2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2024)

Waikoloa, Hawaii, USA
3-8 January 2024

Pages 1-663

IEEE Catalog Number: CFP24082-POD
ISBN: 979-8-3503-1893-7

Copyright © 2024 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved

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

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

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

IEEE Catalog Number:	CFP24082-POD
ISBN (Print-On-Demand):	979-8-3503-1893-7
ISBN (Online):	979-8-3503-1892-0
ISSN:	2472-6737

Additional Copies of This Publication Are Available From:

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

2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)

WACV 2024

Table of Contents

Message from the General and Program Chairs ..... **xcvii**
Organizing Committee and Area Chairs ..... **xcviii**
Sponsors ..... **xcix**

Machine Learning: Architectures, Formulations, Algorithms I

Asymmetric Image Retrieval with Cross Model Compatible Ensembles 1
Alon Shoshan (Amazon), Ori Linial (Technion - Israel Institute of Technology), Nadav Bhonker (Amazon), Elad Hirsch (Technion - Israel Institute of Technology), Lior Zamir (Amazon), Igor Kviatkovsky (Amazon), and Gérard Medioni (Amazon)

Cross-Feature Contrastive Loss for Decentralized Deep Learning on Heterogeneous Data 12
Sai Aparna Aketi (Purdue University, USA) and Kaushik Roy (Purdue Uniiversity, USA)

Learning Generalizable Perceptual Representations for Data-Efficient No-Reference Image Quality Assessment 22
Suhas Srinath (Indian Institute of Science, India), Shankhanil Mitra (Indian Institute of Science, India), Shika Rao (Indian Institute of Science, India), and Rajiv Soundararajan (Indian Institute of Science, India)

Robust Feature Learning and Global Variance-Driven Classifier Alignment for Long-Tail Class Incremental Learning 32
Jayateja Kalla (Indian Institute of Science, India) and Soma Biswas (Indian Institute of Science, India)

PECoP: Parameter Efficient Continual Pretraining for Action Quality Assessment 42
Amirhossein Dadashzadeh (University of Bristol, UK), Shuchao Duan (University of Bristol, UK), Alan L Whone (University of Bristol, UK), and Majid Mirmehdi (University of Bristol, UK)

Wino Vidi Vici: Conquering Numerical Instability of 8-Bit Winograd Convolution for Accurate Inference Acceleration on Edge 53
Pierpaolo Mori (Politecnico di Torino, Italy; BMW Group, Germany; Technical University of Munich, Germany), Lukas Frickenstein (BMW Group, Germany), Shambhavi Balamuthu Sampath (BMW Group, Germany), Moritz Thoma (BMW Group, Germany), Nael Fasfous (BMW Group, Germany), Manoj Rohit Vemparala (BMW Group, Germany), Alexander Frickenstein (BMW Group, Germany), Christian Unger (BMW Group, Germany), Walter Stechele (Technical University of Munich, Germany), Daniel Mueller-Gritschneder (Technical University of Munich, Germany), and Claudio Passerone (Politecnico di Torino, Italy)

Data-Centric Debugging: Mitigating Model Failures via Targeted Image Retrieval 63
Sahil Singla (Google Research), Atoosa Malemir Chegini (University of Maryland), Mazda Moayeri (University of Maryland), and Soheil Feizi (University of Maryland)

Distortion-Disentangled Contrastive Learning 75
Jinfeng Wang (Xi'an Jiaotong-Liverpool University, China; University of Science and Technology of China, China), Sifan Song (University of Liverpool based in Xi'an Jiaotong-Liverpool University, China; University of Science and Technology of China, China), Jionglong Su (Xi'an Jiaotong-Liverpool University, China), and S. Kevin Zhou (University of Science and Technology of China, China; Institute of Computing Technology, China)

GTP-ViT: Efficient Vision Transformers via Graph-Based Token Propagation 86
Xuwei Xu (The University of Queensland, Australia; CSIRO Data61, Australia), Sen Wang (The University of Queensland, Australia), Yudong Chen (The University of Queensland, Australia; CSIRO Data61, Australia), Yanping Zheng (Renmin University of China; CSIRO Data61, Australia), Zhewei Wei (Renmin University of China), and Jiajun Liu (The University of Queensland, Australia; CSIRO Data61, Australia)

SequenceMatch: Revisiting the Design of Weak-Strong Augmentations for Semi-Supervised Learning 96
Khanh-Binh Nguyen (Sungkyunkwan University, South Korea)

Stochastic Binary Network for Universal Domain Adaptation 106
Saurabh Kumar Jain (Indian Institute of Technology, India) and Sukhendu Das (Indian Institute of Technology, India)

Location-Aware Self-Supervised Transformers for Semantic Segmentation 116
Mathilde Caron (Google Research), Neil Houlsby (Google Research), and Cordelia Schmid (Google Research)

Image/Video Recognition & Understanding; Low-Level & Physics-Based Vision

EfficientAD: Accurate Visual Anomaly Detection at Millisecond-Level Latencies 127
Kilian Batzner (MVTec Software GmbH), Lars Heckler (MVTec Software GmbH; Technical University of Munich), and Rebecca König (MVTec Software GmbH)

Efficient Semantic Matching with Hypercolumn Correlation 138
Seungwook Kim (Pohang University of Science and Technology, South Korea), Juhong Min (Pohang University of Science and Technology, South Korea), and Minsu Cho (Pohang University of Science and Technology, South Korea)

Contextual Affinity Distillation for Image Anomaly Detection 148
Jie Zhang (Tohoku University), Masanori Suganuma (Tohoku University; RIKEN Center for AIP), and Takayuki Okatani (Tohoku University; RIKEN Center for AIP)

Offline-to-Online Knowledge Distillation for Video Instance Segmentation 158
Hojin Kim (SI-analytics), Seunghun Lee (DGIST), Hyeon Kang (DGIST), and Sunghoon Im (DGIST)

Disentangled Pre-Training for Image Matting 168
Yanda Li (University of Technology Sydney, Australia), Zilong Huang (Tencent), Gang Yu (Tencent), Ling Chen (University of Technology Sydney, Australia), Yunchao Wei (University of Technology Sydney, Australia), and Jianbo Jiao (University of Birmingham, UK)

Conditional Velocity Score Estimation for Image Restoration 178
Ziqiang Shi (Fujitsu R & D Center, China) and Rujie Liu (Fujitsu R&D Center, China)

ARNIQA: Learning Distortion Manifold for Image Quality Assessment 188
Lorenzo Agnolucci (University of Florence - Media Integration and Communication Center (MICC), Italy), Leonardo Galteri (University of Florence - Media Integration and Communication Center (MICC), Italy), Marco Bertini (University of Florence - Media Integration and Communication Center (MICC), Italy), and Alberto Del Bimbo (University of Florence - Media Integration and Communication Center (MICC), Italy)

Hard Sample-Aware Consistency for Low-Resolution Facial Expression Recognition 198
Bokyeung Lee (Korea University), Kyungdeuk Ko (Korea University), Jonghwan Hong (Korea university), and Hanseok Ko (Korea University)

ArcGeo: Localizing Limited Field-of-View Images Using Cross-View Matching 208
Maxim Shugaev (BlueHalo), Ilya Semenov (Blue Halo), Kyle Ashley (BlueHalo), Michael Klaczynski (BlueHalo), Naresh Cuntoor (BlueHalo), Mun Wai Lee (BlueHalo), and Nathan Jacobs (Washington University)

Open-Set Object Detection by Aligning Known Class Representations 218
Hiran Sarkar (Sony Research, India), Vishal Chudasama (Sony Research, India), Naoyuki Onoe (Sony Research, India), Pankaj Wasnik (Sony Research, India), and Vineeth N Balasubramanian (Indian Institute of Technology Hyderabad)

Video-kMaX: A Simple Unified Approach for Online and Near-Online Video Panoptic Segmentation 228
Inkyu Shin (KAIST; Google Research), Dahun Kim (Google Research), Qihang Yu (Google Research), Jun Xie (Google Research), Hong-Seok Kim (Google Research), Bradley Green (Google Research), In So Kweon (KAIST), Kuk-Jin Yoon (KAIST), and Liang-Chieh Chen (Google Research)

DeVos: Flow-Guided Deformable Transformer for Video Object Segmentation 239
Volodymyr Fedynyak (Ukrainian Catholic University), Yaroslav Romanus (Ukrainian Catholic University; ADVA Soft), Bohdan Hlovatskyi (Ukrainian Catholic University), Bohdan Sydor (Ukrainian Catholic University), Oles Dobosevych (Ukrainian Catholic University), Igor Babin (Ukrainian Catholic University; ADVA Soft), and Roman Riazantsev (ADVA Soft)

Image Recognition & Understanding

Semi-Supervised Semantic Depth Estimation Using Symbiotic Transformer and NearFarMix Augmentation 249
Awsafur Rahman (BUET, Bangladesh) and Shaikh Anowarul Fattah (BUET, Bangladesh)

Training Ensembles with Inliers and Outliers for Semi-Supervised Active Learning 259
Vladan Stojnić (Czech Technical University in Prague), Zakaria Laskar (Czech Technical University in Prague), and Giorgos Tolias (Czech Technical University in Prague)

Multi-View Classification Using Hybrid Fusion and Mutual Distillation 269
Samuel Black (Temple University, USA) and Richard Souvenir (Temple University, USA)

Amodal Intra-Class Instance Segmentation: Synthetic Datasets and Benchmark 280
Jiayang Ao (The University of Melbourne), Qiuhong Ke (Monash University), and Krista A. Ehinger (University of Melbourne)

Prompting classes: Exploring the Power of Prompt Class Learning in Weakly Supervised Semantic Segmentation 290
Balamurali Murugesan (ETS Montreal), Rukhshanda Hussain (Jadavpur University), Rajarshi Bhattacharya (Jadavpur University), Ismail Ben Ayed (ETS Montreal), and Jose Dolz (ETS Montreal)

RSMPNet: Relationship Guided Semantic Map Prediction 302
Jingwen Sun (Cardiff University), Jing Wu (Cardiff University), Ze Ji (Cardiff University), and Yu-Kun Lai (Cardiff University)

3SD: Self-Supervised Saliency DetectionWith No Labels 312
Rajeev Yasarla (Johns Hopkins University; AIBEE), Renliang Weng (Amir Sadeghian), Wongun Choi (AIBEE), Vishal M. Patel (Johns Hopkins University), and Amir Sadeghian (AIBEE)

Training-Free Object Counting With Prompts 322
Zenglin Shi (Astar), Ying Sun (Institute for Infocomm Research, ASTAR), and Mengmi Zhang (NTU, ASTAR)

Unsupervised and Semi-Supervised Co-Salient Object Detection via Segmentation Frequency Statistics 331
Souradeep Chakraborty (Stony Brook University; Visual Search & AR, Amazon), Shujon Naha (Visual Search & AR, Amazon; Indiana University), Muhammet Bastan (Visual Search & AR, Amazon), Amit Kumar K C (Visual Search & AR, Amazon), and Dimitris Samaras (Stony Brook University)

Glance to Count: Learning to Rank with Anchors for Weakly-Supervised Crowd Counting 342
Zheng Xiong (South China University of Technology; Singapore Management University), Liangyu Chai (South China University of Technology; Singapore Management University), Wenxi Liu (Fuzhou University), Yongtuo Liu (University of Amsterdam), Sucheng Ren (Singapore Management University), and Shengfeng He (Singapore Management University)

TransRadar: Adaptive-Directional Transformer for Real-Time Multi-View Radar Semantic Segmentation 352
Yahia I Dalbah (Mohamed Bin Zayed University of Artificial Intelligence), Jean Lahoud (MBZUAI), and Hisham Cholakkal (MBZUAI)

Booster-SHOT: Boosting Stacked Homography Transformations for Multiview Pedestrian Detection with Attention 362
Jinwoo Hwang (Seoul National University), Philipp Benz (Deeping Source Inc.), and Pete Kim (Deeping Source Inc.)

360BEV: Panoramic Semantic Mapping for Indoor Bird's-Eye View 372
Zhifeng Teng (Karlsruhe Institute of Technology), Jiaming Zhang (Karlsruhe Institute of Technology), Kailun Yang (Hunan University), Kunyu Peng (Karlsruhe Institute of Technology), Hao Shi (Zhejiang University), Simon Reiß (Karlsruhe Institute of Technology), Ke Cao (Karlsruher Institut für Technologie), and Rainer Stiefelhagen (Karlsruhe Institute of Technology)

Learning Saliency From Fixations 382
Yasser Abdelaziz Dahou Djilali (Dublin City University, Ireland; Technology Innovation Institute, UAE), Kevin Mcguinness (Dublin City University, Ireland), and Noel O'Connor (Dublin City University, Ireland)

Mitigate Domain Shift by Primary-Auxiliary Objectives Association for Generalizing Person ReID 393
Qilei Li (Queen Mary, University of London) and Shaogang Gong (Queen Mary University of London)

MIST: Medical Image Segmentation Transformer with Convolutional Attention Mixing (CAM) Decoder 403
Md Motiur Rahman (Purdue University, USA), Shiva Shokouhmand (Purdue University, USA), Smriti Bhatt (Purdue University, USA), and Miad Faezipour (Purdue University, USA)

Small Objects Matters in Weakly-Supervised Semantic Segmentation 413
Cheolhyun Mun (Samsung Research, Korea), Sanghuk Lee (SOCAR AI Research, Korea), Youngjung Uh (Yonsei University, Korea), Junsuk Choe (Sogang University, Korea), and Hyeran Byun (Yonsei University, Korea)

Gradient-Guided Knowledge Distillation for Object Detectors 423
Qizhen Lan (Bowling Green State University; University of Alabama at Birmingham) and Qing Tian (Bowling Green State University; University of Alabama at Birmingham)

MetaSeg: MetaFormer-Based Global Contexts-Aware Network for Efficient Semantic Segmentation 433
Beoungwoo Kang (Sogang University, Republic of Korea), Seunghun Moon (Sogang University, Republic of Korea), Yubin Cho (Sogang University, Republic of Korea), Hyunwoo Yu (Sogang university, Republic of Korea), and Suk-Ju Kang (Sogang University, Republic of Korea)

Panelformer: Sewing Pattern Reconstruction From 2D Garment Images 443
Cheng-Hsiu Chen (National Tsing Hua University), Jheng-Wei Su (National Tsing Hua University), Min-Chun Hu (National Tsing Hua University), Chih-Yuan Yao (National Taiwan University of Science and Technology), and Hung-Kuo Chu (National Tsing Hua University)

From Denoising Training To Test-Time Adaptation: Enhancing Domain Generalization for Medical Image Segmentation 453
Ruxue Wen (Zhejiang University), Hangjie Yuan (Zhejiang University), Dong Ni (Zhejiang University), Wenbo Xiao (Zhejiang University), and Yaoyao Wu (Zhejiang University)

Guided Distillation for Semi-Supervised Instance Segmentation 464
Tariq Berrada (FAIR, Meta; Univ. Grenoble Alpes, Inria, CNRS), Camille Couprie (FAIR, Meta), Jakob Verbeek (FAIR, Meta), and Karteek Alahari (Univ. Grenoble Alpes, Inria, CNRS, UK)

Real-Time User-Guided Adaptive Colorization With Vision Transformer 473
Gwanghan Lee (Sungkyunkwan University (SKKU)), Saebyeol Shin (Sungkyunkwan University (SKKU)), Taeyoung Na (SK telecom), and Simon S Woo (Sungkyunkwan University (SKKU))

Mining and Unifying Heterogeneous Contrastive Relations for Weakly-Supervised Actor-Action Segmentation 483
Bin Duan (Illinois Institute of Technology), Hao Tang (ETH Zurich), Changchang Sun (Illinois Institute of Technology), Ye Zhu (Princeton University), and Yan Yan (Illinois Institute of Technology)

Graph Neural Networks for End-to-End Information Extraction From Handwritten Documents 493
Yessine Khanfir (InstaDeep), Marwa Dhiaf (Instadeep), Emna Ghodhbani (InstaDeep), Ahmed Cheikh Rouhou (InstaDeep), and Yousri Kessentini (CRNS)

CPSeg: Finer-Grained Image Semantic Segmentation via Chain-of-Thought Language Prompting . 502
Lei Li (University of Copenhagen)

Foundation Model AssistedWeakly Supervised Semantic Segmentation 512
Xiaobo Yang (Zhejiang University, China) and Xiaojin Gong (Zhejiang University, China)

On the Importance of Large Objects in CNN Based Object Detection Algorithms 522
Ahmed Ben Saad (ENS Paris Saclay), Gabriele Facciolo (ENS Paris - Saclay), and Axel Davy (ENS Paris-Saclay)

Deep Metric Learning with Chance Constraints 532
Yeti Z. Gürbüz (MetaDialog), Oğul Can (MetaDialog; Cerebrate AI), and A. Aydın Alatan (OGAM and METU)

TransFed: A Way To Epitomize Focal Modulation Using Transformer-Based Federated Learning .. 543
Tajamul Ashraf (Indian Institute of Technology Delhi), Fuzayil-Bin-Afzal Mir (Indian Institute of Technology Delhi), and Iqra Altaf Gillani (National Institute of Technology, Srinagar)

Learning Better Keypoints for Multi-Object 6DoF Pose Estimation 553
Yangzheng Wu (Queen's University, Canada) and Michael Greenspan (Queen's University, Canada)

Object Aware Contrastive Prior for Interactive Image Segmentation 564
Praful Mathur (Samsung R&D Institute India - Bangalore, India), Shashi Kumar Parwani (Samsung R&D Institute India - Bangalore, India), Mrinmoy Sen (Samsung R&D Institute India - Bangalore, India), Roopa Sheshadri (Samsung R&D Institute India - Bangalore, India), and Aman Sharma (Samsung R&D Institute India - Bangalore, India)

Beyond Classification: Definition and Density-Based Estimation of Calibration in Object Detection 574
Teodora Popordanoska (ESAT-PSI, KU Leuven), Aleksei Tiulpin (University of Oulu), and Matthew B. Blaschko (ESAT-PSI, KU Leuven)

FAKD: Feature Augmented Knowledge Distillation for Semantic Segmentation 584
Jianlong Yuan (Damo Academy, Alibaba Group), Minh Hieu Phan (University of Adelaide), Liyang Liu (University of Adelaide), and Yifan Liu (University of Adelaide)

Efficient MAE Towards Large-Scale Vision Transformers 595
Qiu Han (Nanyang Technological University), Gongjie Zhang (Nanyang Technological University; Black Sesame Technologies), Jiaxing Huang (Nanyang Technological University), Peng Gao (Shanghai Artificial Intelligence Laboratory), Zhang Wei (Shanghai Artificial Intelligence Laboratory), and Shijian Lu (Nanyang Technological University)

MS-EVS: Multispectral Event-Based Vision for Deep Learning Based Face Detection 605
Saad Himmi (ETH Zurich, Switzerland), Vincent Parret (Sony Europe B.V.), Ajad Chhatkuli (ETH Zurich, Switzerland), and Luc Van Gool (ETH Zurich, Switzerland)

Interactive Network Perturbation Between Teacher and Students for Semi-Supervised Semantic Segmentation 615
Hyuna Cho (Pohang University of Science and Technology (POSTECH)), Injun Choi (Pohang University of Science and Technology (POSTECH)), Suha Kwak (Pohang University of Science and Technology (POSTECH)), and Won Hwa Kim (Pohang University of Science and Technology (POSTECH))

Can Vision-Language Models Be a Good Guesser? Exploring VLMs for Times and Location Reasoning 625
Gengyuan Zhang (Ludwig Maximilian University of Munich, Germany; Munich Center for Machine Learning, Germany), Yurui Zhang (Technical University of Munich), Kerui Zhang (Ludwig Maximilian University of Munich, Germany), and Volker Tresp (Ludwig Maximilian University of Munich, Germany; Munich Center for Machine Learning, Germany)

USDN: A Unified Sample-Wise Dynamic Network with Mixed-Precision and Early-Exit 635
Ji-Ye Jeon (Seoul National University), Xuan Truong Nguyen (Seoul National University), Soojung Ryu (SK Telecom), and Hyuk-Jae Lee (Seoul National University)

Beyond Fusion: Modality Hallucination-Based Multispectral Fusion for Pedestrian Detection 644
Qian Xie (University of Oxford), Ta-Ying Cheng (University of Oxford), Jia-Xing Zhong (Department of Computer Science, University of Oxford), Kaichen Zhou (University of Oxford), Andrew Markham (University of Oxford), and Niki Trigoni (University of Oxford)

DocReal: Robust Document Dewarping of Real-Life Images via Attention-Enhanced Control Point Prediction 654
Fangchen Yu (The Chinese University of Hong Kong, China), Yina Xie (vivo AI Lab, China), Lei Wu (vivo AI Lab, China), Yafei Wen (vivo AI Lab, China), Guozhi Wang (vivo AI Lab, China), Shuai Ren (vivo AI Lab, China), Xiaoxin Chen (vivo AI Lab, China), Jianfeng Mao (The Chinese University of Hong Kong, China; Shenzhen Research Institute of Big Data, China), and Wenye Li (The Chinese University of Hong Kong, China; Shenzhen Research Institute of Big Data, China)

Learning to Recognize Occluded and Small Objects with Partial Inputs 664
Hasib Zunair (Concordia University, Canada) and Abdessamad Ben Hamza (Concordia University, Canada)

Temporally-Consistent Video Semantic Segmentation with Bidirectional Occlusion-Guided Feature Propagation 674
Razieh Kaviani Baghbaderani (SOC R&D, Samsung Semiconductor, Inc.), Yuanxin Li (SOC R&D, Samsung Semiconductor, Inc.), Shuangquan Wang (SOC R&D, Samsung Semiconductor, Inc.), and Hairong Qi (The University of Tennessee, USA)

Towards Domain-Aware Knowledge Distillation for Continual Model Generalization 685
Nikhil Reddy (UQ-IIT Delhi Research Academy (UQIDAR)), Mahsa Baktashmotlagh (University of Queensland, Australia), and Chetan Arora (Indian Institute of Technology Delhi, India)

Let's Observe Them Over Time: An Improved Pedestrian Attribute Recognition Approach 697
Kamalakar Vijay Thakare (Indian Institute of Technology, India), Debi Prosad Dogra (Indian Institute of Technology, India), Heeseung Choi (Korea Institute of Science and Technology, Republic of Korea; Yonsei University, Republic of Korea), Haksub Kim (Korea Institute of Science and Technology, Republic of Korea), and Ig-Jae Kim (Korea Institute of Science and Technology, Republic of Korea; Yonsei University, Republic of Korea)

Harnessing the Power of Multi-Lingual Datasets for Pre-Training: Towards Enhancing Text Spotting Performance 707
Alloy Das (Indian Statistical Institute), Sanket Biswas (CVC, Universitat Autonoma de Barcelona), Ayan Banerjee (CVC, Universitat Autonoma de Barcelona), Josep Llados (CVC, Universitat Autonoma de Barcelona), Umapada Pal (Indian Statistical Institute), and Saumik Bhattacharya (ECE, Indian Institute of Technology)

Patch-Based Selection and Refinement for Early Object Detection 718
Tianyi Zhang (University of Minnesota), Kishore Kasichainula (Arizona State University), Yaoxin Zhuo (Arizona State University), Baoxin Li (Arizona State University), Jae-Sun Seo (Cornell Tech), and Yu Cao (University of Minnesota)

Boosting Weakly Supervised Object Detection Using Fusion and Priors From Hallucinated Depth 728
Cagri Gungor (University of Pittsburgh) and Adriana Kovashka (University of Pittsburgh)

C2AIR: Consolidated Compact Aerial Image Haze Removal 738
Ashutosh C Kulkarni (Indian Institute of Technology, India), Shruti S Phutke (Institute for Integrated and Intelligent Systems, Griffith Univeristy, Australia), Santosh Kumar Vipparthi (Indian Institute of Technology, India), and Subrahmanyam Murala (Trinity College Dublin, Ireland)

Favoring One Among Equals – Not a Good Idea: Many-to-One Matching for Robust Transformer Based Pedestrian Detection 748
K.N Ajay Shastry (Indian Institute of Technology), Kuriseti Ravi Sri Teja (Indian Institute of Technology), Aditya Nigam (Indian Institute of Technology), and Chetan Arora (Indian Institute of Technology)

Improving Vision-and-Language Reasoning via Spatial Relations Modeling 758
Cheng Yang (Zhejiang University), Rui Xu (Peking University), Ye Guo (Kuaishou Technology), Peixiang Huang (Peking University), Yiru Chen (Kuaishou), Wenkui Ding (Kuaishou), Zhongyuan Wang (Kuaishou), and Hong Zhou (Zhejiang University)

LP-OVOD: Open-Vocabulary Object Detection by Linear Probing 768
Chau Hai Pham (VinAI Research), Truong Tuan Vu (VinAI Research), and Khoi Nguyen (VinAI Research)

Continuous Adaptation for Interactive Segmentation Using Teacher-Student Architecture 778
Barsegh Atanyan (Picsart AI Research), Levon Khachatryan (Picsart AI Research), Shant Navasardyan (Picsart AI Research), Yunchao Wei (BJTU), and Humphrey Shi (Picsart AI Research; Georgia Tech)

Interpretable Object Recognition by Semantic Prototype Analysis 789
Qiyang Wan (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Ruiping Wang (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China)

RecycleNet: Latent Feature Recycling Leads to Iterative Decision Refinement 799
Gregor Köhler (German Cancer Research Center), Tassilo Wald (DKFZ), Constantin Ulrich (German Cancer Research Center (DKFZ)), David Zimmerer (German Cancer Research Center (DKFZ)), Paul Jäger (German Cancer Research Center), Jörg K.H. Franke (University of Freiburg), Simon Kohl (Latent Labs), Fabian Isensee (German Cancer Research Center (DKFZ)), and Klaus H. Maier-Hein (German Cancer Research Center (DKFZ))

Learning to Detour: Shortcut Mitigating Augmentation for Weakly Supervised Semantic Segmentation 808
Junehyoung Kwon (Chung-Ang University, Korea), Eunju Lee (Chung-Ang University, Korea), Yunsung Cho (Chung-Ang University, Korea), and Youngbin Kim (Chung-Ang University, Korea)

Elusive Images: Beyond Coarse Analysis for Fine-Grained Recognition 818
Connor Anderson (Brigham Young University), Matthew A Gwilliam (University of Maryland), Evelyn Gaskin (Brigham Young University), and Ryan Farrell (Brigham Young University)

Understanding Dark Scenes by Contrasting Multi-Modal Observations 829
Xiaoyu Dong (The University of Tokyo, Japan; RIKEN AIP, Japan) and Naoto Yokoya (The University of Tokyo, Japan; RIKEN AIP, Japan)

MaskConver: Revisiting Pure Convolution Model for Panoptic Segmentation 840
Abdullah Rashwan (University of Waterloo), Jiageng Zhang (Google), Ali Taalimi (Google), Fan Yang (Google), Xingyi Zhou (Google), Chaochao Yan (University of Texas at Arlington), Liang-Chieh Chen (ByteDance Inc.), and Yeqing Li (Google)

Masked Collaborative Contrast for Weakly Supervised Semantic Segmentation 851
Fangwen Wu (Zhejiang Lab), Jingxuan He (Zhejiang University), Yufei Yin (University of Science and Technology of China), Yanbin Hao (University of Science and Technology of China), Gang Huang (Zhejiang University), and Lechao Cheng (Zhejiang Lab)

Universal Semi-Supervised Model Adaptation via Collaborative Consistency Training 861
Zizheng Yan (CUHKSZ), Yushuang Wu (CUHKSZ), Yipeng Qin (Cardiff University), Xiaoguang Han (CUHKSZ), Shuguang Cui (CUHKSZ), and Guanbin Li (Sun Yat-sen University)

STEP – Towards Structured Scene-Text Spotting 872
Sergi Garcia Bordils (Computer Vision Centre, Spain; AllRead MLT), Dimosthenis Karatzas (Computer Vision Centre, Spain), and Marçal Rusiñol (AllRead MLT)

Efficient Feature Distillation for Zero-Shot Annotation Object Detection 882
Zhuoming Liu (University of Southern California), Xuefeng Hu (University of Southern California), and Ram Nevatia (University of Southern California)

Hierarchical Text Spotter for Joint Text Spotting and Layout Analysis 892
Shangbang Long (Google Research), Siyang Qin (Google Research), Yasuhisa Fujii (Google Research), Alessandro Bissacco (Google Research), and Michalis Raptis (Google Research)

iBARLE: imBalance-Aware Room Layout Estimation 903
Taotao Jing (Tulane University), Lichen Wang (Zillow Group), Naji Khosravan (Zillow Group), Zhiqiang Wan (Zillow Group), Zachary Bessinger (Zillow Group), Zhengming Ding (Tulane University), and Sing Bing Kang (Zillow Group)

TSP-Transformer: Task-Specific Prompts Boosted Transformer for Holistic Scene Understanding 914
Shuo Wang (Shanghaitech university), Jing Li (Xiaohongshu Inc.), Zibo Zhao (ShanghaiTech University), Dongze Lian (National University of Singapore), Binbin Huang (Shanghaitech University), Xiaomei Wang (Fudan University), Zhengxin Li (ShanghaiTech University), and Shenghua Gao (Shanghaitech University; Shanghai Engineering Research Center of Intelligent Vision and Imaging; Shanghai Engineering Research Center of Energy Efficient and Custom AI IC)

Implicit Neural Representation for Change Detection 924
Peter J Naylor (RIKEN AIP, Japan), Diego Di Carlo (RIKEN, Japan), Arianna Traviglia (Istituto Italiano di Tecnologia (IIT)), Makoto Yamada (RIKEN AIP, Japan; OIST, Japan), and Marco Fiorucci (Istituto Italiano di Tecnologia, Italy)

Label-Free Synthetic Pretraining of Object Detectors 935
Hei Law (Princeton) and Jia Deng (Princeton University)

Improved Techniques for Quantizing Deep Networks with Adaptive Bit-Widths 946
Ximeng Sun (Boston University), Rameswar Panda (MIT-IBM Watson AI Lab; IBM Research), Chun-Fu Richard Chen (MIT-IBM Watson AI Lab; IBM Research), Naigang Wang (IBM Research), Bowen Pan (MIT), Aude Oliva (MIT-IBM Watson AI Lab; MIT), Rogerio Feris (MIT-IBM Watson AI Lab, IBM Research), and Kate Saenko (Boston University; MIT-IBM Watson AI Lab)

What's Outside the Intersection? Fine-Grained Error Analysis for Semantic Segmentation Beyond IoU 957
Maximilian Bernhard (Ludwig-Maximilians-Universität München; MCML), Roberto Amoroso (University of Modena and Reggio Emilia), Yannic Kindermann (Ludwig-Maximilians-Universität München), Lorenzo Baraldi (University of Modena and Reggio Emilia), Rita Cucchiara (Università di Modena e Reggio Emilia), Volker Tresp (Ludwig-Maximilians-Universität München; MCML), and Matthias Schubert (Ludwig-Maximilians-Universität München; MCML)

Pixel Matching Network for Cross-Domain Few-Shot Segmentation 967
Hao Chen (Zhejiang university), Yonghan Dong (Huawei Technologies Ltd.), Zheming Lu (Zhejiang University), Yunlong Yu (Zhejiang University), and Jungong Han (University of Sheffield)

EResFD: Rediscovery of the Effectiveness of Standard Convolution for Lightweight Face Detection 977
Joonhyun Jeong (NAVER Cloud, ImageVision; KAIST), Beomyoung Kim (NAVER Cloud, ImageVision; KAIST), Joonsang Yu (NAVER Cloud, ImageVision; NAVER AI Lab), and Youngjoon Yoo (NAVER Cloud, ImageVision)

Framework-Agnostic Semantically-Aware Global Reasoning for Segmentation 988
Mir Rayat Imtiaz Hossain (University of British Columbia; Vector Institute for AI), Leonid Sigal (University of British Columbia; Vector Institute for AI; Canada CIFAR AI Chair), and James J Little (University of British Columbia)

High-Fidelity Pseudo-Labels for Boosting Weakly-Supervised Segmentation 999
Arvi Jonnarth (Linköping University, Sweden), Yushan Zhang (Linköping University, Sweden), and Michael Felsberg (Linköping University, Sweden)

Missing Modality Robustness in Semi-Supervised Multi-Modal Semantic Segmentation 1009
Harsh Maheshwari (Georgia Institute of Technology; Avataar.ai), Yen-Cheng Liu (Georgia Institute of Technology), and Zsolt Kira (Georgia Institute of Technology)

Unsupervised Graphic Layout Grouping with Transformers 1020
Jialiang Zhu (Southeast University), Danqing Huang (Microsoft Research asia), Chunyu Wang (Microsoft Research asia), Mingxi Cheng (Microsoft Research asia), Ji Li (Microsoft Research asia), Han Hu (Microsoft Research Asia), Xin Geng (Southeast University), and Baining Guo (Microsoft Research asia; Southeast University)

Contrastive Viewpoint-Aware Shape Learning for Long-Term Person Re-Identification 1030
Vuong Nguyen (University of Houston), Khadija Khaldi (University of Houston), Dung Tien Nguyen (Hanoi University of Science and Technology), Pranav Mantini (University of Houston), and Shishir Shah (University of Houston)

PolyMaX: General Dense Prediction with Mask Transformer 1039
Xuan Yang (Google Research), Liangzhe Yuan (Google Research), Kimberly Wilber (Google Research), Astuti Sharma (Google Research), Xiuye Gu (Google Research), Siyuan Qiao (Google Research), Stephanie Debats (Google Research), Huisheng Wang (Google Research), Hartwig Adam (Google Research), Mikhail Sirotenko (Google Research), and Liang-Chieh Chen (Google Research)

BPKD: Boundary Privileged Knowledge Distillation for Semantic Segmentation 1051
Akide Liu (The University of Adelaide, Australia), Zihan Wang (The University of Queensland, Australia; CSIRO's Data61, Australia), Minh Hieu Phan (The University of Adelaide, Australia), Bowen Zhang (The University of Adelaide, Australia), Jinchao Ge (The University of Adelaide, Australia), and Yifan Liu (The University of Adelaide, Australia)

Label Shift Estimation for Class-Imbalance Problem: A Bayesian Approach 1062
Changkun Ye (Australian National University, Australia; Data61, CSIRO, Australia), Russell Tsuchida (Data61, CSIRO, Australia), Lars Petersson (Data61, CSIRO, Australia), and Nick Barnes (Australian National University, Australia)

Query-Guided Attention in Vision Transformers for Localizing Objects Using a Single Sketch..... 1072
Aditay Tripathi (Indian Institute of Science), Anand Mishra (Indian Institute of Technology), and Anirban Chakraborty (Indian Institute of Science)

PromptAD: Zero-Shot Anomaly Detection using Text Prompts 1082
Yiting Li (Institute for Infocomm Research, Singapore), Adam Goodge (Institute for Infocomm Research, Singapore), Fayao Liu (Institute for Infocomm Research, Singapore), and Chuan Sheng Foo (Institute for Infocomm Research, Singapore; Centre for Frontier AI Research, Singapore)

Learning Quality Labels for Robust Image Classification 1092
Xiaosong Wang (Shanghai AI Laboratory), Ziyue Xu (NVIDIA), Dong Yang (NVIDIA Corporation), Leo K Tam (Amicus AI), Holger R Roth (NVIDIA), and Daguang Xu (NVIDIA Corporation)

Learning Transferable Representations for Image Anomaly Localization Using Dense Pretraining 1102
Haitian He (The University of Melbourne), Sarah Erfani (The University of Melbourne), Mingming Gong (The University of Melbourne), and Qiuhong Ke (Monash University)

SBCFormer: Lightweight Network Capable of Full-Size ImageNet Classification at 1 FPS on Single Board Computers 1112
Xiangyong Lu (Tohoku University), Masanori Suganuma (Tohoku University / RIKEN AIP), and Takayuki Okatani (Tohoku University/RIKEN AIP)

High-Fidelity Zero-Shot Texture Anomaly Localization Using Feature Correspondence Analysis.. 1123
Andrei-Timotei Ardelean (Friedrich-Alexander-Universität Erlangen-Nürnberg) and Tim Weyrich (Friedrich-Alexander-Universität Erlangen-Nürnberg)

Grafting Vision Transformers 1134
Jongwoo Park (Stony Brook University), Kumara Kahatapitiya (Stony Brook University), Donghyun Kim (MIT-IBM Watson AI Lab), Shivchander Sudalairaj (MIT-IBM Watson AI Lab), Quanfu Fan (Amazon), and Michael S Ryoo (Stony Brook University)

Rethinking Knowledge Distillation with Raw Features for Semantic Segmentation 1144
Tao Liu (Hikvision Research Institute, China), Chenshu Chen (Hikvision Research Institute, China), Xi Yang (Hikvision Research Institute, China), and Wenming Tan (Hikvision Research Institute, China)

Efficient Expansion and Gradient Based Task Inference for Replay Free Incremental Learning..... 1154
Soumya Roy (Amazon), Vinay K Verma (Amazon), and Deepak Gupta (Amazon)

CLRerNet: Improving Confidence of Lane Detection with LaneIoU 1165
Hiroto Honda (GO Inc.) and Yusuke Uchida (GO Inc.)

