane University), Bhavya Kailkhura (Lawrence Livermore National Laboratory), Pin-Yu Chen (IBM Research), and Jihun Hamm (Tulane University)</i>	
FaceSec: A Fine-Grained Robustness Evaluation Framework for Face Recognition Systems	13249
<i>Liang Tong (Washington University in Saint Louis), Zhengzhang Chen (NEC Laboratories America, Inc.), Jingchao Ni (NEC Laboratories America), Wei Cheng (NEC Laboratories America), Dongjin Song (University of Connecticut), Haifeng Chen (NEC Labs), and Yevgeniy Vorobeychik (Washington University in St. Louis)</i>	
Rethinking the Heatmap Regression for Bottom-Up Human Pose Estimation	13259
<i>Zhengxiong Luo (Institute of Automation Chinese Academy of Sciences), Zhicheng Wang (Megvii), Yan Huang (Institute of Automation, Chinese Academy of Sciences), Liang Wang (NLPR, China), Tieniu Tan (NLPR, China), and Erjin Zhou (Megvii Research)</i>	
Camera-Space Hand Mesh Recovery via Semantic Aggregation and Adaptive 2D-1D Registration	13269
<i>Xingyu Chen (Kuaishou Technology), Yufeng Liu (Kuaishou Technology), Chongyang Ma (Kuaishou Technology), Jianlong Chang (Huawei Cloud & AI), Huayan Wang (Kuaishou Technology), Tian Chen (Kuaishou Technology), Xiaoyan Guo (Kuaishou Technology), Pengfei Wan (Kuaishou Technology), and Wen Zheng (Kuaishou Technology)</i>	
S3: Neural Shape, Skeleton, and Skinning Fields for 3D Human Modeling	13279
<i>Ze Yang (University of Toronto), Shenlong Wang (Uber ATG, University of Toronto), Sivabalan Manivasagam (University of Toronto), Zeng Huang (Snap Inc.), Wei-Chiu Ma (MIT), Xinchen Yan (Uber ATG), Ersin Yumer (Uber ATG), and Raquel Urtasun (Uber ATG)</i>	
CanonPose: Self-Supervised Monocular 3D Human Pose Estimation in the Wild	13289
<i>Bastian Wandt (Leibniz University Hannover), Marco Rudolph (Leibniz University Hannover), Petriッサ Zell (Leibniz University Hannover), Helge Rhodin (UBC), and Bodo Rosenhahn (Leibniz University Hannover)</i>	
Lipstick Ain't Enough: Beyond Color Matching for In-the-Wild Makeup Transfer	13300
<i>Thao Nguyen (VinAI Research), Anh Tuan Tran (VinAI Research), and Minh Hoai (Stony Brook University)</i>	
Virtual Fully-Connected Layer: Training a Large-Scale Face Recognition Dataset With Limited Computational Resources	13310
<i>Pengyu Li (Alibaba Group), Biao Wang (Alibaba), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)</i>	
Learning From the Master: Distilling Cross-Modal Advanced Knowledge for Lip Reading	13320
<i>Sucheng Ren (South China University of Technology), Yong Du (Ocean University of China), Jianming Lv (South China University of Technology), Guoqiang Han (South China University of Technology), and Shengfeng He (South China University of Technology)</i>	
Watching You: Global-Guided Reciprocal Learning for Video-Based Person Re-Identification	13329
<i>Xuehu Liu (Dalian University of Technology), Pingping Zhang (Dalian University of Technology), Chenyang Yu (Dalian University Of Technology), Huchuan Lu (Dalian University of Technology), and Xiaoyun Yang (Remark Holdings)</i>	

Sparse Multi-Path Corrections in Fringe Projection Profilometry	13339
<i>Yu Zhang (Nanjing University), Daniel Lau (University of Kentucky), and David Wipf (Amazon)</i>	
Attention-Guided Image Compression by Deep Reconstruction of Compressive Sensed Saliency Skeleton	13349
<i>Xi Zhang (Shanghai Jiao Tong University) and Xiaolin Wu (McMaster University)</i>	
Invertible Denoising Network: A Light Solution for Real Noise Removal	13360
<i>Yang Liu (The Australian National University & Data61), Zhenyue Qin (Australian National University), Saeed Anwar (The Australian National University), Pan Ji (OPPO US Research Center), Dongwoo Kim (POSTECH), Sabrina Caldwell (Australian National University), and Tom Gedeon (The Australian National University)</i>	
Multi-Decoding Deraining Network and Quasi-Sparsity Based Training	13370
<i>Yinglong Wang (University of Electronic Science and Technology of China), Chao Ma (Shanghai Jiao Tong University), and Bing Zeng (University of Electronic Science and Technology of China)</i>	
Unsupervised Real-World Image Super Resolution via Domain-Distance Aware Training	13380
<i>Yunxuan Wei (South China University of Technology), Shuhang Gu (ETH Zurich, Switzerland), Yawei Li (ETH Zurich), Radu Timofte (ETH Zurich), Longcun Jin (South China University of Technology), and Hengjie Song (School of Software Engineering, South China University of Technology)</i>	
Single Image Reflection Removal With Absorption Effect	13390
<i>Qian Zheng (Nanyang Technological University), Boxin Shi (Peking University), Jinnan Chen (Nanyang Technological University), Xudong Jiang (Nanyang Technological University), Ling-Yu Duan (Peking University), and Alex C. Kot (Nanyang Technological University)</i>	
Exploiting Aliasing for Manga Restoration	13400
<i>Minshan Xie (The Chinese University of Hong Kong), Menghan Xia (The Chinese University of Hong Kong), and Tien-Tsin Wong (The Chinese University of Hong Kong)</i>	
Rich Context Aggregation With Reflection Prior for Glass Surface Detection	13410
<i>Jiaying Lin (City University of Hong Kong), Zebang He (City university of Hong Kong), and Rynson W.H. Lau (City University of Hong Kong)</i>	
MR Image Super-Resolution With Squeeze and Excitation Reasoning Attention Network	13420
<i>Yulun Zhang (Northeastern University), Kai Li (Northeastern University), Kunpeng Li (Northeastern University), and Yun Fu (Northeastern University)</i>	
Rich Features for Perceptual Quality Assessment of UGC Videos	13430
<i>Yilin Wang (Google Inc.), Junjie Ke (Google), Hossein Talebi (Google), Joong Gon Yim (Google Inc.), Neil Birkbeck (YouTube/Google), Balu Adsumilli (YouTube/Google), Peyman Milanfar (Google), and Feng Yang (Google Research)</i>	
A 3D GAN for Improved Large-Pose Facial Recognition	13440
<i>Richard T. Marriott (Ecole Centrale de Lyon), Sami Romdhani (IDEMIA), and Liming Chen (Ecole Centrale de Lyon)</i>	

Sewer-ML: A Multi-Label Sewer Defect Classification Dataset and Benchmark	13451
<i>Joakim Bruslund Haurum (Aalborg University) and Thomas B. Moeslund (Aalborg University)</i>	
AGORA: Avatars in Geography Optimized for Regression Analysis	13463
<i>Priyanka Patel (Max Planck Institute of Intelligent Systems, Tuebingen), Chun-Hao P. Huang (Max Planck Institute for Intelligent Systems), Joachim Tesch (Max Planck Institute for Intelligent Systems), David T. Hoffmann (University of Freiburg; Bosch Center for Artificial Intelligence), Shashank Tripathi (Max Planck Institute for Intelligent Systems), and Michael J. Black (Max Planck Institute for Intelligent Systems)</i>	
SKFAC: Training Neural Networks With Faster Kronecker-Factored Approximate Curvature	13474
<i>Zedong Tang (Xidian University), Fenlong Jiang (Xidian University), Maoguo Gong (Xidian University), Hao Li (Xidian University), Yue Wu (Xidian university), Fan Yu (Huawei Technologies Co. Ltd), Zidong Wang (Huawei Technologies Co. Ltd), and Min Wang (Huawei Technologies Co. Ltd)</i>	
Tree-Like Decision Distillation	13483
<i>Jie Song (Zhejiang University), Haofei Zhang (Zhejiang University), Xinchao Wang (National University of Singapore), Mengqi Xue (Zhejiang University), Ying Chen (Zhejiang University), Li Sun (Zhejiang University), Dacheng Tao (The University of Sydney), and Mingli Song (Zhejiang University)</i>	
How Does Topology Influence Gradient Propagation and Model Performance of Deep Networks With DenseNet-Type Skip Connections?	13493
<i>Kartikeya Bhardwaj (Arm Inc.), Guihong Li (UT Austin), and Radu Marculescu (The University of Texas at Austin)</i>	
EnD: Entangling and Disentangling Deep Representations for Bias Correction	13503
<i>Enzo Tartaglione (University of Torino), Carlo Alberto Barbano (University of Torino), and Marco Grangetto (University of Torino)</i>	
Learning Decision Trees Recurrently Through Communication	13513
<i>Stephan Alaniz (Max Planck Institute for Informatics), Diego Marcos (Wageningen University), Bernt Schiele (MPI Informatics), and Zeynep Akata (University of Tübingen)</i>	
Neural Response Interpretation Through the Lens of Critical Pathways	13523
<i>Ashkan Khakzar (Technical University of Munich), Soroosh Baselizadeh (Sharif University of Technology), Saurabh Khanduja (Technical University of Munich), Christian Rupprecht (University of Oxford), Seong Tae Kim (Kyung Hee University), and Nassir Navab (TU Munich, Germany)</i>	
Masksembles for Uncertainty Estimation	13534
<i>Nikita Durasov (École polytechnique fédérale de Lausanne), Timur Bagautdinov (Facebook), Pierre Baque (Neural Concept), and Pascal Fua (EPFL, Switzerland)</i>	
Self-Supervised Video Hashing via Bidirectional Transformers	13544
<i>Shuyan Li (Tsinghua University), Xiu Li (Tsinghua University), Jiwen Lu (Tsinghua University), and Jie Zhou (Tsinghua University)</i>	

3D Shape Generation With Grid-Based Implicit Functions	13554
<i>Moritz Ibing (RWTH Aachen University), Isaak Lim (RWTH Aachen University), and Leif Kobbelt (RWTH Aachen University)</i>	
Positional Encoding As Spatial Inductive Bias in GANs	13564
<i>Rui Xu (CUHK), Xintao Wang (Tencent), Kai Chen (SenseTime), Bolei Zhou (CUHK), and Chen Change Loy (Nanyang Technological University)</i>	
Blur, Noise, and Compression Robust Generative Adversarial Networks	13574
<i>Takuhiko Kaneko (The University of Tokyo) and Tatsuya Harada (The University of Tokyo / RIKEN)</i>	
Learning by Planning: Language-Guided Global Image Editing	13585
<i>Jing Shi (University of Rochester), Ning Xu (Adobe Research), Yihang Xu (University of Rochester), Trung Bui (Adobe Research), Franck Dernoncourt (Adobe Research), and Chenliang Xu (University of Rochester)</i>	
Teachers Do More Than Teach: Compressing Image-to-Image Models	13595
<i>Qing Jin (Northeastern University), Jian Ren (Snap Inc.), Oliver J. Woodford (Snap Inc), Jiazhao Wang (Snap Inc.), Geng Yuan (Northeastern University), Yanzhi Wang (Northeastern University), and Sergey Tulyakov (Snap Inc)</i>	
Autoregressive Stylized Motion Synthesis With Generative Flow	13607
<i>Yu-Hui Wen (Tsinghua University), Zhipeng Yang (Institute of Computing Technology, Chinese Academy of Sciences), Hongbo Fu (City University of Hong Kong), Lin Gao (Institute of Computing Technology, Chinese Academy of Sciences), Yanan Sun (Tsinghua University), and Yong-Jin Liu (Tsinghua University)</i>	
MUST-GAN: Multi-Level Statistics Transfer for Self-Driven Person Image Generation	13617
<i>Tianxiang Ma (Institute of Automation, Chinese Academy of Sciences), Bo Peng (Institute of Automation, Chinese Academy of Sciences), Wei Wang (Center for Research on Intelligent Perception and Computing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences), and Jing Dong (Chinese Academy of Sciences)</i>	
House-GAN++: Generative Adversarial Layout Refinement Network towards Intelligent Computational Agent for Professional Architects	13627
<i>Nelson Nauata (Simon Fraser University), Sepidehsadat Hosseini (Seoul National University), Kai-Hung Chang (Autodesk Research), Hang Chu (Autodesk Research), Chin-Yi Cheng (Autodesk Research), and Yasutaka Furukawa (Simon Fraser University)</i>	
Variational Transformer Networks for Layout Generation	13637
<i>Diego Martín Arroyo (Google), Janis Postels (ETH Zurich), and Federico Tombari (Google, TU Munich)</i>	
Motion Representations for Articulated Animation	13648
<i>Aliaksandr Siarohin (University of Trento), Oliver J. Woodford (Snap Inc), Jian Ren (Snap Inc.), Menglei Chai (Snap Inc.), and Sergey Tulyakov (Snap Inc)</i>	

