

2025 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW 2025)

**Honolulu, Hawaii, USA
19-20 October 2025**

Pages 1-705



**IEEE Catalog Number: CFP2591A-POD
ISBN: 979-8-3315-8989-9**

**Copyright © 2025 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:	CFP2591A-POD
ISBN (Print-On-Demand):	979-8-3315-8989-9
ISBN (Online):	979-8-3315-8988-2
ISSN:	2473-9936

Additional Copies of This Publication Are Available From:

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

CURRAN ASSOCIATES INC.
proceedings
.com

2025 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW) ICCVW 2025

Table of Contents

Welcome Message cxviii

Affective & Behavior Analysis in-the-wild

From Emotions to Violence: Multimodal Fine-Grained Behavior Analysis at the 9th ABAW 1
Dimitrios Kollias (Queen Mary University London), Stefanos Zafeiriou (Imperial College London, UK), Irene Kotsia (Cogitat, UK), Greg Slabaugh (Queen Mary University of London), Damith Chamalke Senadeera (Queen Mary University of London), Jianian Zheng (University College London), Kaushal Kumar Keshlal Yadav (Queen Mary University of London), Chunchang Shao (Queen Mary University of London), and Guanyu Hu (Queen Mary University of London)

On the Use of Hierarchical Vision Foundation Models for Low-Cost Human Mesh Recovery and Pose Estimation 13
Shuhei Tarashima (NTT DOCOMO BUSINESS, Inc. ;Tokyo Metropolitan University), Yushan Wang (Tokyo Metropolitan University), and Norio Tagawa (Tokyo Metropolitan University)

Valeo Near-Field: a novel dataset for pedestrian intent detection 25
Antonyo Musabini (Valeo, France), Rachid Benmokhtar (Valeo, France), Xavier Perrotton (Valeo, France), Jagdish Bhanushali (Valeo, United States), Victor Galizzi (Universite Paris-Saclay, France), and Bertrand Luvison (Universite Paris-Saclay, France)

Sweating the Stress: GSR Insights into Cognitive Load and Distraction 33
Santiago Clemente (George Walton Comprehensive High School, GA), Layla Varghese (George Walton Comprehensive High School, GA), Valentina Nino (Kennesaw State University, GA), and Maria Valero (Kennesaw State University, GA)

A Multimodal Dataset of Viewer Responses to Japanese Manzai Comedy 42
Kazuki Kawamura (Sony CSL Kyoto ;The University of Tokyo), Kengo Nakai (Yoshimoto Kogyo Holdings Co.,Ltd.), and Jun Rekimoto (Sony CSL Kyoto ;The University of Tokyo)

DEAP DIVE: Dataset Investigation with Vision transformers for EEG evaluation	51
<i>Annemarie Hoffsommer (Karlsruhe Institute of Technology (KIT), Germany), Helen Schneider (Karlsruhe Institute of Technology (KIT), Germany), Svetlana Paolitska (Karlsruhe Institute of Technology (KIT), Germany);FZI Research Center for Information Technology, Germany), and J. Marius Zöllner (Karlsruhe Institute of Technology (KIT), Germany ;FZI Research Center for Information Technology, Germany)</i>	
Dynamic Temporal Gating Networks for Cross-Modal Valence-Arousal Estimation	61
<i>Yubeen Lee (Sungkyunkwan University), Sangeun Lee (Electronics and Telecommunications Research Institute), Chaewon Park (Sungkyunkwan University), Junyeop Cha (Sungkyunkwan University), and Eunil Park (Sungkyunkwan University)</i>	
Zero-Shot Multimodal Compound Expression Recognition Approach using Off-the-Shelf Large Visual-Language Models	71
<i>Elena Ryumina (Russian Academy of Sciences, Russia), Maxim Markitantov (Russian Academy of Sciences, Russia), Alexandr Axyonov (Russian Academy of Sciences, Russia), Dmitry Ryumin (Russian Academy of Sciences, Russia), Mikhail Dolgushin (Russian Academy of Sciences, Russia), and Alexey Karpov (ITMO University, Russia)</i>	
Face Video Steganography for Privacy-protection Automatic Depression Assessment	80
<i>Xinyi Ni (University of Exeter, United Kingdom), Zijian Wu (Nanjing University of Science and Technology, China), Lu Liu (University of Exeter, United Kingdom), and Siyang Song (University of Exeter, United Kingdom)</i>	
VelocityNet: Real-Time Crowd Anomaly Detection via Person-Specific Velocity Analysis	87
<i>Fatima Alghamdi (Saudi Data and Artificial Intelligence Authority (SDAIA)), Omar Alharbi (Saudi Data and Artificial Intelligence Authority (SDAIA)), Abdullah Aldwyish (Saudi Data and Artificial Intelligence Authority (SDAIA)), Raied Aljadaany (Saudi Data and Artificial Intelligence Authority (SDAIA)), Muhammad Kamran J Khan (Saudi Data and Artificial Intelligence Authority (SDAIA)), and Huda Alamri (Saudi Data and Artificial Intelligence Authority (SDAIA))</i>	
Multi-task Learning For Joint Action and Gesture Recognition	95
<i>Konstantinos Spathis (Robotics Institute, Greece), Nikolaos Kardaris (National Technical University of Athens, Greece), and Petros Maragos (Robotics Institute, Greece ;National Technical University of Athens, Greece ;HERON - Hellenic Robotics Center of Excellence, Greece)</i>	
Experimental Evaluation of the Impact of Robot Path Shape and Speed on Human Affective States in a Hallway-Passing Scenario	105
<i>Benjamin Greenberg (Rutgers University, NJ), Michael Grossi (Rutgers University, NJ), Sai Likhith Karri (Rutgers University, NJ), Jingang Yi (Rutgers University, NJ), Jacob Feldman (Rutgers University, NJ), and Karin Stromswold (Rutgers University, NJ)</i>	
Aligning Multimodal Data for Fine-Grained Video Understanding via Cross-Attentive Recurrent Fusion	113
<i>Nam-Ho Kim (KBS(Korean Broadcasting System), Korea) and Jun-Hwa Kim (Konyang University, Korea)</i>	