Multi-Modal Gaze Following in Conversational Scenarios 1175
Yuqi Hou (University of Birmingham, United Kingdom), Zhongqun Zhang (University of Birmingham, United Kingdom), Nora Horanyi (University of Birmingham, United Kingdom), Jaewon Moon (KETI, Korea), Yihua Cheng (University of Birmingham, United Kingdom), and Hyung Jin Chang (University of Birmingham, United Kingdom)

Enhancing Multi-View Pedestrian Detection Through Generalized 3D Feature Pulling 1185
Sithu Aung (Korea Institute of Science and Technology, Republic of Korea; UST, Republic of Korea), Haesol Park (Korea Institute of Science and Technology, Republic of Korea), Hyungjoo Jung (Korea Institute of Science and Technology, Republic of Korea), and Junghyun Cho (Korea Institute of Science and Technology, Republic of Korea)

Self-Supervised Learning for Visual Relationship Detection Through Masked Bounding Box Reconstruction 1195
Zacharias Anastasakis (National Technical University of Athens), Dimitrios Mallis (CVI2 - University of Luxembourg), Markos Diomataris (ETH Zurich), Georgios Alexandridis (National Tecnhical University of Athens), Stefanos Kollias (NTUA), and Vassilis Pitsikalis (DeepLab)

The Background Also Matters: Background-Aware Motion-Guided Objects Discovery 1205
Sandra Kara (CEA), Hejer Ammar (CEA), Florian Chabot (CEA), and Quoc-Cuong Pham (CEA)

Semi-Supervised Scene Change Detection by Distillation From Feature-Metric Alignment 1215
Seonhoon Lee (KAIST) and Jong-Hwan Kim (KAIST)

OmniVec: Learning Robust Representations with Cross Modal Sharing 1225
Siddharth Srivastava (TensorTour Inc.) and Gaurav Sharma (TensorTour Inc.)

Cross-Attention Between Satellite and Ground Views for Enhanced Fine-Grained Robot Geo-Localization 1238
Dong Yuan (Queensland University of Technology, Australia), Frederic Maire (Queensland University of Technology, Australia), and Feras Dayoub (The University of Adelaide, Australia)

Data Augmentation for Object Detection via Controllable Diffusion Models 1246
Haoyang Fang (AWS AI, US), Boran Han (AWS AI, US), Shuai Zhang (AWS AI, US), Su Zhou (AWS AI, US), Cuixiong Hu (AWS AI, US), and Wenming Ye (AWS AI, US)

Physical-Space Multi-Body Mesh Detection Achieved by Local Alignment and Global Dense Learning 1256
Haoye Dong (Carnegie Mellon University), Tiange Xiang (Stanford University), Sravan Chittupalli (Carnegie Mellon University), Jun Liu (Carnegie Mellon University), and Dong Huang (Carnegie Mellon University)

Multi-Source Domain Adaptation for Object Detection with Prototype-Based Mean Teacher 1266
Atif Belal (ETS Montreal, Canada), Akhil Pilakkatt Meethal (ETS Montreal, Canada), Francisco Perdigon Romero (Ericsson, Canada), Marco Pedersoli (ETS Montreal, Canada), and Eric Granger (ETS Montreal, Canada)

Beyond Self-Attention: Deformable Large Kernel Attention for Medical Image Segmentation 1276
Reza Azad (RWTH Aachen University), Leon Niggemeier (RWTH Aachen University), Michael Hüttemann (RWTH Aachen University), Amirhossein Kazerouni (Iran University of Science and Technology), Ehsan Khodapanah Aghdam (Shahid Beheshti University), Yury Velichko (Northwestern University), Ulas Bagci (Northwestern University), and Dorit Merhof (University of Regensburg)

INCODE: Implicit Neural Conditioning with Prior Knowledge Embeddings 1287
Amirhossein Kazerouni (Iran University of Science and Technology), Reza Azad (RWTH Aachen University), Alireza Hosseini (University of Tehran), Dorit Merhof (University of Regensburg), and Ulas Bagci (Northwestern University)

ProcSim: Proxy-Based Confidence for Robust Similarity Learning 1297
Oriol Barbany (Institut de Robòtica i Informàtica Industrial, CSIC-UPC), Xiaofan Lin (Amazon), Muhammet Bastan (Amazon), and Arnab Dhua (Amazon)

Refine and Redistribute: Multi-Domain Fusion and Dynamic Label Assignment for Unbiased Scene Graph Generation 1307
Yujie Zang (Xi'an Jiaotong University), Yaochen Li (Xi'an Jiaotong University), Yuan Gao (Xi'an Jiaotong University), Yimou Guo (Xi'an Jiaotong University), Wenneng Tang (Xi'an Jiaotong University), Yanxue Li (Xi'an Jiaotong University), and Meklit Mesfin Atlaw (Xian Jiaotong University)

Joint Depth Prediction and Semantic Segmentation With Multi-View SAM 1317
Mykhailo Shvets (University of North Carolina at Chapel Hill), Dongxu Zhao (UNC-Chapel Hill), Marc Niethammer (UNC), Roni Sengupta (University of North Carolina at Chapel Hill), and Alexander C Berg (University of North Carolina, USA)

Self-Supervised Relation Alignment for Scene Graph Generation 1328
Bicheng Xu (University of British Columbia; Vector Institute for AI), Renjie Liao (University of British Columbia; Vector Institute for AI; Canada CIFAR AI Chair), and Leonid Sigal (University of British Columbia; Vector Institute for AI; Canada CIFAR AI Chair)

Semantic Transfer From Head to Tail: Enlarging Tail Margin for Long-Tailed Visual Recognition 1339
Shan Zhang (Australian National University), Yao Ni (Australian National University), Jinhao Du (Peking University), Yanxia Liu (Beijing Union University), and Piotr Koniusz (Australian National University; CSIRO)

PatchRefineNet: Improving Binary Segmentation by Incorporating Signals From Optimal Patch-Wise Binarization 1350
Savinay Nagendra (The Pennsylvania State University) and Daniel Kifer (The Pennsylvania State University)

Adaptive Deep Neural Network Inference Optimization with EENet 1362
Fatih Ilhan (Georgia Institute of Technology, USA), Ka-Ho Chow (Georgia Institute of Technology, USA), Sihao Hu (Georgia Institute of Technology, USA), Tiansheng Huang (georgia Institute of technology, USA), Selim F Tekin (Georgia Institute of Technology, USA), Wenqi Wei (Georgia Institute of Technology, USA), Yanzhao Wu (Georgia Institute of Technology, USA), Myungjin Lee (CISCO Research, USA), Ramana Kompella (CISCO Research, USA), Hugo M Latapie (CISCO Research, USA), Gaowen Liu (CISCO Research, USA), and Ling Liu (Georgia Institute of Technology, USA)

Token Fusion: Bridging the Gap Between Token Pruning and Token Merging 1372
Minchul Kim (Michigan State University), Shangqian Gao (University of Pittsburgh), Yen-Chang Hsu (Samsung Research America), Yilin Shen (Samsung Research America), and Hongxia Jin (Samsung Research America)

Pruning From Scratch via Shared Pruning Module and Nuclear Norm-Based Regularization 1382
Donghyeon Lee (Chungbuk National University), Eunho Lee (Chungbuk National University), and Youngbae Hwang (Chungbuk National University)

CLIP-DIY: CLIP Dense Inference Yields Open-Vocabulary Semantic Segmentation For-Free 1392
Monika Wysoczanska (Warsaw University of Technology), Michael Ramamonjisoa (Meta AI), Tomasz Trzcinski (Warsaw University of Technology; Tooploox; IDEAS NCBR), and Oriane Simeoni (Valeo.ai)

Layer-Wise Auto-Weighting for Non-Stationary Test-Time Adaptation 1403
Junyoung Park (Yonsei University), Jin Kim (Yonsei University), Hyeongjun Kwon (Yonsei University), Ilhoon Yoon (Yonsei university), and Kwanghoon Sohn (Yonsei University; Korea Institute of Science and Technology (KIST))

Uncertainty Estimation in Instance Segmentation with Star-Convex Shapes 1413
Qasim M. K. Siddiqui (Hemholtz-Zentrum Dresden-Rossendorf, Germany), Sebastian Starke (Hemholtz-Zentrum Dresden-Rossendorf, Germany), and Peter Steinbach (Hemholtz-Zentrum Dresden-Rossendorf, Germany)

CamoFocus: Enhancing Camouflage Object Detection With Split-Feature Focal Modulation and Context Refinement 1423
Abbas Khan (Mohamed Bin Zayed University of Artificial Intelligence), Mustaqeem Khan (Mohamed Bin Zayed University of Artificial Intelligence), Wail Gueaieb (Ottawa University, CANADA), Abdulmotaleb Ei Saddik (University of Ottawa), Giulia De Masi (TII), and Fakhri Karray (University of Waterloo)

HalluciDet: Hallucinating RGB Modality for Person Detection Through Privileged Information.... 1433
Heitor Rapela Medeiros (LIVIA, Dept. of Systems Engineering, ETS Montreal, Canada), Fidel A Guerrero Pena (LIVIA, Dept. of Systems Engineering, ETS Montreal, Canada), Masih Aminbeidokhti (LIVIA, Dept. of Systems Engineering, ETS Montreal, Canada), Thomas Dubail (LIVIA, Dept. of Systems Engineering, ETS Montreal, Canada), Eric Granger (LIVIA, Dept. of Systems Engineering, ETS Montreal, Canada), and Marco Pedersoli (LIVIA, Dept. of Systems Engineering, ETS Montreal, Canada)

Spectroformer: Multi-Domain Query Cascaded Transformer Network for Underwater Image Enhancement 1443
Raqib Khan (Trinity College Dublin, Ireland), Priyanka Mishra (Indian Institute of Technology Ropar, India), Nancy Mehta (CAIDAS, IFI, University of Wurzburg, Germany), Shruti S Phutke (Griffith Univeristy, Australia), Santosh Kumar Vipparthi (Indian Institute of Technology Ropar, India), Sukumar Nandi (Indian Institute of Technology Guwahati, India), and Subrahmanyam Murala (Trinity College Dublin, Ireland)

FOSSIL: Free Open-Vocabulary Semantic Segmentation through Synthetic References Retrieval .. 1453
Luca Barsellotti (University of Modena and Reggio Emilia, Italy), Roberto Amoroso (University of Modena and Reggio Emilia, Italy), Lorenzo Baraldi (University of Modena and Reggio Emilia, Italy), and Rita Cucchiara (University of Modena and Reggio Emilia, Italy; IIT-CNR, Italy)

Low-Level & Physics-Based Vision

Self-Supervised Denoising Transformer with Gaussian Process 1463
Rajeev Yasarla (Johns Hopkins University, USA), Jeya Maria Jose Valanarasu (Johns Hopkins University, USA), Vishwanath Sindagi (Johns Hopkins University, USA), and Vishal Patel (Johns Hopkins University, USA)

Scene Text Image Super-Resolution Based on Text-Conditional Diffusion Models 1474
Chihiro Noguchi (Toyota Motor Corporation, Japan), Shun Fukuda (Toyota Motor Corporation, Japan), and Masao Yamanaka (Toyota Motor Corporation, Japan)

Meta-Learned Kernel for Blind Super-Resolution Kernel Estimation 1485
Royson Lee (University of Cambridge, UK; Samsung AI Center, UK), Rui Li (Samsung AI Center, UK), Stylianos Venieris (Samsung AI Center, UK), Timothy Hospedales (Samsung AI Center, UK; Edinburgh University, UK), Ferenc Huszar (University of Cambridge, UK), and Nicholas Lane (University of Cambridge, UK; Flower Labs)

PhISH-Net: Physics Inspired System for High Resolution Underwater Image Enhancement 1495
Aditya Chandrasekar (Indian Institute of Science), Manogna Sreenivas (Indian Institute of Science), and Soma Biswas (Indian Institute of Science)

Leveraging Bitstream Metadata for Fast, Accurate, Generalized Compressed Video Quality Enhancement 1506
Max Ehrlich (NVIDIA), Jon Barker (NVIDIA), Namitha Padmanabhan (University of Maryland), Larry Davis (University of Maryland), Andrew Tao (NVIDIA), Bryan Catanzaro (NVIDIA), and Abhinav Shrivastava (University of Maryland)

Image Denoising and the Generative Accumulation of Photons 1517
Alexander Krull (University of Birmingham, UK; Universities of Nottingham and Birmingham, UK), Hector Basevi (Universities of Nottingham and Birmingham, UK; University of Birmingham, UK), Benjamin Salmon (University of Birmingham, UK), Andre Zeug (Medizinische Hochschule Hannover, Germany), Franziska E. Müller (Medizinische Hochschule Hannover, Germany), Samuel M Tonks (University of Birmingham, UK), Leela S Muppala (The University of Birmingham, UK), and Ales Leonardis (University of Birmingham, UK)

Deep Plug-and-Play Nighttime Non-Blind Deblurring with Saturated Pixel Handling Schemes ... 1527
Hung-Yu Shu (National Taiwan University), Yi-Hsien Lin (National Taiwan University), and Yi-Chang Lu (National Taiwan University)

Best of Both Worlds: Learning Arbitrary-Scale Blind Super-Resolution via Dual Degradation Representations and Cycle-Consistency 1536
Shao-Yu Weng (National Yang Ming Chiao Tung University), Hsuan Yuan (National Yang Ming Chiao Tung University), Yu-Syuan Xu (MediaTek Inc.), Ching-Chun Huang (National Yang Ming Chiao Tung University), and Wei-Chen Chiu (National Chiao Tung University)

ICF-SRSR: Invertible Scale-Conditional Function for Self-Supervised Real-World Single Image Super-Resolution 1546
Reyhaneh Neshatavar (Dept. of ECE & ASRI), Mohsen Yavartanoo (Dept. of ECE & ASRI), Sanghyun Son (Dept. of ECE & ASRI), and Kyoung Mu Lee (Dept. of ECE & ASRI; IPAI, Seoul National University, Korea)

A Neural Height-Map Approach for the Binocular Photometric Stereo Problem 1557
Fotios Logothetis (Toshiba Europe Ltd.), Ignas Budvytis (University of Cambridge, UK), and Roberto Cipolla (University of Cambridge, UK)

Estimating Fog Parameters From an Image Sequence Using Non-Linear Optimisation 1567
Yining Ding (Heriot-Watt University, UK), Andrew Wallace (Heriot Watt University, UK), and Sen Wang (Imperial College London, UK)

4K-Resolution Photo Exposure Correction at 125 FPS with ~8K Parameters 1576
Yijie Zhou (Nankai University), Chao Li (Nankai university), Jin Liang (Nankai University), Tianyi Xu (Nankai University), Xin Liu (Tianjin University; Lappeenranta-Lahti University of Technology), and Jun Xu (Nankai University; Guangdong Provincial Key Laboratory of Big Data Computing)

UGPNet: Universal Generative Prior for Image Restoration 1587
Hwayoon Lee (GENGENAI), Kyoungkook Kang (POSTECH), Hyeongmin Lee (POSTECH), Seung-Hwan Baek (POSTECH), and Sunghyun Cho (POSTECH)

Fully-Automatic Reflection Removal for 360-Degree Images 1598
Jonghyuk Park (Ulsan National Institute of Science and Technology, Republic of Korea), Hyeona Kim (Ulsan National Institute of Science and Technology, Republic of Korea), Eunpil Park (Ulsan National Institute of Science and Technology, Republic of Korea), and Jae-Young Sim (Ulsan National Institute of Science and Technology, Republic of Korea)

PETIT-GAN: Physically Enhanced Thermal Image-Translating Generative Adversarial Network . 1607
Omri Berman (Tel Aviv University, Israel; Agricultural Research Organization - Volcani Institute, Israel), Navot Oz (Tel-Aviv University, Israel; Agricultural Research Organization - Volcani Institute, Israel), David Mendlovic (Tel Aviv University, Israel), Nir Sochen (Tel-Aviv University, Israel), Yafit Cohen (Agricultural Research Organization - Volcani Institute, Israel), and Iftach Klapp (Agricultural Research Organization - Volcani Institute, Israel)

Bridging the Gap Between Multi-Focus and Multi-Modal: A Focused Integration Framework for Multi-Modal Image Fusion 1617
Xilai Li (Foshan University), Xiaosong Li (Foshan University), Tao Ye (China University of Mining and Technology), Xiaoqi Cheng (Foshan University), Wuyang Liu (Foshan University), and Haishu Tan (Foshan University)

BoostRad: Enhancing Object Detection by Boosting Radar Reflections 1627
Yuval Haitman (General Motors, Technical Center, Israel) and Oded Bialer (General Motors, Technical Center, Israel)

RankDVQA: Deep VQA Based on Ranking-Inspired Hybrid Training 1637
Chen Feng (University of Bristol, UK), Duolikun Danier (University of Bristol, UK), Fan Zhang (University of Bristol, UK), and David Bull (University of bristol, UK)

Reference-Based Restoration of Digitized Analog Videotapes 1648
Lorenzo Agnolucci (University of Florence - Media Integration and Communication Center (MICC), Italy), Leonardo Galteri (University of Florence - Media Integration and Communication Center (MICC), Italy), Marco Bertini (University of Florence - Media Integration and Communication Center (MICC), Italy), and Alberto Del Bimbo (University of Florence - Media Integration and Communication Center (MICC), Italy)

Fixed Pattern Noise Removal for Multi-View Single-Sensor Infrared Camera 1658
Arnaud Barral (ENS Paris Saclay, France), Pablo Arias (Universitat Pompeu Fabra, Dept. of Information and Communication Technologies, Spain), and Axel Davy (ENS Paris-Saclay, France)

Machine Learning: Architectures, Formulations, Algorithms II

Efficient Transferability Assessment for Selection of Pre-Trained Detectors 1668
Zhao Wang (The Chinese University of Hong Kong), Aoxue Li (Huawei Noah's Ark Lab), Zhenguo Li (Huawei Noah's Ark Lab), and Dou Qi (The Chinese University of Hong Kong)

Plasticity-Optimized Complementary Networks for Unsupervised Continual Learning 1679
Alex Gomez-Villa (Computer vision center, Spain; Universitat Autonoma de Barcelona, Spain), Bartlomiej Twardowski (Computer vision center, Spain; Universitat Autonoma de Barcelona, Spain; IDEAS NCBR, Poland), Kai Wang (Computer vision center, Spain; Universitat Autonoma de Barcelona, Spain), and Joost Van De Weijer (Computer vision center, Spain; Universitat Autonoma de Barcelona, Spain)

Continual Test-Time Domain Adaptation via Dynamic Sample Selection 1690
Yanshuo Wang (Australia National University; Data61-CSIRO, Australia), Jie Hong (Australia National University; Data61-CSIRO, Australia), Ali Cheraghian (Australia National University; Data61-CSIRO, Australia), Shafin Rahman (North South University, Bangladesh), David Ahmedt-Aristizabal (Data61-CSIRO, Australia), Lars Petersson (Data61-CSIRO, Australia), and Mehrtash Harandi (Data61-CSIRO, Australia; Monash University, Australia)

Source-Guided Similarity Preservation for Online Person Re-Identification 1700
Hamza Rami (Institut Polytechnique de Paris; Atos), Jhony H. Giraldo (Institut Polytechnique de Paris), Nicolas Winckler (Atos), and Stéphane Lathuilière (Institut Polytechnique de Paris)

CAILA: Concept-Aware Intra-Layer Adapters for Compositional Zero-Shot Learning 1710
Zhaoheng Zheng (University of Southern California), Haidong Zhu (University of Southern California), and Ram Nevatia (U of Southern California)

Mini but Mighty: Finetuning ViTs with Mini Adapters 1721
Imad Eddine Marouf (Institut Polytechnique de Paris, France), Enzo Tartaglione (Institut Polytechnique de Paris, France), and Stéphane Lathuilière (Institut Polytechnique de Paris, France)

Recognition of Unseen Bird Species by Learning from Field Guides 1731
Andres C Rodriguez (ETH Zurich, Switzerland), Stefano D'aronco (ETH Zurich, Switzerland), Rodrigo Caye Daudt (ETH Zurich, Switzerland), Jan Dirk Wegner (ETH Zurich, Switzerland; University of Zurich, Switzerland), and Konrad Schindler (ETH Zurich, Switzerland)

Single Domain Generalization via Normalised Cross-Correlation Based Convolutions 1741
Wei Qin Chuah (RMIT University, Australia), Ruwan Tennakoon (RMIT University, Australia), Reza Hoseinnezhad (RMIT University, Australia), David Suter (Edith Cowan University, Australia), and Alireza Bab-Hadiashar (RMIT University, Australia)

MoP-CLIP: A Mixture of Prompt-Tuned CLIP Models for Domain Incremental Learning 1751
Julien Nicolas (ETS Montreal), Florent Chiaroni (Thales Digital Solutions), Imtiaz Masud Ziko (Thales Digital Solutions), Ola Ahmad (Thales Digital Solutions), Christian Desrosiers (ETS Montreal), and Jose Dolz (ETS Montreal)

Self-Supervised Representation Learning with Cross-Context Learning Between Global and Hypercolumn Features 1762
Zheng Gao (Queen Mary University of London), Chen Feng (Queen Mary University of London), and Ioannis Patras (Queen Mary University of London)

Revisiting Pixel-Level Contrastive Pre-Training on Scene Images 1773
Zongshang Pang (Osaka University), Yuta Nakashima (Osaka University), Mayu Otani (CyberAgent, Inc.), and Hajime Nagahara (Osaka University)

Improving Graph Networks Through Selection-Based Convolution 1783
David M Hart (East Carolina University) and Bryan S Morse (Brigham Young University)

Exploiting CLIP for Zero-Shot HOI Detection Requires Knowledge Distillation at Multiple Levels 1794
Bo Wan (KU Leuven) and Tinne Tuytelaars (KU Leuven)

Learning Class and Domain Augmentations for Single-Source Open-Domain Generalization 1805
Prathmesh P Bele (Indian Institute of Technology Bombay, India), Valay M Bundele (Indian Institute of Technology Bombay, India), Avigyan Bhattacharya (Indian Institute of Technology Bombay, India), Ankit Jha (Indian Institute of Technology Bombay, India), Gemma Roig (Goethe University Frankfurt, Germany), and Biplab Banerjee (Indian Institute of Technology Bombay, India)

Dynamic Token-Pass Transformers for Semantic Segmentation 1816
Yuang Liu (East China Normal University), Qiang Zhou (Alibaba Group), Jing Wang (Alibaba Group), Zhibin Wang (Alibaba Group), Fan Wang (Alibaba Group), Jun Wang (ECNU), and Wei Zhang (East China Normal University)

An Analysis of Initial Training Strategies for Exemplar-Free Class-Incremental Learning 1826
Grégoire Petit (Université Paris-Saclay, France; Univ Gustave Eiffel, France), Michaël Soumm (Université Paris-Saclay, France), Eva Feillet (Université Paris-Saclay, France), Adrian Popescu (Université Paris-Saclay, France), Bertrand Delezoide (Amanda, 34 Avenue Des Champs Elysees, France), David Picard (Univ Gustave Eiffel, France), and Céline Hudelot (Université Paris-Saclay, France)

Few-Shot Shape Recognition by Learning Deep Shape-Aware Features 1837
Wenlong Shi (southeast university), Changsheng Lu (The Australian National University), Ming Shao (University of Massachusetts Dartmouth), Yinjie Zhang (Southeast University), Siyu Xia (Southeast University), and Piotr Koniusz (The Australian National University; Data61/CSIRO)

Active Learning for Single-Stage Object Detection in UAV Images 1849
Asma Z Yamani (King Fahd University of Petroleum and Minerals), Albandari Abdullah (kfupm), Hamzah Luqman (King Fahd University of Petroleum & Minerals), Bernard Ghanem (KAUST), and Silvio Giancola (KAUST)

Learning Intra-Class Multimodal Distributions with Orthonormal Matrices 1859
Jumpei Goto (Panasonic Holdings Corporation, Japan), Yohei Nakata (Panasonic Holdings Corporation, Japan), Kiyofumi Abe (Panasonic Holdings Corporation, Japan), Yasunori Ishii (Panasonic Holdings Corporation, Japan), and Takayoshi Yamashita (Chubu University, Japan)

Active Transfer Learning for Efficient Video-Specific Human Pose Estimation 1869
Hiromu Taketsugu (Toyota Technological Institute, Japan) and Norimichi Ukita (Toyota Technological Institute, Japan)

Understanding Hyperbolic Metric Learning Through Hard Negative Sampling 1880
Yun Yue (Worcester Polytechnic Institute, USA), Fangzhou Lin (Worcester Polytechnic Institute, USA; Tohoku University, Japan), Guanyi Mou (Worcester Polytechnic Institute, USA), and Ziming Zhang (Worcester Polytechnic Institute, USA)

Aligning Non-Causal Factors for Transformer-Based Source-Free Domain Adaptation 1893
Sunandini Sanyal (Indian Institute of Science), Ashish Ramayee Asokan (Indian Institute of Science), Suvaansh Bhambri (Indian Institute of Science), Pradyumna Ym (Indian Institute of Science), Akshay R Kulkarni (Indian Institute of Science), Jogendra Nath Kundu (Indian Institute of Science), and Venkatesh Babu Radhakrishnan (Indian Institute of Science)

Diverse Imagenet Models Transfer Better 1903
Niv Nayman (Technion - Israel Institute of Technology), Avram Golbert (Google Research), Asaf Noy (Mobileye), and Lihi Zelnik-Manor (Technion - Israel Institute of Technology)

Design Choices for Enhancing Noisy Student Self-Training 1915
Aswathnarayan Radhakrishnan (Ohio State University), Jim Davis (Ohio State University), Zachary C Rabin (The Ohio State University), Benjamin Lewis (AFRL), Matthew Scherreik (AFRL), and Roman Ilin (Air Force Research Laboratory)

Bag of Tricks for Fully Test-Time Adaptation 1925
Saypraseuth Mounsaveng (ETS Montreal, Canada), Florent Chiaroni (ETS Montreal, Canada), Malik Boudiaf (ÉTS Montreal, Canada), Marco Pedersoli (ÉTS Montreal, Canada), and Ismail Ben Ayed (ÉTS Montreal, Canada)

Gradual Source Domain Expansion for Unsupervised Domain Adaptation 1935
Thomas Westfechtel (The University of Tokyo), Hao-Wei Yeh (The University of Tokyo), Dexuan Zhang (the university of tokyo), and Tatsuya Harada (The University of Tokyo / RIKEN)

OOD Aware Supervised Contrastive Learning 1945
Soroush Seifi (Toyota Motor Europe), Daniel Olmeda Reino (Toyota Motor Europe), Nikolay Chumerin (Toyota Motor Europe), and Rahaf Aljundi (Toyota Motor Europe)

Expanding Hyperspherical Space for Few-Shot Class-Incremental Learning 1956
Yao Deng (Huazhong University of Science and Technology) and Xiang Xiang (Huazhong University of Science and Technology)

Adapt Your Teacher: Improving Knowledge Distillation for Exemplar-Free Continual Learning .. 1966
Filip Szatkowski (Warsaw University of Technlogy; IDEAS NCBR), Mateusz Pyla (IDEAS NCBR; Jagiellonian University), Marcin Przewiezlikowski (IDEAS NCBR; Jagiellonian University), Sebastian Cygert (Gdansk University of Technology; IDEAS-NCBR), Bartlomiej Twardowski (IDEAS-NCBR; Autonomous University of Barcelona; Computer Vision Center), and Tomasz Trzcinski (Warsaw University of Technology; IDEAS NCBR; Tooploox)

Tunable Hybrid Proposal Networks for the Open World 1977
Matthew J Inkawhich (Duke University), Nathan Inkawhich (Air Force Research Laboratory), Hai Li (Duke University), and Yiran Chen (Duke University)

Domain Generalization with Correlated Style Uncertainty 1989
Zheyuan Zhang (Northwestern University, USA), Bin Wang (Northwestern University, USA), Debesh Jha (Northwestern University, USA), Ugur Demir (Northwestern University, USA), and Ulas Bagci (Northwestern University, USA)

Correlation-Aware Active Learning for Surgery Video Segmentation 1999
Fei Wu (University of Bern), Pablo Márquez Neila (University of Bern), Mingyi Zheng (Auris Health. Inc), Hedyeh Rafii-Tari (JNJ), and Raphael Sznitman (University of Bern)

A Multimodal Benchmark and Improved Architecture for Zero Shot Learning 2010
Keval Doshi (Amazon Prime Video), Amanmeet Garg (Amazon Prime Video), Burak Uzkent (Amazon Prime Video), Xiaolong Wang (Amazon Prime Video), and Mohamed Omar (Amazon Prime Video)

TCP: Triplet Contrastive-Relationship Preserving for Class-Incremental Learning 2020
Shiyao Li (Tsinghua University, China), Xuefei Ning (Tsinghua University, China), Shanghang Zhang (Peking University, China), Lidong Guo (Tsinghua University, China), Tianchen Zhao (Tsinghua University, China), Huazhong Yang (Tsinghua University, China), and Yu Wang (Tsinghua University, China)

Improving Normalization with the James-Stein Estimator 2030
Seyedalireza Khoshsirat (University of Delaware) and Chandra Kambhamettu (University of Delaware)

ReConPatch: Contrastive Patch Representation Learning for Industrial Anomaly Detection 2041
Jeeho Hyun (LG AI Research), Sangyun Kim (LG AI Research), Giyoung Jeon (LG AI Research), Seung Hwan Kim (LG AI Research), Kyunghoon Bae (LG AI Research), and Byung Jun Kang (LG AI Research)

REALM: Robust Entropy Adaptive Loss Minimization for Improved Single-Sample Test-Time Adaptation 2051
Skyler Seto (Apple), Barry-John Theobald (Apple), Federico Danieli (Apple), Navdeep Jaitly (Apple), and Dan Busbridge (Apple)

On the Quantification of Image Reconstruction Uncertainty Without Training Data 2061
Jiaxin Zhang (Intuit AI Research), Sirui Bi (Walmart Global Tech), and Victor Fung (Georgia Institute of Technology)

Hyperbolic vs Euclidean Embeddings in Few-Shot Learning: Two Sides of the Same Coin 2071
Gabriel Moreira (Carnegie Mellon University), Manuel Marques (Institute for Systems and Robotics, Instituto Superior Tecnico), Joao Paulo Costeira (Instituto Superior Tecnico), and Alexander Hauptmann (Carnegie Mellon University)

Effective Restoration of Source Knowledge in Continual Test Time Adaptation 2080
Fahim Faisal Niloy (University of California, Riverside), Sk Miraj Ahmed (University of California, Riverside), Dripta S. Raychaudhuri (AWS AI Labs), Samet Oymak (University of Michigan, Ann Arbor), and Amit K. Roy-Chowdhury (University of California, Riverside)

AMEND: Adaptive Margin and Expanded Neighborhood for Efficient Generalized Category Discovery 2090
Anwesha Banerjee (Indian Institute of Science), Liyana Sahir K (Indian Institute of Science), and Soma Biswas (Indian Institute of Science, Bangalore)

Optical Flow Domain Adaptation via Target Style Transfer 2100
Jeongbeen Yoon (POSTECH, South Korea), Sanghyun Kim (POSTECH, South Korea), Suha Kwak (POSTECH, South Korea), and Minsu Cho (POSTECH, South Korea)

Out-of-Distribution Detection With Logical Reasoning 2111
Konstantin Kirchheim (OvGU Magdeburg), Tim Gonschorek (OvGU Magdeburg), and Frank Ortmeier (Herr/Herrn)

Evidential Uncertainty Quantification: A Variance-Based Perspective 2121
Ruxiao Duan (Johns Hopkins University), Brian Caffo (Johns Hopkins University), Harrison Bai (Johns Hopkins University School of Medicine), Haris Sair (Johns Hopkins University School of Medicine), and Craig Jones (Johns Hopkins University)

Self-Supervised Learning of Semantic Correspondence Using Web Videos 2131
Donghyeon Kwon (POSTECH), Minsu Cho (POSTECH), and Suha Kwak (POSTECH)

Opinion Unaware Image Quality Assessment via Adversarial Convolutional Variational Autoencoder 2142
Ankit Shukla (Bennett University), Avinash Upadhyay (Bennett University), Swati Bhugra (IIT Delhi), and Manoj Sharma (Bennett University)