Pareto Self-Supervised Training for Few-Shot Learning	13658
<i>Zhengyu Chen (Westlake University), Jixie Ge (Westlake University), Heshen Zhan (Westlake University), Siteng Huang (Westlake University), and Donglin Wang (Westlake University)</i>	
RaScaNet: Learning Tiny Models by Raster-Scanning Images	13668
<i>Jaehyung Yoo (Samsung), Dongwook Lee (Samsung), Changyong Son (Samsung), Sangil Jung (Samsung), ByungIn Yoo (Samsung Advanced Institute of Technology), Changkyu Choi (Samsung), Jae-Joon Han (Samsung), and Bohyung Han (Seoul National University)</i>	
AlphaMatch: Improving Consistency for Semi-Supervised Learning With Alpha-Divergence	13678
<i>Chengyue Gong (UT Austin), Dilin Wang (Facebook), and Qiang Liu (UT Austin)</i>	
Nearest Neighbor Matching for Deep Clustering	13688
<i>Zhiyuan Dang (JD Tech), Cheng Deng (Xidian University), Xu Yang (Xidian University), Kun Wei (Xidian University), and Heng Huang (University of Pittsburgh & JD Tech)</i>	
DeepACG: Co-Saliency Detection via Semantic-Aware Contrast Gromov-Wasserstein Distance .	13698
<i>Kaihua Zhang (NUIST), Mingliang Dong (nuist), Bo Liu (JD.com), Xiao-Tong Yuan (Nanjing University of Information Science and Technology), and Qingshan Liu (Nanjing University of Information Science & Technology)</i>	
Coordinate Attention for Efficient Mobile Network Design	13708
<i>Qibin Hou (National University of Singapore), Daquan Zhou (National University of Singapore), and Jiashi Feng (NUS)</i>	
Landmark Regularization: Ranking Guided Super-Net Training in Neural Architecture Search ..	13718
<i>Kaicheng Yu (EPFL), René Ranftl (Intel Labs), and Mathieu Salzmann (EPFL)</i>	
RepVGG: Making VGG-Style ConvNets Great Again	13728
<i>Xiaohan Ding (Tsinghua University), Xiangyu Zhang (Megvii Technology), Ningning Ma (hkust), Jungong Han (Aberystwyth University), Guiguang Ding (Tsinghua University, China), and Jian Sun (Megvii Technology)</i>	
3D Graph Anatomy Geometry-Integrated Network for Pancreatic Mass Segmentation, Diagnosis, and Quantitative Patient Management	13738
<i>Tianyi Zhao (Stony Brook University), Kai Cao (Changhai Hospital), Jiawen Yao (PAII, INC), Isabella Nagues (Harvard University), Le Lu (PAII Inc.), Lingyun Huang (PingAn Technology), Jing Xiao (Ping An Insurance, (Group) Company of China), Zhaozheng Yin (Stony Brook University), and Ling Zhang (PAII Inc.)</i>	
Exploring and Distilling Posterior and Prior Knowledge for Radiology Report Generation	13748
<i>Fenglin Liu (Peking University), Xian Wu (Tencent Medical AI Lab), Shen Ge (Tencent Medical AI Lab), Wei Fan (Tencent), and Yuexian Zou (Peking University)</i>	

Towards More Flexible and Accurate Object Tracking With Natural Language: Algorithms and Benchmark	13758
<i>Xiao Wang (Peng Cheng Laboratory), Xiujun Shu (Peng Cheng Laboratory), Zhipeng Zhang (Chinese Academy of Sciences), Bo Jiang (Anhui University), Yaowei Wang (PengCheng Laboratory), Yonghong Tian (Peking University), and Feng Wu (University of Science and Technology of China)</i>	
STMTrack: Template-Free Visual Tracking With Space-Time Memory Networks	13769
<i>Zhihong Fu (State Key Laboratory of Virtual Reality Technology and System, Beihang University, China), Qingjie Liu (State Key Laboratory of Virtual Reality Technology and System, Beihang University, China), Zehua Fu (Hangzhou Innovation Institute, Beihang University), and Yunhong Wang (State Key Laboratory of Virtual Reality Technology and System, Beihang University, China)</i>	
DyGLIP: A Dynamic Graph Model With Link Prediction for Accurate Multi-Camera Multiple Object Tracking	13779
<i>Kha Gia Quach (Concordia University), Pha Nguyen (VinAI Research), Huu Le (Chalmers University of Technology), Thanh-Dat Truong (University of Arkansas), Chi Nhan Duong (Concordia University), Minh-Triet Tran (University of Science, VNU-HCM), and Khoa Luu (University of Arkansas)</i>	
SuperMix: Supervising the Mixing Data Augmentation	13789
<i>Ali Dabouei (West Virginia university), Sobhan Soleymani (West Virginia University), Fariborz Taherkhani (West Virginia University), and Nasser M. Nasrabadi (West Virginia University)</i>	
Monte Carlo Scene Search for 3D Scene Understanding	13799
<i>Shreyas Hampali (TU Graz), Sinisa Stekovic (Graz University of Technology), Sayan Deb Sarkar (Graz University of Technology), Chetan S. Kumar (Graz University of Technology), Friedrich Fraundorfer (Graz University of Technology), and Vincent Lepetit (Ecole des Ponts ParisTech)</i>	
MAZE: Data-Free Model Stealing Attack Using Zeroth-Order Gradient Estimation	13809
<i>Sanjay Kariyappa (Georgia Institute of Technology), Atul Prakash (University of Michigan), and Moinuddin K Qureshi (Georgia Institute of Technology)</i>	
Visualizing Adapted Knowledge in Domain Transfer	13819
<i>Yunzhong Hou (Australian National University) and Liang Zheng (Australian National University)</i>	
Prototypical Cross-Domain Self-Supervised Learning for Few-Shot Unsupervised Domain Adaptation	13829
<i>Xiangyu Yue (University of California, Berkeley), Zangwei Zheng (Nanjing University), Shanghang Zhang (UC Berkeley), Yang Gao (Tsinghua University), Trevor Darrell (UC Berkeley), Kurt Keutzer (EECS, UC Berkeley), and Alberto Sangiovanni Vincentelli (University of California, Berkeley)</i>	
KSM: Fast Multiple Task Adaption via Kernel-Wise Soft Mask Learning	13840
<i>Li Yang (Arizona State University), Zhezhi He (Arizona State University), Junshan Zhang (Arizona State University), and Deliang Fan (Arizona State University)</i>	

Picasso: A CUDA-Based Library for Deep Learning Over 3D Meshes	13849
<i>Huan Lei (The University of Western Australia), Naveed Akhtar (The University of Western Australia), and Ajmal Mian (University of Western Australia)</i>	
Efficient Feature Transformations for Discriminative and Generative Continual Learning	13860
<i>Vinay Kumar Verma (Duke University), Kevin J Liang (Duke University), Nikhil Mehta (Duke University), Piyush Rai (IIT Kanpur), and Lawrence Carin (CS)</i>	
Spatial Assembly Networks for Image Representation Learning	13871
<i>Yang Li (University of Missouri), Shichao Kan (Beijing Jiaotong University), Jianhe Yuan (University of Missouri-Columbia), Wenming Cao (Shenzhen University), and Zhihai He (University of Missouri Columbia)</i>	
Self-Supervised Video Representation Learning by Context and Motion Decoupling	13881
<i>Lianghua Huang (Alibaba Group), Yu Liu (Alibaba Group), Bin Wang (alibaba group), Pan Pan (Alibaba Group), Yinghui Xu (Alibaba DAMO Academy), and Rong Jin (alibaba group)</i>	
Learning Probabilistic Ordinal Embeddings for Uncertainty-Aware Regression	13891
<i>Wanhua Li (Tsinghua University), Xiaoke Huang (Beijing Normal University), Jiwen Lu (Tsinghua University), Jianjiang Feng (Tsinghua University), and Jie Zhou (Tsinghua University)</i>	
CFNet: Cascade and Fused Cost Volume for Robust Stereo Matching	13901
<i>Zhelun Shen (Peking University), Yuchao Dai (Northwestern Polytechnical University), and Zhibo Rao (Northwestern Polytechnical University)</i>	
Depth Completion Using Plane-Residual Representation	13911
<i>Byeong-Uk Lee (KAIST), Kyunghyun Lee (KAIST), and In So Kweon (KAIST)</i>	
Look Closer To Segment Better: Boundary Patch Refinement for Instance Segmentation	13921
<i>Chufeng Tang (Tsinghua University), Hang Chen (Tsinghua University), Xiao Li (Tsinghua University), Jianmin Li (Tsinghua University), Zhaoxiang Zhang (Chinese Academy of Sciences, China), and Xiaolin Hu (Tsinghua University)</i>	
Energy-Based Learning for Scene Graph Generation	13931
<i>Mohammed Suhail (University of British Columbia), Abhay Mittal (Amazon), Behjat Siddiquie (Amazon), Chris Broaddus (Amazon), Jayan Eledath (Amazon), Gerard Medioni (USC), and Leonid Sigal (University of British Columbia)</i>	
Heterogeneous Grid Convolution for Adaptive, Efficient, and Controllable Computation	13941
<i>Ryuhei Hamaguchi (National Institute of Advanced Industrial Science and Technology), Yasutaka Furukawa (Simon Fraser University), Masaki Onishi (National Institute of Advanced Industrial Science and Technology), and Ken Sakurada (National Institute of Advanced Industrial Science and Technology)</i>	

DCNAS: Densely Connected Neural Architecture Search for Semantic Image Segmentation	13951
<i>Xiong Zhang (JOYY Inc.), Hongmin Xu (Y-Lab, Kwai), Hong Mo (State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University), Jianchao Tan (Kwai Inc.), Cheng Yang (JOYY Inc.), Lei Wang (JOYY Inc), and Wenqi Ren (Institute of Information Engineering, Chinese Academy of Sciences)</i>	
Weakly Supervised Instance Segmentation for Videos With Temporal Mask Consistency	13963
<i>Qing Liu (Johns Hopkins University), Vignesh Ramanathan (Facebook), Dhruv Mahajan (Facebook), Alan Yuille (Johns Hopkins University), and Zhenheng Yang (Facebook AI)</i>	
Few-Shot Segmentation Without Meta-Learning: A Good Transductive Inference Is All You Need?	13974
<i>Malik Boudiaf (ÉTS Montreal), Hoel Kervadec (ÉTS Montréal), Ziko Intiaz Masud (ETS Montreal), Pablo Piantanida (CentraleSupélec - CNRS - L2S), Ismail Ben Ayed (ETS Montreal), and Jose Dolz (ETS Montreal)</i>	
Conditional Bures Metric for Domain Adaptation	13984
<i>You-Wei Luo (Sun Yat-sen University) and Chuan-Xian Ren (Sun Yat-Sen University)</i>	
Relative Order Analysis and Optimization for Unsupervised Deep Metric Learning	13994
<i>Shichao Kan (Beijing Jiaotong University), Yigang Cen (Beijing Jiaotong University), Yang Li (University of Missouri), Vladimir Mladenovic (University of Kragujevac), and Zhihai He (University of Missouri Columbia)</i>	
MIST: Multiple Instance Self-Training Framework for Video Anomaly Detection	14004
<i>Jia-Chang Feng (Sun Yat-Sen University), Fa-Ting Hong (Sun Yat-Sen University), and Wei-Shi Zheng (Sun Yat-sen University, China)</i>	
Patch-VQ: ‘Patching Up’ the Video Quality Problem	14014
<i>Zhenqiang Ying (The University of Texas at Austin), Maniratnam Mandal (University of Texas at Austin), Deepti Ghadiyaram (Facebook), and Alan Bovik (University of Texas at Austin)</i>	
Boosting Video Representation Learning With Multi-Faceted Integration	14025
<i>Zhaofan Qiu (JD.com), Ting Yao (JD AI Research), Chong-Wah Ngo (Singapore Management University), Xiao-Ping Zhang (Ryerson University), Dong Wu (JD.com), and Tao Mei (AI Research of JD.com)</i>	
Delving Deep Into Many-to-Many Attention for Few-Shot Video Object Segmentation	14035
<i>Haoxin Chen (South China University of Technology), Hanjie Wu (South China University of Technology), Nanxuan Zhao (Chinese University of Hong Kong), Sucheng Ren (South China University of Technology), and Shengfeng He (South China University of Technology)</i>	
FAIEr: Fidelity and Adequacy Ensured Image Caption Evaluation	14045
<i>Sijin Wang (Institute of Computing Technology, Chinese Academy of Sciences), Ziwei Yao (Institute of Computing Technology, Chinese Academy of Sciences), Ruiping Wang (Institute of Computing Technology, Chinese Academy of Sciences), Zhongqin Wu (Tomorrow Advancing Life), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences)</i>	

Iterative Shrinking for Referring Expression Grounding Using Deep Reinforcement Learning	14055
<i>Mingjie Sun (Xi'an Jiaotong-Liverpool University), Jimin Xiao (Xi'an Jiaotong-Liverpool University), and Eng Gee Lim (Xian jiaotong-liverpool University)</i>	
Repetitive Activity Counting by Sight and Sound	14065
<i>Yunhua Zhang (University of Amsterdam), Ling Shao (Inception Institute of Artificial Intelligence), and Cees G. M. Snoek (University of Amsterdam)</i>	
Audio-Driven Emotional Video Portraits	14075
<i>Xinya Ji (Nanjing University), Hang Zhou (The Chinese University of Hong Kong), Kaisiyuan Wang (the University of Sydney), Wayne Wu (SenseTime Research), Chen Change Loy (Nanyang Technological University), Xun Cao (Nanjing University), and Feng Xu (Tsinghua University)</i>	
Improving Weakly Supervised Visual Grounding by Contrastive Knowledge Distillation	14085
<i>Liwei Wang (Tencent), Jing Huang (UIUC), Yin Li (University of Wisconsin-Madison), Kun Xu (Tencent), Zhengyuan Yang (University of Rochester), and Dong Yu (Tencent AI Lab)</i>	
Hierarchical and Partially Observable Goal-Driven Policy Learning With Goals Relational Graph	14096
<i>Xin Ye (Arizona State University) and Yezhou Yang (Arizona State University)</i>	
KRISP: Integrating Implicit and Symbolic Knowledge for Open-Domain Knowledge-Based VQA	14106
<i>Kenneth Marino (CMU), Xinlei Chen (FAIR), Devi Parikh (Georgia Tech & Facebook AI Research), Abhinav Gupta (CMU/FAIR), and Marcus Rohrbach (Facebook AI Research)</i>	
Focus on Local: Detecting Lane Marker From Bottom Up via Key Point	14117
<i>Zhan Qu (Noah's Ark Lab, Huawei Technologies Co., Ltd.), Huan Jin (Noah's Ark Lab, Huawei Technologies Co., Ltd.), Yang Zhou (Noah's Ark Lab, Huawei Technologies Co., Ltd.), Zhen Yang (Huawei Noah's Ark Lab), and Wei Zhang (Noah's Ark Lab, Huawei Technologies)</i>	
VITON-HD: High-Resolution Virtual Try-On via Misalignment-Aware Normalization	14126
<i>Seunghwan Choi (Korea Advanced Institute of Science and Technology), Sunghyun Park (KAIST), Minsoo Lee (KAIST), and Jaegul Choo (Korea Advanced Institute of Science and Technology)</i>	
Patch-NetVLAD: Multi-Scale Fusion of Locally-Global Descriptors for Place Recognition	14136
<i>Stephen Hausler (Queensland University of Technology), Sourav Garg (Queensland University of Technology), Ming Xu (Queensland University of Technology), Michael Milford (ACRV and QUT, Australia), and Tobias Fischer (Queensland University of Technology)</i>	
DeRF: Decomposed Radiance Fields	14148
<i>Daniel Rebaun (University of British Columbia), Wei Jiang (University of British Columbia), Soroosh Yazdani (Google Inc.), Ke Li (Simon Fraser University), Kwang Moo Yi (University of British Columbia), and Andrea Tagliasacchi (Google Inc.)</i>	