Cheating Depth: Enhancing 3D Surface Anomaly Detection via Depth Simulation 2153
Vitjan Zavrtanik (University of Ljubljana), Matej Kristan (University of Ljubljana), and Danijel Skocaj (University of Ljubljana)

HELA-VFA: A Hellinger Distance-Attention-Based Feature Aggregation Network for Few-Shot Classification 2162
Gao Yu Lee (Nanyang Technological University, Singapore), Tanmoy Dam (Nanyang Technological University, Singapore), Daniel Puiu Poenar (Nanyang Technological University, Singapore), Vu Duong (Nanyang Technological University, Singapore), and Md Meftahul Ferdaus (University of New Orleans, USA)

Late to the Party? On-Demand Unlabeled Personalized Federated Learning 2173
Ohad Amosy (Bar Ilan University, Israel), Gal Eyal (Bar Ilan University, Israel), and Gal Chechik (NVIDIA Research, Israel)

GazeGNN: A Gaze-Guided Graph Neural Network for Chest X-Ray Classification 2183
Bin Wang (Northwestern University), Hongyi Pan (Northwestern University), Armstrong Aboah (University of Arizona), Zheyuan Zhang (Northwestern University), Elif Keles (Northwestern University), Drew Torigian (University of Pennsylvania), Baris Turkbey (National Cancer Institute), Elizabeth Krupinski (Emory University), Jayaram Udupa (University of Pennsylvania), and Ulas Bagci (Northwestern University)

Towards Better Structured Pruning Saliency by Reorganizing Convolution 2193
Xinglong Sun (Stanford University; Georgia Tech & UIUC & U of Oregon) and Humphrey Shi (Georgia Tech & UIUC & U of Oregon; Picsart AI Research (PAIR))

Domain Generalization by Rejecting Extreme Augmentations 2204
Masih Aminbeidokhti (École de technologie supérieure, Canada), Fidel A Guerrero Pena (École de technologie supérieure, Canada), Heitor Rapela Medeiros (École de technologie supérieure, Canada), Thomas Dubail (École de technologie supérieure, Canada), Eric Granger (École de technologie supérieure, Canada), and Marco Pedersoli (École de technologie supérieure, Canada)

Wakening Past Concepts Without Past Data: Class-Incremental Learning From Online Placebos .. 2215
Yaoyao Liu (Johns Hopkins University; Max Planck Institute for Informatics, Saarland Informatics Campus), Yingying Li (University of Illinois Urbana-Champaign), Bernt Schiele (MPI Informatics), and Qianru Sun (Singapore Management University)

MICS: Midpoint Interpolation to Learn Compact and Separated Representations for Few-Shot Class-Incremental Learning 2225
Solang Kim (Ulsan National Institute of Science and Technology (UNIST), Republic of Korea), Yuho Jeong (Nuvilab Inc., Republic of Korea), Joon Sung Park (Ulsan National Institute of Science & Technology UNIST), Republic of Korea), and Sung Whan Yoon (Ulsan National Institute of Science and Technology (UNIST), Republic of Korea)

Group-Wise Contrastive Bottleneck for Weakly-Supervised Visual Representation Learning 2235
Boon Peng Yap (Nanyang Technological University, Singapore) and Beng Koon Ng (Nanyang Technological University, Singapore)

Training-Based Model Refinement and Representation Disagreement for Semi-Supervised Object Detection 2245
S. Mojtaba Marvasti Zadeh (University of Alberta), Nilanjan Ray (University of Alberta), and Nadir Erbilgin (University of Alberta)

DPPMask: Masked Image Modeling with Determinantal Point Processes 2255
Junde Xu (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Zhejiang lab), Zikai Lin (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Donghao Zhou (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Yaodong Yang (The Chinese University of Hong Kong), Xiangyun Liao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Qiong Wang (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Bian Wu (Zhejiang lab), Guangyong Chen (Zhejiang Lab), and Pheng-Ann Heng (The Chinese University of Hong Kong)

Frequency Attention for Knowledge Distillation 2266
Cuong Van Pham (Monash University, Australia), Van Anh Nguyen (Monash University, Australia), Trung Le (Monash University, Australia), Dinh Q Phung (Monash University, Australia), Gustavo Carneiro (University of Surrey, United Kingdom), and Thanh-Toan Do (Monash University, Australia)

Minimizing Layerwise Activation Norm Improves Generalization in Federated Learning 2276
M Yashwanth (Indian Institute of Science), Gaurav Kumar Nayak (University of Central Florida), Harsh Rangwani (Indian Institute of Science), Arya Singh (BITS Pilani), Venkatesh Babu Radhakrishnan (Indian Institute of Science), and Anirban Chakraborty (Indian Institute of Science)

Adaptive Manifold for Imbalanced Transductive Few-Shot Learning 2286
Michalis Ml Lazarou (Imperial College London), Yannis Avrithis (IARAI), and Tania Stathaki (Imperial College London)

Cross-Domain Few-Shot Incremental Learning for Point-Cloud Recognition 2296
Yuwen Tan (Huazhong University of Science and Technology) and Xiang Xiang (Huazhong University of Science and Technology)

Randomized Adversarial Style Perturbations for Domain Generalization 2306
Taehoon Kim (Seoul National University) and Bohyung Han (Seoul National University)

Shape-Biased CNNs are Not Always Superior in Out-of-Distribution Robustness 2315
Xinkuan Qiu (Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Meina Kan (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Yongbin Zhou (Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China; Nanjing University of Science and Technology, China), Yanchao Bi (Beijing Normal University, China), and Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China; Peng Cheng Laboratory, China)

Partial Binarization of Neural Networks for Budget-Aware Efficient Learning 2325
Udbhav Bamba (Transmute AI Lab, India), Neeraj Anand (Nyun AI, India), Saksham Aggarwal (Transmute AI Lab, India), Dilip K Prasad (UiT The Arctic University of Norway, Norway), and Deepak K Gupta (Transmute AI Lab, India)

Monocular 3D Object Detection with LiDAR Guided Semi Supervised Active Learning 2335
Aral Hekimoglu (Techincal University of Munich, Germany), Michael Schmidt (BMW Group, Germany), and Alvaro Marcos-Ramiro (BMW Group, Germany)

Improving Open-Set Semi-Supervised Learning with Self-Supervision 2345
Erik Wallin (Saab AB; Chalmers University of Technology), Lennart Svensson (Chalmers University of Technology), Fredrik Kahl (Chalmers University of Technology), and Lars Hammarstrand (Chalmers university of technology)

Evolve: Enhancing Unsupervised Continual Learning with Multiple Experts 2355
Xiaofan Yu (University of California San Diego), Tajana Rosing (University of California San Diego), and Yunhui Guo (University of Texas at Dallas)

Masked Event Modeling: Self-Supervised Pretraining for Event Cameras 2367
Simon Klenk (Technical University of Munich; Munich Center for Machine Learning), David Bonello (Technical University of Munich), Lukas Koestler (Technical University of Munich; Munich Center for Machine Learning), Nikita Araslanov (Technical University of Munich; Munich Center for Machine Learning), and Daniel Cremers (Technical University of Munich; Munich Center for Machine Learning)

Overcoming Catastrophic Forgetting for Multi-Label Class-Incremental Learning 2378
Xiang Song (XI'AN JIAOTONG UNIVERSITY), Kuang Shu (Southern University of Science and Technology), Songlin Dong (Xi'an Jiaotong University), Jie Cheng (Huawei Technology Co., Ltd.), Xing Wei (Xi'an Jiaotong University), and Yihong Gong (Xi'an Jiaotong University)

Unsupervised Domain Adaptation for Semantic Segmentation with Pseudo Label Self-Refinement..... 2388
Xingchen Zhao (Northeastern University, USA), Niluthpol C Mithun (SRI International, USA), Abhinav Rajvanshi (SRI International, USA), Han-Pang Chiu (SRI International, USA), and Supun Samarasekera (SRI International, USA)

HyperMix: Out-of-Distribution Detection and Classification in Few-Shot Settings 2399
Nikhil Mehta (Duke University; Meta), Kevin J Liang (Meta), Jing Huang (Meta), Fu-Jen Chu (Meta), Li Yin (Meta), and Tal Hassner (Meta)

PrivObfNet: A Weakly Supervised Semantic Segmentation Model for Data Protection 2410
Chiat Pin Tay (ASTAR), Vigneshwaran Subbaraju (ASTAR Singapore), and Thivya Kandappu (Singapore Management University, Singapore)

DISCO: Distributed Inference With Sparse Communications 2421
Minghai Qin (Western Digital Research), Chao Sun (Western Digital Corp.), Jaco Hofmann (Western Digital), and Dejan Vucinic (Western Digital Corp.)

Debiasing, Calibrating, and Improving Semi-Supervised Learning Performance via Simple Ensemble Projector 2430
Khanh-Binh Nguyen (Sungkyunkwan University, South Korea)

Reducing the Side-Effects of Oscillations in Training of Quantized YOLO Networks 2440
Kartik Gupta (Seeing Machines, Australia) and Akshay Asthana (seeing machines, Australia)

Robust Unsupervised Domain Adaptation Through Negative-View Regularization 2450
Joonhyeok Jang (Korea Advanced Institute of Science and Technology), Sunhyeok Lee (Korea Advanced Institute of Science and Technology), Seonghak Kim (Korea Advanced Institute of Science and Technology), Jung-un Kim (Shinsegae I&C), Seonghyun Kim (SHINSEGAE I&C), and Daeshik Kim (Korea Advanced Institute of Science and Technology)

Link Prediction for Flow-Driven Spatial Networks 2460
Bastian Wittmann (University of Zurich), Johannes C. Paetzold (Technical University of Munich), Chinmay Prabhakar (University of Zurich), Daniel Rueckert (Technische Universität München), and Bjoern Menze (University of Zurich)

FLORA: Fine-Grained Low-Rank Architecture Search for Vision Transformer 2470
Chi-Chih Chang (National Yang Ming Chiao Tung University), Yuan-Yao Sung (National Yang-Ming Chiao-Tung University), Shixing Yu (University of Texas at Austin), Ning-Chi Huang (National Chiao Tung University), Diana Marculescu (The University of Texas at Austin), and Kai-Chiang Wu (National Chiao Tung University)

CL-MAE: Curriculum-Learned Masked Autoencoders 2480
Neelu Madan (Aalborg University, Denmark), Nicolae Catalin Ristea (University Politehnica of Bucharest, Romania; University of Bucharest, Romania), Kamal Nasrollahi (Aalborg University, Denmark; Milestone Systems, Denmark), Thomas B. Moeslund (Aalborg University, Denmark), and Radu Tudor Ionescu (University of Bucharest, Denmark; SecurifAI, Romania)

Active Learning with Task Consistency and Diversity in Multi-Task Networks 2491
Aral Hekimoglu (Techincal University of Munich, Germany), Michael Schmidt (BMW Group, Germany), and Alvaro Marcos-Ramiro (BMW Group, Germany)

Enhancing Diverse Intra-Identity Representation for Visible-Infrared Person Re-Identification 2501
Sejun Kim (Seokyeong University, Korea), Soonyong Gwon (Seokyeong University, Korea), and Kisung Seo (Seokyeong University, Korea)

Steering Prototypes with Prompt-Tuning for Rehearsal-Free Continual Learning 2511
Zhuowei Li (Rutgers University), Long Zhao (Google Research), Zizhao Zhang (Google Cloud AI), Han Zhang (Google Research), Di Liu (Rutgers University), Ting Liu (Google Research), and Dimitris N. Metaxas (Rutgers University)

Active Batch Sampling for Multi-Label Classification with Binary User Feedback 2522
Debanjan Goswami (Florida State University) and Shayok Chakraborty (Florida State University)

M^33D: Learning 3D Priors Using Multi-Modal Masked Autoencoders for 2D Image and Video Understanding 2532
Muhammad Abdullah Jamal (Intuitive Surgical Inc.) and Omid Mohareri (Intuitive Surgical Inc.)

Universal Test-Time Adaptation through Weight Ensembling, Diversity Weighting, and Prior Correction 2543
Robert Marsden (University of Stuttgart), Mario Döbler (University of Stuttgart), and Bin Yang (University of Stuttgart)

SLoSH: Set Locality Sensitive Hashing via Sliced-Wasserstein Embeddings 2554
Yuzhe Lu (Carnegie Mellon University), Xinran Liu (Vanderbilt University), Andrea Soltoggio (Loughborough University), and Soheil Kolouri (Vanderbilt University)

D3GU: Multi-Target Active Domain Adaptation via Enhancing Domain Alignment 2565
Lin Zhang (Robotics Institute, Carnegie Mellon University), Linghan Xu (Robotics Institute, Carnegie Mellon University), Saman Motamed (Robotics Institute, Carnegie Mellon University), Shayok Chakraborty (Florida State University), and Fernando De La Torre (Robotics Institute, Carnegie Mellon University)

MetaVers: Meta-Learned Versatile Representations for Personalized Federated Learning 2575
Jin Hyuk Lim (Ulsan National Institute of Science and Technology), Seungbum Ha (Ulsan National Institute of Science and Technology), and Sung Whan Yoon (Ulsan National Institute of Science and Technology (UNIST))

Instruct Me More! Random Prompting for Visual In-Context Learning 2585
Jiahao Zhang (Osaka University), Bowen Wang (Osaka University), Liangzhi Li (Osaka University), Yuta Nakashima (Osaka University), and Hajime Nagahara (Osaka University)

SimA: Simple Softmax-Free Attention for Vision Transformers 2595
Soroush Abbasi Koohpayegani (University of California) and Hamed Pirsiavash (University of California)

Guided Cluster Aggregation: A Hierarchical Approach to Generalized Category Discovery 2606
Jona Otholt (Hasso Plattner Institute), Christoph Meinel (Hasso Plattner Institute, Potsdam Germany), and Haojin Yang (Hasso-Plattner-Institut für Digital Engineering gGmbH)

Hardware Aware Evolutionary Neural Architecture Search using Representation Similarity Metric 2616
Nilotpal Sinha (SnT, University of Luxembourg), Abd El Rahman Shabayek (SnT, University of Luxembourg), Anis Kacem (SnT, University of Luxembourg), Peyman Rostami (SnT, University of Luxembourg), Carl Shneider (SnT, University of Luxembourg), and Djamila Aouada (SnT, University of Luxembourg)

Using Early Readouts to Mediate Featural Bias in Distillation 2626
Rishabh Tiwari (Google Research, India), Durga Sivasubramanian (Indian Institute of Technology Bombay), Anmol Reddy Mekala (University of Massachusetts Amherst), Ganesh Ramakrishnan (Indian Institute of Technology Bombay), and Pradeep Shenoy (Google Research, India)

Gradient Coreset for Federated Learning 2636
Durga Sivasubramanian (Indian Institute of Technology Bombay), Lokesh Nagalapatti (Indian Institute of Technology Bombay), Rishabh Iyer (University of Texas at Dallas), and Ganesh Ramakrishnan (Indian Institute of Technology Bombay)

Revisiting Token Pruning for Object Detection and Instance Segmentation 2646
Yifei Liu (University of Zurich, Switzerland), Mathias Gehrig (University of Zurich, Switzerland), Nico Messikommer (University of Zurich, Switzerland), Marco Cannici (University of Zurich, Switzerland), and Davide Scaramuzza (University of Zurich, Switzerland)

LatentDR: Improving Model Generalization Through Sample-Aware Latent Degradation and Restoration 2657
Ran Liu (Georgia Institute of Technology), Sahil S Khose (Georgia Institute of Technology), Jingyun Xiao (Georgia Institute of Technology), Lakshmi Sathidevi (Georgia Institute of Technology), Keerthan Ramnath (Georgia Institute of Technology), Zsolt Kira (Georgia Institute of Technology), and Eva L. Dyer (Georgia Institute of Technology)

Fixing Overconfidence in Dynamic Neural Networks 2668
Lassi Meronen (Aalto University; Saab Finland Oy), Martin Trapp (Aalto University), Andrea Pilzer (NVIDIA), Le Yang (Xi'an Jiaotong University), and Arno Solin (Aalto University)

Empowering Unsupervised Domain Adaptation With Large-Scale Pre-Trained Vision-Language Models 2679
Zhengfeng Lai (University of California, Davis), Haoping Bai (Apple), Haotian Zhang (Apple AI/ML), Xianzhi Du (Google Brain), Jiulong Shan (Apple Inc.), Yinfei Yang (Apple), Chen-Nee Chuah (University of California Davis), and Meng Cao (Apple)

pSTarC: Pseudo Source Guided Target Clustering for Fully Test-Time Adaptation 2690
Manogna Sreenivas (Indian Institute of Science), Goirik Chakrabarty (Indian Institute of Science, Education and Research), and Soma Biswas (Indian Institute of Science)

Torque Based Structured Pruning for Deep Neural Network 2699
Arshita Gupta (Samsung Research America), Tien C Bau (Samsung Research America), Joonsoo Kim (Samsung Research America), Zhe Zhu (Samsung Research America), Sumit Jha (Samsung Research America), and Hrishikesh Garud (Google)

Meta-Learned Attribute Self-Interaction Network for Continual and Generalized Zero-Shot Learning 2709
Vinay Verma (Duke University), Nikhil Mehta (Duke University), Kevin J Liang (FAIR, Meta), Aakansha Mishra (IITG), and Lawrence Carin Duke (KAUST)

Letting 3D Guide the Way: 3D Guided 2D Few-Shot Image Classification 2720
Jiajing Chen (Syracuse University, USA), Minmin Yang (Syracuse University, USA), and Senem Velipasalar (Syracuse University, USA)

Robust Learning via Conditional Prevalence Adjustment 2729
Minh Nguyen (Cornell University), Alan Q Wang (Cornell University), Heejong Kim (Weill Cornell Medicine), and Mert Sabuncu (Cornell University; Weill Cornell Medicine)

Learning to Compose SuperWeights for Neural Parameter Allocation Search 2739
Piotr Teterwak (Boston University), Soren Nelson (Boston University), Nikoli Dryden (ETH Zurich), Dina Bashkirova (Boston University), Kate Saenko (Boston University), and Bryan Plummer (Boston University)

Masking Improves Contrastive Self-Supervised Learning for ConvNets, and Saliency Tells You Where 2749
Zhi-Yi Chin (National Yang Ming Chiao Tung University), Chieh-Ming Jiang (National Yang Ming Chiao Tung University), Ching-Chun Huang (National Yang Ming Chiao Tung University), Pin-Yu Chen (IBM Research), and Wei-Chen Chiu (National Chiao Tung University)

Critical Gap Between Generalization Error and Empirical Error in Active Learning 2759
Yusuke Kanebako (Ricoh)

Appearance-Based Curriculum for Semi-Supervised Learning with Multi-Angle Unlabeled Data . 2768
Yuki Tanaka (NEC Corporation), Shuhei M Yoshida (NEC Corporation), Takashi Shibata (NEC Corporation), Makoto Terao (NEC Corporation), Takayuki Okatani (Tohoku University/RIKEN AIP), and Masashi Sugiyama (RIKEN AIP/The University of Tokyo)

Domain Generalisation via Risk Distribution Matching 2778
Toan Nguyen (Deakin University, Australia), Kien Duc Do (Deakin Unviersity, Australia), Bao Duong (Deakin University, Australia), and Thin Nguyen (Deakin University, Australia)

MixtureGrowth: Growing Neural Networks by Recombining Learned Parameters 2788
Chau Pham (Boston University), Piotr Teterwak (Boston University), Soren Nelson (Physical Sciences Inc), and Bryan Plummer (Boston University)

Increasing Biases Can Be More Efficient Than Increasing Weights 2798
Carlo Metta (ISTI CNR), Marco Fantozzi (None), Andrea Papini (Scuola Normale Superiore Pisa), Gianluca Amato (Università di Chieti-Pescara), Matteo Bergamaschi (Università di Padova), Silvia G Galfre (University of Pisa), Alessandro Marchetti (Università di Chieti-Pescara), Michelangelo Vegliò (Università di Chieti-Pescara), Maurizio Parton (Università di Chieti-Pescara), and Francesco Morandin (Università di Parma)

Deep Subdomain Alignment for Cross-Domain Image Classification 2808
Yewei Zhao (Institute of Computing Technology, Chinese Academy of Sciences), Hu Han (Institute of Computing Technology, Chinese Academy of Sciences), Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences), and Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences)

Generalization by Adaptation: Diffusion-Based Domain Extension for Domain-Generalized Semantic Segmentation 2818
Joshua Niemeijer (DLR), Manuel Schwonberg (CARIAD SE), Jan-Aike Termöhlen (Technische Universität Braunschweig), Nico M Schmidt (CARIAD SE), and Tim Fingscheidt (Technische Universität Braunschweig)

Kaizen: Practical Self-Supervised Continual Learning with Continual Fine-Tuning 2829
Chi Ian Tang (Nokia Bell Labs, UK; University of Cambridge, UK), Lorena Qendro (Nokia Bell Labs, UK), Dimitris Spathis (Nokia Bell Labs, UK), Fahim Kawsar (Nokia Bell Labs, UK), Cecilia Mascolo (University of Cambridge, UK), and Akhil Mathur (Nokia Bell Labs, UK)

Learning Low-Rank Latent Spaces with Simple Deterministic Autoencoder: Theoretical and Empirical Insights 2839
Alokendu Mazumder (Indian Institute of Science), Tirthajit Baruah (Indian Institute of Science), Bhartendu Kumar (TCS Research), Rishab Rajesh Sharma (Dayananda Sagar College of Engineering), Vishwajeet Pattanaik (Indian Institute of Science), and Punit Rathore (Indian Institute of Science)

CycleCL: Self-Supervised Learning for Periodic Videos 2849
Matteo Destro (Cerrion, Inc.) and Michael Gygli (Cerrion, Inc.)

Textron: Weakly Supervised Multilingual Text Detection Through Data Programming 2859
Dhruv Kudale (IIT Bombay, India), Badri Vishal Kasuba (IIT Bombay, India), Venkatapathy Subramanian (IITB, India), Parag Chaudhuri (IIT Bombay, India), and Ganesh Ramakrishnan (IIT Bombay, India)

Beyond Active Learning: Leveraging the Full Potential of Human Interaction via Auto-Labeling, Human Correction, and Human Verification 2869
Nathan A Beck (University of Texas), Krishnateja Killamsetty (University of Texas), Suraj Kothawade (University of Texas), and Rishabh Iyer (University of Texas)

Generative Models; Visualization

TIAM – A Metric for Evaluating Alignment in Text-to-Image Generation 2878
Paul Grimal (Universite Paris-Saclay, CEA, France), Herve Le Borgne (Universite Paris-Saclay, CEA, France), Olivier Ferret (Universite Paris-Saclay, CEA, France), and Julien Tourille (Universite Paris-Saclay, CEA, France)

Controlling Rate, Distortion, and Realism: Towards a Single Comprehensive Neural Image Compression Model 2888
Shoma Iwai (Tohoku University, Japan), Tomo Miyazaki (Tohoku University, Japan), and Shinichiro Omachi (Tohoku University, Japan)

Improving the Fairness of the Min-Max Game in GANs Training 2898
Zhaoyu Zhang (Queen's University Belfast), Yang Hua (Queen's University Belfast), Hui Wang (Queen's University Belfast), and Seán F Mcloone (Queen's University Belfast)

Synthesizing Coherent Story with Auto-Regressive Latent Diffusion Models 2908
Xichen Pan (New York University), Pengda Qin (Alibaba Group), Yuhong Li (Alibaba Group), Hui Xue (Alibaba Group), and Wenhu Chen (University of Waterloo)

Textual Alchemy: CoFormer for Scene Text Understanding 2919
Gayatri S Deshmukh (Independent Researcher), Onkar Susladkar (Independent Researcher), Dhruv Makwana (Independent Researcher), Sparsh Mittal (IIT Roorkee), and Sai Chandra Teja R (Green PMU Semi Pvt Ltd)

3D; Adversarial & Defense; Explainable, Fair, Privacy, & Ethical; Vision + X

Synergizing Contrastive Learning and Optimal Transport for 3D Point Cloud Domain Adaptation 2930
Siddharth Katageri (International Institute of Information Technology, India), Arkadipta De (Fujitsu Research of India), Chaitanya Devaguptapu (Fujitsu Research of India), Vssvprasad Kasu (Fujitsu Research of India), Charu Sharma (International Institute of Information Technology, India), and Manohar Kaul (Fujitsu Research of India)

Assessing Neural Network Robustness via Adversarial Pivotal Tuning 2940
Peter Ebert Christensen (University of Copenhagen), Vésteinn Snæbjarnarson (University of Copenhagen), Andrea Dittadi (Technical University of Denmark), Serge Belongie (University of Copenhagen), and Sagie Benaim (Hebrew University of Jerusalem)

Beyond Document Page Classification: Design, Datasets, and Challenges 2950
Jordy Van Landeghem (KU Leuven, Contract.fit), Sanket Biswas (Computer Vision Centre, Universitat Autonoma de Barcelona), Matthew B. Blaschko (KU Leuven), and Marie-Francine Moens (KU Leuven)

InfraParis: A Multi-Modal and Multi-Task Autonomous Driving Dataset 2961
Gianni Franchi (ENSTA Paris), Marwane Hariat (ENSTA Paris), Xuanlong Yu (SATIE, Paris-Saclay University, ENSTA Paris), Nacim Belkhir (Safrantech, Safran Group), Antoine Manzanera (ENSTA Paris), and David Filliat (ENSTA Paris)

Discovering and Mitigating Biases in CLIP-Based Image Editing 2972
Md Mehrab Tanjim (Adobe Research), Krishna Kumar Singh (Adobe Research), Kushal Kafle (Adobe Research), Ritwik Sinha (Adobe Research), and Garrison W Cottrell (University of California)

ReCLIP: Refine Contrastive Language Image Pre-Training with Source Free Domain Adaptation . 2982
Xuefeng Hu (University of Southern California), Ke Zhang (Amazon), Lu Xia (Amazon), Albert Y Chen (Amazon), Jiajia Luo (Amazon), Yuyin Sun (Amazon), Ken Wang (Amazon), Nan Qiao (Amazon), Xiao Zeng (Amazon), Min Sun (NTHU), Cheng-Hao Kuo (Amazon), and Ram Nevatia (University of Southern California)

UNSPAT: Uncertainty-Guided SpatioTemporal Transformer for 3D Human Pose and Shape Estimation on Videos 2992
Minsoo Lee (LG AI Research), Hyunmin Lee (POSTECH), Bumsoo Kim (LG AI Research), and Seung Hwan Kim (LG AI Research)

Visually Guided Audio Source Separation with Meta Consistency Learning 3002
Md Amirul Islam (Huawei Noah's Ark Lab), Seyed Shahabeddin Nabavi (Huawei Noah's Ark Lab), Irina Kezele (Huawei Noah's Ark Lab), Yang Wang (Concordia University), Yuanhao Yu (Huawei Noah's Ark Lab), and Jin Tang (Huawei Noah's Ark Lab)

Sign Language Production With Latent Motion Transformer 3012
Pan Xie (Beihang University), Peng Taiying (Beihang University), Yao Du (Beihang University), and Qipeng Zhang (Beihang University)

3D Computer Vision

Task-Oriented Human-Object Interactions Generation with Implicit Neural Representations 3023
Quanzhou Li (Nanyang Technological University), Jingbo Wang (Shanghai AI Laboratory), Chen Change Loy (Nanyang Technological University), and Bo Dai (Shanghai AI Laboratory)

Depth from Asymmetric Frame-Event Stereo: A Divide-and-Conquer Approach 3033
Xihao Chen (University of Science and Technology of China), Wenming Weng (University of Science and Technology of China), Yueyi Zhang (University of Science and Technology of China), and Zhiwei Xiong (University of Science and Technology of China)

TriPlaneNet: An Encoder for EG3D Inversion 3043
Ananta R Bhattarai (Technical University of Munich), Matthias Niessner (Technical University of Munich), and Artem Sevastopolskiy (Technical University of Munich)

Attentive Prototypes for Source-Free Unsupervised Domain Adaptive 3D Object Detection 3054
Deepti B Hegde (Johns Hopkins University) and Vishal Patel (Johns Hopkins University)

FIRe: Fast Inverse Rendering using Directional and Signed Distance Functions 3065
Tarun Yenamandra (TU Munich, MCML), Ayush Tewari (Massachusetts Institute of Technology), Nan Yang (TU Munich, MCML), Florian Bernard (University of Bonn), Christian Theobalt (MPI Informatik, SIC), and Daniel Cremers (TU Munich, MCML)

A Generic and Flexible Regularization Framework for NeRFs 3076
Thibaud Ehret (Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, France), Roger Marí (Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, France), and Gabriele Facciolo (Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, France)

Multi-View 3D Object Reconstruction and Uncertainty Modelling with Neural Shape Prior 3086
Ziwei Liao (University of Toronto) and Steven L Waslander (University of Toronto)

Neural Textured Deformable Meshes for Robust Analysis-by-Synthesis 3096
Angtian Wang (Johns Hopkins University), Wufei Ma (Johns Hopkins University), Alan Yuille (Johns Hopkins University), and Adam Kortylewski (2University of Freiburg; Max Planck Institute for Informatics)

Ray Deformation Networks for Novel View Synthesis of Refractive Objects 3106
Weijian Deng (Australian National University), Dylan Campbell (Australian National University), Chunyi Sun (Australian National University), Shubham S Kanitkar (RIOS Intelligent Machines), Matthew E Shaffer (RIOS Intelligent Machines), and Stephen Gould (Australian National University)

Registered and Segmented Deformable Object Reconstruction From a Single View Point Cloud ... 3117
Pit Henrich (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Balazs Gyenes (Karlsruhe Institute of Technology, Germany), Paul M Scheikl (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Gerhard Neumann (Karlsruhe Institute of Technology, Germany), and Franziska Mathis-Ullrich (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)

SupeRVol: Super-Resolution Shape and Reflectance Estimation in Inverse Volume Rendering 3127
Mohammed Brahimi (Technical University of Munich; Munich Center for Machine Learning), Bjoern Haefner (Technical University of Munich; Munich Center for Machine Learning), Tarun Yenamandra (Technical University of Munich; Munich Center for Machine Learning), Bastian Goldluecke (University of Konstanz), and Daniel Cremers (Technical University of Munich; Munich Center for Machine Learning)

RIMeshGNN: A Rotation-Invariant Graph Neural Network for Mesh Classification 3138
Bahareh Shakibajahromi (Drexel University, USA), Edward Kim (Drexel University, USA), and David E. Breen (Drexel University, USA)

OptFlow: Fast Optimization-Based Scene Flow Estimation without Supervision 3149
Rahul Ahuja (ISEE AI), Chris Baker (ISEE AI), and Wilko Schwarting (ISEE AI)

Point-DynRF: Point-Based Dynamic Radiance Fields from a Monocular Video 3159
Byeongjun Park (KAIST) and Changick Kim (KAIST)

LensNeRF: Rethinking Volume Rendering Based on Thin-Lens Camera Model 3170
Min-Jung Kim (Korea Advanced Institute of Science and Technology), Gyojung Gu (Korea Advanced Institute of Science and Technology), and Jaegul Choo (Korea Advanced Institute of Science and Technology)

Domain Adaptive 3D Shape Retrieval From Monocular Images 3180
Harsh Pal (Indian Institute of Technology Bombay), Ritwik Khandelwal (IIT Bombay), Shivam Pande (IIT Bombay), Biplab Banerjee (Indian Institute of Technology, Bombay), and Srikrishna Karanam (Adobe Research)

HD-Fusion: Detailed Text-to-3D Generation Leveraging Multiple Noise Estimation 3190
Jinbo Wu (Department of Computer Vision Technology (VIS), Baidu Inc., China), Xiaobo Gao (Department of Computer Vision Technology (VIS), Baidu Inc., China), Xing Liu (Department of Computer Vision Technology (VIS), Baidu Inc., China), Zhengyang Shen (Department of Computer Vision Technology (VIS), Baidu Inc., China), Chen Zhao (Department of Computer Vision Technology (VIS), Baidu Inc., China), Haocheng Feng (Department of Computer Vision Technology (VIS), Baidu Inc., China), Jingtuo Liu (Department of Computer Vision Technology (VIS), Baidu Inc., China), and Errui Ding (Department of Computer Vision Technology (VIS), Baidu Inc., China)

Sparse Convolutional Networks for Surface Reconstruction From Noisy Point Clouds 3200
Tao Wang (Cardiff university), Jing Wu (Cardiff University), Ze Ji (Cardiff University), and Yu-Kun Lai (Cardiff University)

Single Frame Semantic Segmentation Using Multi-Modal Spherical Images 3210
Suresh Guttikonda (German Research Center for Artificial Intelligence (DFKI)) and Jason Rambach (German Research Center for Artificial Intelligence (DFKI))

Context-Based Interpretable Spatio-Temporal Graph Convolutional Network for Human Motion Forecasting 3220
Edgar Eduardo Medina Castañeda (QualityMinds GmbH), Leyong Loh (QualityMinds GmbH), Namrata Gurung (QualityMinds GmbH), Kyung Hun Oh (QualityMinds GmbH), and Niels Heller (QualityMinds GmbH)

GC-MVSNet: Multi-View, Multi-Scale, Geometrically-Consistent Multi-View Stereo 3230
Vibhas K. Vats (Indiana University Bloomington), Sripad Joshi (Indiana University Bloomington), David J. Crandall (Indiana University Bloomington), Alimoor Reza (Drake University), and Soon-Heung Jung (ETRI)

SAM Fewshot Finetuning for Anatomical Segmentation in Medical Images 3241
Weiyi Xie (Stryker AI Research), Nathalie Willems (Stryker AI Research), Shubham Patil (Stryker AI Research), Yang Li (Stryker AI Research), and Mayank Kumar (Stryker AI Research)

MSCC: Multi-Scale Transformers for Camera Calibration 3250
Xu Song (Fujitsu R&D Center Co., Ltd.), Hao Kang (Fujitsu R&D Center Co., Ltd.), Atsunori Moteki (Fujitsu Limited), Genta Suzuki (Fujitsu Limited), Yoshie Kobayashi (Fujitsu Limited), and Zhiming Tan (Fujitsu R&D Center Co., Ltd.)

A Geometry Loss Combination for 3D Human Pose Estimation 3260
Ai Matsune (Fujitsu R&D Center Co., Ltd.), Shichen Hu (Fujitsu R&D Center Co., Ltd.), Guangquan Li (Fujitsu R&D Center Co., Ltd.), Sihan Wen (Fujitsu R&D Center Co., Ltd.), Xiantan Zhu (Fujitsu R&D Center Co., Ltd.), and Zhiming Tan (Fujitsu R&D Center Co., Ltd.)

A Robust Diffusion Modeling Framework for Radar Camera 3D Object Detection 3270
Zizhang Wu (zongmu tech.com), Yunzhe Wu (ZongmuTech), Wang Xiaouqan (zongmutech), Yuanzhu Gan (Zongmu Technology), and Jian Pu (Fudan University)

WalkFormer: Point Cloud Completion via Guided Walks 3281
Mohang Zhang (Xi'an Jiaotong-liverpool University; University of Liverpool), Yushi Li (Xi'an Jiaotong-Liverpool University), Rong Chen (Dalian Maritime University), Yushan Pan (Xi'an Jiaotong-Liverpool University), Jia Wang (Xi'an Jiaotong-Liverpool University), Yunzhe Wang (Suzhou University of Science and Technology), and Rong Xiang (The Hong Kong Polytechnic University)

MGM-AE: Self-Supervised Learning on 3D Shape Using Mesh Graph Masked Autoencoders 3291
Zhangsihao Yang (Arizona State University), Kaize Ding (Northwestern University), Huan Liu (Arizona State University), and Yalin Wang (Arizona State University)

Unsupervised 3D Pose Estimation With Non-Rigid Structure-From-Motion Modeling 3302
Haorui Ji (The Australian National University), Hui Deng (Northwestern Polytechnical University), Yuchao Dai (Northwestern Polytechnical University), and Hongdong Li (ANU)

Residual Graph Convolutional Network for Bird's-Eye-View Semantic Segmentation 3312
Qiuxiao Chen (utah state university) and Xiaojun Qi (USU)

3D Human Pose Estimation With Two-Step Mixed-Training Strategy 3320
Yingfeng Wang (Centre for Intelligent Multidimensional Data Analysis Limited), Zhengwei Wang (Centre for Intelligent Multidimensional Data Analysis Limited), Muyu Li (Center for Intelligent Multidimensional Data Analysis Limited), and Hong Yan (City University of Hong Kong)

RGB-D Mapping and Tracking in a Plenoxel Radiance Field 3330
Andreas L. Teigen (Norwegian University of Science and Technology (NTNU), Norway), Yeonsoo Park (Mobiltech, Republic of Korea), Annette Stahl (Norwegian University of Science and Technology (NTNU), Norway), and Rudolf Mester (Norwegian University of Science and Technology (NTNU), Norway)

SOAP: Cross-Sensor Domain Adaptation for 3D Object Detection Using Stationary Object Aggregation Pseudo-Labelling 3340
Chengjie Huang (University of Waterloo), Vahdat Abdelzad (University of Waterloo), Sean Sedwards (University of Waterloo), and Krzysztof Czarnecki (University of Waterloo)

BALF: Simple and Efficient Blur Aware Local Feature Detector 3350
Zhenjun Zhao (The Chinese University of Hong Kong)

MACP: Efficient Model Adaptation for Cooperative Perception 3361
Yunsheng Ma (Purdue University, USA), Juanwu Lu (Purdue University, USA), Can Cui (Purdue University, USA), Sicheng Zhao (Tsinghua University, China), Xu Cao (University of Illinois Urbana-Champaign, USA; PediaMed AI, China), Wenqian Ye (University of Virginia, USA; PediaMed AI, China), and Ziran Wang (Purdue University, USA)

MAELi: Masked Autoencoder for Large-Scale LiDAR Point Clouds 3371
Georg Krispel (Graz University of Technology), David Schinagl (Graz University of Technology; Christian Doppler Laboratory for Embedded Machine Learning), Christian Fruhwirth-Reisinger (Graz University of Technology; Christian Doppler Laboratory for Embedded Machine Learning), Horst Possegger (Graz University of Technology), and Horst Bischof (Graz University of Technology; Christian Doppler Laboratory for Embedded Machine Learning)

HDMNet: A Hierarchical Matching Network With Double Attention for Large-Scale Outdoor LiDAR Point Cloud Registration 3381
Weiyi Xue (Tongji University), Fan Lu (Tongji University), and Guang Chen (Tongji University)

SGRec3D: Self-Supervised 3D Scene Graph Learning via Object-Level Scene Reconstruction 3392
Sebastian Koch (Bosch Center for Artificial Intelligence; Robert Bosch Corporate Research; University of Ulm), Pedro Hermosilla (TU Wien), Narunas Vaskevicius (Bosch Center for Artificial Intelligence; Robert Bosch Corporate Research), Mirco Colosi (Robert Bosch Corporate Research), and Timo Ropinski (University of Ulm)

MPT: Mesh Pre-Training with Transformers for Human Pose and Mesh Reconstruction 3403
Kevin Lin (Microsoft), Chung-Ching Lin (Microsoft), Lin Liang (Microsoft), Zicheng Liu (Microsoft), and Lijuan Wang (Microsoft)

LInKs "Lifting Independent Keypoints" – Partial Pose Lifting for Occlusion Handling with Improved Accuracy in 2D-3D Human Pose Estimation 3414
Peter Hardy (University of Southampton) and Hansung Kim (University Of Southampton)

ECSIC: Epipolar Cross Attention for Stereo Image Compression 3424
Matthias Wödlinger (TU Wien), Jan Kotera (UTIA CAS, Czechia), Manuel Keglevic (TU Wien), Jan Xu (Deep Render), and Robert Sablatnig (TU Wien)

Robust Category-Level 3D Pose Estimation From Diffusion-Enhanced Synthetic Data 3434
Jiahao Yang (Peking University), Wufei Ma (Johns Hopkins University), Angtian Wang (Johns Hopkins University), Xiaoding Yuan (Johns Hopkins University), Alan Yuille (Johns Hopkins University), and Adam Kortylewski (University of Freiburg; Max Planck Institute for Informatics)

Open-NeRF: Towards Open Vocabulary NeRF Decomposition 3444
Hao Zhang (University of Illinois Urbana-Champaign), Fang Li (University of Illinois at Urbana-Champaign), and Narendra Ahuja (University of Illinois at Urbana-Champaign)

HAMMER: Learning Entropy Maps to Create Accurate 3D Models in Multi-View Stereo 3454
Rafael Weilharter (Graz University of Technology) and Friedrich Fraundorfer (Graz University of Technology)

Polarimetric PatchMatch Multi-View Stereo 3464
Jinyu Zhao (Tokyo Institute of Technology), Jumpei Oishi (Tokyo Institute of Technology), Yusuke Monno (Tokyo Institute of Technology), and Masatoshi Okutomi (Tokyo Institute of Technology)

Solving the Plane-Sphere Ambiguity in Top-Down Structure-From-Motion 3473
Lars Haalck (University of Münster, Germany) and Benjamin Risse (University of Münster, Germany)

Self-Annotated 3D Geometric Learning for Smeared Points Removal 3482
Miaowei Wang (University of Edinburgh) and Daniel Morris (Michigan State University)

3D Face Style Transfer With a Hybrid Solution of NeRF and Mesh Rasterization 3492
Jianwei Feng (Amazon) and Prateek Singhal (Amazon)

A Sequential Learning-Based Approach for Monocular Human Performance Capture 3502
Jianchun Chen (Carnegie Mellon University), Jayakorn Vongkulbhisal (Independent Researcher), and Fernando De la Torre Frade (Carnegie Mellon University)

ScanEnts3D: Exploiting Phrase-to-3D-Object Correspondences for Improved Visio-Linguistic Models in 3D Scenes 3512
Ahmed Abdelreheem (King Abdullah University of Science and Technology (KAUST); Snap Inc.), Kyle Olszewski (Snap Inc.), Hsin-Ying Lee (Snap Inc.), Peter Wonka (King Abdullah University of Science and Technology (KAUST)), and Panos Achlioptas (Snap Inc.)

Global Occlusion-Aware Transformer for Robust Stereo Matching 3523
Zihua Liu (Tokyo Institute of Technology, Japan), Yizhou Li (Tokyo Institute of Technology, Japan), and Masatoshi Okutomi (Tokyo Institute of Technology, Japan)

MoRF: Mobile Realistic Fullbody Avatars From a Monocular Video 3533
Renat Bashirov (Samsung Research), Alexey Larionov (Samsung Research), Evgeniya Ustinova (Samsung Research), Mikhail Sidorenko (Samsung Research), David Svitov (Samsung Research), Ilya Zakharkin (ZERO10), and Victor Lempitsky (Cinemersive Labs)

PointCT: Point Central Transformer Network forWeakly-supervised Point Cloud Semantic Segmentation 3544
Anh-Thuan Tran (Pukyong National University, South Korea), Hoanh-Su Le (University of Economics and Law; Vietnam National University, Vietnam), Suk-Hwan Lee (Dong-A University, South Korea), and Ki-Ryong Kwon (Pukyong National University, South Korea)

Top-Down Beats Bottom-Up in 3D Instance Segmentation 3554
Maksim Kolodiazhnyi (Samsung Research), Anna Vorontsova (Samsung Research), Anton S. Konushin (Samsung AI Center Moscow), and Danila Rukhovich (Samsung AI Center Moscow)

Longformer: Longitudinal Transformer for Alzheimer's Disease Classification With Structural MRIs 3563
Qiuhui Chen (Shanghai Jiao Tong University), Qiang Fu (Shanghai Jiao Tong University), Bai Hao (Shanghai Jiao Tong University), and Yi Hong (Shanghai Jiao Tong University)

TEGLO: High Fidelity Canonical Texture Mapping From Single-View Images 3573
Vishal Vinod (University of California; Google Research), Tanmay Shah (Google), and Dmitry Lagun (Google)

DiffCLIP: Leveraging Stable Diffusion for Language Grounded 3D Classification 3584
Sitian Shen (Beijing Institute of Technology, China), Zilin Zhu (Beijing Institute of Technology, China), Linqian Fan (Beijing Jiaotong University, China), Harry Zhang (Massachusetts Institute of Technology, USA), and Xinxiao Wu (Beijing Institute of Technology, China)

FocusTune: Tuning Visual Localization Through Focus-Guided Sampling 3594
Son Tung Nguyen (Queensland University of Technology), Alejandro Fontan (Queensland University of Technology), Michael Milford (ACRV and QUT, Australia), and Tobias Fischer (Queensland University of Technology)

Indoor Visual Localization Using Point and Line Correspondences in Dense Colored Point Cloud 3604
Yuya Matsumoto (NEC Corporation), Gaku Nakano (NEC Corporation), and Kazumine Ogura (NEC Corporation)

Fast Sun-Aligned Outdoor Scene Relighting Based on TensoRF 3614
Yeonjin Chang (Seoul National University), Yearim Kim (Seoul National University), Seunghyeon Seo (Seoul National University), Jung Yi (Yonsei University), and Nojun Kwak (Seoul National University)

MonoProb: Self-Supervised Monocular Depth Estimation with Interpretable Uncertainty 3625
Rémi Marsal (Université Paris-Saclay, CEA; Sorbonne University, CNRS), Florian Chabot (Université Paris-Saclay, CEA), Angelique Loesch (Université Paris-Saclay, CEA), William Grolleau (Université Paris-Saclay, CEA), and Hichem Sahbi (Sorbonne University, CNRS)

AvatarOne: Monocular 3D Human Animation 3635
Akash Karthikeyan (University of Toronto), Robert Ren (University of Toronto), Yash Kant (University of Toronto), and Igor Gilitschenski (University of Toronto; Vector Institute for AI)

Deblur-NSFF: Neural Scene Flow Fields for Blurry Dynamic Scenes 3646
Achleshwar Luthra (Carnegie Mellon University), Shiva Souhith Gantha (Georgia Institute of Technology), Xiyun Song (FUTUREWEITECH), Heather Yu (Futurewei Technologies), Zongfang Lin (Futurewei Technologies), and Liang Peng (Futurewei)

SimpliMix: A Simplified Manifold Mixup for Few-Shot Point Cloud Classification 3656
Minmin Yang (Syracuse University), Weiheng Chai (Syracuse University), Jiyang Wang (Syracuse University), and Senem Velipasalar (Syracuse University)

PMVC: Promoting Multi-View Consistency for 3D Scene Reconstruction 3666
Chushan Zhang (The Australian National University), Jinguang Tong (The Australian National University; Data61, CSIRO), Tao Jun Lin (The Australian National University), Chuong Nguyen (The Australian National University; Data61, CSIRO), and Hongdong Li (The Australian National University)

Hyb-NeRF: A Multiresolution Hybrid Encoding for Neural Radiance Fields 3677
Yifan Wang (Southern University of Science and Technology), Yuan Zeng (Southern University of Science and Technology), and Yi Gong (Southern University of Science and Technology)

Joint 3D Shape and Motion Estimation From Rolling Shutter Light-Field Images 3687
Hermes Mcgriff (Université de Bourgognegne; Université de Franche-Comté), Renato Martins (Université de Bourgogne; Université de Lorraine), Nicolas Andreff (Université de Franche-Comté), and Cédric Demonceaux (Université de Bourgogne; Université de Lorraine)

Sharp-NeRF: Grid-Based Fast Deblurring Neural Radiance Fields Using Sharpness Prior 3697
Byeonghyeon Lee (Sungkyunkwan University), Howoong Lee (Sungkyunkwan University; Hanwha Vision), Usman Ali (Sungkyunkwan University), and Eunbyung Park (Sungkyunkwan University)

When 3D Bounding-Box Meets SAM: Point Cloud Instance Segmentation With Weak-and-Noisy Supervision 3707
Qingtao Yu (ANU), Heming Du (Australian National University), Chen Liu (The Univeristy of Queensland), and Xin Yu (University of Queensland)

Auto-BPA: An Enhanced Ball-Pivoting Algorithm with Adaptive Radius Using Contextual Bandits 3717
Houda Saffi (University Mohammed VI Polytechnic), Naima Otberdout (University Mohammed VI Polytechnic), Youssef Hmamouche (University Mohammed VI Polytechnic), and Amal El Fallah Seghrouchni (University Mohammed VI Polytechnic; Sorbonne Université)

Towards Realistic Generative 3D Face Models 3726
Aashish Rai (Carnegie Mellon University), Hiresh Gupta (Carnegie Mellon University), Ayush Pandey (Carnegie Mellon University), Francisco Vicente Carrasco (Carnegie Mellon University), Shingo Jason Takagi (Meta Reality Labs), Amaury Aubel (Meta Reality Labs), Daeil Kim (Meta Reality Labs), Aayush Prakash (Meta Reality Labs), and Fernando De la Torre (Carnegie Mellon University)

Camera-Independent Single Image Depth Estimation From Defocus Blur 3737
Lahiru Wijayasingha (University of Virginia, USA), Homa Alemzadeh (University of Virginia, USA), and John A. Stankovic (University of Virginia, USA)

U3DS3: Unsupervised 3D Semantic Scene Segmentation 3747
Jiaxu Liu (Durham University, UK), Zhengdi Yu (Durham University, UK), Toby P. Breckon (Durham University, UK), and Hubert P. H. Shum (Durham University, UK)

SSP: Semi-Signed Prioritized Neural Fitting for Surface Reconstruction From Unoriented Point Clouds 3757
Runsong Zhu (The Chinese University of Hong Kong), Di Kang (Tencent AI Lab), Ka-Hei Hui (The Chinese University of Hong Kong), Yue Qian (Tencent AI Lab), Shi Qiu (The Chinese Univsersity of Hong Kong), Zhen Dong (Wuhan University), Linchao Bao (Tencent AI Lab), Pheng-Ann Heng (The Chinese Univsersity of Hong Kong), and Chi-Wing Fu (The Chinese University of Hong Kong)

Adversarial Learning, Adversarial Attack & Defense Methods

Adversarial Likelihood EstimationWith One-Way Flows 3767
Omri Ben-Dov (Max Planck Institute for Intelligent Systems, Germany), Pravir Singh Gupta (Perceive Inc.), Victoria Abrevaya (Max Planck Institute for Intelligent Systems, Germany), Michael J. Black (Max Planck Institute for Intelligent Systems, Germany), and Partha Ghosh (Max Planck Institute for Intelligent Systems, Germany)

NCIS: Neural Contextual Iterative Smoothing for Purifying Adversarial Perturbations 3777
Sungmin Cha (New York University), Naeun Ko (NAVER Cloud), Heewoong Choi (ASRI, INMC, Seoul National University), Youngjoon Yoo (NAVER Cloud; NAVER AI Lab), and Taesup Moon (ASRI, INMC, Seoul National University)

On the Fly Neural Style Smoothing for Risk-Averse Domain Generalization 3788
Akshay Mehra (Tulane University), Yunbei Zhang (Tulane University), Bhavya Kailkhura (Lawrence Livermore National Laboratory), and Jihun Hamm (Tulane University)

D4: Detection of Adversarial Diffusion Deepfakes Using Disjoint Ensembles 3800
Ashish Hooda (University of Wisconsin–Madison), Neal Mangaokar (University of Michigan), Ryan Feng (University of Michigan), Kassem Fawaz (University of Wisconsin-Madison), Somesh Jha (University of Wisconsin-Madison), and Atul Prakash (University of Michigan)

Army of Thieves: Enhancing Black-Box Model Extraction via Ensemble Based Sample Selection .. 3811
Akshit Jindal (IIIT Delhi), Vikram Goyal (IIIT Delhi, India), Saket Anand (Indraprastha Institute of Information Technology Delhi), and Chetan Arora (Indian Institute of Technology Delhi)

Assist Is Just As Important as the Goal: Image Resurfacing To Aid Model's Robust Prediction 3821
Abhijith Sharma (University of British Columbia), Phil Munz (TrojAI), and Apurva Narayan (University of Western Ontario)

CLIPAG: Towards Generator-Free Text-to-Image Generation 3831
Roy Ganz (Technion, Israel) and Michael Elad (Technion, Israel)

Defending Object Detection Models Against Image Distortions 3842
Mark Ofori-Oduro (Concordia University) and Maria Amer (Concordia University)

ATS: Adaptive Temperature Scaling for Enhancing Out-of-Distribution Detection Methods 3852
Gerhard Krumpl (Graz University of Technology; KESTRELEYE GmbH, Austria), Henning Avenhaus (KESTRELEYE GmbH, Austria), Horst Possegger (Graz University of Technology, Austria), and Horst Bischof (Graz University of Technology, Austria)

A Closer Look at Robustness of Vision Transformers to Backdoor Attacks 3862
Akshayvarun Subramanya (University of Maryland), Soroush Abbasi Koohpayegani (University of California), Aniruddha Saha (University of Maryland), Ajinkya Tejankar (University of California), and Hamed Pirsiavash (University of California)

Maximum Knowledge Orthogonality Reconstruction with Gradients in Federated Learning 3872
Feng Wang (Syracuse University), Senem Velipasalar (Syracuse University), and M. Cenk Gursoy (Syracuse University)

Learning to Generate Training Datasets for Robust Semantic Segmentation 3882
Marwane Hariat (U2IS, ENSTA Paris, Institut Polytechnique de Paris), Olivier Laurent (U2IS, ENSTA Paris, Institut Polytechnique de Paris; SATIE, Université Paris-Saclay), Rémi Kazmierczak (U2IS, ENSTA Paris, Institut Polytechnique de Paris), Shihao Zhang (National University of Singapore), Andrei Bursuc (valeo.ai), Angela Yao (National University of Singapore), and Gianni Franchi (U2IS, ENSTA Paris, Institut Polytechnique de Paris)

Uncertainty-Weighted Loss Functions for Improved Adversarial Attacks on Semantic Segmentation 3894
Kira Maag (Technical University of Berlin) and Asja Fischer (Ruhr University Bochum)

Natural Light Can Also Be Dangerous: Traffic Sign Misinterpretation Under Adversarial Natural Light Attacks 3903
Teng-Fang Hsiao (National Yang Ming Chiao Tung University, Taiwan), Bo-Lun Huang (National Yang Ming Chiao Tung University, Taiwan), Zi-Xiang Ni (National Yang Ming Chiao Tung University, Taiwan), Yan-Ting Lin (National Yang Ming Chiao Tung Univerisity, Taiwan), Hong-Han Shuai (National Yang Ming Chiao Tung University, Taiwan), Yung-Hui Li (Hon Hai Research Institute), and Wen-Huang Cheng (Information Engineering Graduate Institute of Taiwan University)

Diffusion Models Meet Image Counter-Forensics 3913
Matías Tailanián (Universidad de la República; Digital Sense), Marina Gardella (École normale supérieure Paris-Saclay, Université Paris-Saclay), Alvaro Pardo (Digital Sense; Universidad Católica del Uruguay), and Pablo Musé (Universidad de la República)

Discriminator-Free Unsupervised Domain Adaptation for Multi-Label Image Classification 3924
Inder Pal Singh (University of Luxembourg), Enjie Ghorbel (SnT, University of Luxembourg), Anis Kacem (SnT, University of Luxembourg), Arunkumar Rathinam (SnT, University of Luxembourg), and Djamila Aouada (SnT, University of Luxembourg)

Few-Shot Generative Model for Skeleton-Based Human Action Synthesis Using Cross-Domain Adversarial Learning 3934
Kenichiro Fukushi (NEC), Yoshitaka Nozaki (NEC), Kosuke Nishihara (NEC), and Kentaro Nakahara (NEC)

Mixing Gradients in Neural Networks as a Strategy to Enhance Privacy in Federated Learning..... 3944
Shaltiel Eloul (CTO, JPMorgan Chase), Fran Silavong (CTO, JPMorgan Chase), Sanket Kamthe (CTO, JPMorgan Chase), Antonios Georgiadis (CTO, JPMorgan Chase), and Sean J. Moran (CTO, JPMorgan Chase)

Neural Style Protection: Counteracting Unauthorized Neural Style Transfer 3954
Yaxin Li (Michigan State University), Jie Ren (Michigan State University), Han Xu (Michigan State University), and Hui Liu (Michigan State University)

Exploring Adversarial Robustness of Vision Transformers in the Spectral Perspective 3964
Gihyun Kim (Yonsei University, Republic of Korea), Juyeop Kim (Yonsei University, Republic of Korea), and Jong-Seok Lee (Yonsei University, Republic of Korea)

Hard-Label Based Small Query Black-Box Adversarial Attack 3974
Jeonghwan Park (Queens University Belfast, United Kingdom), Paul Miller (Queen's University Belfast, United Kingdom), and Niall Mclaughlin (Queen's University Belfast, United Kingdom)

Simple Post-Training Robustness Using Test Time Augmentations and Random Forest 3984
Gilad Cohen (Tel Aviv University) and Raja Giryes (Tel Aviv University)

Biometrics, Face, Gesture, Body Pose I

Co-Speech Gesture Detection Through Multi-Phase Sequence Labeling 3995
Esam Ghaleb (University of Amsterdam), Ilya Burenko (ScaDS.AI Dresden/Leipzig; TU Dresden), Marlou Rasenberg (Meertens Institute; Max Planck Institute for Psycholinguistics), Wim Pouw (Radboud University), Peter Uhrig (ScaDS.AI Dresden/Leipzig; TU Dresden), Judith Holler (Radboud University; Max Planck Institute for Psycholinguistics), Ivan Toni (Radboud University), Aslı Özyürek (Radboud University; Max Planck Institute for Psycholinguistics), and Raquel Fernández (University of Amsterdam)

Computational Photography, Image and Video Synthesis

Learning Residual Elastic Warps for Image Stitching Under Dirichlet Boundary Condition 4004
Minsu Kim (DGIST), Yongjun Lee (Korea University), Woo Kyoung Han (DGIST), and Kyong Hwan Jin (Korea University)

Exploiting the Signal-Leak Bias in Diffusion Models 4013
Martin Nicolas Everaert (EPFL, Switzerland), Athanasios Fitsios (EPFL; Largo.ai, Switzerland), Marco Bocchio (Largo.ai, Switzerland), Sami Arpa (Largo.ai, Switzerland), Sabine Süsstrunk (EPFL, Switzerland), and Radhakrishna Achanta (EPFL, Switzerland)

Synthesizing Anyone, Anywhere, in Any Pose 4023
Håkon Hukkelås (Norwegian University of Science and Technology) and Frank Lindseth (Norwegian University of Science and Technology)

Specular Object Reconstruction Behind Frosted Glass by Differentiable Rendering 4035
Takafumi Iwaguchi (Kyushu Univ.), Hiroyuki Kubo (Chiba University), and Hiroshi Kawasaki (Kyushu univ.)

Multi-Level Attention Aggregation for Aesthetic Face Relighting 4045
Hemanth Pidaparthy (Samsung R&D Institute Bangalore, India), Abhay Chauhan (Samsung R&D Institute Bangalore, India), and Pavan Sudheendra (Samsung R&D Institute Bangalore, India)

Deep Image Fingerprint: Towards Low Budget Synthetic Image Detection and Model Lineage Analysis 4055
Sergey Sinitsa (Reichman University) and Ohad Fried (Reichman University)

StyleGenes: Discrete and Efficient Latent Distributions for GANs 4065
Evangelos Ntavelis (ETH Zurich; CSEM SA), Mohamad Shahbazi (ETH Zurich), Iason Kastanis (CSEM SA), Martin Danelljan (ETH Zurich), and Luc Van Gool (ETH Zurich)

Implicit Neural Image Stitching With Enhanced and Blended Feature Reconstruction 4075
Minsu Kim (DGIST), Jaewon Lee (DGIST), Byeonghun Lee (Korea University), Sunghoon Im (DGIST), and Kyong Hwan Jin (Korea University)

PDA-RWSR: Pixel-Wise Degradation Adaptive Real-World Super-Resolution 4085
Andreas Aakerberg (Aalborg University, Denmark), Majed El Helou (ETH Zürich, Switzerland), Kamal Nasrollahi (Aalborg University; Milestone Systems, Denmark), and Thomas Moeslund (Aalborg University, Denmark)

Real Time GAZED: Online Shot Selection and Editing of Virtual Cameras From Wide-Angle Monocular Video Recordings 4096
Sudheer Achary (International Institute of Information Technology, Hyderabad), Rohit Girmaji (International Institute of Information Technology , Hyderabad), Adhiraj Anil Deshmukh (IIIT Hyderabad), and Vineet Gandhi (IIIT Hyderabad)

Unsupervised Model-Based Learning for Simultaneous Video Deflickering and Deblotching 4105
Anuj Fulari (IIT Bombay, India), Satish Mulleti (IIT Bombay, India), and Ajit Rajwade (IIT Bombay, India)

Differentiable JPEG: The Devil Is in the Details 4114
Christoph Reich (Technische Universität Darmstadt), Biplob Debnath (NEC Labs), Deep A Patel (NEC Laboratories America), and Srimat Chakradhar (nec labs)

TSA2: Temporal Segment Adaptation and Aggregation for Video Harmonization 4124
Zeyu Xiao (University of Science and Technology of China), Yurui Zhu (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), and Zhiwei Xiong (University of Science and Technology of China)

Diffusion in the Dark: A Diffusion Model for Low-Light Text Recognition 4134
Cindy M. Nguyen (Stanford University), Eric R. Chan (Stanford University), Alexander W. Bergman (Stanford University), and Gordon Wetzstein (Stanford University)

Lightweight Portrait Matting via Regional Attention and Refinement 4146
Yatao Zhong (Microsoft) and Ilya Zharkov (Microsoft)

RADIO: Reference-Agnostic Dubbing Video Synthesis 4156
Dongyeun Lee (Klleon AI Research), Chaewon Kim (Klleon AI Research), Sangjoon Yu (Klleon AI Research), Jaejun Yoo (UNIST), and Gyeong-Moon Park (Kyung Hee University)

Unsupervised Event-Based Video Reconstruction 4167
Gereon Fox (MPI Informatik, Saarland Informatics Campus), Xingang Pan (Nanyang Technological University), Ayush Tewari (MIT), Mohamed Elgharib (MPI Informatik, Saarland Informatics Campus), and Christian Theobalt (MPI Informatik, Saarland Informatics Campus)

Neural Image Compression Using Masked Sparse Visual Representation 4177
Wei Jiang (Futurewei Technologies Inc.), Wei Wang (Futurewei Technologies Inc.), and Yue Chen (Futurewei Technologies Inc.)

Shape-Guided Diffusion with Inside-Outside Attention 4186
Dong Huk Park (UC Berkeley), Grace Luo (UC Berkeley), Clayton Toste (UC Berkeley), Samaneh Azadi (Meta AI), Xihui Liu (UC Berkeley; The University of Hong Kong), Maka Karalashvili (BMW Group), Anna Rohrbach (UC Berkeley), and Trevor Darrell (UC Berkeley)

Collage Diffusion 4196
Vishnu Sarukkai (Stanford), Linden Li (Stanford University), Arden Ma (Stanford University), Christopher Ré (Stanford University), and Kayvon Fatahalian (Stanford University)

Learning-Based Spotlight Position Optimization for Non-Line-of-Sight Human Localization and Posture Classification 4206
Sreenithy Chandran (Arizona State University), Tatsuya Yatagawa (Hitotsubashi University), Hiroyuki Kubo (Chiba University), and Suren Jayasuriya (Arizona State University)

Scale-Adaptive Feature Aggregation for Efficient Space-Time Video Super-Resolution 4216
Zhewei Huang (Megvii Technology), Ailin Huang (Megvii Technology), Xiaotao Hu (Megvii Technology; Nankai University), Chen Hu (Megvii Technology), Jun Xu (Nankai University; Guangdong Provincial Key Laboratory of Big Data Computing, CUHK), and Shuchang Zhou (Megvii Technology)

What Decreases Editing Capability? Domain-Specific Hybrid Refinement for Improved GAN Inversion 4228
Pu Cao (Beijing University of Posts and Telecommunications; Metavatar), Lu Yang (Beijing University of Posts and Telecommunications), Dongxv Liu (Beijing University of Posts and Telecommunications), Xiaoya Yang (Beijing University of Posts and Telecommunications), Tianrui Huang (Beijing University of Posts and Telecommunications), and Qing Song (Beijing University of Posts and Telecommunications)

Latent-Guided Exemplar-Based Image Re-Colorization 4238
Wenjie Yang (Shanghai Jiao Tong University), Ning Xu (n/a), and Yifei Fan (Adobe; Academy of Art University)

Stereo Conversion with Disparity-Aware Warping, Compositing and Inpainting 4248
Lukas Mehl (University of Stuttgart), Andrés Bruhn (University of Stuttgart), Markus Gross (ETH Zurich; Disney Research, Studios, Switzerland), and Christopher Schroers (Disney Research, Studios, Switzerland)

Arbitrary-Resolution and Arbitrary-Scale Face Super-Resolution with Implicit Representation Networks 4258
Yi Ting Tsai (National Yang Ming Chiao Tung University), Yu Wei Chen (National Yang Ming Chiao Tung University), Hong-Han Shuai (National Yang Ming Chiao Tung University), and Ching-Chun Huang (National Yang Ming Chiao Tung University)

Blurry Video Compression: A Trade-Off Between Visual Enhancement and Data Compression 4268
Dawit Mureja Argaw (KAIST), Junsik Kim (Harvard University), and In So Kweon (KAIST)

ProxEdit: Improving Tuning-Free Real Image Editing with Proximal Guidance 4279
Ligong Han (Rutgers University), Song Wen (Rutgers University), Qi Chen (Laval University), Zhixing Zhang (Rutgers University), Kunpeng Song (Rutgers University), Mengwei Ren (New York University), Ruijiang Gao (UT Austin), Anastasis Stathopoulos (Rutgers University), Xiaoxiao He (Rutgers University), Yuxiao Chen (Rutgers University), Di Liu (Rutgers University), Qilong Zhangli (Rutgers University), Jindong Jiang (Rutgers University), Zhaoyang Xia (Rutgers University), Akash Srivastava (MIT-IBM AI Lab), and Dimitris Metaxas (Rutgers)

VCISR: Blind Single Image Super-Resolution with Video Compression Synthetic Data 4290
Boyang Wang (University of Michigan, Ann Arbor), Bowen Liu (University of Michigan, Ann Arbor), Shiyu Liu (University of Michigan, Ann Arbor), and Fengyu Yang (University of Michigan, Ann Arbority)

Latent Feature-Guided Diffusion Models for Shadow Removal 4301
Kangfu Mei (Johns Hopkins University), Luis Figueroa (Adobe Research), Zhe Lin (Adobe Research), Zhihong Ding (Adobe Research), Scott Cohen (Adobe Research), and Vishal Patel (Johns Hopkins University)

MobileNVC: Real-Time 1080p Neural Video Compression on a Mobile Device 4311
Ties van Rozendaal (Qualcomm AI Research), Tushar Singhal (Qualcomm AI Research), Hoang Le (Qualcomm AI Research), Guillaume Sautiere (Qualcomm AI Research), Amir Said (Qualcomm AI Research), Krishna Buska (Qualcomm AI Research), Anjuman Raha (Qualcomm AI Research), Dimitris Kalatzis (Qualcomm AI Research), Hitarth Mehta (Qualcomm AI Research), Frank Mayer (Qualcomm AI Research), Liang Zhang (Qualcomm AI Research), Markus Nagel (Qualcomm AI Research), and Auke Wiggers (Qualcomm AI Research)

LatentPaint: Image Inpainting in Latent Space With Diffusion Models 4322
Ciprian A Corneanu (Amazon), Raghudeep Gadde (Amazon), and Aleix M Martinez (Amazon)

Datasets and Evaluations

Beyond RGB: A RealWorld Dataset for Multispectral Imaging in Mobile Devices 4332
Ortal Glatt (Samsung Israel R&D Center, Israel), Yotam Ater (Samsung Israel R&D Center, Israel), Woo-Shik Kim (Samsung Advanced Institute of Technology, Korea), Shira Werman (Samsung Israel R&D Center, Israel), Oded Berby (Samsung Israel R&D Center, Israel), Yael Zini (Samsung Israel R&D Center, Israel), Shay Zelinger (Samsung Israel R&D Center, Israel), Sangyoon Lee (Samsung Advanced Institute of Technology, Korea), Heejin Choi (Samsung Advanced Institute of Technology, Korea), and Evgeny Soloveichik (Samsung Israel R&D Center, Israel)

IKEA Ego 3D Dataset: Understanding Furniture Assembly Actions From Ego-View 3D Point Clouds 4343
Yizhak Ben-Shabat (Australian National University; Technion, Israel Institute of Technology), Jonathan Paul (Technion, Israel Institute of Technology), Eviatar Segev (Technion, Israel Institute of Technology), Oren Shrout (Technion, Israel Institute of Technology), and Stephen Gould (Australian National University)

IndustReal: A Dataset for Procedure Step Recognition Handling Execution Errors in Egocentric Videos in an Industrial-Like Setting 4353
Tim J. Schoonbeek (Eindhoven University of Technology, Netherlands), Tim Houben (Eindhoven University of Technology, Netherlands), Hans Onvlee (ASML Research, Netherlands), Peter H.N. de With (Eindhoven University of Technology, Netherlands), and Fons van der Sommen (Eindhoven University of Technology, Netherlands)

Ego2HandsPose: A Dataset for Egocentric Two-Hand 3D Global Pose Estimation 4363
Fanqing Lin (Magic Leap, Inc., USA) and Tony Martinez (Brigham Young University)