Session 11

POSEFusion: Pose-Guided Selective Fusion for Single-View Human Volumetric Capture	14157
<i>Zhe Li (Tsinghua University), Tao Yu (Tsinghua University), Zerong Zheng (Tsinghua University), Kaiwen Guo (Google), and Yebin Liu (Tsinghua University)</i>	
FESTA: Flow Estimation via Spatial-Temporal Attention for Scene Point Clouds	14168
<i>Haiyan Wang (The City College of New York), Jiahao Pang (InterDigital Communications, Inc.), Muhammad A. Lodhi (Interdigital), Yingli Tian (City University of New York), and Dong Tian (InterDigital)</i>	
Isometric Multi-Shape Matching	14178
<i>Maolin Gao (TUM), Zorah Löhner (University of Siegen), Johan Thunberg (Halmstad University), Daniel Cremers (TU Munich), and Florian Bernard (TUM)</i>	
PatchmatchNet: Learned Multi-View Patchmatch Stereo	14189
<i>Fangjinhua Wang (ETH Zurich), Silvano Galliani (Microsoft), Christoph Vogel (Microsoft MixedReality&AI), Pablo Speciale (Microsoft), and Marc Pollefeys (ETH Zurich / Microsoft)</i>	
Point 4D Transformer Networks for Spatio-Temporal Modeling in Point Cloud Videos	14199
<i>Hehe Fan (NUS), Yi Yang (UTS), and Mohan Kankanhalli (National University of Singapore)</i>	
Learning To Aggregate and Personalize 3D Face From In-the-Wild Photo Collection	14209
<i>Zhenyu Zhang (Tencent), Yanhao Ge (Tencent YouTu), Renwang Chen (Tencent), Ying Tai (Tencent YouTu), Yan Yan (Nanjing University of Science and Technology), Jian Yang (Nanjing University of Science and Technology), Chengjie Wang (Tencent), Jilin Li (Tencent), and Feiyue Huang (Tencent)</i>	
MagFace: A Universal Representation for Face Recognition and Quality Assessment	14220
<i>Qiang Meng (AIBee), Shichao Zhao (Aibee), Zhida Huang (Aibee), and Feng Zhou (Aibee)</i>	
Event-Based Synthetic Aperture Imaging With a Hybrid Network	14230
<i>Xiang Zhang (Wuhan University), Wei Liao (WuHan University), Lei Yu (Wuhan University), Wen Yang (Wuhan University), and Gui-Song Xia (Wuhan University)</i>	
GLEAN: Generative Latent Bank for Large-Factor Image Super-Resolution	14240
<i>Kelvin C.K. Chan (Nanyang Technological University), Xintao Wang (Tencent), Xiangyu Xu (Nanyang Technological University), Jinwei Gu (SenseBrain), and Chen Change Loy (Nanyang Technological University)</i>	
NPAS: A Compiler-Aware Framework of Unified Network Pruning and Architecture Search for Beyond Real-Time Mobile Acceleration	14250
<i>Zhengang Li (Northeastern University), Geng Yuan (Northeastern University), Wei Niu (William & Mary), Pu Zhao (Northeastern University), Yanyu Li (Northeastern University), Yuxuan Cai (Northeastern University), Xuan Shen (Northeastern University), Zheng Zhan (Northeastern University), Zhenglun Kong (Northeastern University), Qing Jin (Northeastern University), Zhiyu Chen (Rice University), Sijia Liu (Michigan State University), Kaiyuan Yang (Rice University), Bin Ren (William & Mary), Yanzhi Wang (Northeastern University), and Xue Lin (Northeastern University)</i>	

Privacy-Preserving Image Features via Adversarial Affine Subspace Embeddings	14262
<i>Mihai Dusmanu (ETH Zurich), Johannes L. Schönberger (Microsoft), Sudipta N. Sinha (Microsoft), and Marc Pollefeys (ETH Zurich / Microsoft)</i>	
Image Generators With Conditionally-Independent Pixel Synthesis	14273
<i>Ivan Anokhin (Skolkovo Institute of Science and Technology), Kirill Demochkin (Skolkovo Institute of Science and Technology), Taras Khakhulin (Skolkovo Institute of Science and Technology), Gleb Sterkin (Samsung AI Center, Moscow), Victor Lempitsky (Samsung), and Denis Korzhenkov (Samsung AI Center, Moscow)</i>	
CoMoGAN: Continuous Model-Guided Image-to-Image Translation	14283
<i>Fabio Pizzati (Inria / Vislab), Pietro Cerri (Vislab), and Raoul de Charette (Inria)</i>	
Positive-Congruent Training: Towards Regression-Free Model Updates	14294
<i>Sijie Yan (CUHK), Yuanjun Xiong (Amazon), Kaustav Kundu (Amazon), Shuo Yang (Amazon), Siqi Deng (Amazon), Meng Wang (Amazon), Wei Xia (Amazon), and Stefano Soatto (AWS Amazon ML)</i>	
Capsule Network Is Not More Robust Than Convolutional Network	14304
<i>Jindong Gu (University of Munich), Volker Tresp (Siemens AG and Ludwig Maximilian University of Munich), and Han Hu (Microsoft Research Asia)</i>	
Dual-Stream Multiple Instance Learning Network for Whole Slide Image Classification With Self-Supervised Contrastive Learning	14313
<i>Bin Li (University of Wisconsin-Madison), Yin Li (University of Wisconsin-Madison), and Kevin W. Eliceiri (University of Wisconsin - Madison)</i>	
Probabilistic Tracklet Scoring and Inpainting for Multiple Object Tracking	14324
<i>Fatemeh Saleh (Australian National University, (ANU)), Sadegh Aliakbarian (Microsoft), Hamid Reza Tofighi (Monash University), Mathieu Salzmann (EPFL), and Stephen Gould (Australian National University, Australia)</i>	
Adaptive Methods for Real-World Domain Generalization	14335
<i>Abhimanyu Dubey (Massachusetts Institute of Technology), Vignesh Ramanathan (Facebook), Alex Pentland (MIT), and Dhruv Mahajan (Facebook)</i>	
Self-Supervised Geometric Perception	14345
<i>Heng Yang (MIT), Wei Dong (Carnegie Mellon University), Luca Carlone (Massachusetts Institute of Technology), and Vladlen Koltun (Intel Labs)</i>	
HITNet: Hierarchical Iterative Tile Refinement Network for Real-time Stereo Matching	14357
<i>Vladimir Tankovich (Google), Christian Häne (Google), Yinda Zhang (Google), Adarsh Kowdle (Google), Sean Fanello (Google), and Sofien Bouaziz (Google)</i>	
Bidirectional Projection Network for Cross Dimension Scene Understanding	14368
<i>Wenbo Hu (The Chinese University of Hong Kong), Hengshuang Zhao (University of Oxford), Li Jiang (The Chinese University of Hong Kong), Jiaya Jia (Chinese University of Hong Kong), and Tien-Tsin Wong (The Chinese University of Hong Kong)</i>	

A Fourier-Based Framework for Domain Generalization	14378
<i>Qinwei Xu (Cooperative Medianet Innovation Center, Shang hai Jiao Tong University), Ruipeng Zhang (Cooperative Medianet Innovation Center, Shang hai Jiao Tong University), Ya Zhang (Cooperative Medianet Innovation Center, Shang hai Jiao Tong University), Yanfeng Wang (Cooperative medianet innovation center of Shanghai Jiao Tong University), and Qi Tian (Huawei Cloud & AI)</i>	
Open-Vocabulary Object Detection Using Captions	14388
<i>Alireza Zareian (Columbia University), Kevin Dela Rosa (Snapchat Inc.), Derek Hao Hu (Snap Inc.), and Shih-Fu Chang (Columbia University)</i>	
MP3: A Unified Model To Map, Perceive, Predict and Plan	14398
<i>Sergio Casas (Uber ATG / University of Toronto), Abbas Sadat (Uber ATG), and Raquel Urtasun (Uber ATG)</i>	
Hierarchical Lovász Embeddings for Proposal-Free Panoptic Segmentation	14408
<i>Tommi Kerola (Preferred Networks, Inc.), Jie Li (Toyota Research Institute), Atsushi Kanehira (Preferred Networks, Inc.), Yasunori Kudo (Preferred Networks, Inc.), Alexis Vallet (Preferred Networks, Inc.), and Adrien Gaidon (Toyota Research Institute)</i>	
Accurate Few-Shot Object Detection With Support-Query Mutual Guidance and Hybrid Loss ...	14419
<i>Lu Zhang (Fudan University), Shuigeng Zhou (Fudan University), Jihong Guan (Tongji University), and Ji Zhang (Zhejiang Lab)</i>	
GLAVNet: Global-Local Audio-Visual Cues for Fine-Grained Material Recognition	14428
<i>Fengmin Shi (Nanjing University), Jie Guo (Nanjing University), Haonan Zhang (Nanjing University), Shan Yang (Nanjing University), Xiyang Wang (IQIYI Intelligence), and Yanwen Guo (Nanjing University)</i>	
Multi-Scale Aligned Distillation for Low-Resolution Detection	14438
<i>Lu Qi (The Chinese University of Hong Kong), Jason Kuen (Adobe Research), Jiuxiang Gu (Adobe Research), Zhe Lin (Adobe Research), Yi Wang (Chinese University of Hong Kong), Yukang Chen (The Chinese University of Hong Kong), Yanwei Li (The Chinese University of Hong Kong), and Jiaya Jia (Chinese University of Hong Kong)</i>	
Sparse R-CNN: End-to-End Object Detection With Learnable Proposals	14449
<i>Peize Sun (The University of Hong Kong), Rufeng Zhang (Tongji University), Yi Jiang (Bytedance), Tao Kong (Bytedance), Chenfeng Xu (UC Berkeley), Wei Zhan (University of California, Berkeley), Masayoshi Tomizuka (University of California, Berkeley), Lei Li (ByteDance AI Lab), Zehuan Yuan (Bytedance.Inc), Changhu Wang (ByteDance.Inc), and Ping Luo (The University of Hong Kong)</i>	
Learning View Selection for 3D Scenes	14459
<i>Yifan Sun (The University of Texas at Austin), Qixing Huang (The University of Texas at Austin), Dun-Yu Hsiao (Wormpex AI Research), Li Guan (Wormpex AI Research), and Gang Hua (Wormpex AI Research)</i>	
Multi-Person Implicit Reconstruction From a Single Image	14469
<i>Armin Mustafa (University of Surrey), Akin Caliskan (Center for Vision Speech and Signal Processing - University of Surrey), Lourdes Agapito (University College London), and Adrian Hilton (University of Surrey)</i>	

Neural Descent for Visual 3D Human Pose and Shape	14479
<i>Andrei Zanfir (Google), Eduard Gabriel Bazavan (Google), Mihai Zanfir (IMAR), William T. Freeman (Google), Rahul Sukthankar (Google), and Cristian Sminchisescu (Lund University)</i>	
SE-SSD: Self-Ensembling Single-Stage Object Detector From Point Cloud	14489
<i>Wu Zheng (The Chinese University of Hong Kong), Weiliang Tang (The Chinese University of Hong Kong), Li Jiang (The Chinese University of Hong Kong), and Chi-Wing Fu (The Chinese University of Hong Kong)</i>	
SCF-Net: Learning Spatial Contextual Features for Large-Scale Point Cloud Segmentation	14499
<i>Siqi Fan (School of Artificial Intelligence, University of Chinese Academy of Sciences; Institute of Automation, Chinese Academy of Sciences), Qiulei Dong (NLPR-IA-CAS), Fenghua Zhu (Institute of Automation, Chinese Academy of Sciences), Yisheng Lv (Institute of Automation, Chinese Academy of Sciences), Peijun Ye (Institute of Automation, Chinese Academy of Sciences), and Fei-Yue Wang (Institute of Automation, Chinese Academy of Sciences)</i>	
Equivariant Point Network for 3D Point Cloud Analysis	14509
<i>Haiwei Chen (University of Southern California), Shichen Liu (University of Southern California), Weikai Chen (Tencent America), Hao Li (Pinscreen / UC Berkeley), and Randall Hill (USC ICT)</i>	
DeepSurfels: Learning Online Appearance Fusion	14519
<i>Marko Mihajlovic (ETH Zurich), Silvan Weder (ETH Zürich), Marc Pollefeys (ETH Zurich / Microsoft), and Martin R. Oswald (ETH Zurich)</i>	
Efficient Deformable Shape Correspondence via Multiscale Spectral Manifold Wavelets Preservation	14531
<i>Ling Hu (Central South University), Qinsong Li (Central South University), Shengjun Liu (Central South University), and Xinru Liu (Central South University)</i>	
Efficient Initial Pose-Graph Generation for Global SfM	14541
<i>Daniel Barath (ETH Zürich), Dmytro Mishkin (Eötvös Loránd University), Iván Eichhardt (Czech Technical University in Prague), Ilia Shipachev (CVUT, CMP), and Jiří Matas (CMP CTU FEE)</i>	
AutoInt: Automatic Integration for Fast Neural Volume Rendering	14551
<i>David B. Lindell (Stanford University), Julien N. P. Martel (Stanford University), and Gordon Wetzstein (Stanford University)</i>	
Extreme Rotation Estimation Using Dense Correlation Volumes	14561
<i>Ruojin Cai (Cornell University), Bharath Hariharan (Cornell University), Noah Snavely (Cornell University and Google AI), and Hadar Averbuch-Elor (Cornell University)</i>	
A Quasiconvex Formulation for Radial Cameras	14571
<i>Carl Olsson (Lund University, Sweden), Viktor Larsson (ETH Zurich), and Fredrik Kahl (Chalmers)</i>	
ReAgent: Point Cloud Registration Using Imitation and Reinforcement Learning	14581
<i>Dominik Bauer (TU Wien), Timothy Patten (TU Wien), and Markus Vincze (TU Wien)</i>	

Monocular Depth Estimation via Listwise Ranking Using the Plackett-Luce Model	14590
<i>Julian Lienen (Paderborn University), Eyke Hüllermeier (University of Munich), Ralph Ewerth (TIB - Leibniz Information Center for Science and Technology), and Nils Nommensen (TIB - Leibniz Information Center for Science and Technology)</i>	
HVPR: Hybrid Voxel-Point Representation for Single-Stage 3D Object Detection	14600
<i>Jongyoun Noh (Yonsei University), Sanghoon Lee (Yonsei University), and Bumsub Ham (Yonsei University)</i>	
3DloUMatch: Leveraging IoU Prediction for Semi-Supervised 3D Object Detection	14610
<i>He Wang (Stanford University), Yezhen Cong (Tsinghua University), Or Litany (NVIDIA), Yue Gao (Tsinghua University), and Leonidas J. Guibas (Stanford University)</i>	
Multi-Label Activity Recognition Using Activity-Specific Features and Activity Correlations	14620
<i>Yanyi Zhang (Rutgers University), Xinyu Li (Amazon), and Ivan Marsic (Rutgers University)</i>	
LaPred: Lane-Aware Prediction of Multi-Modal Future Trajectories of Dynamic Agents	14631
<i>ByeoungDo Kim (Hanyang University), Seong Hyeon Park (Hanyang University), Seokhwan Lee (Hanyang university), Elbek Khoshimjonov (Hanyang University), Dongsuk Kum (Korea Advanced Institute of Science and Technology), Junsoo Kim (Hyundai Motor Company), Jeong Soo Kim (Hyundai Motor Company), and Jun Won Choi (Hanyang University)</i>	
Detecting Human-Object Interaction via Fabricated Compositional Learning	14641
<i>Zhi Hou (The University of Sydney), Baosheng Yu (The University of Sydney), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Xiaojiang Peng (Shenzhen Technology University), and Dacheng Tao (The University of Sydney)</i>	
Understanding the Robustness of Skeleton-Based Action Recognition Under Adversarial Attack.	14651
<i>He Wang (Leeds University), Feixiang He (the University of Leeds), Zhexi Peng (Zhejiang University), Tianjia Shao (Zhejiang University), Yong-Liang Yang (University of Bath), Kun Zhou (Zhejiang University), and David Hogg (University of Leeds)</i>	
Invisible Perturbations: Physical Adversarial Examples Exploiting the Rolling Shutter Effect	14661
<i>Athena Sayles (University of Wisconsin Madison), Ashish Hooda (UW-Madison), Mohit Gupta (University of Wisconsin-Madison, USA), Rahul Chatterjee (University of Wisconsin--Madison), and Earlence Fernandes (University of Wisconsin-Madison)</i>	
Bottom-Up Human Pose Estimation via Disentangled Keypoint Regression	14671
<i>Zigang Geng (University of Science and Technology of China), Ke Sun (University of Science and Technology of China), Bin Xiao (Microsoft), Zhaoxiang Zhang (Chinese Academy of Sciences, China), and Jingdong Wang (Microsoft)</i>	
Semi-Supervised 3D Hand-Object Poses Estimation With Interactions in Time	14682
<i>Shaowei Liu (UCSD), Hanwen Jiang (UCSD), Jiarui Xu (UCSD), Sifei Liu (NVIDIA), and Xiaolong Wang (UCSD)</i>	

Inverse Simulation: Reconstructing Dynamic Geometry of Clothed Humans via Optimal Control	14693
<i>Jingfan Guo (University of Minnesota), Jie Li (University of Minnesota), Rahul Narain (IIT DELHI), and Hyun Soo Park (The University of Minnesota)</i>	
Populating 3D Scenes by Learning Human-Scene Interaction	14703
<i>Mohamed Hassan (Max Planck Institute for Intelligent Systems), Partha Ghosh (Max Planck Inst. for Int. Sys), Joachim Tesch (Max Planck Institute for Intelligent Systems), Dimitrios Tzionas (Max Planck Institute for Intelligent Systems), and Michael J. Black (Max Planck Institute for Intelligent Systems)</i>	
Towards High Fidelity Face Relighting With Realistic Shadows	14714
<i>Andrew Hou (Michigan State University), Ze Zhang (Michigan State University), Michel Sarkis (Qualcomm Technologies Inc.), Ning Bi (Qualcomm), Yiyang Tong (Michigan State University), and Xiaoming Liu (Michigan State University)</i>	
VirFace: Enhancing Face Recognition via Unlabeled Shallow Data	14724
<i>Wenyu Li (Harbin Institute of Technology), Tianchu Guo (Artificial Intelligence Center, DAMO Academy, Alibaba Group, Hangzhou, China), Pengyu Li (Alibaba Group), Binghui Chen (Alibaba), Biao Wang (Alibaba), Wangmeng Zuo (Harbin Institute of Technology, China), and Lei Zhang (Hong Kong Polytechnic University, Hong Kong, China)</i>	
Birds of a Feather: Capturing Avian Shape Models From Images	14734
<i>Yufu Wang (University of Pennsylvania), Nikos Kolotouros (University of Pennsylvania), Kostas Daniilidis (University of Pennsylvania), and Marc Badger (University of Pennsylvania)</i>	
Unsupervised Pre-Training for Person Re-Identification	14745
<i>Dengpan Fu (University of Science and Technology of China), Dongdong Chen (Microsoft Cloud AI), Jianmin Bao (Microsoft Research Asia), Hao Yang (Microsoft Research Asia), Lu Yuan (Microsoft), Lei Zhang (Microsoft), Houqiang Li (University of Science and Technology of China), and Dong Chen (Microsoft Research Asia)</i>	
Indoor Lighting Estimation Using an Event Camera	14755
<i>Zehao Chen (Zhejiang University), Qian Zheng (Nanyang Technological University), Peisong Niu (Zhejiang University), Huajin Tang (Zhejiang University), and Gang Pan (Zhejiang University)</i>	
Checkerboard Context Model for Efficient Learned Image Compression	14766
<i>Dailan He (SenseTime), Yaoyan Zheng (SenseTime), Baocheng Sun (SenseTime), Yan Wang (SenseTime), and Hongwei Qin (Sensetime)</i>	
Neighbor2Neighbor: Self-Supervised Denoising From Single Noisy Images	14776
<i>Tao Huang (Renmin University of China), Songjiang Li (Huawei Noah's Ark Lab), Xu Jia (Dalian University of Technology), Huchuan Lu (Dalian University of Technology), and Jianzhuang Liu (Huawei Noah's Ark Lab)</i>	
From Rain Generation to Rain Removal	14786
<i>Hong Wang (Xi'an Jiaotong University), Zongsheng Yue (Xi'an Jiaotong University), Qi Xie (Xi'an Jiaotong University), Qian Zhao (Xi'an Jiaotong University), Yefeng Zheng (Tencent), and Deyu Meng (Xi'an Jiaotong University)</i>	

Rank-One Prior: Toward Real-Time Scene Recovery	14797
<i>Jun Liu (Northeast Normal University), Wen Liu (Wuhan University of Technology), Jianing Sun (Northeast Normal University), and Tiejong Zeng (The Chinese University of Hong Kong)</i>	
Robust Reflection Removal With Reflection-Free Flash-Only Cues	14806
<i>Chenyang Lei (HKUST) and Qifeng Chen (HKUST)</i>	
Multi-Stage Progressive Image Restoration	14816
<i>Syed Waqas Zamir (IIAI), Aditya Arora (IIAI), Salman Khan (Australian National University, (ANU)), Munawar Hayat (Monash University), Fahad Shahbaz Khan (MBZUAI), Ming-Hsuan Yang (University of California at Merced), and Ling Shao (Inception Institute of Artificial Intelligence)</i>	
Shape From Sky: Polarimetric Normal Recovery Under the Sky	14827
<i>Tomoki Ichikawa (Kyoto University), Matthew Purri (Rutgers University), Ryo Kawahara (Kyoto University), Shohei Nobuhara (Kyoto University), Kristin Dana (Rutgers University), and Ko Nishino (Kyoto University)</i>	
Cross-MPI: Cross-Scale Stereo for Image Super-Resolution Using Multiplane Images	14837
<i>Yuemei Zhou (Tsinghua University), Gaochang Wu (Northeastern University), Ying Fu (Beijing Institute of Technology), Kun Li (Tianjin University), and Yebin Liu (Tsinghua University)</i>	
Deep Perceptual Preprocessing for Video Coding	14847
<i>Aaron Chadha (iSIZE Technologies) and Yiannis Andreopoulos (iSize Technologies)</i>	
StyleMix: Separating Content and Style for Enhanced Data Augmentation	14857
<i>Minui Hong (Seoul National University), Jinwoo Choi (Seoul National University), and Gunhee Kim (Seoul National University)</i>	
Spoken Moments: Learning Joint Audio-Visual Representations From Video Descriptions	14866
<i>Mathew Monfort (MIT), SouYoung Jin (MIT), Alexander Liu (MIT), David Harwath (MIT CSAIL), Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research), James Glass (Massachusetts Institute of Technology), and Aude Oliva (MIT)</i>	
Spatially-Adaptive Pixelwise Networks for Fast Image Translation	14877
<i>Tamar Rott Shaham (Technion), Michaël Gharbi (Adobe Research), Richard Zhang (Adobe), Eli Shechtman (Adobe Research, US), and Tomer Michaeli (Technion)</i>	
No Frame Left Behind: Full Video Action Recognition	14887
<i>Xin Liu (Delft University of Technology), Silvia L. Pinteá (TU Delft), Fatemeh Karimi Nejadasl (TomTom), Olaf Booij (TomTom), and Jan C. van Gemert (Delft University of Technology)</i>	
Multiresolution Knowledge Distillation for Anomaly Detection	14897
<i>Mohammadreza Salehi (Sharif University of Technology), Niousha Sadjadi (Sharif University of Technology), Soroosh Baselizadeh (Sharif University of Technology), Mohammad H. Rohban (Sharif University of Technology), and Hamid R. Rabiee (Sharif University of Technology)</i>	
Convolutional Neural Network Pruning With Structural Redundancy Reduction	14908
<i>Zi Wang (UTK), Chengcheng Li (The University of Tennessee, Knoxville), and Xiangyang Wang (Sun Yat-sen University)</i>	

Representative Forgery Mining for Fake Face Detection	14918
<i>Chengrui Wang (Beijing University of Posts and Telecommunications) and Weihong Deng (Beijing University of Posts and Telecommunications)</i>	
Neural Prototype Trees for Interpretable Fine-Grained Image Recognition	14928
<i>Meike Nauta (University of Twente), Ron van Bree (University of Twente), and Christin Seifert (University of Duisburg-Essen)</i>	
Relevance-CAM: Your Model Already Knows Where To Look	14939
<i>Jeong Ryong Lee (Yonsei university), Sewon Kim (Yonsei University), Inyong Park (Yonsei University), Taejoon Eo (Yonsei University), and Dosik Hwang (Yonsei University)</i>	
Adaptive Cross-Modal Prototypes for Cross-Domain Visual-Language Retrieval	14949
<i>Yang Liu (University of Oxford), Qingchao Chen (University of Oxford), and Samuel Albanie (University of Oxford)</i>	
Efficient Object Embedding for Spliced Image Retrieval	14960
<i>Bor-Chun Chen (Facebook AI), Zuxuan Wu (UMD), Larry S. Davis (University of Maryland), and Ser-Nam Lim (Facebook AI)</i>	
Generative PointNet: Deep Energy-Based Learning on Unordered Point Sets for 3D Generation, Reconstruction and Classification	14971
<i>Jianwen Xie (Baidu Research), Yifei Xu (UCLA), Zilong Zheng (UCLA), Song-Chun Zhu (UCLA), and Ying Nian Wu (University of California, Los Angeles)</i>	
Anycost GANs for Interactive Image Synthesis and Editing	14981
<i>Ji Lin (MIT), Richard Zhang (Adobe), Frieder Ganz (Adobe), Song Han (MIT), and Jun-Yan Zhu (Carnegie Mellon University)</i>	
Ensembling With Deep Generative Views	14992
<i>Lucy Chai (MIT), Jun-Yan Zhu (Carnegie Mellon University), Eli Shechtman (Adobe Research, US), Phillip Isola (MIT), and Richard Zhang (Adobe)</i>	
Continuous Face Aging via Self-Estimated Residual Age Embedding	15003
<i>Zeqi Li (ModiFace Inc.), Ruowei Jiang (ModiFace Inc.), and Parham Aarabi (ModiFace Inc.)</i>	
ReMix: Towards Image-to-Image Translation With Limited Data	15013
<i>Jie Cao (Institute of Automation, Chinese Academy of Sciences), Luanxuan Hou (Institute of Automation, Chinese Academy of Sciences), Ming-Hsuan Yang (University of California at Merced), Ran He (Institute of Automation, Chinese Academy of Sciences), and Zhenan Sun (Chinese of Academy of Sciences)</i>	
Unbalanced Feature Transport for Exemplar-Based Image Translation	15023
<i>Fangneng Zhan (Nanyang Technological University), Yingchen Yu (Nanyang Technological University), Kaiwen Cui (Nanyang Technology University), Gongjie Zhang (Nanyang Technological University), Shijian Lu (Nanyang Technological University), Jianxiong Pan (Alibaba Group), Changgong Zhang (Alibaba Group), Feiyang Ma (alibaba), Xuansong Xie (Alibaba), and Chunyan Miao (NTU)</i>	