ISAR: A Benchmark for Single- and Few-Shot Object Instance Segmentation and Re-Identification 4372
Nicolas J Gorlo (ETH Zurich), Kenneth Blomqvist (ETH Zurich), Francesco Milano (ETH Zurich), and Roland Siegwart (ETH Zürich, Autonomous Systems Lab)

Towards Addressing the Misalignment of Object Proposal Evaluation for Vision-Language Tasks via Semantic Grounding 4385
Joshua Feinglass (Arizona State University) and Yezhou Yang (Arizona State University)

SphereCraft: A Dataset for Spherical Keypoint Detection, Matching and Camera Pose Estimation 4396
Christiano Gava (DFKI), Yunmin Cho (Aimmo Germany GmbH), Federico Raue (DFKI), Sebastian Palacio (DFKI), Alain Pagani (DFKI), and Andreas Dengel (DFKI; University of Kaiserslautern-Landau)

Benchmark Generation Framework With Customizable Distortions for Image Classifier Robustness 4406
Soumyendu Sarkar (Hewlett Packard Enterprise), Ashwin Ramesh Babu (Hewlett Packard Enterprise Labs), Sajad Mousavi (Hewlett Packard Enterprise), Zachariah Carmichael (University of Notre Dame), Vineet Gundecha (Hewlett Packard Enterpise), Sahand Ghorbanpour (Hewlett Packard Enterprise), Ricardo Luna Gutierrez (Hewlett Packard Enterprise), Antonio Guillen (Hewlett Packard Enterprise), and Avisek Naug (Hewlett Packard Enterprise)

UOW-Vessel: A Benchmark Dataset of High-Resolution Optical Satellite Images for Vessel Detection and Segmentation 4416
Ly Bui (University of Wollongong), Son Lam Phung (University of Wollongong, VinAI), Yang Di (University of Wollongong), Thanh Le (University of Wollongong), Tran Thanh Phong Nguyen (University of Wollongong), Sandy Burden (Australian Department of Defence), and Abdesselam Bouzerdoum (UoW, Wollongong, Australia)

NITEC: Versatile Hand-Annotated Eye Contact Dataset for Ego-Vision Interaction 4425
Thorsten Hempel (Otto von Guericke University), Magnus Jung (OVGU), Ahmed Abdelrahman (Otto von Guericke University Magdeburg), and Ayoub Al-Hamadi (University of Magdeburg)

Time to Shine: Fine-Tuning Object Detection Models with Synthetic Adverse Weather Images 4435
Thomas Rothmeier (University of Applied Sciences Ingolstadt), Werner Huber (University of Applied Sciences Ingolstadt), and Alois C. Knoll (Technical University of Munich)

Effects of Markers in Training Datasets on the Accuracy of 6D Pose Estimation 4445
Janis Rosskamp (University of Bremen), Gabriel Zachmann (University of Bremen), and Rene Weller (University of Bremen)

VEATIC: Video-Based Emotion and Affect Tracking in Context Dataset 4455
Zhihang Ren (University of California, Berkeley), Jefferson Ortega (University of California, Berkeley), Yifan Wang (University of California, Berkeley), Zhimin Chen (University of California, Berkeley), Yunhui Guo (University of Texas at Dallas), Stella X. Yu (University of California, Berkeley; University of Michigan, Ann Arbor), and David Whitney (University of California, Berkeley)

Tackling Data Bias in MUSIC-AVQA: Crafting a Balanced Dataset for Unbiased Question-Answering 4466
Xiulong Liu (University of Washington), Zhikang Dong (Stony Brook University), and Peng Zhang (Bytedance Inc)

An Empirical Investigation Into Benchmarking Model Multiplicity for Trustworthy Machine Learning: A Case Study on Image Classification 4476
Prakhar Ganesh (Mila, Quebec AI Insitute)

Can You Even Tell Left From Right? Presenting a New Challenge for VQA 4486
Sai Raam Venkataraman (Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam, India), Rishi Sridhar Rao (Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam, India), S. Balasubramanian (Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam, India), R. Raghunatha Sarma (Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam, India), and Chandra Sekhar Vorugunti (Samsung Research, India)

MuSHRoom: Multi-Sensor Hybrid Room Dataset for Joint 3D Reconstruction and Novel View Synthesis 4496
Xuqian Ren (Tampere University, Finland), Wenjia Wang (The University of Hong Kong, China), Dingding Cai (Tampere University, Finland), Tuuli Tuominen (Tampere University, Finland), Juho Kannala (Aalto University, Finland), and Esa Rahtu (Tampere University, Finland)

RobustCLEVR: A Benchmark and Framework for Evaluating Robustness in Object-Centric Learning 4506
Nathan Drenkow (The Johns Hopkins University; The Johns Hopkins University Applied Physics Laboratory) and Mathias Unberath (The Johns Hopkins University)

So You Think You Can Track? 4516
Derek Gloudemans (Vanderbilt University), Gergely Zachár (Vanderbilt University), Yanbing Wang (Vanderbilt University), Junyi Ji (Vanderbilt University), Matt Nice (Vanderbilt University), Matt Bunting (Vanderbilt University), William W. Barbour (Vanderbilt University), Jonathan Sprinkle (Vanderbilt University), Benedetto Piccoli (Rutgers University - Camden), Alexandre Bayen (UC Berkeley), Maria Laura Delle Monache (UC Berkeley, University Ave and Oxford St), Benjamin Seibold (Temple University), and Daniel B. Work (Vanderbilt University)

Estimating Blood Alcohol Level Through Facial Features for Driver Impairment Assessment 4527
Ensiyeh Keshtkaran (Edith Cowan University, Western Australia), Brodie Von Berg (MiX Telematics, Western Australia), Grant Regan (MiX Telematics, Western Australia), David Suter (Edith Cowan University, Western Australia), and Syed Zulqarnain Gilani (Edith Cowan University; Nutrition and Health Innovation Research Institute, ECU, Western Australia)

ENIGMA-51: Towards a Fine-Grained Understanding of Human Behavior in Industrial Scenarios 4537
Francesco Ragusa (FPV@IPLab - University of Catania; Next Vision s.r.l. - Spinoff of the University of Catania, Italy), Rosario Leonardi (FPV@IPLab - University of Catania, Italy), Michele Mazzamuto (FPV@IPLab - University of Catania; Next Vision s.r.l. - Spinoff of the University of Catania, Italy), Claudia Bonanno (FPV@IPLab - University of Catania; Next Vision s.r.l. - Spinoff of the University of Catania, Italy), Rosario Scavo (FPV@IPLab - University of Catania, Italy), Antonino Furnari (FPV@IPLab - University of Catania; Next Vision s.r.l. - Spinoff of the University of Catania, Italy), and Giovanni Maria Farinella (FPV@IPLab - University of Catania; Next Vision s.r.l. - Spinoff of the University of Catania, Italy)

SciOL and MuLMS-Img: Introducing a Large-Scale Multimodal Scientific Dataset and Models for Image-Text Tasks in the Scientific Domain 4548
Tim Tarsi (Bosch Center for Artificial Intelligence, Germany), Heike Adel (Hochschule der Medien, Germany), Jan Hendrik Metzen (Bosch Center for Artificial Intelligence, Germany), Dan Zhang (Bosch Center for Artificial Intelligence, Germany), Matteo Finco (Robert Bosch GmbH, Germany), and Annemarie Friedrich (University of Augsburg, Germany)

HaGRID — HAnd Gesture Recognition Image Dataset 4560
Kapitanov Alexander (SaluteDevices, Russia), Kvanchiani Karina (SaluteDevices, Russia), Nagaev Alexander (SaluteDevices, Russia), Kraynov Roman (SaluteDevices, Russia), and Makhliarchuk Andrei (SaluteDevices, Russia)

Identifying Label Errors in Object Detection Datasets by Loss Inspection 4570
Marius Schubert (University of Wuppertal, Germany), Tobias Riedlinger (University of Wuppertal, Germany), Karsten Kahl (University of Wuppertal, Germany), Daniel Kröll (ControlExpert GmbH, Germany), Sebastian Schoenen (ControlExpert GmbH, Germany), Siniša Šegvić (University of Zagreb, Croatia), and Matthias Rottmann (University of Wuppertal, Germany)

Exploring the Impact of Rendering Method and Motion Quality on Model Performance When Using Multi-View Synthetic Data for Action Recognition 4580
Stanislav Panev (Carnegie Mellon University, USA), Emily Kim (Carnegie Mellon University, USA), Sai Abhishek Siddhartha Namburu (Carnegie Mellon University, USA), Desislava V. Nikolova (Technical University of Sofia, Bulgaria), Celso DeMelo (DEVCOM Army Research Laboratory, USA), Fernando De LaTorre (Carnegie Mellon University, USA), and Jessica Hodgins (Carnegie Mellon University, USA)

PsyMo: A Dataset for Estimating Self-Reported Psychological Traits From Gait 4591
Adrian Cosma (University Politehnica of Bucharest, Romania) and Emilian Radoi (University Politehnica of Bucharest, Romania)

IDD-AW: A Benchmark for Safe and Robust Segmentation of Drive Scenes in Unstructured Traffic and AdverseWeather 4602
Furqan Ahmed Shaik (IIIT Hyderabad), Abhishek Reddy Malreddy (IIIT Hyderabad), Nikhil Reddy Billa (IIIT Hyderabad), Kunal Chaudhary (DYSL AI, DRDO), Sunny Manchanda (DYSL AI, DRDO), and Girish Varma (IIIT Hyderabad)

CrashCar101: Procedural Generation for Damage Assessment 4612
Jens Parslov (Technical University of Denmark), Erik Riise (Technical University of Denmark), and Dim P. Papadopoulos (Technical University of Denmark; Pioneer Center for AI)

ZRG: A Dataset for Multimodal 3D Residential Rooftop Understanding 4623
Isaac A Corley (University of Texas at San Antonio), Jonathan R Lwowski (Zeitview), and Peyman Najafirad (University of Texas at San Antonio)

Explainable, Fair, Accountable, Privacy-Preserving, Ethical Computer Vision

Taming Normalizing Flows 4632
Shimon Malnick (Tel Aviv University), Shai Avidan (Tel Aviv University), and Ohad Fried (Reichman University)

Improving Fairness in Deepfake Detection 4643
Yan Ju (University at Buffalo, State University of New York), Shu Hu (University at Buffalo, State University of New York), Shan Jia (University at Buffalo, State University of New York), George H. Chen (Carnegie Mellon University), and Siwei Lyu (University at Buffalo, State University of New York)

Causal Feature Alignment: Learning to Ignore Spurious Background Features 4654
Rahul Venkataramani (Indian Institute of Science, Bangalore), Parag Dutta (Indian Institute of Science, Bangalore), Vikram Melapudi (GE Healthcare, Bangalore), and Ambedkar Dukkipati (Indian Institute of Science, Bangalore)

Localization and Manipulation of Immoral Visual Cues for Safe Text-to-Image Generation 4663
Seongbeom Park (CSE, Korea University), Suhong Moon (EECS, UC Berkeley), Seunghyun Park (NAVER Cloud AI), and Jinkyu Kim (CSE, Korea University)

Causal Analysis for Robust Interpretability of Neural Networks 4673
Ola Ahmad (Thales Digital Solutions, CortAIx, Canada), Nicolas Béreux (Paris-Saclay University, France), Loïc Baret (Thales Digital Solutions, CortAIx, Canada), Vahid Hashemi (AUDI AG, Germany), and Freddy Lecue (Inria, Sophia Antipolis, France)

Improving Fairness Using Vision-Language Driven Image Augmentation 4683
Moreno D'incà (University of Trento), Christos Tzelepis (City, University of London), Ioannis Patras (Queen Mary University of London), and Nicu Sebe (University of Trento)

Linking Convolutional Kernel Size to Generalization Bias in Face Analysis CNNs 4693
Hao Liang (Rice University, USA), Josue Ortega Caro (Yale University, USA), Vikram Maheshri (Houston University, USA), Ankit B. Patel (Rice University, USA), and Guha Balakrishnan (Rice University, USA)

PATROL: Privacy-Oriented Pruning for Collaborative Inference Against Model Inversion Attacks 4704
Shiwei Ding (Michigan Tech University), Lan Zhang (Michigan Technological University), Miao Pan (University of Houston), and Xiaoyong Yuan (Michigan Technological University)

Towards Visual Saliency Explanations of Face Verification 4714
Yuhang Lu (EPFL, Switzerland), Zewei Xu (EPFL, Switzerland), and Touradj Ebrahimi (EPFL, Switzerland)

Efficient Explainable Face Verification Based on Similarity Score Argument Backpropagation..... 4724
Marco Huber (Fraunhofer Institute for Computer Graphics Research IGD; TU Darmstadt, Germany), Anh Thi Luu (Fraunhofer Institute for Computer Graphics Research IGD, Germany), Philipp Terhörst (Paderborn University, Germany), and Naser Damer (Fraunhofer Institute for Computer Graphics Research IGD; TU Darmstadt, Germany)

SynthProv: Interpretable Framework for Profiling Identity Leakage 4734
Jaisidh Singh (IIT Jodhpur, India), Harshil Bhatia (IIT jodhpur, India), Mayank Vatsa (IIT Jodhpur, India), Richa Singh (IIT Jodhpur, India), and Aparna Bharati (Lehigh University, USA)

Text-to-Image Models for Counterfactual Explanations: A Black-Box Approach 4745
Guillaume Jeanneret (Normandy University, ENSICAEN, UNICAEN, CNRS, GREYC, France), Loïc Simon (Normandy University, ENSICAEN, UNICAEN, CNRS, GREYC, France), and Frédéric Jurie (Normandy University, ENSICAEN, UNICAEN, CNRS, GREYC, France)

Pixel-Grounded Prototypical Part Networks 4756
Zachariah Carmichael (University of Notre Dame; Mitsubishi Electric Research Laboratories), Suhas Lohit (Mitsubishi Electric Research Laboratories), Anoop Cherian (Mitsubishi Electric Research Laboratories), Michael J. Jones (Mitsubishi Electric Research Laboratories), and Walter J. Scheirer (University of Notre Dame)

How Do Deepfakes Move? Motion Magnification for Deepfake Source Detection 4768
İlke Demir (Intel Labs) and Umur Aybars Çiftçi (Binghamton University)

Hybrid Sample Synthesis-Based Debiasing of Classifier in Limited Data Setting 4779
Piyush Arora (Indian Institute of Technology Jodhpur, India) and Pratik Mazumder (Indian Institute of Technology Jodhpur, India)

Analyzing the Domain Shift Immunity of Deep Homography Estimation 4788
Mingzhen Shao (University of Utah), Tolga Tasdizen (University of Utah, USA), and Sarang Joshi (University of Utah, USA)

RPCANet: Deep Unfolding RPCA Based Infrared Small Target Detection 4797
Fengyi Wu (University of Electronic Science and Technology of China), Tianfang Zhang (University of Electronic Science and Technology of China), Lei Li (University of Copenhagen), Yian Huang (University of Electronic Science and Technology of China), and Zhenming Peng (University of Electronic Science and Technology of China)

Learn to Unlearn for Deep Neural Networks: Minimizing Unlearning Interference with Gradient Projection 4807
Tuan Hoang (Deakin University), Santu Rana (Deakin University), Sunil Gupta (Deakin University), and Svetha Venkatesh (Deakin University)

Occlusion Sensitivity Analysis with Augmentation Subspace Perturbation in Deep Feature Space 4817
Pedro H. V. Valois (University of Tsukuba), Koichiro Niinuma (Fujitsu Research of America), and Kazuhiro Fukui (University of Tsukuba)

Generated Distributions Are All You Need for Membership Inference Attacks Against Generative Models 4827
Minxing Zhang (CISPA Helmholtz Center for Information Security), Ning Yu (Salesforce Research), Rui Wen (CISPA Helmholtz Center for Information Security), Michael Backes (CISPA Helmholtz Center for Information Security), and Yang Zhang (CISPA Helmholtz Center for Information Security)

Do We Still Need Non-Maximum Suppression? Accurate Confidence Estimates and Implicit Duplication Modeling with IoU-Aware Calibration 4838
Johannes Gilg (Technical University of Munich), Torben Teepe (Technical University of Munich), Fabian Herzog (Technical University of Munich), Philipp Wolters (Technical University Munich), and Gerhard Rigoll (Technical University Munich)

Towards More Realistic Membership Inference Attacks on Large Diffusion Models 4848
Jan Dubiński (Warsaw University of Technology; IDEAS NCBR), Antoni Kowalczuk (Warsaw University of Technology), Stanisław Pawlak (Warsaw University of Technology), Przemyslaw Rokita (Warsaw University of Technology), Tomasz Trzcinski (Warsaw University of Technology; IDEAS NCBR; Tooploox), and Paweł Morawiecki (Polish Academy of Sciences)

ClusterFix: A Cluster-Based Debiasing Approach Without Protected-Group Supervision 4858
Giacomo Capitani (Università degli Studi di Modena e Reggio Emilia, Italy), Federico Bolelli (Università degli Studi di Modena e Reggio Emilia, Italy), Angelo Porrello (Università degli Studi di Modena e Reggio Emilia, Italy), Simone Calderara (Università degli Studi di Modena e Reggio Emilia, Italy), and Elisa Ficarra (Università degli Studi di Modena e Reggio Emilia, Italy)

Concept-Centric Transformers: Enhancing Model Interpretability through Object-Centric Concept Learning within a Shared GlobalWorkspace 4868
Jinyung Hong (Arizona State University), Keun Hee Park (Arizona State University), and Theodore P. Pavlic (Arizona State University)

Membership Inference Attack Using Self Influence Functions 4880
Gilad Cohen (Tel Aviv University) and Raja Giryes (Tel Aviv University)

CARE: Counterfactual-Based Algorithmic Recourse for Explainable Pose Correction 4890
Bhat Dittakavi (Indian Institute of Technology Hyderabad), Bharathi Callepalli (Variance.ai), Aleti Vardhan (Manipal Institute of Technology), Sai Vikas Desai (Indian Institute of Technology Hyderabad), and Vineeth N Balasubramanian (Indian Institute of Technology Hyderabad)

Generative Models For Image, Video, 3D, etc.

Spiking Denoising Diffusion Probabilistic Models 4900
Jiahang Cao (The Hong Kong University of Science and Technology (Guangzhou)), Ziqing Wang (The Hong Kong University of Science and Technology (Guangzhou); North Carolina State University), Hanzhong Guo (Renmin University of China), Hao Cheng (The Hong Kong University of Science and Technology (Guangzhou)), Qiang Zhang (The Hong Kong University of Science and Technology (Guangzhou)), and Renjing Xu (The Hong Kong University of Science and Technology (Guangzhou))

Improving the Effectiveness of Deep Generative Data 4910
Ruyu Wang (Bosch Center for Artificial Intelligence; University of Stuttgart, Germany), Sabrina Schmedding (Bosch Center for Artificial Intelligence, Germany), and Marco F. Huber (University of Stuttgart; Fraunhofer Institute for Manufacturing Engineering and Automation (IPA), Germany)

Customizing 360-Degree Panoramas Through Text-to-Image Diffusion Models 4921
Hai Wang (University College London), Xiaoyu Xiang (Meta Reality Labs), Yuchen Fan (Meta Reality Labs), and Jing-Hao Xue (University College London)

Label Augmentation As Inter-Class Data Augmentation for Conditional Image Synthesis with Imbalanced Data 4932
Kai Katsumata (The University of Tokyo, Japan), Duc Minh Vo (The University of Tokyo, Japan), and Hideki Nakayama (The University of Tokyo, Japan)

3D-Aware Talking-Head Video Motion Transfer 4942
Haomiao Ni (The Pennsylvania State University, USA), Jiachen Liu (The Pennsylvania State University, USA), Yuan Xue (The Ohio State University, USA), and Sharon X. Huang (The Pennsylvania State University, USA)

GRIT: GAN Residuals for Paired Image-to-Image Translation 4953
Saksham Suri (University of Maryland, College Park), Moustafa Meshry (University of Maryland, College Park), Larry S. Davis (University of Maryland, College Park), and Abhinav Shrivastava (University of Maryland, College Park)

Multimodality-Guided Image Style Transfer Using Cross-Modal GAN Inversion 4964
Hanyu Wang (University of Maryland, College Park), Pengxiang Wu (Snap Inc.), Kevin Dela Rosa (Snapchat Inc.), Chen Wang (Snap Inc.), and Abhinav Shrivastava (University of Maryland, College Park)

ZIGNeRF: Zero-Shot 3D Scene Representation with Invertible Generative Neural Radiance Fields 4974
Kanghyeok Ko (Chung-Ang University, South Korea) and Minhyeok Lee (Chung-Ang University, South Korea)

GraphFill: Deep Image Inpainting Using Graphs 4984
Shashikant Verma (Indian Institute of Technology Gandhinagar, India), Aman Sharma (Samsung R&D Institute Bangalore, India), Roopa Sheshadri (Samsung R&D Institute Bangalore, India), and Shanmuganathan Raman (Indian Institute of Technology (IIT) Gandhinagar, India)

Nested Diffusion Processes for Anytime Image Generation 4995
Noam Elata (Technion – Israel Institute of Technology), Bahjat Kawar (Technion – Israel Institute of Technology), Tomer Michaeli (Technion – Israel Institute of Technology), and Michael Elad (Technion – Israel Institute of Technology)

PoseDiff: Pose-Conditioned Multimodal Diffusion Model for Unbounded Scene Synthesis From Sparse Inputs 5005
Seoyoung Lee (Seoul National University) and Joonseok Lee (Google Research & Seoul National University)

Expanding Expressiveness of Diffusion Models with Limited Data via Self-Distillation Based Fine-Tuning 5016
Jiwan Hur (KAIST, South Korea), Jaehyun Choi (KAIST, South Korea), Gyojin Han (KAIST, South Korea), Dong-Jae Lee (KAIST, South Korea), and Junmo Kim (KAIST, South Korea)

One Style Is All You Need to Generate a Video 5026
Sandeep Manandhar (IBENS, Ecole Normale Supérieure, France) and Auguste Genovesio (IBENS, Ecole Normale Supérieure, France)

Consistent Multimodal Generation via a Unified GAN Framework 5036
Zhen Zhu (UIUC), Yijun Li (Adobe Inc.), Weijie Lyu (UIUC), Krishna Kumar Singh (Adobe Inc.), Zhixin Shu (Adobe Inc.), Söeren Pirk (Adobe Inc.), and Derek Hoiem (UIUC)

Unsupervised Co-Generation of Foreground-Background Segmentation From Text-to-Image Synthesis 5046
Yeruru Asrar Ahmed (Indian Institute of Technology Madras) and Anurag Mittal (Indian Institute of Technology Madras)

MotionGPT: Human Motion Synthesis With Improved Diversity and Realism via GPT-3 Prompting ... 5058
Jose Ribeiro-Gomes (Instituto Superior Tecnico), Tianhui Cai (Carnegie Mellon University), Zoltán Á Milacski (Carnegie Mellon University), Chen Henry Wu (Carnegie Mellon University), Aayush Prakash (Meta), Shingo J Takagi (Meta), Amaury Aubel (Meta), Daeil Kim (Meta), Alexandre Bernardino (missing), and Fernando De La Torre (Carnegie Mellon)

Human Motion Aware Text-to-Video Generation with Explicit Camera Control 5069
Taehoon Kim (Pukyong National University), ChanHee Kang (Sogang University), Jaehyuk Park (Pukyong National University), Daun Jeong (Pukyong National University), Changhee Yang (Sogang University), Suk-Ju Kang (Sogang University), and Kyeongbo Kong (Pusan National University)

Second-Order Graph ODEs for Multi-Agent Trajectory Forecasting 5079
Song Wen (Rutgers University), Hao Wang (Rutgers University), Di Liu (Rutgers University), Qilong Zhangli (Rutgers University), and Dimitris Metaxas (Rutgers University)

Diffused Heads: Diffusion Models Beat GANs on Talking-Face Generation 5089
Michał Stypułkowski (University of Wroclaw), Konstantinos Vougioukas (Imperial College London), Sen He (University of Surrey), Maciej Zieba (Wroclaw University of Science and Technology), Stavros Petridis (Imperial College London), and Maja Pantic (Imperial College London)

Unified Concept Editing in Diffusion Models 5099
Rohit Gandikota (Northeastern University), Hadas Orgad (Technion), Yonatan Belinkov (Technion), Joanna Materzyńska (Massachusetts Institute of Technology), and David Bau (Northeastern University)

SpectralCLIP: Preventing Artifacts in Text-Guided Style Transfer From a Spectral Perspective 5109
Zipeng Xu (University of Trento, Italy), Songlong Xing (University of Trento, Italy), Enver Sangineto (University of Modena and Reggio Emilia, Italy), and Nicu Sebe (University of Trento, Italy)

Diffusion-Based Generation of Histopathological Whole Slide Images at a Gigapixel Scale 5119
Robert Harb (Graz University of Technology; Medical University of Graz, Austria), Thomas Pock (Graz University of Technology, Austria), and Heimo Müller (Medical University Graz, Austria)

Painterly Image Harmonization via Adversarial Residual Learning 5129
Xudong Wang (Shanghai Jiao Tong University), Li Niu (Shanghai Jiao Tong University), Junyan Cao (Shanghai Jiao Tong University), Yan Hong (Shanghai Jiao Tong University), and Liqing Zhang (Shanghai Jiao Tong University)

Training-Free Content Injection Using H-Space in Diffusion Models 5139
Jaeseok Jeong (Yonsei University, Republic of Korea), Mingi Kwon (Yonsei University, Republic of Korea), and Youngjung Uh (Yonsei University, Republic of Korea)

ENTED: Enhanced Neural Texture Extraction and Distribution for Reference-Based Blind Face Restoration 5150
Yuen-Fui Lau (HKUST), Tianjia Zhang (HKUST), Zhefan Rao (HKUST), and Qifeng Chen (HKUST)

PIDiffu: Pixel-Aligned Diffusion Model for High-Fidelity Clothed Human Reconstruction 5160
Jungeun Lee (Korea Institute of Science and Technology), Sanghun Kim (Korea Institute of Science and Technology), Hansol Lee (KIST), Tserendorj Adiya (Korea Institute of Science and Technology), and Hwasup Lim (Korea Institute of Science and Technology)

PathLDM: Text Conditioned Latent Diffusion Model for Histopathology 5170
Srikar Yellapragada (Stony Brook University), Alexandros Graikos (Stony Brook University), Prateek Prasanna (Stony Brook University), Tahsin Kurc (Stony Brook University), Joel Saltz (Stony Brook University), and Dimitris Samaras (Stony Brook University)

Content-Aware Image Color Editing with Auxiliary Color Restoration Tasks 5180
Yixuan Ren (University of Maryland, College Park), Jing Shi (Adobe Research), Zhifei Zhang (Adobe Research), Yifei Fan (Adobe Research), Zhe Lin (Adobe Research), Bo He (University of Maryland, College Park), and Abhinav Shrivastava (University of Maryland, College Park)

On Manipulating Scene Text in theWild with Diffusion Models 5190
Joshua Santoso (Independent Researcher), Christian Simon (The Australian National University), and Williem Williem (Independent Researcher)

CXR-IRGen: An Integrated Vision and Language Model for the Generation of Clinically Accurate Chest X-Ray Image-Report Pairs 5200
Junjie Shentu (Durham University) and Noura Al Moubayed (University of Durham)

Face Identity-Aware Disentanglement in StyleGAN 5210
Adrian Suwała (Jagiellonian University), Bartosz Wójcik (Jagiellonian University), Magdalena Proszewska (Jagiellonian University), Jacek Tabor (Jagiellonian University), Przemysław Spurek (Jagiellonian University), and Marek Smieja (Jagiellonian University)

Text-to-Image Editing by Image Information Removal 5220
Zhongping Zhang (Boston University), Jian Zheng (Amazon Alexa AI), Jacob Zhiyuan Fang (Amazon Alexa AI), and Bryan A. Plummer (Boston University)

Preserving Image Properties Through Initializations in Diffusion Models 5230
Jeffrey Zhang (n/a), Shao-Yu Chang (n/a), Kedan Li (n/a), and David Forsyth (n/a)

GC-VTON: Predicting Globally Consistent and Occlusion Aware Local Flows with Neighborhood Integrity Preservation for Virtual Try-On 5239
Hamza Rawal (Motive), Muhammad Junaid Ahmad (Motive), and Farooq Zaman (Information Technology University, Pakistan)

Generation of Upright Panoramic Image From Non-Upright Panoramic Image 5249
Jingguo Liu (Southwest University), Heyu Chen (Southwest University), Shigang Li (Hiroshima City University), and Jianfeng Li (Southwest University)

Fast Diffusion EM: A Diffusion Model for Blind Inverse Problems with Application to Deconvolution 5259
Charles Laroche (GoPro & MAP5), Andrés Almansa (CNRS & Université Paris Cité), and Eva Coupete (GoPro)

Enforcing Sparsity on Latent Space for Robust and Explainable Representations 5270
Hanao Li (Stevens Institute of Technology) and Tian Han (Stevens Institute of Technology)

Diff2Lip: Audio Conditioned Diffusion Models for Lip-Synchronization 5280
Soumik Mukhopadhyay (University of Maryland, College Park), Saksham Suri (University of Maryland, College Park), Ravi Teja Gadde (New York University), and Abhinav Shrivastava (University of Maryland, College Park)

3D Reconstruction of Interacting Multi-Person in Clothing From a Single Image 5291
Junuk Cha (UNIST), Hansol Lee (UNIST; KIST), Jaewon Kim (UNIST), Nhat Nguyen Bao Truong (UNIST; KAIST), Jae Shin Yoon (Adobe Research), and Seungryul Baek (UNIST; KAIST)

Revisiting Latent Space of GAN Inversion for Robust Real Image Editing 5301
Kai Katsumata (The University of Tokyo), Duc Minh Vo (The University of Tokyo), Bei Liu (Microsoft Research), and Hideki Nakayama (The University of Tokyo)

Soft Curriculum for Learning Conditional GANs with Noisy-Labeled and Uncurated Unlabeled Data 5311
Kai Katsumata (The University of Tokyo), Duc Minh Vo (The University of Tokyo), Tatsuya Harada (The University of Tokyo; RIKEN), and Hideki Nakayama (The University of Tokyo)

Bipartite Graph Diffusion Model for Human Interaction Generation 5321
Baptiste Chopin (Inria, Université Côte d'Azur, France), Hao Tang (Carnegie Mellon University, USA), and Mohamed Daoudi (Univ. Lille, CNRS, Centrale Lille, Institut Mines-Télécom; IMT Nord Europe, Institut Mines-Télécom, Univ. Lille, France)

Training-Free Layout Control with Cross-Attention Guidance 5331
Minghao Chen (University of Oxford), Iro Laina (University of Oxford), and Andrea Vedaldi (University of Oxford)

Controllable Image Synthesis of Industrial Data Using Stable Diffusion 5342
Gabriele Valvano (Baker Hughes), Antonino Agostino (Baker Hughes), Giovanni De Magistris (Baker Hughes), Antonino Graziano (Baker Hughes), and Giacomo Veneri (Baker Hughes)

Removing the Quality Tax in Controllable Face Generation 5352
Yiwen Huang (Brown University), Zhiqiu Yu (Brown University), Xinjie Yi (Brown University), Yue Wang (Stony Brook University), and James Tompkin (Brown University)

Hierarchical Diffusion Autoencoders and Disentangled Image Manipulation 5362
Zeyu Lu (Shanghai Jiao Tong University; Shanghai Artificial Intelligence Laboratory), Chengyue Wu (The University of Hong Kong), Xinyuan Chen (Shanghai Artificial Intelligence Laboratory), Yaohui Wang (Shanghai Artificial Intelligence Laboratory), Lei Bai (Shanghai Artificial Intelligence Laboratory), Yu Qiao (Shanghai Artificial Intelligence Laboratory), and Xihui Liu (The University of Hong Kong)

FacadeNet: Conditional Facade Synthesis via Selective Editing 5372
Yiangos Georgiou (Univesity of Cyprus/CYENS CoE, Cyprus), Marios Loizou (Univesity of Cyprus/CYENS CoE, Cyprus), Tom Kelly (KAUST, Saudi Arabia), and Melinos Averkiou (Univesity of Cyprus/CYENS CoE, Cyprus)

TPSeNCE: Towards Artifact-Free Realistic Rain Generation for Deraining and Object Detection in Rain 5382
Shen Zheng (Carnegie Mellon University), Changjie Lu (University of Illinois Urbana-Champaign), and Srinivasa G. Narasimhan (Carnegie Mellon University)

Common Diffusion Noise Schedules and Sample Steps Are Flawed 5392
Shanchuan Lin (ByteDance), Bingchen Liu (Bytedance Inc.), Jiashi Li (ByteDance Inc), and Xiao Yang (Bytedance AI Lab)

Improving the Leaking of Augmentations in Data-Efficient GANs via Adaptive Negative Data Augmentation 5400
Zhaoyu Zhang (Queen's University Belfast), Yang Hua (Queen's University Belfast), Guanxiong Sun (Queen's University; Huawei UKRD Belfast), Hui Wang (Queen's University Belfast), and Seán Mcloone (Queen's University Belfast)

P2D: Plug and Play Discriminator for Accelerating GAN Frameworks 5410
Min Jin Chong (ByteDance Inc.), Krishna Kumar Singh (Adobe Research), Yijun Li (Adobe Research), Jingwan Lu (Adobe Research), and David Forsyth (UIUC)

Personalized Face Inpainting with Diffusion Models by Parallel Visual Attention 5420
Jianjin Xu (Carnegie Mellon University), Saman Motamed (Carnegie Mellon University; INSAIT, Sofia University), Praneetha Vaddamanu (Carnegie Mellon University), Chen Henry Wu (Carnegie Mellon University), Christian Haene (Independent Researcher), Jean-Charles Bazin (Independent Researcher), and Fernando de la Torre (Carnegie Mellon University)

Semantic Generative Augmentations for Few-Shot Counting 5431
Perla Doubinsky (CEDRIC (EA4329), Cnam Paris, France), Nicolas Audebert (CEDRIC (EA4329), Cnam Paris, France), Michel Crucianu (CEDRIC (EA4329), Cnam Paris, France), and Hervé Le Borgne (Université Paris-Saclay, CEA List, France)

StyleGAN-Fusion: Diffusion Guided Domain Adaptation of Image Generators 5441
Kunpeng Song (Rutgers University), Ligong Han (Rutgers University), Bingchen Liu (Bytedance Inc.), Dimitris Metaxas (Rutgers University), and Ahmed Elgammal (Rutgers University; Playform AI)

Vision + Language and/or Other Modalities

Zero-Shot Video Moment Retrieval From Frozen Vision-Language Models 5452
Dezhao Luo (Queen Mary University of London), Jiabo Huang (Queen Mary University of London), Shaogang Gong (Queen Mary University of London), Hailin Jin (Adobe Research), and Yang Liu (Peking University)