Pose-Guided Human Animation From a Single Image in the Wild	15034
<i>Jae Shin Yoon (U of Minnesota), Lingjie Liu (Max Planck Institute for Informatics), Vladislav Golyanik (MPI for Informatics), Kripasindhu Sarkar (Max Plank Institute for Informatics), Hyun Soo Park (The University of Minnesota), and Christian Theobalt (MPI Informatik)</i>	
Context-Aware Layout to Image Generation With Enhanced Object Appearance	15044
<i>Sen He (University of Surrey), Wentong Liao (Leibniz University Hannover), Michael Ying Yang (University of Twente), Yongxin Yang (University of Surrey), Yi-Zhe Song (University of Surrey), Bodo Rosenhahn (Leibniz University Hannover), and Tao Xiang (University of Surrey)</i>	
SetVAE: Learning Hierarchical Composition for Generative Modeling of Set-Structured Data	15054
<i>Jinwoo Kim (KAIST), Jaehoon Yoo (KAIST), Juho Lee (KAIST), and Seunghoon Hong (KAIST)</i>	
Are Labels Always Necessary for Classifier Accuracy Evaluation?	15064
<i>Weijian Deng (Australian National University) and Liang Zheng (Australian National University)</i>	
Graph-Based High-Order Relation Discovery for Fine-Grained Recognition	15074
<i>Yifan Zhao (Beihang University), Ke Yan (Tencent), Feiyue Huang (Tencent), and Jia Li (Beihang University)</i>	
Long-Tailed Multi-Label Visual Recognition by Collaborative Training on Uniform and Re-Balanced Samplings	15084
<i>Hao Guo (University of South Carolina) and Song Wang (University of South Carolina)</i>	
SimPLE: Similar Pseudo Label Exploitation for Semi-Supervised Classification	15094
<i>Zijian Hu (University of Southern California), Zhengyu Yang (University of Southern California), Xuefeng Hu (University of Southern California), and Ram Nevatia (U of Southern California)</i>	
Cluster-Wise Hierarchical Generative Model for Deep Amortized Clustering	15104
<i>Huafeng Liu (Beijing Jiaotong University), Jiaqi Wang (Beijing Jiaotong University), and Liping Jing (Beijing Jiaotong University)</i>	
From Semantic Categories to Fixations: A Novel Weakly-Supervised Visual-Auditory Saliency Detection Approach	15114
<i>Guotao Wang (Beihang University), Chenglizhao Chen (Qingdao University), Deng-Ping Fan (Inception Institute of Artificial Intelligence), Aimin Hao (Beihang University), and Hong Qin (Stony Brook University)</i>	
Gaussian Context Transformer	15124
<i>Dongsheng Ruan (Zhejiang University), Daiyin Wang (Zhejiang University), Yuan Zheng (Zhejiang University), Nenggan Zheng (Zhejiang University), and Min Zheng (Zhejiang University)</i>	
FP-NAS: Fast Probabilistic Neural Architecture Search	15134
<i>Zhicheng Yan (Facebook AI), Xiaoliang Dai (Facebook), Peizhao Zhang (Facebook), Yuandong Tian (Facebook), Bichen Wu (Facebook Research), and Matt Feiszli (Facebook Research)</i>	
Time Adaptive Recurrent Neural Network	15144
<i>Anil Kag (Boston University) and Venkatesh Saligrama (Boston University)</i>	

Deep Lesion Tracker: Monitoring Lesions in 4D Longitudinal Imaging Studies	15154
<i>Jinzheng Cai (PAII Inc.), Youbao Tang (PAII Inc.), Ke Yan (PAII Inc.), Adam P. Harrison (PAII Inc.), Jing Xiao (Ping An Insurance, (Group) Company of China), Gigin Lin (Chang Gung Memorial Hospital), and Le Lu (PAII Inc.)</i>	
Reciprocal Landmark Detection and Tracking With Extremely Few Annotations	15165
<i>Jianzhe Lin (University of British Columbia), Ghazal Sahebzamani (The University of British Columbia), Christina Luong (Vancouver General Hospital), Fatemeh Taheri Dezaki (University of British Columbia), Mohammad Jafari (University of British Columbia), Purang Abolmaesumi (The Univ. of British Columbia), and Teresa Tsang (Vancouver General Hospital)</i>	
LightTrack: Finding Lightweight Neural Networks for Object Tracking via One-Shot Architecture Search	15175
<i>Bin Yan (Dalian University of Technology), Houwen Peng (Microsoft Research), Kan Wu (Sun Yat-Sen University), Dong Wang (Dalian University of Technology), Jianlong Fu (Microsoft Research), and Huchuan Lu (Dalian University of Technology)</i>	
TesseTrack: End-to-End Learnable Multi-Person Articulated 3D Pose Tracking	15185
<i>N Dinesh Reddy (Carnegie Mellon University), Laurent Guigues (Amazon), Leonid Pishchulin (Amazon Go), Jayan Eledath (Amazon), and Srinivasa G. Narasimhan (Carnegie Mellon University, USA)</i>	
Learning Optical Flow From Still Images	15196
<i>Filippo Aleotti (University of Bologna), Matteo Poggi (University of Bologna), and Stefano Mattoccia (University of Bologna)</i>	
Towards Robust Classification Model by Counterfactual and Invariant Data Generation	15207
<i>Chun-Hao Chang (University of Toronto), George Alexandru Adam (University of Toronto), and Anna Goldenberg</i>	
StablePose: Learning 6D Object Poses From Geometrically Stable Patches	15217
<i>Yifei Shi (National University of Defense Technology), Junwen Huang (National University of Defense Technology), Xin Xu (National University of Defense Technology), Yifan Zhang (National University of Defense Technology), and Kai Xu (National University of Defense Technology)</i>	
The Translucent Patch: A Physical and Universal Attack on Object Detectors	15227
<i>Alon Zolfi (Ben-Gurion University of the Negev), Moshe Kravchik (Ben-Gurion University of the Negev), Yuval Elovici (Ben-Gurion University of the Negev), and Asaf Shabtai (Ben-Gurion University of the Negev)</i>	
Dynamic Weighted Learning for Unsupervised Domain Adaptation	15237
<i>Ni Xiao (Chongqing University) and Lei Zhang (Chongqing University)</i>	
DRANet: Disentangling Representation and Adaptation Networks for Unsupervised Cross-Domain Adaptation	15247
<i>Seunghun Lee (DGIST), Sunghyun Cho (POSTECH), and Sunghoon Im (DGIST)</i>	
Natural Adversarial Examples	15257
<i>Dan Hendrycks (UC Berkeley), Kevin Zhao (University of Washington), Steven Basart (University of Chicago), Jacob Steinhardt (UC Berkeley), and Dawn Song (UC Berkeley)</i>	

Fast End-to-End Learning on Protein Surfaces	15267
<i>Freyr Sverrisson (EPFL), Jean Feydy (Imperial College London), Bruno E. Correia (EPFL), and Michael M. Bronstein (Imperial College London / Twitter)</i>	
Rectification-Based Knowledge Retention for Continual Learning	15277
<i>Pravendra Singh (Indian Institute of Technology Kanpur), Pratik Mazumder (Indian Institute of Technology, Kanpur), Piyush Rai (IIT Kanpur), and Vinay P. Namboodiri (University of Bath)</i>	
Cross-Domain Similarity Learning for Face Recognition in Unseen Domains	15287
<i>Masoud Faraki (NEC Labs), Xiang Yu (NEC Labs), Yi-Hsuan Tsai (NEC Labs America), Yumin Suh (NEC Labs America), and Manmohan Chandraker (UC San Diego)</i>	
Sequence-to-Sequence Contrastive Learning for Text Recognition	15297
<i>Aviad Aberdam (Technion), Ron Litman (Amazon), Shahar Tsiper (Amazon), Oron Anshel (AWS), Ron Slossberg (AWS), Shai Mazor (Amazon), R. Manmatha (Amazon), and Pietro Perona (California Institute of Technology)</i>	
MOOD: Multi-Level Out-of-Distribution Detection	15308
<i>Ziqian Lin (University of Wisconsin–Madison), Sreya Dutta Roy (University of Wisconsin-Madison), and Yixuan Li (University of Wisconsin-Madison)</i>	
DeepVideoMVS: Multi-View Stereo on Video With Recurrent Spatio-Temporal Fusion	15319
<i>Arda Düzçeker (ETH Zürich), Silvano Galliani (Microsoft), Christoph Vogel (Microsoft MixedReality&AI), Pablo Speciale (Microsoft), Mihai Dusmanu (ETH Zurich), and Marc Pollefeys (ETH Zurich / Microsoft)</i>	
Boundary IoU: Improving Object-Centric Image Segmentation Evaluation	15329
<i>Bowen Cheng (UIUC), Ross Girshick (FAIR), Piotr Dollár (FAIR), Alexander C. Berg (University of North Carolina, USA), and Alexander Kirillov (Facebook AI Reserach)</i>	
A2-FPN: Attention Aggregation Based Feature Pyramid Network for Instance Segmentation	15338
<i>Miao Hu (Department of Electronic Engineering, Tsinghua University), Yali Li (Tsinghua University), Lu Fang (Tsinghua University), and Shengjin Wang (Tsinghua University)</i>	
SSLLayout360: Semi-Supervised Indoor Layout Estimation From 360° Panorama	15348
<i>Phi Vu Tran (Flyreel)</i>	
Complete & Label: A Domain Adaptation Approach to Semantic Segmentation of LiDAR Point Clouds	15358
<i>Li Yi (Google Research), Boqing Gong (Google), and Thomas Funkhouser (Google Research)</i>	
Improved Image Matting via Real-Time User Clicks and Uncertainty Estimation	15369
<i>Tianyi Wei (University of Science and Technology of China), Dongdong Chen (Microsoft Cloud AI), Wenbo Zhou (University of Science and Technology of China), Jing Liao (City University of Hong Kong), Hanqing Zhao (University of Science and Technology of China), Weiming Zhang (University of Science and Technology of China), and Nenghai Yu (University of Science and Technology of China)</i>	
Self-Supervised Augmentation Consistency for Adapting Semantic Segmentation	15379
<i>Nikita Araslanov (TU Darmstadt) and Stefan Roth (TU Darmstadt)</i>	

Few-Shot Object Detection via Classification Refinement and Distractor Retreatment	15390
<i>Yiting Li (NUS), Haiyue Zhu (Singapore Institute of Manufacturing Technology), Yu Cheng (National University of Singapore), Wenxin Wang (National University of Singapore), Chek Sing Teo (Singapore Institute of Manufacturing Technology), Cheng Xiang (National University of Singapore), Prahlad Vadakkepat (NUS), and Tong Heng Lee (National University of Singapore)</i>	
Counterfactual Zero-Shot and Open-Set Visual Recognition	15399
<i>Zhongqi Yue (Nanyang Technological University), Tan Wang (Nanyang Technological University), Qianru Sun (Singapore Management University), Xian-Sheng Hua (Damo Academy, Alibaba Group), and Hanwang Zhang (Nanyang Technological University)</i>	
Learning Deep Latent Variable Models by Short-Run MCMC Inference With Optimal Transport Correction	15410
<i>Dongsheng An (Stony Brook University), Jianwen Xie (Baidu Research), and Ping Li (Baidu Research)</i>	
Learning Normal Dynamics in Videos With Meta Prototype Network	15420
<i>Hui Lv (Nanjing University of Science and Technology), Chen Chen (University of North Carolina at Charlotte), Zhen Cui (Nanjing University of Science and Technology), Chunyan Xu (Nanjing University of Science and Technology), Yong Li (Nanjing University of Science and Technology), and Jian Yang (Nanjing University of Science and Technology)</i>	
MotionRNN: A Flexible Model for Video Prediction With Spacetime-Varying Motions	15430
<i>Haixu Wu (Tsinghua University), Zhiyu Yao (Tsinghua University), Jianmin Wang (Tsinghua University), and Mingsheng Long (Tsinghua University, China)</i>	
Learning To Recommend Frame for Interactive Video Object Segmentation in the Wild	15440
<i>Zhaoyuan Yin (Hunan University), Jia Zheng (Manycore, (Kujiale)), Weixin Luo (Meituan), Shenhan Qian (ShanghaiTech University), Hanling Zhang (Hunan University), and Shenghua Gao (ShanghaiTech University)</i>	
Reciprocal Transformations for Unsupervised Video Object Segmentation	15450
<i>Sucheng Ren (South China University of Technology), Wenxi Liu (Fuzhou University), Yongtuo Liu (South China University of Technology), Haoxin Chen (South China University of Technology), Guoqiang Han (South China University of Technology), and Shengfeng He (South China University of Technology)</i>	
RSTNet: Captioning With Adaptive Attention on Visual and Non-Visual Words	15460
<i>Xuying Zhang (Xiamen University), Xiaoshuai Sun (Xiamen University), Yunpeng Luo (Xiamen University), Jiayi Ji (Xiamen University), Yiyi Zhou (Xiamen University), Yongjian Wu (Tencent Technology, (Shanghai) Co., Ltd), Feiyue Huang (Tencent), and Rongrong Ji (Xiamen University, China)</i>	
Revamping Cross-Modal Recipe Retrieval With Hierarchical Transformers and Self-Supervised Learning	15470
<i>Amaia Salvador (Amazon), Erhan Gundogdu (Amazon), Loris Bazzani (Amazon), and Michael Donoser (Amazon)</i>	

Visually Informed Binaural Audio Generation without Binaural Audios	15480
<i>Xudong Xu (the Chinese University of Hong Kong), Hang Zhou (The Chinese University of Hong Kong), Ziwei Liu (Nanyang Technological University), Bo Dai (Nanyang Technological University), Xiaogang Wang (Chinese University of Hong Kong, Hong Kong), and Dahua Lin (The Chinese University of Hong Kong)</i>	
VisualVoice: Audio-Visual Speech Separation With Cross-Modal Consistency	15490
<i>Ruohan Gao (Stanford University) and Kristen Grauman (Facebook AI Research & UT Austin)</i>	
Encoder Fusion Network With Co-Attention Embedding for Referring Image Segmentation	15501
<i>Guang Feng (Dalian University of Technology), Zhiwei Hu (Dalian University of Technology), Lihe Zhang (Dalian University of Technology), and Huchuan Lu (Dalian University of Technology)</i>	
Semantic Audio-Visual Navigation	15511
<i>Changan Chen (University of Texas at Austin), Ziad Al-Halah (UT Austin), and Kristen Grauman (Facebook AI Research & UT Austin)</i>	
Bridge To Answer: Structure-Aware Graph Interaction Network for Video Question Answering	15521
<i>Jungin Park (Yonsei University), Jiyoung Lee (Yonsei University), and Kwanghoon Sohn (Yonsei Univ.)</i>	
Projecting Your View Attentively: Monocular Road Scene Layout Estimation via Cross-View Transformation	15531
<i>Weixiang Yang (Fuzhou University), Qi Li (Fuzhou University), Wenxi Liu (Fuzhou University), Yuanlong Yu (Fuzhou University), Yuexin Ma (ShanghaiTech University), Shengfeng He (South China University of Technology), and Jia Pan (The University of Hong Kong)</i>	
Toward Accurate and Realistic Outfits Visualization With Attention to Details	15541
<i>Kedan Li (University of Illinois at Urbana-Champaign), Min Jin Chong (Univeristy of Illinois at Urbana-Champaign), Jeffrey Zhang (UIUC), and Jingen Liu (JD AI Research)</i>	
Interpretable Social Anchors for Human Trajectory Forecasting in Crowds	15551
<i>Parth Kothari (Ecole Polytechnique Fédérale de Lausanne), Brian Sifringer (EPFL), and Alexandre Alahi (EPFL)</i>	

Session 12

Deep Polarization Imaging for 3D Shape and SVBRDF Acquisition	15562
<i>Valentin Deschaintre (Imperial College London), Yiming Lin (Imperial College London), and Abhijeet Ghosh (Imperial College London)</i>	
Self-Point-Flow: Self-Supervised Scene Flow Estimation From Point Clouds With Optimal Transport and Random Walk	15572
<i>Ruibo Li (Nanyang Technological University), Guosheng Lin (Nanyang Technological University), and Lihua Xie (Nanyang Technological University)</i>	

Exploring Data-Efficient 3D Scene Understanding With Contrastive Scene Contexts	15582
<i>Ji Hou (Technical University of Munich), Benjamin Graham (Facebook Research), Matthias Nießner (Technical University of Munich), and Saining Xie (Facebook AI Research)</i>	
NeuralRecon: Real-Time Coherent 3D Reconstruction From Monocular Video	15593
<i>Jiaming Sun (SenseTime), Yiming Xie (SenseTime), Linghao Chen (Zhejiang University), Xiaowei Zhou (Zhejiang University), and Hujun Bao (Zhejiang University)</i>	
FrameExit: Conditional Early Exiting for Efficient Video Recognition	15603
<i>Amir Ghodrati (Qualcomm AI Research), Babak Ehteshami Bejnordi (Qualcomm AI Reseach), and Amirhossein Habibian (Qualcomm AI Research)</i>	
Inverting Generative Adversarial Renderer for Face Reconstruction	15614
<i>Jingtian Piao (CUHK MMLab), Keqiang Sun (Chinese University of Hong Kong), Quan Wang (SenseTime), Kwan-Yee Lin (SenseTime Research), and Hongsheng Li (Chinese University of Hong Kong)</i>	
Spherical Confidence Learning for Face Recognition	15624
<i>Shen Li (National University of Singapore), Jianqing Xu (tencent), Xiaqing Xu (Aibee), Pengcheng Shen (Tencent), Shaoxin Li (Tencent), and Bryan Hooi (National University of Singapore)</i>	
Event-Based Bispectral Photometry Using Temporally Modulated Illumination	15633
<i>Tsuyoshi Takatani (University of Tsukuba), Yuzuha Ito (University of Tsukuba), Ayaka Ebisu (University of Tsukuba), Yinqiang Zheng (The University of Tokyo), and Takahito Aoto (University of Tsukuba)</i>	
AdderSR: Towards Energy Efficient Image Super-Resolution	15643
<i>Dehua Song (Huawei Noah's Ark Lab), Yunhe Wang (Huawei Technologies), Hanting Chen (Peking University), Chang Xu (University of Sydney), Chunjing Xu (Huawei Noah's Ark Lab), and Dacheng Tao (The University of Sydney)</i>	
Diversifying Sample Generation for Accurate Data-Free Quantization	15653
<i>Xiangguo Zhang (Beihang University), Haotong Qin (Beihang University), Yifu Ding (Beihang University), Ruihao Gong (SenseTime), Qinghua Yan (Beihang University), Renshuai Tao (Beihang University), Yuhang Li (Yale University), Fengwei Yu (SenseTime Research), and Xianglong Liu (Beihang University)</i>	
How Privacy-Preserving Are Line Clouds? Recovering Scene Details From 3D Lines	15663
<i>Kunal Chelani (Chalmers University of Technology), Fredrik Kahl (Chalmers), and Torsten Sattler (Czech Technical University in Prague)</i>	
Unsupervised Learning of Depth and Depth-of-Field Effect From Natural Images With Aperture Rendering Generative Adversarial Networks	15674
<i>Takuhiro Kaneko (NTT Corporation)</i>	
Stylized Neural Painting	15684
<i>Zhengxia Zou (University of Michigan), Tianyang Shi (NetEase Fuxi AI Lab), Shuang Qiu (University of Michigan), Yi Yuan (NetEase Fuxi AI Lab), and Zhenwei Shi (Beihang University)</i>	

Permute, Quantize, and Fine-Tune: Efficient Compression of Neural Networks	15694
<i>Julieta Martinez (Uber ATG), Jashan Shewakramani (University of Waterloo), Ting Wei Liu (University of Waterloo), Ioan Andrei Bârsan (Uber ATG, University of Toronto), Wenyuan Zeng (Uber ATG, University of Toronto), and Raquel Urtasun (Uber ATG)</i>	
Amalgamating Knowledge From Heterogeneous Graph Neural Networks	15704
<i>Yongcheng Jing (The University of Sydney), Yiding Yang (Stevens Institute of Technology), Xinchao Wang (National University of Singapore), Mingli Song (Zhejiang University), and Dacheng Tao (The University of Sydney)</i>	
XProtoNet: Diagnosis in Chest Radiography With Global and Local Explanations	15714
<i>Eunji Kim (Seoul National University), Siwon Kim (Seoul National University), Minji Seo (Seoul National University), and Sungroh Yoon (Seoul National University)</i>	
LQF: Linear Quadratic Fine-Tuning	15724
<i>Alessandro Achille (Amazon Web Services), Aditya Golatkar (University of California, Los Angeles), Avinash Ravichandran (Amazon), Marzia Polito (Amazon Web Services), and Stefano Soatto (UCLA)</i>	
DECOR-GAN: 3D Shape Detailization by Conditional Refinement	15735
<i>Zhiqin Chen (Simon Fraser University), Vladimir G. Kim (Adobe), Matthew Fisher (Adobe Research), Noam Aigerman (Adobe), Hao Zhang (Simon Fraser University), and Siddhartha Chaudhuri (Adobe Research)</i>	
Exploring Simple Siamese Representation Learning	15745
<i>Xinlei Chen (FAIR) and Kaiming He (Facebook AI Research)</i>	
Camera Pose Matters: Improving Depth Prediction by Mitigating Pose Distribution Bias	15754
<i>Yunhan Zhao (University of California, Irvine), Shu Kong (Carnegie Mellon University), and Charless Fowlkes (UC Irvine)</i>	
DANNet: A One-Stage Domain Adaptation Network for Unsupervised Nighttime Semantic Segmentation	15764
<i>Xinyi Wu (University of South Carolina), Zhenyao Wu (University of South Carolina), Hao Guo (University of South Carolina), Lili Ju (University of South Carolina), and Song Wang (University of South Carolina)</i>	
OTCE: A Transferability Metric for Cross-Domain Cross-Task Representations	15774
<i>Yang Tan (Tsinghua-Berkeley Shenzhen Institute, Tsinghua University), Yang Li (Tsinghua-Berkeley Shenzhen Institute, Tsinghua University), and Shao-Lun Huang (Tsinghua-Berkeley Shenzhen Institute, Tsinghua University)</i>	
Learning the Best Pooling Strategy for Visual Semantic Embedding	15784
<i>Jiacheng Chen (USC), Hexiang Hu (University of Southern California), Hao Wu (Fudan University), Yuning Jiang (Bytedance AI Lab), and Changhu Wang (ByteDance.Inc)</i>	
Divide-and-Conquer for Lane-Aware Diverse Trajectory Prediction	15794
<i>Sriram Narayanan (NEC Laboratories America), Ramin Moslemi (NEC Laboratories America, Inc.), Francesco Pittaluga (NEC Laboratories America), Buyu Liu (NEC Labs), and Manmohan Chandraker (UC San Diego)</i>	

<p>Magic Layouts: Structural Prior for Component Detection in User Interface Designs 15804 <i>Dipu Manandhar (University of Surrey), Hailin Jin (Adobe Research), and John Collomosse (Adobe Research)</i></p>	
<p>Dense Label Encoding for Boundary Discontinuity Free Rotation Detection 15814 <i>Xue Yang (Shanghai Jiao Tong University), Liping Hou (University of Chinese Academy of Sciences), Yue Zhou (Shanghai Jiao Tong University), Wentao Wang (Shanghai Jiao Tong University), and Junchi Yan (Shanghai Jiao Tong University)</i></p>	
<p>MetaHTR: Towards Writer-Adaptive Handwritten Text Recognition 15825 <i>Ayan Kumar Bhunia (University of Surrey), Shuvojit Ghose (Institute of Engineering and Management), Amandeep Kumar (West Bengal University of Technology), Pinaki Nath Chowdhury (University of Surrey), Aneeshan Sain (University of Surrey), and Yi-Zhe Song (University of Surrey)</i></p>	
<p>PSRR-MaxpoolNMS: Pyramid Shifted MaxpoolNMS With Relationship Recovery 15835 <i>Tianyi Zhang (Institute for Infocomm Research, A*star), Jie Lin (Institute for Infocomm Research, I2R, Singapore), Peng Hu (College of Computer Science, Sichuan University), Bin Zhao (Nil), and Mohamed M. Sabry Sabry Aly (Nanyang Technological University)</i></p>	
<p>End-to-End Object Detection With Fully Convolutional Network 15844 <i>Jianfeng Wang (MEGVII Technology), Lin Song (Xi'an Jiaotong University), Zeming Li (Megvii(Face++) Inc), Hongbin Sun (Xi'an Jiaotong University), Jian Sun (Megvii Technology), and Nanning Zheng (Xi'an Jiaotong University)</i></p>	
<p>PointDSC: Robust Point Cloud Registration Using Deep Spatial Consistency 15854 <i>Xuyang Bai (HKUST), Zixin Luo (HKUST), Lei Zhou (HKUST), Hongkai Chen (HKUST), Lei Li (Ecole Polytechnique), Zeyu Hu (Hong Kong University of Science and Technology), Hongbo Fu (City University of Hong Kong), and Chiew-Lan Tai (Hong Kong University of Science & Technology)</i></p>	
<p>Wide-Depth-Range 6D Object Pose Estimation in Space 15865 <i>Yinlin Hu (EPFL), Sébastien Speierer (EPFL), Wenzel Jakob (EPFL), Pascal Fua (EPFL, Switzerland), and Mathieu Salzmann (EPFL)</i></p>	
<p>Fostering Generalization in Single-View 3D Reconstruction by Learning a Hierarchy of Local and Global Shape Priors 15875 <i>Jan Bechtold (Bosch Center for Artificial Intelligence), Maxim Tatarchenko (University of Freiburg), Volker Fischer (Bosch Center for Artificial Intelligence), and Thomas Brox (University of Freiburg)</i></p>	
<p>View-Guided Point Cloud Completion 15885 <i>Xuancheng Zhang (Tsinghua University), Yutong Feng (Tsinghua University), Siqi Li (Tsinghua University), Changqing Zou (Huawei Technologies Canada), Hai Wan (Tsinghua University), Xibin Zhao (Tsinghua University), Yandong Guo (MSR), and Yue Gao (Tsinghua University)</i></p>	
<p>Regularization Strategy for Point Cloud via Rigidly Mixed Sample 15895 <i>Dogyoon Lee (Yonsei University), Jaeha Lee (Yonsei University), Junhyeop Lee (Yonsei University), Hyeongmin Lee (Yonsei University), Minhyeok Lee (Yonsei University), Sungmin Woo (Yonsei University), and Sangyoun Lee (Yonsei University)</i></p>	

PWCLO-Net: Deep LiDAR Odometry in 3D Point Clouds Using Hierarchical Embedding Mask Optimization	15905
<i>Guangming Wang (Shanghai Jiao Tong University), Xinrui Wu (Shanghai Jiao Tong University), Zhe Liu (University of Cambridge), and Hesheng Wang (SJTU)</i>	
Co-Attention for Conditioned Image Matching	15915
<i>Olivia Wiles (DeepMind), Sébastien Ehrhardt (University of Oxford), and Andrew Zisserman (University of Oxford)</i>	
A Dual Iterative Refinement Method for Non-Rigid Shape Matching	15925
<i>Rui Xiang (University of California, Irvine), Rongjie Lai (Rensselaer Polytechnic Institute), and Hongkai Zhao (Duke University)</i>	
NeRD: Neural 3D Reflection Symmetry Detector	15935
<i>Yichao Zhou (UC Berkeley), Shichen Liu (University of Southern California), and Yi Ma (UC Berkeley)</i>	
Deep Lucas-Kanade Homography for Multimodal Image Alignment	15945
<i>Yiming Zhao (Worcester Polytechnic Institute), Xinming Huang (Worcester Polytechnic Institute), and Ziming Zhang (Worcester Polytechnic Institute)</i>	
DeepI2P: Image-to-Point Cloud Registration via Deep Classification	15955
<i>Jiaxin Li (Bytedance, TikTok) and Gim Hee Lee (National University of Singapore)</i>	
NormalFusion: Real-Time Acquisition of Surface Normals for High-Resolution RGB-D Scanning	15965
<i>Hyunho Ha (KAIST), Joo Ho Lee (University of Tuebingen), Andreas Meuleman (KAIST), and Min H. Kim (KAIST)</i>	
LASR: Learning Articulated Shape Reconstruction From a Monocular Video	15975
<i>Gengshan Yang (Carnegie Mellon University), Deqing Sun (Google), Varun Jampani (Google), Daniel Vlasic (Google), Forrester Cole (Google Research), Huiwen Chang (Google), Deva Ramanan (Carnegie Mellon University), William T. Freeman (Google), and Ce Liu (Google)</i>	
Mirror3D: Depth Refinement for Mirror Surfaces	15985
<i>Jiaqi Tan (Simon Fraser University), Weijie Lin (Simon Fraser University), Angel X. Chang (Simon Fraser University), and Manolis Savva (Simon Fraser University)</i>	
To the Point: Efficient 3D Object Detection in the Range Image With Graph Convolution Kernels	15995
<i>Yuning Chai (Alphabet), Pei Sun (Waymo), Jiquan Ngiam (Google Brain), Weiyue Wang (Waymo), Benjamin Caine (Google), Vijay Vasudevan (Google Brain), Xiao Zhang (Google), and Dragomir Anguelov (Waymo)</i>	
CoLA: Weakly-Supervised Temporal Action Localization With Snippet Contrastive Learning	16005
<i>Can Zhang (Peking University), Meng Cao (Peking University), Dongming Yang (peking university), Jie Chen (Peking University), and Yuexian Zou (Peking University)</i>	
MoViNets: Mobile Video Networks for Efficient Video Recognition	16015
<i>Dan Kondratyuk (Google), Liangzhe Yuan (Google Research), Yandong Li (University of Central Florida), Li Zhang (Google), Mingxing Tan (Google Brain), Matthew Brown (Google), and Boqing Gong (Google)</i>	