Investigating the Role of Attribute Context in Vision-Language Models for Object Recognition and Detection 5462
Kyle Buettner (University of Pittsburgh, USA) and Adriana Kovashka (University of Pittsburgh, USA)

Benchmarking Out-of-Distribution Detection in Visual Question Answering 5473
Xiangxi Shi (Oregon State University) and Stefan Lee (Oregon State University)

Sound3DVDet: 3D Sound Source Detection Using Multiview Microphone Array and RGB Images 5484
Yuhang He (University of Oxford, UK), Sangyun Shin (University of Oxford, UK), Anoop Cherian (Mitsubishi Electric Research Labs, US), Niki Trigoni (University of Oxford, UK), and Andrew Markham (University of Oxford, UK)

LAVSS: Location-Guided Audio-Visual Spatial Audio Separation 5496
Yuxin Ye (Tsinghua University, China), Wenming Yang (Tsinghua University, China), and Yapeng Tian (The University of Texas at Dallas, USA)

Augment the Pairs: Semantics-Preserving Image-Caption Pair Augmentation for Grounding-Based Vision and Language Models 5508
Jingru Yi (Amazon), Burak Uzkent (Stanford University), Oana Ignat (University of Michigan), Zili Li (Amazon), Amanmeet Garg (Amazon), Xiang Yu (Amazon), and Linda Liu (Amazon)

CLID: Controlled-Length Image Descriptions with Limited Data 5519
Elad Hirsch (Technion – Israel Institute of Technology) and Ayellet Tal (Technion – Israel Institute of Technology)

STYLIP: Multi-Scale Style-Conditioned Prompt Learning for CLIP-Based Domain Generalization. 5530
Shirsha Bose (Technical University of Munich, Germany), Ankit Jha (Indian Institute of Technology Bombay, India), Enrico Fini (University of Trento, Italy), Mainak Singha (Indian Institute of Technology Bombay, India), Elisa Ricci (University of Trento, Italy), and Biplab Banerjee (Indian Institute of Technology Bombay, India)

THInImg: Cross-Modal Steganography for Presenting Talking Heads in Images 5541
Lin Zhao (Nankai University), Hongxuan Li (Nankai University), Xuefei Ning (Tsinghua University), and Xinru Jiang (University of British Columbia)

Enhancing Multimodal Compositional Reasoning of Visual Language Models with Generative Negative Mining 5551
Ugur Sahin (Technical University of Munich), Hang Li (LMU Munich; Siemens AG), Qadeer Khan (Technical University of Munich; Munich Center for Machine Learning), Daniel Cremers (Technical University of Munich; Munich Center for Machine Learning), and Volker Tresp (LMU Munich; Munich Center for Machine Learning)

Temporal Context Enhanced Referring Video Object Segmentation 5562
Xiao Hu (University of Ottawa, Canada), Basavaraj Hampiholi (Ulm University, Germany), Heiko Neumann (Ulm University, Germany), and Jochen Lang (University of Ottawa, Canada)

Fine-Grained Alignment for Cross-Modal Recipe Retrieval 5572
Muntasir Wahed (Virginia Tech), Xiaona Zhou (Virginia Tech), Tianjiao Yu (Virginia Tech), and Ismini Lourentzou (Virginia Tech)

Learning To Adapt CLIP for Few-Shot Monocular Depth Estimation 5582
Xueting Hu (Southern University of Science and Technology), Ce Zhang (Carnegie Mellon University), Yi Zhang (Southern University of Science and Technology), Bowen Hai (Southern University of Science and Technology), Ke Yu (Southern University of Science and Technology), and Zhihai He (Southern University of Science and Technology)

Annotation-Free Audio-Visual Segmentation 5592
Jinxiang Liu (Shanghai Jiao Tong University), Yu Wang (Shanghai Jiao Tong University), Chen Ju (Cooperative Medianet Innovation Center, Shang hai Jiao Tong University), Chaofan Ma (Shanghai Jiao Tong University), Ya Zhang (Cooperative Medianet Innovation Center, Shang hai Jiao Tong University), and Weidi Xie (Shanghai Jiao Tong University)

Rethink Cross-Modal Fusion inWeakly-Supervised Audio-Visual Video Parsing 5603
Yating Xu (National University of Singapore), Conghui Hu (National University of Singapore), and Gim Hee Lee (National University of Singapore)

SDNet: An Extremely Efficient Portrait Matting Model via Self-Distillation 5613
Ziwen Li (OPPO Research Institute), Bo Xu (OPPO Research Institute), Jiake Xie (PicUP.Ai), Yong Tang (PicUP.Ai), and Cheng Lu (XPENG)

FELGA: Unsupervised Fragment Embedding for Fine-Grained Cross-Modal Association 5623
Yaoxin Zhuo (Arizona State University) and Baoxin Li (Arizona State University)

Modality-Aware Representation Learning for Zero-Shot Sketch-Based Image Retrieval 5634
Eunyi Lyou (Seoul National University), Doyeon Lee (Seoul National University), Jooeun Kim (Seoul National University), and Joonseok Lee (Seoul National University; Google Research)

Multitask Vision-Language Prompt Tuning 5644
Sheng Shen (University of California, Berkeley), Shijia Yang (University of California, Berkeley), Tianjun Zhang (University of California, Berkeley), Bohan Zhai (University of California, Berkeley), Joseph E. Gonzalez (University of California, Berkeley), Kurt Keutzer (University of California, Berkeley), and Trevor Darrell (University of California, Berkeley)

EASUM: Enhancing Affective State Understanding Through Joint Sentiment and Emotion Modeling for Multimodal Tasks 5656
Yewon Hwang (KAIST) and Jong-Hwan Kim (KAIST)

Complementary-Contradictory Feature Regularization Against Multimodal Overfitting 5667
Antonio Tejero-de-Pablos (CyberAgent, Japan)

FuseCap: Leveraging Large Language Models for Enriched Fused Image Captions 5677
Noam Rotstein (Technion - Israel Institute of Technology), David Bensaïd (Technion - Israel Institute of Technology), Shaked Brody (Technion - Israel Institute of Technology), Roy Ganz (Technion - Israel Institute of Technology), and Ron Kimmel (Technion - Israel Institute of Technology)

Describe Images in a BoringWay: Towards Cross-Modal Sarcasm Generation 5689
Jie Ruan (Peking University), Yue Wu (Peking University), Xiaojun Wan (Peking University), and Yuesheng Zhu (Peking University)

Can CLIP Help Sound Source Localization? 5699
Sooyoung Park (ETRI), Arda Senocak (KAIST), and Joon Son Chung (KAIST)

Domain Aligned CLIP for Few-Shot Classification 5709
Muhammad Waleed Gondal (Amazon), Jochen Gast (Amazon), Inigo Alonso Ruiz (Amazon), Richard Droste (Amazon), Tommaso Macri (Amazon), Suren Kumar (Amazon), and Luitpold Staudigl (Amazon)

SCoRD: Subject-Conditional Relation Detection with Text-Augmented Data 5719
Ziyan Yang (Rice University), Kushal Kafle (Adobe Research), Zhe Lin (Adobe Research), Scott Cohen (Adobe Research), Zhihong Ding (Adobe Research), and Vicente Ordonez (Rice University)

Simple Token-Level Confidence Improves Caption Correctness 5730
Suzanne Petryk (UC Berkeley; Meta), Spencer Whitehead (Meta AI), Joseph E. Gonzalez (U.C. Berkeley), Trevor Darrell (UC Berkeley), Anna Rohrbach (TU Darmstadt), and Marcus Rohrbach (Meta)

Bi-Directional Training for Composed Image Retrieval via Text Prompt Learning 5741
Zheyuan Liu (Australian National University), Weixuan Sun (Australian National University), Yicong Hong (Australian National University), Damien Teney (Idiap Research Institute; University of Adelaide), and Stephen Gould (Australian National University)

MOPA: Modular Object Navigation with PointGoal Agents 5751
Sonia Raychaudhuri (Simon Fraser University), Tommaso Campari (University of Padova; FBK), Unnat Jain (Meta AI), Manolis Savva (Simon Fraser University), and Angel X. Chang (Simon Fraser University; Amii)

GIPCOL: Graph-Injected Soft Prompting for Compositional Zero-Shot Learning 5762
Guangyue Xu (Michigan State University), Joyce Chai (University of Michigan), and Parisa Kordjamshidi (Michigan State University)

Text-Guided Face Recognition Using Multi-Granularity Cross-Modal Contrastive Learning 5772
Md Mahedi Hasan (West Virginia University), Shoaib Meraj Sami (West Virginia University), and Nasser Nasrabadi (West Virginia University)

Leveraging Task-Specific Pre-Training to Reason Across Images and Videos 5782
Arka Sadhu (University of Southern California) and Ram Nevatia (University of Southern California)

VD-GR: Boosting Visual Dialog with Cascaded Spatial-Temporal Multi-Modal Graphs 5793
Adnen Abdessaied (University of Stuttgart, Germany), Lei Shi (University of Stuttgart, Germany), and Andreas Bulling (University of Stuttgart, Germany)

TriCoLo: Trimodal Contrastive Loss for Text to Shape Retrieval 5803
Yue Ruan (Simon Fraser University), Han-Hung Lee (Simon Fraser University), Yiming Zhang (Simon Fraser University), Ke Zhang (Simon Fraser University), and Angel X. Chang (Simon Fraser University; Alberta Machine Intelligence Institute (Amii))

Visualization

3D Super-Resolution Model for Vehicle Flow Field Enrichment 5814
Thanh Luan Trinh (Mobility & AI Laboratory, Nissan Motor Co., Ltd.; Yokohama National University, Japan), Fangge Chen (Mobility & AI Laboratory, Nissan Motor Co., Ltd., Japan), Takuya Nanri (Mobility & AI Laboratory, Nissan Motor Co., Ltd., Japan), and Kei Akasaka (Nissan Motor Co., Ltd, Japan)

AssemblyNet: A Point Cloud Dataset and Benchmark for Predicting Part Directions in an Exploded Layout 5824
Jesper Gaarsdal (SynergyXR ApS; Aalborg University, Denmark), Joakim Bruslund Haurum (Aalborg University & Pioneer Centre for AI, Denmark), Sune Wolff (SynergyXR ApS, Denmark), and Claus Brøndgaard Madsen (Aalborg University, Denmark)

EvDNeRF: Reconstructing Event Data with Dynamic Neural Radiance Fields 5834
Anish Bhattacharya (University of Pennsylvania), Ratnesh Madaan (Microsoft), Fernando Cladera (University of Pennsylvania), Sai Vemprala (Scaled Foundations), Rogerio Bonatti (Microsoft), Kostas Daniilidis (University of Pennsylvania), Ashish Kapoor (Scaled Foundations), Vijay Kumar (University of Pennsylvania), Nikolai Matni (University of Pennsylvania), and Jayesh K. Gupta (Poly Corporation)

LIVENet: A Novel Network for Real-World Low-Light Image Denoising and Enhancement 5844
Dhruv Makwana (Independent Researcher), Gayatri Deshmukh (Independent Researcher), Onkar Susladkar (Independent Researcher), Sparsh Mittal (IIT Roorkee), and Sai Chandra Teja R (Green PMU Semi Pvt Ltd)

Controlling Virtual Try-On Pipeline Through Rendering Policies 5854
Kedan Li (n/a), Jeffrey Zhang (n/a), Shao-Yu Chang (n/a), and David Forsyth (n/a)

Automated Camera Calibration via Homography Estimation With GNNs 5864
Giacomo D'amicantonio (Eindhoven University of Technology), Egor Bondarev (TU Eindhoven), and P. H. N. De With (Eindhoven University of Technology)

WaveMixSR: Resource-Efficient Neural Network for Image Super-Resolution 5872
Pranav Jeevan (IIT Bombay), Akella Vs Srinidhi (IIT Bombay), Pasunuri Prathiba (IITB), and Amit Sethi (Indian Institute of Technology Bombay)

People; Automotive; Medical; Environmental; Robotics

You Can Run but Not Hide: Improving Gait Recognition with Intrinsic Occlusion Type Awareness 5881
Ayush Gupta (Johns Hopkins University) and Rama Chellappa (Johns Hopkins University)

Rethinking Visibility in Human Pose Estimation: Occluded Pose Reasoning via Transformers 5891
Pengzhan Sun (National University of Singapore), Kerui Gu (National University of Singapore), Yunsong Wang (National University of Singapore), Linlin Yang (Communication University of China), and Angela Yao (National University of Singapore)

RMFER: Semi-Supervised Contrastive Learning for Facial Expression Recognition with Reaction Mashup Video 5901
Yunseong Cho (UNIST; SNOW Corp.), Chanwoo Kim (UNIST), Hoseong Cho (UNIST), Yunhoe Ku (UNIST), Eunseo Kim (UNIST), Muhammadjon Boboev (UNIST), Joonseok Lee (Seoul National University), and Seungryul Baek (UNIST)

CSAM: A 2.5D Cross-Slice Attention Module for Anisotropic Volumetric Medical Image Segmentation 5911
Alex Hung (UCLA), Haoxin Zheng (UCLA), Kai Zhao (UCLA), Xiaoxi Du (UCLA), Kaifeng Pang (UCLA), Qi Miao (UCLA), Steven S. Raman (UCLA), Demetri Terzopoulos (University of California, Los Angeles), and Kyung Hyun Sung (UCLA)

Motion Matters: Neural Motion Transfer for Better Camera Physiological Measurement 5921
Akshay Paruchuri (UNC Chapel Hill), Xin Liu (University of Washington), Yulu Pan (UNC Chapel Hill), Shwetak Patel (University of Washington), Daniel Mcduff (University of Washington), and Soumyadip Sengupta (UNC Chapel Hill)

Image Labels Are All You Need for Coarse Seagrass Segmentation 5931
Scarlett Raine (Queensland University of Technology; CSIRO Data61, Australia), Ross Marchant (Image Analytics, Australia), Brano Kusy (CSIRO Data61, Australia), Frederic Maire (Queensland University of Technology, Australia), and Tobias Fischer (Queensland University of Technology, Australia)

WildlifeDatasets: An Open-Source Toolkit for Animal Re-Identification 5941
Vojtěch Čermák (Czech Technical University in Prague), Lukas Picek (University of West Bohemia; INRIA), Lukáš Adam (Czech Technical University in Prague), and Kostas Papafitsoros (Queen Mary University of London)

HALSIE: Hybrid Approach to Learning Segmentation by Simultaneously Exploiting Image and Event Modalities 5952
Shristi Das Biswas (Purdue University, USA), Adarsh K Kosta (Purdue University, USA), Chamika Liyanagedera (Purdue university, USA), Marco Apolinario (Purdue University, USA), and Kaushik Roy (Purdue Uniiversity, USA)

ParticleNeRF: A Particle-Based Encoding for Online Neural Radiance Fields 5963
Jad Abou-Chakra (Queensland University of Technology), Feras Dayoub (University of Adelaide), and Niko Süenderhauf (Queensland University of Technology)

Rotation-Constrained Cross-View Feature Fusion for Multi-View Appearance-Based Gaze Estimation 5973
Yoichiro Hisadome (The University of Tokyo, Japan), Tianyi Wu (The University of Tokyo, Japan), Jiawei Qin (The University of Tokyo, Japan), and Yusuke Sugano (The University of Tokyo, Japan)

SICKLE: A Multi-Sensor Satellite Imagery Dataset Annotated with Multiple Key Cropping Parameters 5983
Depanshu Sani (Indraprastha Institute of Information Technology, India), Sandeep Mahato (MS Swaminathan Research Foundation, India), Sourabh Saini (Indraprastha Institute of Information Technology, India), Harsh Kumar Agarwal (Indraprastha Institute of Information Technology, India), Charu Chandra Devshali (MS Swaminathan Research Foundation, India), Saket Anand (Indraprastha Institute of Information Technology, India), Gaurav Arora (Indraprastha Institute of Information Technology, India), and Jayaraman Thiagarajan (MS Swaminathan Research Foundation, India)

Re-Evaluating LiDAR Scene Flow 5993
Nathaniel Chodosh (Carnegie Mellon University), Deva Ramanan (Carnegie Mellon University), and Simon Lucey (University of Adelaide)

Semantic Labels-Aware Transformer Model for Searching Over a Large Collection of Lecture-Slides 6004
K. V. Jobin (IIIT Hyderabad), Anand Mishra (IIT Jodhpur), and C.V. Jawahar (IIIT-Hyderabad)

ConeQuest: A Benchmark for Cone Segmentation on Mars 6014
Mirali Purohit (Arizona State University), Jacob Adler (Arizona State University), and Hannah Kerner (Arizona State University)

FastSR-NeRF: Improving NeRF Efficiency on Consumer Devices with a Simple Super-Resolution Pipeline 6024
Chien-Yu Lin (University of Washington, USA), Qichen Fu (Apple, Inc.), Thomas Merth (Apple, Inc.), Karren Yang (Apple, Inc.), and Anurag Ranjan (Apple, Inc.)

Biometrics, Face, Gesture, Body Pose II

HashReID: Dynamic Network With Binary Codes for Efficient Person Re-Identification 6034
Kshitij N Nikhal (University of Nebraska Lincoln), Yujunrong Ma (University of Maryland), Shuvra Bhattacharyya (University of Maryland), and Benjamin Riggan (University of Nebraska-Lincoln)

RGBT-Dog: A Parametric Model and Pose Prior for Canine Body Analysis Data Creation 6044
Jake Deane (University of Bath), Sinéad Kearney (University of Bath), Kwang In Kim (POSTECH), and Darren Cosker (University of Bath)

FPGAN-Control: A Controllable Fingerprint Generator for Training with Synthetic Data 6055
Alon Shoshan (Amazon), Nadav Bhonker (Amazon), Emanuel Ben Baruch (Amazon), Ori Nizan (Technion - Israel Institute of Technology), Igor Kviatkovsky (Amazon), Joshua J Engelsma (Amazon), Manoj Aggarwal (Amazon), and Gérard Medioni (Amazon)

Multimodal Channel-Mixing: Channel and Spatial Masked AutoEncoder on Facial Action Unit Detection 6065
Xiang Zhang (State University of New York at Binghamton), Huiyuan Yang (Rice University), Taoyue Wang (State Univerisity of New York at Binghamton), Xiaotian Li (State University of New York at Binghamton), and Lijun Yin (State University of New York at Binghamton)

ProS: Facial Omni-Representation Learning via Prototype-Based Self-Distillation 6075
Xing Di (Protagolabs Inc.), Yiyu Zheng (ProtagoLabs Inc.), Xiaoming Liu (Michigan State University), and Yu Cheng (The Chinese University of Hong Kong)

FG-Net: Facial Action Unit Detection with Generalizable Pyramidal Features 6087
Yufeng Yin (University of Southern California), Di Chang (University of Southern California), Guoxian Song (Bytedance), Shen Sang (ByteDance), Tiancheng Zhi (ByteDance), Jing Liu (Bytedance), Linjie Luo (ByteDance), and Mohammad Soleymani (University of Southern California)

Watch Where You Head: A View-Biased Domain Gap in Gait Recognition and Unsupervised Adaptation 6097
Gavriel Habib (OriginAI, Israel), Noa Barzilay (OriginAI, Israel), Or Shimshi (OriginAI, Israel), Rami Ben-Ari (OriginAI, Israel), and Nir Darshan (OriginAI, Israel)

Intrinsic Hand Avatar: Illumination-Aware Hand Appearance and Shape Reconstruction From Monocular RGB Video 6108
Pratik Kalshetti (Indian Institute of Technology Bombay) and Parag Chaudhuri (Indian Institute of Technology Bombay)

CVTHead: One-Shot Controllable Head Avatar With Vertex-Feature Transformer 6119
Haoyu Ma (University of California, Irvine), Tong Zhang (UC Irvine), Shanlin Sun (University of California, Irvine), Xiangyi Yan (University of California, Irvine), Kun Han (University of California Irvine), and Xiaohui Xie (University of California, Irvine)

Back to Optimization: Diffusion-Based Zero-Shot 3D Human Pose Estimation 6130
Zhongyu Jiang (University of Washington), Zhuoran Zhou (University of Washington), Lei Li (University of Copenhagen), Wenhao Chai (University of Washington), Cheng-Yen Yang (University of Washington), and Jenq-Neng Hwang (University of Washington)

POISE: Pose Guided Human Silhouette Extraction Under Occlusions 6141
Arindam Dutta (University of California, Riverside), Rohit Lal (University of California, Riverside), Dripta S. Raychaudhuri (University of California, Riverside; AWS AI Labs), Calvin-Khang Ta (University of California, Riverside), and Amit K. Roy-Chowdhury (University of California, Riverside)

Incorporating Physics Principles for Precise Human Motion Prediction 6152
Yufei Zhang (Rensselaer Polytechnic Institute), Jeffrey O. Kephart (IBM Research), and Qiang Ji (Renselaer Polytechnic Institute)

Fingervein Verification Using Convolutional Multi-Head Attention Network 6163
Raghavendra Ramachandra (NTNU, Norway) and Sushma Krupa Venkatesh (Aiba)

Multispectral Imaging for Differential Face Morphing Attack Detection: A Preliminary Study..... 6173
Raghavendra Ramachandra (NTNU, Norway), Sushma Krupa Venkatesh (Aiba), Naser Damer (Fraunhofer Institute for Computer Graphics Research IGD and TU Darmstadt), Narayan Vetrekar (Goa University), and Rajendra Gad (UoG, India)

Controlling Character Motions Without Observable Driving Source 6182
Weiyuan Li (Xiaobing.AI), Bin Dai (Xiaobing.ai), Ziyi Zhou (Xiaobing.ai), Qi Yao (Xiaobing.ai), and Baoyuan Wang (Xiaobing.AI)

DR2: Disentangled Recurrent Representation Learning for Data-Efficient Speech Video Synthesis 6192
Chenxu Zhang (ByteDance Inc), Chao Wang (Bytedance Inc), Yifan Zhao (Peking University), Shuo Cheng (Georgia Institute of Technology), Linjie Luo (ByteDance Inc), and Xiaohu Guo (The University of Texas at Dallas)

Bias and Diversity in Synthetic-Based Face Recognition 6203
Marco Huber (Fraunhofer Institute for Computer Graphics Research IGD, Germany), Anh Thi Luu (Fraunhofer Institute for Computer Graphics Research IGD, Germany), Fadi Boutros (Fraunhofer Institute for Computer Graphics Research IGD, Germany), Arjan Kuijper (Fraunhofer Institute for Computer Graphics Research IGD; TU Darmstadt, Germany), and Naser Damer (Fraunhofer Institute for Computer Graphics Research IGD; TU Darmstadt, Germany)

FarSight: A Physics-Driven Whole-Body Biometric System at Large Distance and Altitude 6215
Feng Liu (Michigan State University), Ryan C Ashbaugh (Michigan State University), Nicholas M Chimitt (Purdue University), Najmul Hassan (University of Oregon), Ali Hassani (University of Oregon), Ajay Jaiswal (UT Austin), Minchul Kim (Michigan State University), Zhiyuan Mao (Purdue University), Christopher C Perry (Michigan State University), Zhiyuan Ren (Michigan State University), Yiyang Su (Michigan State University), Pegah Varghaei (Michigan State University), Kai Wang (University of Oregon), Xingguang Zhang (Purdue University), Stanley Chan (Purdue University, USA), Arun Ross (Michigan State University), Humphrey Shi (U of Oregon; UIUC; PAIR), Zhangyang Wang (University of Texas at Austin), Anil Jain (Michigan State University), and Xiaoming Liu (Michigan State University)

AU-Aware Dynamic 3D Face Reconstruction From Videos with Transformer 6225
Chenyi Kuang (Rensselaer Polytechnic Institute), Jeffrey O. Kephart (IBM Thomas J. Watson Research Ctr.), and Qiang Ji (Renselaer Polytechnic Institute)

Handformer2T: A Lightweight Regression-Based Model for Interacting Hands Pose Estimation From a Single RGB Image 6236
Pengfei Zhang (University of California, Irvine) and Deying Kong (university of california, irvine)

Weakly-Supervised Deepfake Localization in Diffusion-Generated Images 6246
Dragoș-Constantin Țânțaru (Bitdefender), Elisabeta Oneață (Bitdefender), and Dan Oneață (University Politehnica of Bucharest)

Face Presentation Attack Detection by Excavating Causal Clues and Adapting Embedding Statistics 6257
Meiling Fang (Fraunhofer Institute for Computer Graphics Research IGD) and Naser Damer (Fraunhofer Institute for Computer Graphics Research IGD and TU Darmstadt)

Denoising and Selecting Pseudo-Heatmaps for Semi-Supervised Human Pose Estimation 6268
Zhuoran Yu (AWS AI Labs; The University of Wisconsin–Madison), Manchen Wang (AWS AI Labs), Yanbei Chen (AWS AI Labs), Paolo Favaro (AWS AI Labs), and Davide Modolo (AWS AI Labs)

ShARc: Shape and Appearance Recognition for Person Identification In-the-Wild 6278
Haidong Zhu (University of Southern California), Wanrong Zheng (University of Southern California), Zhaoheng Zheng (University of Southern California), and Ram Nevatia (U of Southern California)

Fast and Interpretable Face Identification for Out-of-Distribution Data Using Vision Transformers 6289
Hai Thanh Phan (Carnegie Mellon University), Cindy Le (Eucli.ai), Vu D Le (Phenikaa University), Yihui He (Carnegie Mellon University), and Anh Nguyen (Auburn University)

SigmML: Metric Meta-Learning for Writer Independent Offline Signature Verification in the Space of SPD Matrices 6300
Alexios Giazitzis (Telsip Laboratory, UniWA, Greece) and Elias N. Zois (Telsip Laboratory, UniWA, Greece)

Progressive Hypothesis Transformer for 3D Human Mesh Recovery 6311
Huang-Ru Liao (Academia Sinica; National Tsing Hua University, Taiwan), Jen-Chun Lin (Academia Sinica, Taiwan), and Chun-Yi Lee (National Tsing Hua University, Taiwan)

DiffBody: Diffusion-Based Pose and Shape Editing of Human Images 6321
Yuta Okuyama (University of Tsukuba), Yuki Endo (University of Tsukuba), and Yoshihiro Kanamori (University of Tsukuba)

Diffuse and Restore: A Region-Adaptive Diffusion Model for Identity-Preserving Blind Face Restoration 6331
Maitreya Suin (Johns Hopkins University), Nithin Gopalakrishnan Nair (Johns Hopkins University), Chun Pong Lau (Johns Hopkins University), Vishal Patel (Johns Hopkins University), and Rama Chellappa (Johns Hopkins University)

HMP: Hand Motion Priors for Pose and Shape Estimation From Video 6341
Enes Duran (MPI for Intelligent Systems; University of Tüebingen, Germany), Muhammed Kocabas (MPI for Intelligent Systems, Germany; ETH Zurich, Switzerland), Vasileios Choutas (MPI for Intelligent Systems, Germany), Zicong Fan (MPI for Intelligent Systems, Germany; ETH Zurich, Switzerland), and Michael J. Black (MPI for Intelligent Systems, Germany)

Approximating Intersections and Differences Between Linear Statistical Shape Models Using Markov Chain Monte Carlo 6352
Maximilian Weiherer (Friedrich-Alexander-Universität Erlangen-Nürnberg), Finn Klein (Friedrich-Alexander-Universität Erlangen-Nürnberg), and Bernhard Egger (Friedrich-Alexander-Universität Erlangen-Nürnberg)

Robust Eye Blink Detection Using Dual Embedding Video Vision Transformer 6362
Jeongmin Hong (Hanyang University), Joseph Shin (Hanyang University), Juhee Choi (Hanyang University), and Minsam Ko (Hanyang University)

EmoStyle: One-Shot Facial Expression Editing Using Continuous Emotion Parameters 6373
Bita Azari (Simon Fraser University, Canada) and Angelica Lim (Simon Fraser University, Canada)

Vikriti-ID: A Novel Approach for Real Looking Fingerprint Data-Set Generation 6383
Rishabh Shukla (Indian Institute of Technology Jammu), Aditya Manish Sinha (Indian Institute of Technology, Jammu), Vansh Singh (IIT JAMMU), and Harkeerat Kaur (IIT Jammu India)

LaughTalk: Expressive 3D Talking Head Generation with Laughter 6392
Kim Sung-Bin (POSTECH), Lee Hyun (POSTECH), Da Hye Hong (Sookmyung Women's University), Suekyeong Nam (KRAFTON), Janghoon Ju (KRAFTON), and Tae-Hyun Oh (Dept. of Electrical Engineering and Grad. School of Artificial Intelligence, POSTECH; Yonsei University)

Video Recognition & Understanding

Detecting Content Segments From Online Sports Streaming Events: Challenges and Solutions 6402
Zongyi Liu (Customer Experience and Business Trends, Amazon.com), Yarong Feng (Customer Experience and Business Trends, Amazon.com), Shunyan Luo (Customer Experience and Business Trends, Amazon.com), Yuan Ling (Customer Experience and Business Trends, Amazon.com), Shujing Dong (Customer Experience and Business Trends, Amazon.com), and Shuyi Wang (Customer Experience and Business Trends, Amazon.com)

Permutation-Aware Activity Segmentation via Unsupervised Frame-To-Segment Alignment 6412
Quoc-Huy Tran (Retrocausal, Inc.), Ahmed Mehmood (Retrocausal, Inc.), Muhammad Ahmed (Retrocausal, Inc.), Muhammad Naufil (Retrocausal, Inc.), Anas Zafar (Retrocausal, Inc.), Andrey Konin (Retrocausal, Inc.), and M. Zeeshan Zia (Retrocausal, Inc.)