Few-Shot Transformation of Common Actions Into Time and Space	16026
<i>Pengwan Yang (Peking University), Pascal Mettes (University of Amsterdam), and Cees G. M. Snoek (University of Amsterdam)</i>	
Learning Asynchronous and Sparse Human-Object Interaction in Videos	16036
<i>Romero Morais (Deakin University), Vuong Le (Deakin University), Svetha Venkatesh (Deakin University), and Truyen Tran (Deakin University)</i>	
Adversarial Imaging Pipelines	16046
<i>Buu Phan (Algolux), Fahim Mannan (Algolux), and Felix Heide (Princeton / Algolux)</i>	
Adversarial Laser Beam: Effective Physical-World Attack to DNNs in a Blink	16057
<i>Ranjie Duan (Swinburne University of Technology), Xiaofeng Mao (Alibaba Group), A. K. Qin (Swinburne University of Technology), Yuefeng Chen (Alibaba Group), Shaokai Ye (EPFL), Yuan He (Alibaba Group), and Yun Yang (Swinburne University of Technology)</i>	
ViPNAS: Efficient Video Pose Estimation via Neural Architecture Search	16067
<i>Lumin Xu (The Chinese University of Hong Kong), Yingda Guan (SenseTime), Sheng Jin (SenseTime Research), Wentao Liu (SenseTime), Chen Qian (SenseTime), Ping Luo (The University of Hong Kong), Wanli Ouyang (The University of Sydney), and Xiaogang Wang (Chinese University of Hong Kong, Hong Kong)</i>	
SCALE: Modeling Clothed Humans with a Surface Codec of Articulated Local Elements	16077
<i>Qianli Ma (Max Planck Institute for Intelligent Systems), Shunsuke Saito (Facebook), Jinlong Yang (Max Planck Institute for Intelligent Systems), Siyu Tang (ETH Zurich), and Michael J. Black (Max Planck Institute for Intelligent Systems)</i>	
Probabilistic 3D Human Shape and Pose Estimation From Multiple Unconstrained Images in the Wild	16089
<i>Akash Sengupta (University of Cambridge), Ignas Budvytis (Department of Engineering, University of Cambridge), and Roberto Cipolla (University of Cambridge)</i>	
Graph Stacked Hourglass Networks for 3D Human Pose Estimation	16100
<i>Tianhan Xu (Osaka University) and Wataru Takano (Osaka University)</i>	
High-Fidelity and Arbitrary Face Editing	16110
<i>Yue Gao (Peking University), Fangyun Wei (Microsoft Research Asia), Jianmin Bao (Microsoft Research Asia), Shuyang Gu (University of Science and Technology of China), Dong Chen (Microsoft Research Asia), Fang Wen (Microsoft Research Asia), and Zhouhui Lian (Peking University)</i>	
IronMask: Modular Architecture for Protecting Deep Face Template	16120
<i>Sunpill Kim (Hanyang University), Yunseong Jeong (Jeonbuk National University), Jinsu Kim (Samsung Electronics), Jungkon Kim (Samsung Electronics), Hyung Tae Lee (Jeonbuk National University), and Jae Hong Seo (Hanyang university)</i>	
Meta-Mining Discriminative Samples for Kinship Verification	16130
<i>Wanhua Li (Tsinghua University), Shiwei Wang (Beijing University of Post and Telecommunication), Jiwen Lu (Tsinghua University), Jianjiang Feng (Tsinghua University), and Jie Zhou (Tsinghua University)</i>	

Generalizable Person Re-Identification With Relevance-Aware Mixture of Experts	16140
<i>Yongxing Dai (Peking University), Xiaotong Li (Peking University), Jun Liu (Singapore University of Technology and Design), Zekun Tong (National University of Singapore), and Ling-Yu Duan (Peking University)</i>	
Time Lens: Event-Based Video Frame Interpolation	16150
<i>Stepan Tulyakov (Huawei Zurich Research Center), Daniel Gehrig (University of Zurich & ETH Zurich), Stamatios Georgoulis (Huawei), Julius Erbach (Huawei), Mathias Gehrig (University of Zurich), Yuanyou Li (Huawei), and Davide Scaramuzza (University of Zurich & ETH Zurich, Switzerland)</i>	
How To Exploit the Transferability of Learned Image Compression to Conventional Codecs	16160
<i>Jan P. Klopp (National Taiwan University), Keng-Chi Liu (Taiwan AI Labs), Liang-Gee Chen (National Taiwan University), and Shao-Yi Chien (National Taiwan University)</i>	
Pseudo 3D Auto-Correlation Network for Real Image Denoising	16170
<i>Xiaowan Hu (Tsinghua ShenZhen International Graduate School), Ruijun Ma (University of Macau), Zhihong Liu (Tsinghua Shenzhen International Graduate School, Tsinghua University), Yuanhao Cai (Tsinghua University, Tsinghua Shenzhen International Graduate School), Xiaole Zhao (University of Electronic Science and Technology of China), Yulun Zhang (Northeastern University), and Haoqian Wang (Tsinghua Shenzhen International Graduate School, Tsinghua University)</i>	
Ultra-High-Definition Image Dehazing via Multi-Guided Bilateral Learning	16180
<i>Zhuoran Zheng (Nanjing University of Science and Technology), Wenqi Ren (Institute of Information Engineering, Chinese Academy of Sciences), Xiaochun Cao (Chinese Academy of Sciences), Xiaobin Hu (Technische Universität München), Tao Wang (Huawei Noah's Ark Lab), Fenglong Song (Huawei Noah's Ark Lab), and Xiuyi Jia (Nanjing University of Science and Technology)</i>	
HLA-Face: Joint High-Low Adaptation for Low Light Face Detection	16190
<i>Wenjing Wang (Peking University), Wenhan Yang (City University of Hong Kong), and Jiaying Liu (Peking University)</i>	
Zero-Shot Single Image Restoration Through Controlled Perturbation of Koschmieder's Model .	16200
<i>Aupendu Kar (Indian Institute of Technology Kharagpur), Sobhan Kanti Dhara (Indian Institute of Technology, Kharagpur, India), Debashis Sen (IIT Kharagpur), and Prabir Kumar Biswas (IIT Khargpur)</i>	
Deep Gaussian Scale Mixture Prior for Spectral Compressive Imaging	16211
<i>Tao Huang (Xidian University), Weisheng Dong (Xidian University), Xin Yuan (Bell Labs), Jinjian Wu (Xidian University), and Guangming Shi (Xidian University)</i>	
Multi-View 3D Reconstruction of a Texture-Less Smooth Surface of Unknown Generic Reflectance	16221
<i>Ziang Cheng (Australian National University), Hongdong Li (Australian National University, Australia), Yuta Asano (National Institute of Informatics), Yinqiang Zheng (The University of Tokyo), and Imari Sato (National Institute of Informatics)</i>	

Tackling the Ill-Posedness of Super-Resolution Through Adaptive Target Generation	16231
<i>Younghyun Jo (Yonsei University), Seoung Wug Oh (Adobe Research), Peter Vajda (Facebook), and Seon Joo Kim (Yonsei University)</i>	
Memory-Efficient Network for Large-Scale Video Compressive Sensing	16241
<i>Ziheng Cheng (Xidian University), Bo Chen (Xidian University), Guanliang Liu (Xidian University), Hao Zhang (Cornell University), Ruiying Lu (Xidian University), Zhengjue Wang (New Jersey Institute of Technology), and Xin Yuan (Bell Labs)</i>	
Troubleshooting Blind Image Quality Models in the Wild	16251
<i>Zhihua Wang (City University of Hong Kong), Haotao Wang (University of Texas at Austin), Tianlong Chen (University of Texas at Austin), Zhangyang Wang (University of Texas at Austin), and Kede Ma (City University of Hong Kong)</i>	
UAV-Human: A Large Benchmark for Human Behavior Understanding With Unmanned Aerial Vehicles	16261
<i>Tianjiao Li (Singapore University of Technology and Design), Jun Liu (Singapore University of Technology and Design), Wei Zhang (Shandong University), Yun Ni (Singapore University of Technology and Design), Wenqian Wang (Shandong University), and Zhiheng Li (Shandong University)</i>	
FBNetV3: Joint Architecture-Recipe Search Using Predictor Pretraining	16271
<i>Xiaoliang Dai (Facebook), Alvin Wan (UC Berkeley), Peizhao Zhang (Facebook), Bichen Wu (Facebook Research), Zijian He (Facebook), Zhen Wei (UNC-Chapel Hill), Kan Chen (Facebook Research), Yuandong Tian (Facebook), Matthew Yu (Facebook), Peter Vajda (Facebook), and Joseph E. Gonzalez (UC Berkeley)</i>	
Domain-Independent Dominance of Adaptive Methods	16281
<i>Pedro Savarese (Toyota Technological Institute at Chicago), David McAllester (TTI Chicago), Sudarshan Babu (Toyota Technological Institute at Chicago), and Michael Maire (University of Chicago)</i>	
Wasserstein Contrastive Representation Distillation	16291
<i>Liqun Chen (Duke University), Dong Wang (Duke University), Zhe Gan (Microsoft), Jingjing Liu (Microsoft), Ricardo Henao (Duke University), and Lawrence Carin (CS)</i>	
The Lottery Tickets Hypothesis for Supervised and Self-Supervised Pre-Training in Computer Vision Models	16301
<i>Tianlong Chen (University of Texas at Austin), Jonathan Frankle (MIT), Shiyu Chang (IBM Research), Sijia Liu (Michigan State University), Yang Zhang (IBM T. J. Watson Research), Michael Carbin (MIT), and Zhangyang Wang (University of Texas at Austin)</i>	
Generalizing Face Forgery Detection With High-Frequency Features	16312
<i>Yuchen Luo (Shanghai Jiao Tong University), Yong Zhang (Tencent AI Lab), Junchi Yan (Shanghai Jiao Tong University), and Wei Liu (Tencent)</i>	

CAMERAS: Enhanced Resolution and Sanity Preserving Class Activation Mapping for Image Saliency	16322
<i>Mohammad A. A. K. Jalwana (University of Western Australia), Naveed Akhtar (The University of Western Australia), Mohammed Bennamoun (University of Western Australia), and Ajmal Mian (University of Western Australia)</i>	
See Through Gradients: Image Batch Recovery via GradInversion	16332
<i>Hongxu Yin (NVIDIA), Arun Mallya (NVIDIA), Arash Vahdat (NVIDIA), Jose M. Alvarez (NVIDIA), Jan Kautz (NVIDIA), and Pavlo Molchanov (NVIDIA)</i>	
Seeking the Shape of Sound: An Adaptive Framework for Learning Voice-Face Association	16342
<i>Peisong Wen (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Qianqian Xu (Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences), Yangbangyan Jiang (Institute of Information Engineering, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Zhiyong Yang (SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences; SCS, University of Chinese Academy of Sciences), Yuan He (Alibaba Group), and Qingming Huang (University of Chinese Academy of Sciences)</i>	
Prototype-Supervised Adversarial Network for Targeted Attack of Deep Hashing	16352
<i>Xunguang Wang (Harbin Institute of TechnologyShenzhen), Zheng Zhang (Harbin Institute of Technology, Shenzhen), Baoyuan Wu (The Chinese University of Hong Kong, Shenzhen; Shenzhen Research Institute of Big Data), Fumin Shen (UESTC), and Guangming Lu (Harbin Institute of Technology, Shenzhen)</i>	
Intrinsic Image Harmonization	16362
<i>Zonghui Guo (Ocean University of China), Haiyong Zheng (Ocean University of China), Yufeng Jiang (Ocean University of China), Zhaorui Gu (Ocean University of China), and Bing Zheng (Ocean University of China)</i>	
DivCo: Diverse Conditional Image Synthesis via Contrastive Generative Adversarial Network ..	16372
<i>Rui Liu (Chinese University of Hong Kong), Yixiao Ge (The Chinese University of Hong Kong), Ching Lam Choi (The Chinese University of Hong Kong, NVIDIA AI Tech Center), Xiaogang Wang (Chinese University of Hong Kong, Hong Kong), and Hongsheng Li (Chinese University of Hong Kong)</i>	
Combining Semantic Guidance and Deep Reinforcement Learning for Generating Human Level Paintings	16382
<i>Jaskirat Singh (Australian National University) and Liang Zheng (Australian National University)</i>	
No Shadow Left Behind: Removing Objects and Their Shadows Using Approximate Lighting and Geometry	16392
<i>Edward Zhang (University of Washington), Ricardo Martin-Brualla (Google), Janne Kontkanen (Google), and Brian L. Curless (University of Washington)</i>	
The Spatially-Correlative Loss for Various Image Translation Tasks	16402
<i>Chuanxia Zheng (Nanyang Technological University), Tat-Jen Cham (Nanyang Technological University), and Jianfei Cai (Monash University)</i>	

Unpaired Image-to-Image Translation via Latent Energy Transport	16413
<i>Yang Zhao (University at Buffalo) and Changyou Chen (University at Buffalo)</i>	
Model-Aware Gesture-to-Gesture Translation	16423
<i>Hezhen Hu (University of Science and Technology of China), Weilun Wang (University of Science and Technology of China), Wengang Zhou (University of Science and Technology of China), Weichao Zhao (University of Science and Technology of China), and Houqiang Li (University of Science and Technology of China)</i>	
Layout-Guided Novel View Synthesis From a Single Indoor Panorama	16433
<i>Jiale Xu (ShanghaiTech University), Jia Zheng (Manycore, (Kujiale)), Yanyu Xu (Institute of High Performance Computing, A*Star), Rui Tang (Kujiale.com), and Shenghua Gao (Shanghaitech University)</i>	
Progressive Temporal Feature Alignment Network for Video Inpainting	16443
<i>Xueyan Zou (university of california, davis), Linjie Yang (ByteDance AI Lab), Ding Liu (Bytedance), and Yong Jae Lee (University of California, Davis)</i>	
On Robustness and Transferability of Convolutional Neural Networks	16453
<i>Josip Djolonga (Google AI, Zurich), Jessica Yung (Google Research), Michael Tschannen (Apple), Rob Romijnders (Google AI), Lucas Beyer (Google Brain), Alexander Kolesnikov (Google Brain), Joan Puigcerver (Google), Matthias Minderer (Google Research), Alexander D'Amour (Google Brain), Dan Moldovan (Google Inc.), Sylvain Gelly (Google Brain), Neil Houlsby (Google), Xiaohua Zhai (Google Brain), and Mario Lucic (Google Brain)</i>	
Boosting Ensemble Accuracy by Revisiting Ensemble Diversity Metrics	16464
<i>Yanzhao Wu (Georgia Institute of Technology), Ling Liu (Georgia Institute of Technology), Zhongwei Xie (Georgia Institute of Technology), Ka-Ho Chow (Georgia Institute of Technology), and Wenqi Wei (Georgia Institute of Technology)</i>	
General Multi-Label Image Classification With Transformers	16473
<i>Jack Lanchantin (University of Virginia), Tianlu Wang (University of Virginia), Vicente Ordonez (University of Virginia), and Yanjun Qi (University of Virginia)</i>	
Improving Calibration for Long-Tailed Recognition	16484
<i>Zhisheng Zhong (Peking University), Jiequan Cui (Chinese University of Hong Kong), Shu Liu (SmartMore), and Jiaya Jia (Chinese University of Hong Kong)</i>	
Hardness Sampling for Self-Training Based Transductive Zero-Shot Learning	16494
<i>Liu Bo (NLPR-IA-CAS), Qiulei Dong (NLPR-IA-CAS), and Zhanyi Hu (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences)</i>	
Saliency-Guided Image Translation	16504
<i>Lai Jiang (BUAA), Mai Xu (BUAA), Xiaofei Wang (BUAA), and Leonid Sigal (University of British Columbia)</i>	

Bottleneck Transformers for Visual Recognition	16514
<i>Aravind Srinivas (UC Berkeley), Tsung-Yi Lin (Google Brain), Niki Parmar (Google), Jonathon Shlens (Google), Pieter Abbeel (UC Berkeley), and Ashish Vaswani (Google Brain)</i>	
One-Shot Neural Ensemble Architecture Search by Diversity-Guided Search Space Shrinking	16525
<i>Minghao Chen (Stony Brook University), Jianlong Fu (Microsoft Research), and Haibin Ling (Stony Brook University)</i>	
Multi-Perspective LSTM for Joint Visual Representation Learning	16535
<i>Alireza Sepas-Moghaddam (Queen’s University), Fernando Pereira (Universidade de Lisboa), Paulo Lobato Correia (Instituto de Telecomunicacoes / Instituto Superior Técnico - Universidade de Lisboa), and Ali Etemad (Queen’s University)</i>	
Multiple Instance Captioning: Learning Representations From Histopathology Textbooks and Articles	16544
<i>Jevgenij Gamper (University of Warwick) and Nasir Rajpoot (University of Warwick)</i>	
clDice – A Novel Topology-Preserving Loss Function for Tubular Structure Segmentation	16555
<i>Suprosanna Shit (TUM), Johannes C. Paetzold (TUM), Anjany Sekuboyina (TUM), Ivan Ezhov (TUM), Alexander Unger (Technische Universität München), Andrey Zhylka (Eindhoven), Josien P. W. Pluim (Eindhoven), Ulrich Bauer (TUM), and Bjoern H. Menze (TUM)</i>	
Learning To Fuse Asymmetric Feature Maps in Siamese Trackers	16565
<i>Wencheng Han (Beijing Institute of Technology), Xingping Dong (Inception Institute of Artificial Intelligence), Fahad Shahbaz Khan (MBZUAI), Ling Shao (Inception Institute of Artificial Intelligence), and Jianbing Shen (Inception Institute of Artificial Intelligence)</i>	
Track, Check, Repeat: An EM Approach to Unsupervised Tracking	16576
<i>Adam W. Harley (Carnegie Mellon University), Yiming Zuo (CMU), Jing Wen (Carnegie Mellon University), Ayush Mangal (IIT Roorkee), Shubhankar Potdar (Carnegie Mellon University), Ritwick Chaudhry (Carnegie Mellon University), and Katerina Fragkiadaki (Carnegie Mellon University)</i>	
Learning Optical Flow From a Few Matches	16587
<i>Shihao Jiang (Australian National University), Yao Lu (Australian National University), Hongdong Li (Australian National University, Australia), and Richard Hartley (Australian National University, Australia)</i>	
AutoDO: Robust AutoAugment for Biased Data With Label Noise via Scalable Probabilistic Implicit Differentiation	16596
<i>Denis Gudovskiy (Panasonic), Luca Rigazio (Panasonic), Shun Ishizaka (Panasonic Corporation), Kazuki Kozuka (Panasonic Corporation), and Sotaro Tsukizawa (Panasonic)</i>	
GDR-Net: Geometry-Guided Direct Regression Network for Monocular 6D Object Pose Estimation.....	16606
<i>Gu Wang (Tsinghua University), Fabian Manhardt (Google), Federico Tombari (Google, TU Munich), and Xiangyang Ji (Tsinghua University)</i>	

Robust and Accurate Object Detection via Adversarial Learning	16617
<i>Xiangning Chen (University of California, Los Angeles), Cihang Xie (University of California, Santa Cruz), Mingxing Tan (Google Brain), Li Zhang (Google), Cho-Jui Hsieh (UCLA), and Boqing Gong (Google)</i>	
Domain Adaptation With Auxiliary Target Domain-Oriented Classifier	16627
<i>Jian Liang (NUS), Dapeng Hu (NUS), and Jiashi Feng (NUS)</i>	
MetaAlign: Coordinating Domain Alignment and Classification for Unsupervised Domain Adaptation	16638
<i>Guoqiang Wei (University of Science and Technology of China), Cuiling Lan (Microsoft Research), Wenjun Zeng (Microsoft Research), and Zhibo Chen (University of Science and Technology of China)</i>	
Partial Feature Selection and Alignment for Multi-Source Domain Adaptation	16649
<i>Yangye Fu (University of Electronic Science and Technology of China), Ming Zhang (University of Electronic Science and Technology of China), Xing Xu (University of Electronic Science and Technology of China), Zuo Cao (MEITUAN), Chao Ma (MEITUAN), Yanli Ji (UESTC), Kai Zuo (MEITUAN), and Huimin Lu (Kyushu Institute of Technology)</i>	
MongeNet: Efficient Sampler for Geometric Deep Learning	16659
<i>Léo Lebrat (CSIRO), Rodrigo Santa Cruz (CSIRO), Clinton Fookes (Queensland University of Technology), and Olivier Salgado (Australian e-Health Research Centre)</i>	
Continual Learning via Bit-Level Information Preserving	16669
<i>Yujun Shi (NUS), Li Yuan (National University of Singapore), Yunpeng Chen (National University of Singapore), and Jiashi Feng (NUS)</i>	
Propagate Yourself: Exploring Pixel-Level Consistency for Unsupervised Visual Representation Learning	16679
<i>Zhenda Xie (Tsinghua University), Yutong Lin (Xi'an Jiaotong University), Zheng Zhang (MSRA), Yue Cao (Microsoft Research), Stephen Lin (Microsoft Research), and Han Hu (Microsoft Research Asia)</i>	
Toward Joint Thing-and-Stuff Mining for Weakly Supervised Panoptic Segmentation	16689
<i>Yunhang Shen (Xiamen University), Liujuan Cao (Xiamen University), Zhiwei Chen (Xiamen University), Feihong Lian (Xiamen University), Baochang Zhang (Beihang University), Chi Su (Kingsoft Cloud), Yongjian Wu (Tencent Technology, (Shanghai) Co., Ltd), Feiyue Huang (Tencent), and Rongrong Ji (Xiamen University, China)</i>	
S3: Learnable Sparse Signal Superdensity for Guided Depth Estimation	16701
<i>Yu-Kai Huang (National Taiwan University), Yueh-Cheng Liu (National Taiwan University), Tsung-Han Wu (National Taiwan University), Hung-Ting Su (National Taiwan University), Yu-Cheng Chang (National Taiwan University), Tsung-Lin Tsou (National Taiwan University), Yu-An Wang (National Taiwan University), and Winston H. Hsu (National Taiwan University)</i>	
Residential Floor Plan Recognition and Reconstruction	16712
<i>Xiaolei Lv (Alibaba Group), Shengchu Zhao (Alibaba Group), Xinyang Yu (Alibaba), and Binqiang Zhao (Alibaba)</i>	

ColorRL: Reinforced Coloring for End-to-End Instance Segmentation	16722
<i>Tran Anh Tuan (Ulsan National Institute of Science and Technology), Nguyen Tuan Khoa (Ulsan National Institute of Science and Technology), Tran Minh Quan (VinBrain), and Won-Ki Jeong (Korea University)</i>	
Harmonious Semantic Line Detection via Maximal Weight Clique Selection	16732
<i>Dongkwon Jin (Korea university), Wonhui Park (Korea university), Seong-Gyun Jeong (42dot.ai), and Chang-Su Kim (Korea university)</i>	
LPSNet: A Lightweight Solution for Fast Panoptic Segmentation	16741
<i>Weixiang Hong (National University of Singapore), Qingpei Guo (Ant Financial Services Group), Wei Zhang (Ant Financial Services Group), Jingdong Chen (Ant Group), and Wei Chu (Ant Group)</i>	
Progressive Semantic Segmentation	16750
<i>Chuonng Huynh (VinAI Research, Vietnam), Anh Tuan Tran (VinAI Research), Khoa Luu (University of Arkansas), and Minh Hoai (Stony Brook University)</i>	
Embedded Discriminative Attention Mechanism for Weakly Supervised Semantic Segmentation	16760
<i>Tong Wu (Beijing Institute of Technology), Junshi Huang (Meituan), Guangyu Gao (Beijing Institute of Technology), Xiaoming Wei (Meituan), Xiaolin Wei (Meituan), Xuan Luo (Meituan), and Chi Harold Liu (Beijing Institute of Technology)</i>	
SelfSAGCN: Self-Supervised Semantic Alignment for Graph Convolution Network	16770
<i>Xu Yang (Xidian University), Cheng Deng (Xidian University), Zhiyuan Dang (JD Tech), Kun Wei (Xidian University), and Junchi Yan (Shanghai Jiao Tong University)</i>	
Wasserstein Barycenter for Multi-Source Domain Adaptation	16780
<i>Eduardo Fernandes Montesuma (Universidade Federal do Ceará) and Fred Maurice Ngolè Mboula (Commissariat à l'énergie atomique, (CEA))</i>	
PiCIE: Unsupervised Semantic Segmentation Using Invariance and Equivariance in Clustering ..	16789
<i>Jang Hyun Cho (The University of Texas at Austin), Utkarsh Mall (Cornell University), Kavita Bala (Cornell University), and Bharath Hariharan (Cornell University)</i>	
Global2Local: Efficient Structure Search for Video Action Segmentation	16800
<i>Shang-Hua Gao (Nankai University), Qi Han (Nankai University), Zhong-Yu Li (Nankai University), Pai Peng (Tencent Youtu Lab), Liang Wang (NLPR, China), and Ming-Ming Cheng (Nankai University)</i>	
Introvert: Human Trajectory Prediction via Conditional 3D Attention	16810
<i>Nasim Shafiee (Northeastern University), Taskin Padir (Northeastern University), and Ehsan Elhamifar (Northeastern University)</i>	
Weakly Supervised Video Salient Object Detection	16821
<i>Wangbo Zhao (Northwestern Polytechnical University), Jing Zhang (Australian National University), Long Li (Northwestern Polytechnical University), Nick Barnes (ANU), Nian Liu (Mohamed bin Zayed University of Artificial Intelligence), and Junwei Han (NWPUI, China)</i>	
Video Object Segmentation Using Global and Instance Embedding Learning	16831
<i>Wenbin Ge (Beijing Institute of Technology), Xiankai Lu (Shandong University), and Jianbing Shen (Inception Institute of Artificial Intelligence)</i>	

Human-Like Controllable Image Captioning With Verb-Specific Semantic Roles	16841
<i>Long Chen (Columbia University), Zhihong Jiang (Zhejiang University), Jun Xiao (Zhejiang University), and Wei Liu (Tencent)</i>	
Read and Attend: Temporal Localisation in Sign Language Videos	16852
<i>Gül Varol (Ecole des Ponts ParisTech), Liliane Momeni (University of Oxford), Samuel Albanie (University of Oxford), Triantafyllos Afouras (University of Oxford), and Andrew Zisserman (University of Oxford)</i>	
Localizing Visual Sounds the Hard Way	16862
<i>Honglie Chen (University of Oxford), Weidi Xie (University of Oxford), Triantafyllos Afouras (University of Oxford), Arsha Nagrani (Oxford University), Andrea Vedaldi (Oxford University), and Andrew Zisserman (University of Oxford)</i>	
Look Before You Speak: Visually Contextualized Utterances	16872
<i>Paul Hongsuck Seo (Google), Arsha Nagrani (Oxford University), and Cordelia Schmid (Google)</i>	
Look Before You Leap: Learning Landmark Features for One-Stage Visual Grounding	16883
<i>Binbin Huang (ShanghaiTech University), Dongze Lian (ShanghaiTech University), Weixin Luo (Meituan), and Shenghua Gao (ShanghaiTech University)</i>	
Visual Navigation With Spatial Attention	16893
<i>Bar Mayo (Technion), Tamir Hazan (Technion), and Ayellet Tal (Technion)</i>	
Perception Matters: Detecting Perception Failures of VQA Models Using Metamorphic Testing .	16903
<i>Yuanyuan Yuan (HKUST), Shuai Wang (HKUST), Mingyue Jiang (Zhejiang Sci-Tech University), and Tsong Yueh Chen (Swinburne University of Technology)</i>	
Pixel-Wise Anomaly Detection in Complex Driving Scenes	16913
<i>Giancarlo Di Biase (ETH Zurich), Hermann Blum (ETH Zurich), Roland Siegwart (ETH Zürich, Autonomous Systems Lab), and Cesar Cadena (ETH Zurich)</i>	
Disentangled Cycle Consistency for Highly-Realistic Virtual Try-On	16923
<i>Chongjian Ge (The University of Hong Kong, (HKU)), Yibing Song (Tencent), Yuying Ge (The University of Hong Kong), Han Yang (ETH Zurich), Wei Liu (Tencent), and Ping Luo (The University of Hong Kong)</i>	
Hyperdimensional Computing as a Framework for Systematic Aggregation of Image Descriptors	16933
<i>Peer Neubert (TU Chemnitz) and Stefan Schubert (TU Chemnitz)</i>	
Author Index	16943