OTAS: Unsupervised Boundary Detection for Object-Centric Temporal Action Segmentation 6423
Yuerong Li (Zhejiang university), Zhengrong Xue (Tsinghua University; Shanghai Qi Zhi Institute; Shanghai AI Lab), and Huazhe Xu (Tsinghua University; Shanghai Qi Zhi Institute; Shanghai AI Lab)

Embodied Human Activity Recognition 6433
Sha Hu (Simon Fraser University, Canada), Yu Gong (Simon Fraser University, Canada), and Greg Mori (Simon Fraser University, Canada)

Semantic-Aware Video Representation for Few-Shot Action Recognition 6444
Yutao Tang (Johns Hopkins University), Benjamin Bejar (Swiss Data Science Center), and Rene Vidal (University of Pennsylvania)

Leveraging the Power of Data Augmentation for Transformer-Based Tracking 6455
Jie Zhao (Dalian University of Technology), Johan Edstedt (Linköping University), Michael Felsberg (Linköping University), Dong Wang (Dalian University of Technology), and Huchuan Lu (Dalian University of Technology)

CAMOT: Camera Angle-Aware Multi-Object Tracking 6465
Felix Limanta (Tokyo Institute of Technology, Japan), Kuniaki Uto (Tokyo Instituite of Technology, Japan), and Koichi Shinoda (Tokyo Institute of Technology, Japan)

Detection Defenses: An Empty Promise Against Adversarial Patch Attacks on Optical Flow 6475
Erik Scheurer (University of Stuttgart), Jenny Schmalfuss (University of Stuttgart), Alexander Lis (University of Stuttgart), and Andrés Bruhn (University of Stuttgart)

Repetitive Action Counting with Motion Feature Learning 6485
Xinjie Li (Pennsylvania State University, USA) and Huijuan Xu (Pennsylvania State University, USA)

United We Stand, Divided We Fall: UnityGraph for Unsupervised Procedure Learning From Videos 6495
Siddhant Bansal (CVIT, IIIT, Hyderabad), Chetan Arora (IIT, Delhi), and C.V. Jawahar (CVIT, IIIT, Hyderabad)

Sequential Transformer for End-to-End Video Text Detection 6506
Junbo Zhang (Institute of Automation of Chinese Academy of Sciences), Mengbiao Zhao (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)

Context in Human Action Through Motion Complementarity 6517
Eadom Dessalene (University of Maryland, College Park), Michael Maynord (University of Maryland College Park), Cornelia Fermuller (University of Maryland, College Park), and Yiannis Aloimonos (University of Maryland, College Park)

Egocentric Action Recognition by Capturing Hand-Object Contact and Object State 6527
Tsukasa Shiota (NTT Corporation), Motohiro Takagi (NTT Corporation), Kaori Kumagai (NTT Corporation), Hitoshi Seshimo (NTT Corporation), and Yushi Aono (NTT Corporation)

MIDAS: Mixing Ambiguous Data with Soft Labels for Dynamic Facial Expression Recognition 6538
Ryosuke Kawamura (Fujitsu Research of America, Inc.), Hideaki Hayashi (Osaka University), Noriko Takemura (Kyushu Institute of Technology), and Hajime Nagahara (Osaka University)

FRoG-MOT: Fast and Robust Generic Multiple-Object Tracking by IoU and Motion-State Associations 6549
Takuya Ogawa (NEC Corporation), Takashi Shibata (NEC Corporation), and Toshinori Hosoi (NEC Corporation)

Density-Based Flow Mask Integration via Deformable Convolution for Video People Flux Estimation 6559
Chang-Lin Wan (National Yang Ming Chiao Tung University), Feng-Kai Huang (National Yang Ming Chiao Tung University), and Hong-Han Shuai (National Yang Ming Chiao Tung University)

ConfTrack: Kalman Filter-Based Multi-Person Tracking by Utilizing Confidence Score of Detection Box 6569
Hyeonchul Jung (HL Klemove), Seokjun Kang (HL Klemove), Takgen Kim (HL Klemove), and Hyeongki Kim (HL Klemove)

CGAPoseNet+GCAN: A Geometric Clifford Algebra Network for Geometry-Aware Camera Pose Regression 6579
Alberto Pepe (University of Cambridge), Sven Buchholz (TH Brandenburg), and Joan Lasenby (University of Cambridge)

Embedding Task Structure for Action Detection 6590
Michael Peven (The Johns Hopkins University) and Gregory D. Hager (The Johns Hopkins University)

RandomWalks for Temporal Action Segmentation with Timestamp Supervision 6600
Roy Hirsch (Veriy AI, Israel), Regev Cohen (Veriy AI, Israel), Tomer Golany (Veriy AI, Israel), Daniel Freedman (Veriy AI, Israel), and Ehud Rivlin (Veriy AI, Israel)

MITFAS: Mutual Information Based Temporal Feature Alignment and Sampling for Aerial Video Action Recognition 6611
Ruiqi Xian (University of Maryland - College Park), Xijun Wang (University of Maryland - College Park), and Dinesh Manocha (University of Maryland - College Park)

Do VSR Models Generalize Beyond LRS3? 6621
Yasser Abdelaziz Dahou Djilali (Technology Innovation Institute, UAE; Dublin City University, Ireland), Sanath Narayan (Technology Innovation Institute, UAE), Eustache Le Bihan (Technology Innovation Institute, UAE), Haithem Boussaid (Technology Innovation Institute, UAE), Ebtesam Almazrouei (Technology Innovation Institute, UAE), and Merouane Debbah (Technology Innovation Institute, UAE)

PGVT: Pose-Guided Video Transformer for Fine-Grained Action Recognition 6631
Haosong Zhang (Nanyang Technological University), Mei Chee Leong (I2R, ASTAR), Liyuan Li (Institute for Infocomm Research), and Weisi Lin (Nanyang Technological University, Singapore)

Differentially Private Video Activity Recognition 6643
Zelun Luo (Stanford University), Yuliang Zou (Virginia Tech), Yijin Yang (Arizona State University), Zane Durante (Stanford University), De-An Huang (NVIDIA), Zhiding Yu (NVIDIA), Chaowei Xiao (NVIDIA; University of Wisconsin–Madison), Li Fei-Fei (Stanford University), and Animashree Anandkumar (NVIDIA; Caltech)

Video Instance Matting 6654
Jiachen Li (SHI Labs @ Georgia Tech & Oregon & UIUC), Roberto Henschel (Picsart AI Research (PAIR)), Vidit Goel (Picsart AI Research (PAIR)), Marianna Ohanyan (Picsart AI Research (PAIR)), Shant Navasardyan (Picsart AI Research (PAIR)), and Humphrey Shi (SHI Labs @ Georgia Tech & Oregon & UIUC; Picsart AI Research (PAIR))

VMFormer: End-to-End Video Matting With Transformer 6664
Jiachen Li (UIUC), Vidit Goel (Indian Institute of Technology, Kharagpur), Marianna Ohanyan (Picsart Inc.), Shant Navasardyan (Picsart Inc.), Yunchao Wei (UTS), and Humphrey Shi (Georgia Tech: Oregon;UIUC;PAIR)

DDAM-PS: Diligent Domain Adaptive Mixer for Person Search 6674
Mohammed K Almansoori (MBZUAI), Mustansar Fiaz (MBZUAI), and Hisham Cholakkal (MBZUAI)

Think Before You Simulate: Symbolic Reasoning to Orchestrate Neural Computation for Counterfactual Question Answering 6684
Adam Ishay (Arizona State University, USA), Zhun Yang (Arizona State University, USA), Joohyung Lee (Arizona State university, USA; Samsung Research, S. Korea), Ilgu Kang (Samsung Research, S. Korea), and Dongjae Lim (Samsung Research, S. Korea)

Separable Self and Mixed Attention Transformers for Efficient Object Tracking 6694
Goutam Yelluru Gopal (Concordia University) and Maria Amer (Concordia University)

Restoring Degraded Old Films with Recursive Recurrent Transformer Networks 6704
Shan Lin (Waseda University) and Edgar Simo-Serra (Waseda University)

Holistic Representation Learning for Multitask Trajectory Anomaly Detection 6715
Alexandros Stergiou (Vrije Universiteit Brussel, Belgium & imec, Belgium), Brent De Weerdt (Vrije Universiteit Brussel, Belgium & imec, Belgium), and Nikos Deligiannis (Vrije Universiteit Brussel, Belgium & imec, Belgium)

Interaction Region Visual Transformer for Egocentric Action Anticipation 6726
Debaditya Roy (Institute of High-Performance Computing, Agency for Science, Technology and Research, Singapore), Ramanathan Rajendiran (Institute of High-Performance Computing, Agency for Science, Technology and Research, Singapore), and Basura Fernando (Institute of High-Performance Computing, Agency for Science, Technology and Research; Centre for Frontier AI Research, Agency for Science, Technology and Research, Singapore)

Object-Centric Video Representation for Long-Term Action Anticipation 6737
Ce Zhang (UNC Chapel Hill), Changcheng Fu (University of Southern California), Shijie Wang (Providence), Nakul Agarwal (Honda Research Institute USA), Kwonjoon Lee (Honda Research Institute USA), Chiho Choi (Samsung Semiconductor US), and Chen Sun (Brown University)

A Hybrid Graph Network for Complex Activity Detection in Video 6748
Salman Khan (Oxford Brookes University), Izzeddin Teeti (Oxford Brookes University), Andrew Bradley (Oxford Brookes University), Mohamed Elhoseiny (KAUST), and Fabio Cuzzolin (Oxford Brookes University)

SSVOD: Semi-Supervised Video Object Detection with Sparse Annotations 6759
Tanvir Mahmud (University of Texas at Austin), Chun-Hao Liu (Amazon Prime Video), Burhaneddin Yaman (Bosch Research North America), and Diana Marculescu (University of Texas at Austin)

Semantic Fusion Augmentation and Semantic Boundary Detection: A Novel Approach to Multi-Target Video Moment Retrieval 6769
Cheng Huang (National Yang Ming Chiao Tung University, Taiwan), Yi-Lun Wu (National Yang Ming Chiao Tung University, Taiwan), Hong-Han Shuai (National Yang Ming Chiao Tung University, Taiwan), and Ching-Chun Huang (National Yang Ming Chiao Tung University, Taiwan)

A Coarse-To-Fine Pseudo-Labeling (C2FPL) Framework for Unsupervised Video Anomaly Detection 6779
Anas Allahham (Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), UAE), Nurbek Tastan (Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), UAE), Muhammad Zaigham Zaheer (Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), UAE), and Karthik Nandakumar (Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), UAE)

PromptonomyViT: Multi-Task Prompt Learning Improves Video Transformers Using Synthetic Scene Data 6789
Roei Herzig (Tel-Aviv University; UC Berkeley; IBM Research), Ofir Abramovich (Reichman University), Elad Ben Avraham (Tel-Aviv University), Assaf Arbelle (IBM Research AI), Leonid Karlinsky (MIT-IBM Watson AI Lab), Ariel Shamir (Reichman University), Trevor Darrell (UC Berkeley), and Amir Globerson (Tel-Aviv University)

GLAD: Global-Local View Alignment and Background Debiasing for Unsupervised Video Domain Adaptation With Large Domain Gap 6802
Hyogun Lee (Kyung Hee University), Kyungho Bae (Kyunghee University), Seong Jong Ha (AI Center, CJ Corporation), Yumin Ko (NCSOFT), Gyeong-Moon Park (Kyung Hee University), and Jinwoo Choi (Kyung Hee University)

Beyond SOT: Tracking Multiple Generic Objects at Once 6812
Christoph Mayer (ETH Zürich), Martin Danelljan (ETH Zürich), Ming-Hsuan Yang (Google Research), Vittorio Ferrari (Google Research), Luc Van Gool (ETH Zürich), and Alina Kuznetsova (Google Research)

MFT: Long-Term Tracking of Every Pixel 6823
Michal Neoral (Czech Technical University in Prague), Jonáš Šerých (Czech Technical University in Prague), and JiřÍ Matas (Czech Technical University in Prague)

Real-Time Weakly Supervised Video Anomaly Detection 6834
Hamza Karim (University of south Florida), Keval Doshi (University of South Florida), and Yasin Yilmaz (University of South Florida)

Single-Image Deblurring, Trajectory and Shape Recovery of Fast Moving Objects with Denoising Diffusion Probabilistic Models 6843
Radim Spetlik (Czech Technical University in Prague), Denys Rozumnyi (Czech Technical University in Prague; ETH Zurich), and JiřÍ Matas (Czech Technical University in Prague)

Contrastive Learning for Multi-Object Tracking with Transformers 6853
Pierre-François De Plaen (ESAT-PSI, KU Leuven, Belgium), Nicola Marinello (ESAT-PSI, KU Leuven, Belgium), Marc Proesmans (ESAT-PSI, KU Leuven, Belgium; TRACE vzw), Tinne Tuytelaars (ESAT-PSI, KU Leuven, Belgium), and Luc Van Gool (ESAT-PSI, KU Leuven, Belgium;CVL, ETH Zürich, Switzerland; TRACE vzw)

Limited Data, Unlimited Potential: A Study on ViTs Augmented by Masked Autoencoders 6864
Srijan Das (UNC Charlotte), Tanmay Jain (Delhi Technological University), Dominick Reilly (UNC Charlotte), Pranav Balaji (BITS Pilani Hyderabad), Soumyajit Karmakar (Indian Institute of Information Technology Guwahati), Shyam Marjit (Indian Institute of Information Technology Guwahati), Xiang Li (Stony Brook University), Abhijit Das (BITS Pilani Hyderabad), and Michael S. Ryoo (Stony Brook University)

JOADAA: Joint Online Action Detection and Action Anticipation 6875
Mohammed Guermal (INRIA), Abid Ali (INRIA, Sophia Antipolis), Rui Dai (INRIA), and Francois Bremond (Inria Sophia Antipolis, France)

CCMR: High Resolution Optical Flow Estimation via Coarse-To-Fine Context-Guided Motion Reasoning 6885
Azin Jahedi (VIS, University of Stuttgart), Maximilian Luz (RLL, University of Freiburg), Marc Rivinius (SEC, University of Stuttgart), and Andrés Bruhn (VIS, University of Stuttgart)

Weakly-Supervised Representation Learning for Video Alignment and Analysis 6895
Guy Bar-Shalom (Verily), George Leifman (Verily), and Michael Elad (Verily)

MotionAGFormer: Enhancing 3D Human Pose Estimation With a Transformer-GCNFormer Network 6905
Soroush Mehraban (University of Toronto), Vida Adeli (University of Toronto), and Babak Taati (University Health Network)

Leveraging Synthetic Data to Learn Video Stabilization Under Adverse Conditions 6916
Abdulrahman Kerim (Lancaster University; University for the Creative Arts, UK), Washington L. S. Ramos (UFMG, Brazil), Leandro Soriano Marcolino (Lancaster University, UK), Erickson R. Nascimento (UFMG, Brazil; Microsoft), and Richard Jiang (Lancaster University, UK)

What's in the Flow? Exploiting Temporal Motion Cues for Unsupervised Generic Event Boundary Detection 6926
Sourabh Vasant Gothe (Samsung R&D Institute Bangalore, India), Vibhav Agarwal (Samsung R&D Institute India, India), Sourav Ghosh (Samsung R&D Institute Bangalore, India), Jayesh Rajkumar Vachhani (Samsung R&D Institute Bengaluru, India), Pranay Kashyap (Samsung Research Institute Bangalore, India), and Barath Raj Kandur Raja (Samsung R&D Institute Bangalore, India)

Learning the What and How of Annotation in Video Object Segmentation 6936
Thanos Delatolas (Technical University of Denmark; Pioneer Center for AI), Vicky Kalogeiton (LIX, Ecole Polytechnique, CNRS, Institut Polytechnique de Paris), and Dim P. Papadopoulos (Technical University of Denmark; Pioneer Center for AI)

Lightweight Delivery Detection on Doorbell Cameras 6947
Pirazh Khorramshahi (Qualcomm Technologies), Zhe Wu (Comcast Research Lab), Tianchen Wang (Comcast, Inc), Luke Deluccia (Comcast), and Hongcheng Wang (Amazon)

Spatio-Temporal Filter Analysis Improves 3D-CNN for Action Classification 6957
Takumi Kobayashi (National Institute of Advanced Industrial Science and Technology, Japan) and Jiaxing Ye (National Institute of Advanced Industrial Science and Technology, Japan)

PMI Sampler: Patch Similarity Guided Frame Selection For Aerial Action Recognition 6967
Ruiqi Xian (University of Maryland - College Park), Xijun Wang (University of Maryland - College Park), Divya Kothandaraman (University of Maryland - College Park), and Dinesh Manocha (University of Maryland - College Park)

Optimizing Long-Term Robot Tracking With Multi-Platform Sensor Fusion 6977
Giuliano Albanese (ETH Zürich), Arka Mitra (ETH Zurich), Jan-Nico Zaech (ETH Zürich), Yupeng Zhao (ETH Zurich), Ajad Chhatkuli (ETH Zurich), and Luc Van Gool (ETH Zurich)

Learnable Cube-Based Video Encryption for Privacy-Preserving Action Recognition 6988
Yuchi Ishikawa (LY Corporation, Japan), Masayoshi Kondo (LY Corporation, Japan), and Hirokatsu Kataoka (LINE Corporation, Japan)

A*: Atrous Spatial Temporal Action Recognition for Real Time Applications 6999
Myeongjun Kim (Deeping Source Inc., Republic of Korea), Federica Spinola (Deeping Source Inc., Republic of Korea), Philipp Benz (Deeping Source Inc., Republic of Korea), and Tae-Hoon Kim (Deeping Source Inc., Republic of Korea)

SEMA: Semantic Attention for Capturing Long-Range Dependencies in Egocentric Lifelogs 7010
Pravin Nagar (University of Maryland), K.N Ajay Shastry (Indian Institute of Technology, Delhi), Jayesh Chaudhari (Indian Institute of Technology, Delhi), and Chetan Arora (Indian Institute of Technology, Delhi)

Triplet Attention Transformer for Spatiotemporal Predictive Learning 7021
Xuesong Nie (Zhejiang University), Xi Chen (Hong Kong University), Haoyuan Jin (Zhejiang University), Zhihang Zhu (Zhejiang University), Yunfeng Yan (Zhejiang University), and Donglian Qi (Zhejiang University)

ZEETAD: Adapting Pretrained Vision-Language Model for Zero-Shot End-to-End Temporal Action Detection 7031
Thinh Phan (University of Arkansas), Khoa Hv Vo (University of Arkansas), Anh Duy Le Dinh (FPT Software), Gianfranco Doretto (West Virginia University), Donald Adjeroh (West Virginia University), and Ngan Le (University of Arkansas)

Agriculture

Computer Vision on the Edge: Individual Cattle Identification in Real-Time With ReadMyCow System 7041
Moniek Smink (University of Wisconsin - Madison), Haotian Liu (University of Wisconsin-Madison), Doerte Doepfer (UW-Madison), and Yong Jae Lee (University of Wisconsin-Madison)

PlantPlotGAN: A Physics-Informed Generative Adversarial Network for Plant Disease Prediction 7051
Felipe A. Lopes (Taylor Geospatial Institute; Federal Institute of Alagoas, USA), Vasit Sagan (Taylor Geospatial Institute, USA), and Flavio Esposito (Saint Louis University, USA)

S3AD: Semi-Supervised Small Apple Detection in Orchard Environments 7061
Robert Johanson (University of Hamburg, Germany), Christian Wilms (University of Hamburg, Germany), Ole Johannsen (University of Hamburg, Germany), and Simone Frintrop (University of Hamburg, Germany)

Towards Accurate Disease Segmentation in Plant Images: A Comprehensive Dataset Creation and Network Evaluation 7071
Komuravelli Prashanth (Indian Institute of Technology, Tirupati), Sri Harsha Jaladi (Indian Institute of Technology, Tirupati), Arun Kumar Sivapuram (Indian Institute of Technology, Tirupati), and Srilekha Jaladi (Professor Jayashankar Telangana State Agricultural University, India)

Automated Monitoring of Ear Biting in Pigs by Tracking Individuals and Events 7080
Anicetus Odo (Queen's University Belfast), Niall Mclaughlin (Queen's University Belfast), and Ilias Kyriazakis (Queen's University)

The Growing Strawberries Dataset: Tracking Multiple Objects with Biological Development Over an Extended Period 7089
Junhan Wen (Delft University of Technology), Camiel R. Verschoor (Birds.ai B.V.), Chengming Feng (Delft University of Technology), Irina Mone Epure (Birds.ai B.V.), Thomas Abeel (Delft University of Technology), and Mathijs de Weerdt (Delft University of Technology)

Animals / Insects

Deep Visual-Genetic Biometrics for Taxonomic Classification of Rare Species 7100
Tayfun Karaderi (University of Bristol), Tilo Burghardt (University of Bristol), Raphael Morard (University of Bremen), and Daniela Schmidt (University of Bristol)

Tracking Tiny Insects in Cluttered Natural Environments Using Refinable Recurrent Neural Networks 7111
Lars Haalck (University of Münster, Germany), Sebastian Thiele (University of Münster, Germany), and Benjamin Risse (University of Münster, Germany)

BirdSAT: Cross-View Contrastive Masked Autoencoders for Bird Species Classification and Mapping 7121
Srikumar Sastry (Washington University in St. Louis), Subash Khanal (Washington University in St. Louis), Aayush Dhakal (Washington University in St Louis), Di Huang (Washington University in St. Louis), and Nathan Jacobs (Washington University in St. Louis)

SeaTurtleID2022: A Long-Span Dataset for Reliable Sea Turtle Re-Identification 7131
Lukáš Adam (Czech Technical University), Vojtěch Čermák (Czech Technical University), Kostas Papafitsoros (Queen Mary University of London), and Lukas Picek (UWB and INRIA)

Towards a Dynamic Vision Sensor-Based Insect Camera Trap 7142
Eike Gebauer (University of Münster), Sebastian Thiele (University of Münster), Pierre A Ouvrard (SLU), Adrien Sicard (Swedish University of Agricultural Sciences), and Benjamin Risse (University of Münster)

FishTrack23: An Ensemble Underwater Dataset for Multi-Object Tracking 7152
Matthew Dawkins (Kitware Inc.), Jack Prior (NOAA NMFS, USA), Bryon Lewis (Kitware Inc.), Robin Faillettaz (DECOD, Ifremer, INRAE, L'Institut Agro, France), Thompson Banez (California Department of Fish and Wildlife, USA), Mary Salvi (Kitware Inc.), Audrey Rollo (NOAA NMFS, USA), Julien Simon (DECOD, Ifremer, INRAE, L'Institut Agro, France), Matthew Campbell (NOAA NMFS, USA), Matthew Lucero (California Department of Fish and Wildlife, USA), Aashish Chaudhary (Kitware Inc.), Benjamin Richards (NOAA FisheriesNOAA NMFS, USA), and Anthony Hoogs (Kitware Inc.)

Arts / Games / Social Media

Let the Beat Follow You – Creating Interactive Drum Sounds From Body Rhythm 7162
Xiulong Liu (University of Washington), Kun Su (University of Washington), and Eli Shlizerman (University of Washington)

Interactive Segmentation for Diverse Gesture Types Without Context 7173
Josh Myers-Dean (University of Colorado Boulder), Yifei Fan (Adobe Research), Brian Price (Adobe Research), Wilson Chan (Adobe Research), and Danna Gurari (University of Colorado Boulder; University of Texas at Austin)

SemST: Semantically Consistent Multi-Scale Image Translation via Structure-Texture Alignment 7184
Ganning Zhao (University of Southern California, USA), Wenhui Cui (University of Southern California, USA), Suya You (DEVCOM Army Research Laboratory, USA), and C.-C. Jay Kuo (University of Southern California, USA)

Visual Narratives: Large-Scale Hierarchical Classification of Art-Historical Images 7195
Matthias Springstein (TIB – Leibniz Information Centre for Science and Technology; Leibniz University Hannover, Germany), Stefanie Schneider (Ludwig Maximilian University of Munich, Germany), Javad Rahnama (Reply GmbH, Germany), Julian Stalter (Ludwig Maximilian University of Munich, Germany), Maximilian Kristen (Ludwig Maximilian University of Munich, Germany), Eric Müller-Budack (TIB - Leibniz Information Centre for Science and Technology; Leibniz University Hannover, Germany), and Ralph Ewerth (TIB - Leibniz Information Center for Science and Technology; Leibniz University Hannover, Germany)

Composite Diffusion: Whole >= ?parts 7206
Vikram Jamwal (TCS Research, India) and Ramaneswaran S (NVIDIA, India)

The Paleographer's Eye ex machina: Using Computer Vision To Assist Humanists in Scribal Hand Identification 7216
Samuel Grieggs (Indiana University of Pennsylvania), C. E. M. Henderson (University of Toronto), Sebastian Sobecki (University of Toronto), Alexandra Gillespie (University of Toronto), and Walter Scheirer (University of Notre Dame)

C-CLIP: Contrastive Image-Text Encoders To Close the Descriptive-Commentative Gap 7226
William Theisen (University of Notre Dame) and Walter J Scheirer (University of Notre Dame)

CAD – Contextual Multi-Modal Alignment for Dynamic AVQA 7236
Asmar Nadeem (University of Surrey, United Kingdom), Adrian Hilton (University of Surrey), Robert Dawes (BBC Research and Development, United Kingdom), Graham Thomas (BBC Research and Development, United Kingdom), and Armin Mustafa (University of Surrey, United Kingdom)

ArcAid: Analysis of Archaeological Artifacts Using Drawings 7249
Offry Hayon (Technion), Stefan Münger (U. of Bern), Ilan Shimshoni (U. of Haifa), and Ayellet Tal (Technion)

Movie Genre Classification by Language Augmentation and Shot Sampling 7260
Zhongping Zhang (Boston University), Yiwen Gu (Boston University), Bryan A. Plummer (Boston University), Xin Miao (Kuaishou Technology), Jiayi Liu (Kuaishou Technology), and Huayan Wang (Kuaishou Technology)

Few-Shot Event Classification in Images Using Knowledge Graphs for Prompting 7271
Golsa Tahmasebzadeh (TIB - Leibniz Information Center for Science and Technology; Leibniz University Hannover, Germany), Matthias Springstein (TIB - Leibniz Information Center for Science and Technology; Leibniz University Hannover, Germany), Ralph Ewerth (TIB - Leibniz Information Center for Science and Technology; Leibniz University Hannover, Germany), and Eric Müller-Budack (TIB - Leibniz Information Center for Science and Technology; Leibniz University Hannover, Germany)

Towards Diverse and Consistent Typography Generation 7281
Wataru Shimoda (CyberAgent, Japan), Daichi Haraguchi (Kyushu University, Japan), Seiichi Uchida (Kyushu University, Japan), and Kota Yamaguchi (CyberAgent, Japan)

NeRFEditor: Differentiable Style Decomposition for 3D Scene Editing 7291
Chunyi Sun (Australian National University), Yanbin Liu (The University of Western Australia), Junlin Han (CSIRO), and Stephen Gould (Australian National University, Australia)

FastCLIPstyler: Optimisation-Free Text-Based Image Style Transfer Using Style Representations 7301
Ananda Padhmanabhan Suresh (Sertis Vision Lab), Sanjana Jain (Sertis Vision Lab), Pavit Noinongyao (Sertis Vision Lab), Ankush Ganguly (Sertis Vision Lab), Ukrit Watchareeruetai (Sertis Vision Lab), and Aubin Samacoits (Sertis Vision Lab)

ArtQuest: Countering Hidden Language Biases in ArtVQA 7311
Sedigheh Eslami (Hasso Plattner Institute), Tibor Bleidt (Hasso Plattner Institut), and Gerard De Melo (HPI)

Autonomous Driving

2D Feature Distillation forWeakly- and Semi-Supervised 3D Semantic Segmentation 7321
Ozan Unal (ETH Zurich), Dengxin Dai (Huawei Technologies), Lukas Hoyer (ETH Zurich), Yigit Baran Can (ETH Zurich), and Luc Van Gool (ETH Zurich; KU Leuven; INSAIT)

Driving Through the Concept Gridlock: Unraveling Explainability Bottlenecks in Automated Driving 7331
Jessica M Echterhoff (University of California, San Diego), An Yan (UC San Diego), Kyungtae Han (Toyota InfoTech Labs), Amr Abdelraouf (Toyota InfoTech Labs), Rohit Gupta (Toyota InfoTech Labs), and Julian Mcauley (UCSD)

StreamMapNet: Streaming Mapping Network for Vectorized Online HD Map Construction 7341
Tianyuan Yuan (Tsinghua University), Yicheng Liu (Tsinghua University), Yue Wang (University of Southern California), Yilun Wang (Tsinghua University), and Hang Zhao (Tsinghua University)

RGB-X Object Detection via Scene-Specific Fusion Modules 7351
Sri Aditya Deevi (California Institute of Technology), Connor T Lee (California Institute of Technology), Lu Gan (California Institute of Technology), Sushruth Nagesh (Ford Motor Company), Gaurav Pandey (Ford Motor Company), and Soon-Jo Chung (Caltech)

NVAutoNet: Fast and Accurate 360° 3D Visual Perception for Self Driving 7361
Trung Pham (Nvidia), Mehran Maghoumi (Nvidia), Wanli Jiang (Nvidia), Bala Siva Sashank Jujjavarapu (Nvidia), Mehdi Sajjadi (Nvidia), Xin Liu (Nvidia), Hsuan-Chu Lin (Nvidia), Bor-Jeng Chen (Nvidia), Giang Truong (Nvidia), Chao Fang (Nvidia), Junghyun Kwon (Nvidia), and Minwoo Park (Nvidia)

MagneticPillars: Efficient Point Cloud Registration Through Hierarchized Birds-Eye-View Cell Correspondence Refinement 7371
Kai Fischer (Valeo Schalter und Sensoren GmbH; Ilmenau University of Technology), Martin Simon (Valeo Schalter und Sensoren GmbH), Stefan Milz (Spleenlab GmbH; Ilmenau University of Technology), and Patrick Mäeder (Ilmenau University of Technology)

Have We Ever Encountered This Before? Retrieving Out-of-Distribution Road Obstacles From Driving Scenes 7381
Youssef Shoeb (Continental AG), Robin Chan (Bielefeld University), Gesina Schwalbe (Continental Automotive GmbH), Azarm Nowzad (Continental AG), Fatma Guney (Koc University), and Hanno Gottschalk (University of Wuppertal)

OVeNet: Offset Vector Network for Semantic Segmentation 7392
Stamatis Alexandropoulos (Princeton University), Christos Sakaridis (ETH Zürich), and Petros Maragos (National Technical University of Athens)

BEVMap: Map-Aware BEV Modeling for 3D Perception 7404
Mincheol Chang (Korea University), Seokha Moon (Korea University), Reza Mahjourian (University of Texas at Austin), and Jinkyu Kim (Korea University)

Evaluation of Video Masked Autoencoders' Performance and Uncertainty Estimations for Driver Action and Intention Recognition 7414
Koen Vellenga (VCC / HIS), H. Joe Steinhauer (HIS), Göran Falkman (HIS), and Tomas Björklund (VCC)

LidarCLIP or: How I Learned to Talk to Point Clouds 7423
Georg Hess (Zenseact; Chalmers University of Technology), Adam Tonderski (Zenseact; Lund University), Christoffer Petersson (Zenseact; Chalmers University of Technology), Karl Åström (Lund University), and Lennart Svensson (Chalmers University of Technology)

Self-Supervised Learning for Place Representation Generalization Across Appearance Changes... 7433
Mohamed Adel Musallam (SnT, University of Luxembourg), Vincent Gaudillière (SnT, University of Luxembourg), and Djamila Aouada (SnT, University of Luxembourg)

PAIR: Perception Aided Image Restoration for Natural Driving Conditions 7444
Pranjay Shyam (Faurecia IRYStec) and Hyunjin Yoo (Faurecia IRYStec)

Lightweight Thermal Super-Resolution and Object Detection for Robust Perception in Adverse Weather Conditions 7456
Pranjay Shyam (Faurecia IRYStec Inc., Canada) and Hyunjin Yoo (Faurecia IRYStec Inc., Canada)

Concurrent Band Selection and Traversability Estimation From Long-Wave Hyperspectral Imagery in Off-Road Settings 7468
Florence Yellin (Kitware), Scott Mccloskey (Kitware), Cole Hill (University of South Florida), Eric Smith (Kitware Inc.), and Brian S Clipp (Kitware)

RS2G: Data-Driven Scene-Graph Extraction and Embedding for Robust Autonomous Perception and Scenario Understanding 7478
Junyao Wang (University of California, Irvine, United States), Arnav Vaibhav Malawade (University of California, Irvine, United States), Junhong Zhou (University of California, Irvine, United States), Shih-Yuan Yu (University of California, Irvine, United States), and Mohammad Abdullah Al Faruque (University of California, Irvine, United States)

Re-VoxelDet: Rethinking Neck and Head Architectures for High-Performance Voxel-Based 3D Detection 7488
Jae-Keun Lee (FutureDrive, Kyungpook National University (KNU)), Jin-Hee Lee (DGIST), Joohyun Lee (DGIST), Soon Kwon (DGIST), and Heechul Jung (Kyungpook National University)

Rank2Tell: A Multimodal Driving Dataset for Joint Importance Ranking and Reasoning 7498
Enna Sachdeva (Honda Research Institute, USA), Nakul Agarwal (Honda Research Institute, USA), Suhas Chundi (Stanford University), Sean Roelofs (Stanford University), Jiachen Li (Stanford University), Mykel Kochenderfer (Stanford University), Chiho Choi (Honda Research Institute, USA), and Behzad Dariush (Honda Research Institute, USA)

Robust Object Detection in Challenging Weather Conditions 7508
Himanshu Gupta (Örebro Universitet), Oleksandr Kotlyar (Örebro University), Henrik Andreasson (AASS, Örebro University), and Achim J Lilienthal (TU Munich)

Graph(Graph): A Nested Graph-Based Framework for Early Accident Anticipation 7518
Nupur Thakur (Arizona State University), Prasanthsai Gouripeddi (ASU), and Baoxin Li (Arizona State University)

Alleviating Foreground Sparsity for Semi-Supervised Monocular 3D Object Detection 7527
Weijia Zhang (University of Sydney), Dongnan Liu (University of Sydney), Chao Ma (Shanghai Jiao Tong University), and Weidong Cai (University of Sydney)

Biomedical / Healthcare / Medicine

Learning Robust Deep Visual Representations From EEG Brain Recordings 7538
Prajwal Singh (CVIG Lab, IIT Gandhinagar, India), Dwip Dalal (CVIG Lab, IIT Gandhinagar, India), Gautam Vashishtha (CVIG Lab, IIT Gandhinagar, India), Krishna Miyapuram (BRAIN Lab, IIT Gandhinagar, India), and Shanmuganathan Raman (CVIG Lab, IIT Gandhinagar, India)

Continual Atlas-Based Segmentation of Prostate MRI 7548
Amin Ranem (Technical University of Darmstadt), Camila Gonzalez (Stanford University), Daniel Pinto Dos Santos (University of Cologne), Andreas M. Bucher (University of Frankfurt), Ahmed E. Othman (University Medical Center Mainz), and Anirban Mukhopadhyay (Technical University of Darmstadt)

Brainomaly: Unsupervised Neurologic Disease Detection Utilizing Unannotated T1-Weighted Brain MR Images 7558
Md Mahfuzur Rahman Siddiquee (Arizona State University; ASU-Mayo Center for Innovative Imaging), Jay Shah (Arizona State University; ASU-Mayo Center for Innovative Imaging), Teresa Wu (Arizona State University; ASU-Mayo Center for Innovative Imaging), Catherine Chong (ASU-Mayo Center for Innovative Imaging; Mayo Clinic), Todd J. Schwedt (ASU-Mayo Center for Innovative Imaging; Mayo Clinic), Gina Dumkrieger (Mayo Clinic), Simona Nikolova (Mayo Clinic), and Baoxin Li (Arizona State University; ASU-Mayo Center for Innovative Imaging)

PHG-Net: Persistent Homology Guided Medical Image Classification 7568
Yaopeng Peng (University of Notre Dame), Hongxiao Wang (University of Notre Dame), Milan Sonka (U Iowa, USA), and Danny Z Chen (University of Notre Dame)

AnyStar: Domain Randomized Universal Star-Convex 3D Instance Segmentation 7578
Neel Dey (MIT CSAIL), S. Mazdak Abulnaga (MIT CSAIL), Benjamin Billot (MIT CSAIL), Esra Abaci Turk (Boston Children's Hospital, Harvard Medical School), P. Ellen Grant (Boston Children's Hospital, Harvard Medical School), Adrian V. Dalca (MIT CSAIL; Martinos Center, Massachusetts General Hospital), and Polina Golland (MIT CSAIL)

Adaptive Latent Diffusion Model for 3D Medical Image to Image Translation: Multi-Modal Magnetic Resonance Imaging Study 7589
Jonghun Kim (Sungkyunkwan University, Korea) and Hyunjin Park (Sungkyunkwan University, Korea)

Slice and Conquer: A Planar-to-3D Framework for Efficient Interactive Segmentation of Volumetric Images 7599
Wonwoo Cho (Korea Advanced Institute of Science and Technology), Dongmin Choi (Korea Advanced Institute of Science and Technology), Hyesu Lim (Korea Advanced Institute of Science and Technology), Jinho Choi (KAIST), Saemee Choi (KAIST), Hyun-Seok Min (Tomocube), Sungbin Lim (Korea University), and Jaegul Choo (Korea Advanced Institute of Science and Technology)

Attention-Guided Prototype Mixing: Diversifying Minority Context on Imbalanced Whole Slide Images Classification Learning 7609
Farchan Hakim Raswa (National Central University, Taiwan), Chun-Shien Lu (IIS, Academia Sinica, Taiwan), and Jia-Ching Wang (IIS, Academia Sinica, Taiwan)

Are Natural Domain Foundation Models Useful for Medical Image Classification? 7619
Joana Palés Huix (KTH Royal Institute of Technology; Science for Life Laboratory; AstraZeneca, Sweden), Adithya Raju Ganeshan (KTH Royal Institute of Technology; Science for Life Laboratory, Sweden), Johan Fredin Haslum (KTH Royal Institute of Technology; Science for Life Laboratory; AstraZeneca, Sweden), Magnus Söderberg (AstraZeneca, Sweden), Christos Matsoukas (KTH Royal Institute of Technology; Science for Life Laboratory; AstraZeneca, Sweden), and Kevin Smith (KTH Royal Institute of Technology; Science for Life Laboratory, Sweden)

Unsupervised Exemplar-Based Image-to-Image Translation and Cascaded Vision Transformers for Tagged and Untagged Cardiac Cine MRI Registration 7629
Meng Ye (Rutgers University), Mikael Kanski (New York University School of Medicine), Dong Yang (NVIDIA), Leon Axel (New York University School of Medicine), and Dimitris Metaxas (Rutgers University)

IR-FRestormer: Iterative Refinement With Fourier-Based Restormer for Accelerated MRI Reconstruction 7640
Mohammad Zalbagi Darestani (Rice University), Vishwesh Nath (NVIDIA), Wenqi Li (NVIDIA), Yufan He (NVidia Corporation), Holger R Roth (NVIDIA), Ziyue Xu (NVIDIA), Daguang Xu (NVIDIA Corporation), Reinhard Heckel (TUM), and Can Zhao (Nvidia)

Constrained Probabilistic Mask Learning for Task-Specific Undersampled MRI Reconstruction ... 7650
Tobias Weber (LMU Munich; Munich Center for Machine Learning (MCML)), Michael Ingrisch (LMU Munich; Munich Center for Machine Learning (MCML)), Bernd Bischl (LMU Munich; Munich Center for Machine Learning (MCML)), and David Rügamer (LMU Munich; Munich Center for Machine Learning (MCML))

Automated Sperm Assessment Framework and Neural Network Specialized for Sperm Video Recognition 7660
Takuro Fujii (Yokohama National University), Hayato Nakagawa (Yokohama National University), Teppei Takeshima (Yokohama City University Medical Center), Yasushi Yumura (Yokohama City University Medical Center), and Tomoki Hamagami (Yokohama National University)

Generalizing to Unseen Domains in Diabetic Retinopathy Classification 7670
Chamuditha Jayanga Galappaththige (Mohamed Bin Zayed University of Artificial Intelligence), Gayal Kuruppu (Mohamed bin Zayed University of Artificial Intelligence), and Muhammad Haris Khan (Mohamed Bin Zayed University of Artificial Intelligence)

Dynamic Multimodal Information Bottleneck for Multimodality Classification 7681
Yingying Fang (National Heart and Lung Institute, Imperial College London, UK), Shuang Wu (Black Sesame Technologies, Singapore), Sheng Zhang (National Heart and Lung Institute, Imperial College London, UK), Chaoyan Huang (The Chinese University of Hong Kong, Hong Kong), Tieyong Zeng (The Chinese University of Hong Kong, Hong Kong), Xiaodan Xing (National Heart and Lung Institute, Imperial College London, UK), Simon Walsh (National Heart and Lung Institute, Imperial College London, UK), and Guang Yang (National Heart and Lung Institute, Imperial College London; Bioengineering Department and Imperial-X, Imperial College London, UK)

Hybrid Neural Diffeomorphic Flow for Shape Representation and Generation via Triplane 7692
Kun Han (University of California, USA), Shanlin Sun (University of California, USA), Thanh-Tung Le (University of California, USA), Xiangyi Yan (University of California, USA), Haoyu Ma (University of California, USA), Chenyu You (Yale University, USA), and Xiaohui Xie (University of California, USA)

Unsupervised Domain Adaptation of MRI Skull-Stripping Trained on Adult Data to Newborns .. 7703
Abbas Omidi (University of Calgary), Aida Mohammadshahi (University of Calgary), Neha Gianchandani (University of Calgary), Regan King (University of Calgary), Lara Leijser (University of Calgary), and Roberto Souza (University of Calgary)

G-CASCADE: Efficient Cascaded Graph Convolutional Decoding for 2D Medical Image Segmentation 7713
Md Mostafijur Rahman (The University of Texas at Austin) and Radu Marculescu (The University of Texas at Austin)

Bridging Generalization Gaps in High Content Imaging Through Online Self-Supervised Domain Adaptation 7723
Johan Fredin Haslum (KTH Royal Institute of Technology; Science for Life Laboratory; AstraZeneca, Sweden), Christos Matsoukas (KTH Royal Institute of Technology; Science for Life Laboratory; AstraZeneca, Sweden), Karl-Johan Leuchowius (AstraZeneca, Sweden), and Kevin Smith (KTH Royal Institute of Technology; Science for Life Laboratory, Sweden)

DR10K: Transfer Learning UsingWeak Labels for Grading Diabetic Retinopathy on DR10K Dataset 7733
Mohamed ElHabebe (Alexandria University; Applied Innovation Center, MCIT, Egypt), Shereen Elkordi (Applied Innovation Center, MCIT, Egypt), Ahmed Gamal ElDin (Applied Innovation Center, MCIT, Egypt), Noha Adly (Alexandria University; Applied Innovation Center, MCIT, Egypt), Marwan Torki (Alexandria University; Applied Innovation Center, MCIT, Egypt), Ahmed ElMassry (Alexandria University, Egypt), and Islam Sh Ahmed (Alexandria University, Egypt)

SCUNet++: Swin-UNet and CNN Bottleneck Hybrid Architecture With Multi-Fusion Dense Skip Connection for Pulmonary Embolism CT Image Segmentation 7744
Yifei Chen (Hangzhou Dianzi University), Binfeng Zou (Hangzhou Dianzi University), Zhaoxin Guo (hangzhou dianzi University), Yiyu Huang (Hangzhou Dianzi University), Yifan Huang (Hangzhou Dianzi University), Feiwei Qin (Hangzhou Dianzi University), Changmiao Wang (Shenzhen Research Institute of Big Data), and Qinhai Li (Second Affiliated Hospital of the Chinese University of Hong Kong, Shenzhen)

SynergyNet: Bridging the Gap Between Discrete and Continuous Representations for Precise Medical Image Segmentation 7753
Vandan Gorade (Indian Institute of Technology Roorkee, India), Sparsh Mittal (Indian Institute of Technology Roorkee, India), Debesh Jha (Northwestern University), and Ulas Bagci (Northwestern University)

Reverse Knowledge Distillation: Training a Large Model Using a Small One for Retinal Image Matching on Limited Data 7763
Sahar Almahfouz Nasser (Indian Institute of Technology Bombay, India), Nihar Gupte (IIT Bombay, India), and Amit Sethi (Indian Institute of Technology Bombay, India)

Activity-Based Early Autism Diagnosis Using a Multi-Dataset Supervised Contrastive Learning Approach 7773
Asha Rani (Indian Institute of Technology Jodhpur) and Yashaswi Verma (IIT Jodhpur)

Convolutional Masked Image Modeling for Dense Prediction Tasks on Pathology Images 7783
Yan Yang (The Australian National University), Liyuan Pan (The Australian National University), Liu Liu (ANU (Australian National University)), and Eric A Stone (The Australian National University)

Real-Time Polyp Detection in Colonoscopy Using Lightweight Transformer 7794
Youngbeom Yoo (KAIST, South Korea), Jae Young Lee (KAIST, South Korea), Dong-Jae Lee (KAIST, South Korea), Jiwoon Jeon (AINex, South Korea), and Junmo Kim (KAIST, South Korea)

Self-Supervised Learning with Masked Autoencoders for Teeth Segmentation From Intra-Oral 3D Scans 7805
Amani Almalki (Temple University, USA) and Longin Jan Latecki (Temple University, USA)

Self-Supervised Edge Detection Reconstruction for Topology-Informed 3D Axon Segmentation and Centerline Detection 7816
Alec S. Xu (MIT Lincoln Laboratory, USA), Nina I. Shamsi (MIT Lincoln Laboratory, USA), Lars A. Gjesteby (MIT Lincoln Laboratory, USA), and Laura J. Brattain (MIT Lincoln Laboratory, USA)

Robust Source-Free Domain Adaptation for Fundus Image Segmentation 7825
Lingrui Li (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Yanfeng Zhou (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), and Ge Yang (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences)

I-AI: A Controllable & Interpretable AI System for Decoding Radiologists' Intense Focus for Accurate CXR Diagnoses 7835
Trong Thang Pham (University of Arkansas), Jacob Michael Brecheisen (University of Arkansas), Anh Nguyen (University of Liverpool), Hien V Nguyen (University of Houston), and Ngan Le (University of Arkansas)

FreMIM: Fourier Transform Meets Masked Image Modeling for Medical Image Segmentation 7845
Wenxuan Wang (University of Science and Technology Beijing), Jing Wang (University of Science and Technology Beijing), Chen Chen (University of Central Florida), Jianbo Jiao (University of Birmingham), Yuanxiu Cai (University of Science and Technology Beijing), Shanshan Song (University of Science and Technology Beijing), and Jiangyun Li (University of Science and Technology Beijing)

Med-DANet V2: A Flexible Dynamic Architecture for Efficient Medical Volumetric Segmentation 7856
Haoran Shen (University of Science and Technology Beijing), Yifu Zhang (University of Science and Technology Beijing), Wenxuan Wang (University of Science and Technology Beijing), Chen Chen (University of Central Florida), Jing Liu (National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences), Shanshan Song (University of Science and Technology Beijing), and Jiangyun Li (University of Science and Technology Beijing)

Ordinal Classification With Distance Regularization for Robust Brain Age Prediction 7867
Jay Shah (Arizona State University), Md Mahfuzur Rahman Siddiquee (Arizona State University), Yi Su (Banner Alzheimer's Institute), Teresa Wu (Arizona State University), and Baoxin Li (Arizona State University)

CryoRL: Reinforcement Learning Enables Efficient Cryo-EM Data Collection 7877
Quanfu Fan (Amazon), Yilai Li (University of Michigan), Yuguang Yao (Michigan State University), John Cohn (IBM), Sijia Liu (Michigan State University), Ziping Xu (Harvard University), Seychelle Vos (MIT), and Michael Cianfrocco (University of Michigan)

Prototype Learning for Explainable Brain Age Prediction 7888
Linde S. Hesse (University of Oxford, UK), Nicola K. Dinsdale (University of Oxford, UK), and Ana I.L. Namburete (University of Oxford, UK)

BigSmall: Efficient Multi-Task Learning for Disparate Spatial and Temporal Physiological Measurements 7899
Girish Narayanswamy (University of Washington), Yujia Liu (University of Washington), Yuzhe Yang (Massachusetts Institute of Technology), Chengqian Ma (University of Washington), Xin Liu (University of Washington), Daniel Mcduff (University of Washington), and Shwetak Patel (University of Washington)

Self-Sampling Meta SAM: Enhancing Few-Shot Medical Image Segmentation With Meta-Learning 7910
Tianang Leng (Huazhong University of Science and Technology), Yiming Zhang (Tokyo Institute of Technology), Kun Han (University of California Irvine), and Xiaohui Xie (University of California, Irvine)

Controllable Text-to-Image Synthesis for Multi-Modality MR Images 7921
Kyuri Kim (Seoul National University), Yoonho Na (Seoul National University), Sung-Joon Ye (Seoul National University), Jimin Lee (Ulsan National Institute of Science & Technology), Sung Soo Ahn (Yonsei University College of Medicine), Ji Eun Park (University of Ulsan College of Medicine), and Hwiyoung Kim (Yonsei University College of Medicine)

SC-MIL: Supervised Contrastive Multiple Instance Learning for Imbalanced Classification in Pathology 7931
Dinkar Juyal (PathAI), Siddhant N Shingi (University of Massachusetts Amherst), Syed Ashar Javed (PathAI), Harshith Padigela (PathAI), Chintan Shah (PathAI, Inc), Anand Sampat (PathAI), Archit Khosla (PathAI), John Abel (PathAI), and Amaro Taylor-Weiner (PathAI)

FATE: Feature-Agnostic Transformer-Based Encoder for Learning Generalized Embedding Spaces in Flow Cytometry Data 7941
Lisa Weijler (TU Wien), Florian Kowarsch (TU Wien), Michael Reiter (TU Wien), Pedro Hermosilla (TU Wien), Margarita Maurer-Granofszky (St. Anna CCRI), and Michael Dworzak (St. Anna CCRI)

Dual Domain Diffusion Guidance for 3D CBCT Metal Artifact Reduction 7950
Yongjin Choi (Korea University), Doeyoung Kwon (Korea University), and Seung Baek (Korea University)

AFTer-SAM: Adapting SAM With Axial Fusion Transformer for Medical Imaging Segmentation . 7960
Xiangyi Yan (University of California, Irvine), Shanlin Sun (University of California, Irvine), Kun Han (University of California Irvine), Thanh-Tung Le (University of California, Irvine), Haoyu Ma (University of California, Irvine), Chenyu You (Yale University), and Xiaohui Xie (University of California, Irvine)

MEGANet: Multi-Scale Edge-Guided Attention Network for Weak Boundary Polyp Segmentation 7970
Nhat-Tan Bui (University of Arkansas, USA), Dinh-Hieu Hoang (University of Science, and John von Neumann Institute, VNU-HCM; Vietnam National University, Vietnam), Quang-Thuc Nguyen (University of Science, and John von Neumann Institute, VNU-HCM; Vietnam National University, Vietnam), Minh-Triet Tran (University of Science, and John von Neumann Institute, VNU-HCM; Vietnam National University, Vietnam), and Ngan Le (University of Arkansas, USA)

Complex Organ Mask Guided Radiology Report Generation 7980
Tiancheng Gu (University of Sydney, Australia), Dongnan Liu (University of Sydney, Australia), Zhiyuan Li (University of Sydney, Australia), and Weidong Cai (University of Sydney, Australia)

Improved Topological Preservation in 3D Axon Segmentation and Centerline Detection Using Geometric Assessment-Driven Topological Smoothing (GATS) 7990
Nina I. Shamsi (MIT Lincoln Laboratory, USA), Alec S. Xu (MIT Lincoln Laboratory, USA), Lars A. Gjesteby (MIT Lincoln Laboratory, USA), and Laura J. Brattain (MIT Lincoln Laboratory, USA)

Commercial / Retail

CoD: Coherent Detection of Entities From Images with Multiple Modalities 8000
Vinay Verma (International Machine Learning (IML) Amazon, India), Dween Sanny (International Machine Learning (IML) Amazon, India), Abhishek Singh (International Machine Learning (IML) Amazon, India), and Deepak Gupta (International Machine Learning (IML) Amazon, India)

DTrOCR: Decoder-Only Transformer for Optical Character Recognition 8010
Masato Fujitake (Fast accounting co., ltd.)

DECDM: Document Enhancement Using Cycle-Consistent Diffusion Models 8021
Jiaxin Zhang (Intuit AI Research), Joy Rimchala (Intuit), Lalla Mouatadid (Intuit), Kamalika Das (Intuit), and Sricharan Kumar (Intuit)

LipAT: Beyond Style Transfer for Controllable Neural Simulation of Lipstick Using Cosmetic Attributes 8031
Amila Silva (The University of Melbourne; Amazon), Olga Moskvyak (Amazon), Alexander Long (Amazon), Ravi Garg (Amazon), Stephen Gould (Amazon; Australian National University), Gil Avraham (Amazon), and Anton Van Den Hengel (Amazon; The University of Adelaide)

Learning Visual Body-Shape-Aware Embeddings for Fashion Compatibility 8041
Kaicheng Pang (The Hong Kong Polytechnic University, Hong Kong SAR), Xingxing Zou (The Hong Kong Polytechnic University, Hong Kong SAR), and Waikeung Wong (The Hong Kong Polytechnic University, Hong Kong SAR)

Lost Your Style? Navigating with Semantic-Level Approach for Text-To-Outfit Retrieval 8051
Junkyu Jang (KAIST, College of Business, Korea), Eugene Hwang (KAIST, College of Business, Korea), and Sung-Hyuk Park (KAIST, College of Business, Korea)

TAMPAR: Visual Tampering Detection for Parcel Logistics in Postal Supply Chains 8061
Alexander Naumann (FZI Research Center for Information Technology), Felix Hertlein (FZI Research Center for Information Technology), Laura Dörr (FZI Research Center for Information Technology), and Kai Furmans (Karlsruhe Institute of Technology)

MAdVerse: A Hierarchical Dataset of Multi-Lingual Ads From Diverse Sources and Categories ... 8072
Amruth Sagar (CVIT, IIIT Hyderabad, India), Rishabh Srivastava (CVIT, IIIT Hyderabad, India), Rakshitha R T (KLE Tech, India), Venkata Kesav Venna (CVIT, IIIT Hyderabad, India), and Ravi Kiran Sarvadevabhatla (CVIT, IIIT Hyderabad, India)

FOUND: Foot Optimization with Uncertain Normals for Surface Deformation Using Synthetic Data 8082
Oliver Boyne (University of Cambridge, U.K), Gwangbin Bae (University of Cambridge, U.K), James Charles (University of Cambridge, U.K), and Roberto Cipolla (University of Cambridge, U.K)

Iterative Multi-Granular Image Editing Using Diffusion Models 8092
K J Joseph (Adobe Research, India), Prateksha Udhayanan (Adobe Research, India), Tripti Shukla (Adobe, India), Aishwarya Agarwal (Adobe Research, India), Srikrishna Karanam (Adobe Research, India), Koustava Goswami (Adobe Research, India), and Balaji Vasan Srinivasan (Adobe Research, India)

MIVC: Multiple Instance Visual Component for Visual-Language Models 8102
Wenyi Wu (Amazon), Qi Li (Amazon), Wenliang Zhong (University of Texas at Arlington), and Junzhou Huang (University of Texas at Arlington)

Education

A One-Shot Learning Approach To Document Layout Segmentation of Ancient Arabic Manuscripts.... 8112
Axel De Nardin (Università di Udine), Silvia Zottin (Università degli Studi di Udine), Claudio Piciarelli (University of Udine, Italy), Emanuela Colombi (Università di Udine), and Gian Luca Foresti (University of Udine, Italy)

Embedded Sensing / Real-Time Techniques

Real-Time 6-DoF Pose Estimation by an Event-Based Camera Using Active LED Markers 8122
Gerald Ebmer (TU Wien), Adam Loch (AIT), Minh Nhat Nhat Vu (TU Wien), Roberto Mecca (Austrian Institute of Technology), Germain Haessig (Prophesee), Christian Hartl-Nesic (TU Wien), Markus Vincze (TU Wien), and Andreas Kugi (TU Wien)

Spiking Neural Networks for Active Time-Resolved SPAD Imaging 8132
Yang Lin (EPFL) and Edoardo Charbon (EPFL)

Robust TRISO-Fueled Pebble Identification by Digit Recognition 8142
Roshan Kenia (Stony Brook University), Jihane Mendil (Missouri University of Science and Technology), Ahmed Jasim (Missouri University of Science and Technology), Muthanna Al-Dahhan (Missouri University of Science and Technology), and Zhaozheng Yin (Stony Brook University)

CATS: Combined Activation and Temporal Suppression for Efficient Network Inference 8151
Zeqi Zhu (GrAI Matter Labs B.V.; Eindhoven University of Technology), Arash Pourtaherian (GrAI Matter Labs B.V.), Luc Waeijen (GrAI Matter Labs B.V.), Ibrahim Batuhan Akkaya (GrAI Matter Labs B.V.), Egor Bondarev (Eindhoven University of Technology), and Orlando Moreira (GrAI Matter Labs B.V.)

Environmental Monitoring / Climate Change / Ecology

Vision Transformer for Multispectral Satellite Imagery: Advancing Landcover Classification..... 8161
Ryan Rad (Northeastern University)

Food Science and Nutrition

FIRE: Food Image to REcipe Generation 8169
Prateek Chhikara (University of Southern California; Information Sciences Institute, USA), Dhiraj Chaurasia (University of Southern California; Information Sciences Institute, USA), Yifan Jiang (University of Southern California; Information Sciences Institute, USA), Omkar Masur (University of Southern California, USA), and Filip Ilievski (Information Sciences Institute, USA; Vrije Universiteit Amsterdam, Netherlands)

Online Class-Incremental Learning for Real-World Food Image Classification 8180
Siddeshwar Raghavan (Purdue University), Jiangpeng He (Purdue University), and Fengqing Maggie Zhu (Purdue University, USA)

Psychology and Cognitive Science

LibreFace: An Open-Source Toolkit for Deep Facial Expression Analysis 8190
Di Chang (University of Southern California), Yufeng Yin (University of Southern California), Zongjian Li (USC Institute for Creative Technologies), Minh Tran (University of Southern California), and Mohammad Soleymani (University of Southern California)

Neural Echos: Depthwise Convolutional Filters Replicate Biological Receptive Fields 8201
Zahra Babaiee (TU Wien), Peyman M Kiasari (University of Waterloo), Daniela Rus (Massachusetts Institute of Technology), and Radu Grosu (TU Wien)

DREAM: Visual Decoding From Reversing Human Visual System 8211
Weihao Xia (University College London), Raoul de Charette (Inria), Cengiz Oztireli (University of Cambridge), and Jing-Hao Xue (University College London)

Remote Sensing

MarsLS-Net: Martian Landslides Segmentation Network and Benchmark Dataset 8221
Sidike Paheding (Fairfield University), Abel A Reyes-Angulo (Michigan Technological University), A Rajaneesh (University of Kerala), Sajin K Ks (University of Kerala), and Thomas Oommen (University of Mississippi)

Attention Modules Improve Image-Level Anomaly Detection for Industrial Inspection: A DifferNet Case Study 8231
André Luiz B Vieira E Silva (Voxar Labs, Centro de Informática, Universidade Federal de Pernambuco), Francisco Paulo Magalhães Simões (Universidade Federal Rural de Pernambuco (UFRPE)), Danny Kowerko (Chemnitz University of Technology), Tobias Schlosser (Chemnitz University of Technology), Felipe M Battisti (UFPE), and Veronica Teichrieb (Voxar Labs, Informatics Center, UFPE)

CHAI: Craters in Historical Aerial Images 8241
Marvin Burges (Computer Vision Lab, TU Wien, Austria), Sebastian Zambanini (Computer Vision Lab, TU Wien, Austria), and Philipp Pirker (Luftbilddatenbank Dr. Carls GmbH, Austria)

ShadowSense: Unsupervised Domain Adaptation and Feature Fusion for Shadow-Agnostic Tree Crown Detection From RGB-Thermal Drone Imagery 8251
Rudraksh Kapil (University of Alberta, Canada), Seyed Mojtaba Marvasti-Zadeh (University of Alberta, Canada), Nadir Erbilgin (University of Alberta, Canada), and Nilanjan Ray (University of Alberta, Canada)

WATCH: Wide-Area Terrestrial Change Hypercube 8262
Connor Greenwell (Kitware, Inc.), Jonathan Crall (Kitware), Matthew Purri (Rutgers University), Kristin Dana (Rutgers University), Nathan Jacobs (Washington University in St. Louis), Armin Hadzic (Johns Hopkins University Applied Physics Laboratory), Scott Workman (DZYNE Technologies), and Matt Leotta (Kitware)

SyntheWorld: A Large-Scale Synthetic Dataset for Land Cover Mapping and Building Change Detection 8272
Jian Song (The University of Tokyo, Japan), Hongruixuan Chen (The University of Tokyo, Japan), and Naoto Yokoya (The University of Tokyo, Japan)

Seeing Stars: Learned Star Localization for Narrow-Field Astrometry 8282
Violet Felt (USSF) and Justin R Fletcher (USSF)

Deep Optics for Optomechanical Control Policy Design 8291
Justin Fletcher (University of Hawai'i)

A Visual Active Search Framework for Geospatial Exploration 8301
Anindya Sarkar (Washington University in St. Louis), Michael Lanier (Washington University In St. Louis), Scott Alfeld (Amherst College), Jiarui Feng (Washington University in St.Louis), Roman Garnett (Washington University in St. Louis), Nathan Jacobs (Washington University in St. Louis), and Yevgeniy Vorobeychik (Washington University in St. Louis)

Revolutionize the Oceanic Drone RGB Imagery With Pioneering Sun Glint Detection and Removal Techniques 8311
Jiangying Qin (Wuhan University), Ming Li (ETH ZurichWuhan University), Jie Zhao (Technical University of Munich), Jiageng Zhong (Wuhan University), and Hanqi Zhang (Wuhan University)

Multimodal Deep Learning for Remote Stress Estimation Using CCT-LSTM 8321
Sayyedjavad Ziaratnia (Shibaura Institute of Technology), Tipporn Laohakangvalvit (Shibaura Institute of Technology), Midori Sugaya (Shibaura Institute of Technology), and Peeraya Sripian (Shibaura Institute of Technology)

Defense Against Adversarial Cloud Attack on Remote Sensing Salient Object Detection 8330
Huiming Sun (Cleveland State University), Lan Fu (University of South Carolina), Jinlong Li (Cleveland State University), Qing Guo (ASTAR), Zibo Meng (Innopeak Technology), Tianyun Zhang (Cleveland State University), Yuewei Lin (Brookhaven National Laboratory), and Hongkai Yu (Cleveland State University)

Segment Anything, From Space? 8340
Simiao Ren (Duke University), Francesco Luzi (Duke University), Saad Lahrichi (Duke Kunshan University), Kaleb Kassaw (Duke University), Leslie M. Collins (Duke University), Kyle Bradbury (Duke University), and Jordan M. Malof (University of Montana)

Prototypical Contrastive Network for Imbalanced Aerial Image Segmentation 8351
Keiller Nogueira (University of Stirling), Mayara Faita (University of Western São Paulo), Ana Paula Ramos (University of Western São Paulo), Wesley Gonçalves (UFMS), José Marcato Junior (UFMS), and Jefersson A Dos Santos (University of Sheffield)

Robotics

Object Re-Identification From Point Clouds 8362
Benjamin Thérien (University of Waterloo), Chengjie Huang (University of Waterloo), Adrian Chow (University of Waterloo), and Krzysztof Czarnecki (University of Waterloo)

Shape From Shading for Robotic Manipulation 8374
Arkadeep Narayan Chaudhury (The Robotics Institute, Carnegie Mellon University, USA), Leonid Keselman (The Robotics Institute, Carnegie Mellon University, USA), and Christopher G. Atkeson (The Robotics Institute, Carnegie Mellon University, USA)

FinderNet: A Data Augmentation Free Canonicalization Aided Loop Detection and Closure Technique for Point Clouds in 6-DOF Separation. 8384
Sudarshan S Harithas (IIIT Hyderabad), Gurkirat Singh (IIIT Hyderabad), Aneesh Chavan (IIIT Hyderabad), Sarthak Sharma (IIIT Hyderabad), Suraj Patni (IIT Delhi), Chetan Arora (IIT Delhi), and Madhava Krishna (IIIT-Hyderabad)

From Chaos to Calibration: A Geometric Mutual Information Approach To Target-Free Camera LiDAR Extrinsic Calibration 8394
John Borer (Spleenlab GmbH), Jeremy Tschirner (Spleenlab GmbH), Florian Ölsner Ölsner (Spleenlab GmbH), and Stefan Milz (Spleenlab GmbH)

Continual Learning of Unsupervised Monocular Depth From Videos 8404
Hemang Chawla (Eindhoven University of Technology (TU/e)), Arnav Varma (TomTom), Elahe Arani (Eindhoven University of Technology (TU/e); Wayve), and Bahram Zonooz (Eindhoven University of Technology (TU/e); TomTom)

Classifying Cable Tendency With Semantic Segmentation by Utilizing Real and Simulated RGB Data 8415
Peichun Chien (Yazaki Corporation), Powei Liao (Waseda University), Eiji Fukuzawa (Waseda University), and Jun Ohya (Waseda University)

Multi-Class Segmentation From Aerial Views Using Recursive Noise Diffusion 8424
Benedikt Kolbeinsson (Imperial College London) and Krystian Mikolajczyk (Imperial College London)

Smartphones / End User Devices

Efficient Layout-Guided Image Inpainting for Mobile Use 8435
Wenbo Li (Samsung Research America), Yi Wei (Samsung Research America), Yilin Shen (Samsung Research America), and Hongxia Jin (Samsung Research America)

Edge Inference with Fully Differentiable Quantized Mixed Precision Neural Networks 8445
Clemens Js Schaefer (University of Notre Dame, USA), Siddharth Joshi (University of Notre Dame, USA), Shan Li (Google LLC, USA), and Raul Blazquez (Google LLC, USA)

POP-VQA – Privacy Preserving, On-Device, Personalized Visual Question Answering 8455
Pragya Paramita Sahu (Samsung Research Institute, India), Abhishek Raut (Samsung Research Institute, India), Jagdish Singh Samant (Samsung Research Institute, India), Mahesh Gorijala (Samsung Research Institute, India), Vignesh Lakshminarayanan (Samsung Research Institute, India), and Pinaki Bhaskar (Samsung Research Institute, India)

Sketch-Based Video Object Localization 8465
Sangmin Woo (KAIST), So-Yeong Jeon (KAIST; Korea Agency for Defense Development), Jinyoung Park (KAIST), Minji Son (LG Electronics), Sumin Lee (KAIST), and Changick Kim (KAIST)

Feed-Forward Latent Domain Adaptation 8475
Ondrej Bohdal (The University of Edinburgh, UK), Da Li (Samsung AI Center Cambridge, UK), Shell Xu Hu (Samsung AI Center Cambridge, UK), and Timothy Hospedales (The University of Edinburgh; Samsung AI Center Cambridge, UK)

BSRAW: Improving Blind RAW Image Super-Resolution 8485
Marcos V. Conde (University of Würzburg), Florin-Alexandru Vasluianu (Computer Vision Lab, University of Wurzburg), and Radu Timofte (University of Wurzburg & ETH Zurich)

Social Good

Tracking Skiers From the Top to the Bottom 8496
Matteo Dunnhofer (University of Udine, Italy), Luca Sordi (University of Udine, Italy), Niki Martinel (University of Udine, Italy), and Christian Micheloni (University of Udine, Italy)

Salient Object Detection for Images Taken by PeopleWith Vision Impairments 8507
Jarek Reynolds (University of Colorado Boulder), Chandra Kanth Nagesh (University of Colorado Boulder), and Danna Gurari (University of Colorado Boulder)

Rethinking Multimodal Content Moderation From an Asymmetric Angle With Mixed-Modality . 8517
Jialin Yuan (Oregon State University), Ye Yu (Microsoft), Gaurav Mittal (Microsoft), Matthew Hall (Microsoft), Sandra Sajeev (Microsoft), and Mei Chen (Microsoft)

Privacy-Enhancing Person Re-Identification Framework – A Dual-Stage Approach 8528
Kajal Kansal (NUS, Singapore), Yongkang Wong (National University of Singapore), and Mohan Kankanhalli (National University of Singapore,)

Toward Planet-Wide Traffic Camera Calibration 8538
Khiem Vuong (Carnegie Mellon University), Robert Tamburo (Carnegie Mellon University), and Srinivasa G. Narasimhan (Carnegie Mellon University)

VideoFACT: Detecting Video Forgeries Using Attention, Scene Context, and Forensic Traces 8548
Tai D. Nguyen (Drexel University, USA), Shengbang Fang (Drexel University, USA), and Matthew C. Stamm (Drexel University, USA)

OE-CTST: Outlier-Embedded Cross Temporal Scale Transformer for Weakly-Supervised Video Anomaly Detection 8559
Snehashis Majhi (INRIA; Côte d'Azur University), Rui Dai (INRIA; Côte d'Azur University), Quan Kong (Woven by Toyota), Lorenzo Garattoni (Toyota Motor Europe), Gianpiero Francesca (Toyota Motor Europe), and François Brémond (INRIA; Côte d'Azur University)

NOMAD: A Natural, Occluded, Multi-Scale Aerial Dataset, for Emergency Response Scenarios ... 8569
Arturo Miguel Russell Bernal (University of Notre Dame, USA), Walter Scheirer (University of Notre Dame, USA), and Jane Cleland-Huang (University of Notre Dame, USA)

PreciseDebias: An Automatic Prompt Engineering Approach for Generative AI to Mitigate Image Demographic Biases 8581
Colton Clemmer (University of North Texas, USA), Junhua Ding (University of North Texas, USA), and Yunhe Feng (University of North Texas, USA)

P-Age: Pexels Dataset for Robust Spatio-Temporal Apparent Age Classification 8591
Abid Ali (INRIA; Université Cote d'Azur, France), Ashish Marisetty (IIIT Naya Raipur, India), and Francois Bremond (INRIA; Université Cote d'Azur, France)

Structural Engineering / Civil Engineering

Learning to Read Analog Gauges from Synthetic Data 8601
Juan C Leon (KAUST), Yazeed Alnumay (Aramco), Cheng Zheng (KAUST), Sahejad Patel (Saudi Aramco), Hassane A Trigui (Aramco), and Bernard Ghanem (KAUST)

dacl10k: Benchmark for Semantic Bridge Damage Segmentation 8611
Johannes Flotzinger (University of the Bundewehr Munich), Philipp J. Rösch (University of the Bundeswehr Munich), and Thomas Braml (University of the Bundeswehr Munich)

Designing a Hybrid Neural System to Learn Real-World Crack Segmentation From Fractal-Based Simulation 8621
Achref Jaziri (Goethe University, Germany), Martin Mundt (TU Darmstadt and hessian.AI, Germany), Andres Fernandez Rodriguez (University of Tübingen, Germany), and Visvanathan Ramesh (Goethe University and hessian.AI, Germany)

Zero-Shot Building Attribute Extraction From Large-Scale Vision and Language Models 8632
Fei Pan (University of Michigan), Sangryul Jeon (University of Michigan), Brian S Wang (University of Michigan), Frank T Mckenna (UC Berkeley), and Stella X Yu (University of Michigan)

Virtual / Augmented Reality

Leveraging Next-Active Objects for Context-Aware Anticipation in Egocentric Videos 8642
Sanket Thakur (Istituto Italiano di Tecnologia (IIT); University of Genoa, Italy), Cigdem Beyan (University of Bergamo, Italy), Pietro Morerio (Istituto Italiano di Tecnologia (IIT)), Vittorio Murino (University of Verona; Istituto Italiano di Tecnologia (IIT), Italy), and Alessio Del Bue (Istituto Italiano di Tecnologia (IIT), Italy)

Volumetric Disentanglement for 3D Scene Manipulation 8652
Sagie Benaim (University of Copenhagen), Frederik Warburg (Technical University of Denmark), Peter Ebert Christensen (University of Copenhagen), and Serge Belongie (University of Copenhagen)

StyleAvatar: Stylizing Animatable Head Avatars 8663
Juan C Perez (KAUST), Thu H Nguyen-Phuoc (Reality Lab Research, Meta), Chen Cao (Facebook Reality Labs), Artsiom Sanakoyeu (Reality Labs), Tomas Simon (Facebook Reality Lab Research), Pablo Arbelaez (Universidad de los Andes), Bernard Ghanem (KAUST), Ali K Thabet (Facebook), and Albert Pumarola (Meta)

Show Your Face: Restoring Complete Facial Images From Partial Observations for VR Meeting ... 8673
Zheng Chen (Indiana University), Zhiqi Zhang (oppo), Junsong Yuan (State University of New York at Buffalo, USA), Yi Xu (OPPO US Research Center), and Lantao Liu (Indiana University, Intelligent Systems Engineering)

PressureVision++: Estimating Fingertip Pressure From Diverse RGB Images 8683
Patrick Grady (Georgia Institute of Technology), Jeremy A. Collins (Georgia Institute of Technology), Chengcheng Tang (Meta Reality Labs), Christopher D. Twigg (Meta Reality Labs), Kunal Aneja (Georgia Institute of Technology), James Hays (Georgia Institute of Technology, USA), and Charles C. Kemp (Georgia Institute of Technology)

A Generative Multi-Resolution Pyramid and Normal-Conditioning 3D Cloth Draping 8694
Hunor Laczkó (Universitat Autònoma de Barcelona; Computer Vision Center), Meysam Madadi (Computer Vision Center; Universitat de Barcelona), Sergio Escalera (Computer Vision Center; Universitat de Barcelona), and Jordi Gonzalez (Universitat Autònoma de Barcelona; Computer Vision Center)

Stereo Matching in Time: 100+ FPS Video Stereo Matching for Extended Reality 8704
Ziang Cheng (Australian National University), Jiayu Yang (The Australian National University), and Hongdong Li (Australian National University, Australia)

Author Index