Camera Calibration and Pose Estimation

Rotation Estimation of Multiple In-Vehicle Cameras from Dual FoV	120
<i>Jun Sato (Nagoya Institute of Technology, Japan), Fumihiko Sakaue (Nagoya Institute of Technology, Japan), Shunya Kumano (SOKEN Inc., Japan), Yusuke Ueda (SOKEN Inc., Japan), and Naoki Kawasaki (SOKEN Inc., Japan)</i>	
Sem-MASt3R: Semantically Guided Feature Matching with MASt3R	130
<i>Dario Tenore (ETH Zurich), Daniel Barath (ETH Zurich; HUN-REN SZTAKI), Marc Pollefeys (ETH Zurich; Microsoft Spatial AI Lab), and Qunjie Zhou (NVIDIA)</i>	
fCOP: Monocular Focal Length Estimation from Category-level Object Priors	140
<i>Xinyue Zhang (ShanghaiTech University), Jiaqi Yang (ShanghaiTech University), Xiangting Meng (ShanghaiTech University), Abdelrahman Mohamed (Rembrand), and Laurent Kneip (ShanghaiTech University)</i>	
Optic Illusion Patterns for Accurate Camera Calibration	149
<i>Gaku Nakano (NEC Corporation) and Takenobu Kiyama (NEC Corporation)</i>	
Model Ensemble to Fuse Geometric and Learning Solutions for Camera Rotation Estimation	159
<i>Bhuvan Aggarwal (Honda R&D, Japan), Amit More (Honda R&D, Japan), Mudit Soni (Honda R&D, Japan), and S Divakar Bhat (Honda R&D, Japan)</i>	
Direct Camera Calibration from Vanishing Points via Polynomial Solvers	169
<i>Norio Kosaka (LY Corporation)</i>	
Generalizable Visual Localization for Gaussian Splatting Scene Representations	178
<i>Fadi Khatib (Weizmann Institute of Science), Dror Moran (Weizmann Institute of Science), Guy Trostianetsky (Weizmann Institute of Science), Yoni Kasten (NVIDIA), Meirav Galun (Weizmann Institute of Science), and Ronen Basri (Weizmann Institute of Science)</i>	
FUSELOC: Fusing Global and Local Descriptors for Fast and Robust 2D-3D Matching in Visual Localization	190
<i>Son Tung Nguyen (Queensland University of Technology, Australia), Alejandro Fontan (Queensland University of Technology, Australia), Michael Milford (Queensland University of Technology, Australia), and Tobias Fischer (Queensland University of Technology, Australia)</i>	
Alignment Scores: Robust Metrics for Multiview Pose Accuracy Evaluation	200
<i>Seong Hun Lee (Sint-Katelijne-Waver, Belgium) and Javier Civera (University of Zaragoza, Spain)</i>	

Instance-Level Recognition and Generation Workshop

Benchmarking Feature Upsampling Methods for Vision Foundation Models using Interactive Segmentation	210
<i>Volodymyr Havrylov (University of Tübingen), Haiwen Huang (University of Tübingen, Tübingen AI Center), Dan Zhang (Bosch Center for Artificial Intelligence), and Andreas Geiger (University of Tübingen, Tübingen AI Center)</i>	

Motion-Refined DINOSAUR for Unsupervised Multi-Object Discovery	220
<i>Xinrui Gong (TU Darmstadt), Oliver Hahn (TU Darmstadt), Christoph Reich (TU Darmstadt ;TU Munich ;MCML ;ELIZA), Krishnakant Singh (TU Darmstadt), Simone Schaub-Meyer (TU Darmstadt ;hessian.AI), Daniel Cremers (TU Munich ;MCML ;ELIZA), and Stefan Roth (TU Darmstadt ;hessian.AI ;ELIZA)</i>	

Systematic Trust in AI Models: Ensuring Fairness, Reliability, Explainability, and Accountability in Machine Learning Frameworks

ImageNet-BG: A Toolkit and Dataset for Evaluating Vision Model Robustness Against Background Variations	231
<i>Lukasz Piekarek (Wrocław University of Science and Technology, Poland) and Kamil Szyk (Wrocław University of Science and Technology, Poland)</i>	
From Global to Local: Social Bias Transfer in CLIP	241
<i>Ryan Ramos (The University of Osaka), Yusuke Hirota (NVIDIA), Yuta Nakashima (The University of Osaka), and Noa Garcia (The University of Osaka)</i>	
Robust Experts: The Effect of Adversarial Training on CNNs with Sparse Mixture-of-Experts Layers	251
<i>Svetlana Pavlitska (Karlsruhe Institute of Technology (KIT), Germany ;FZI Research Center for Information Technology, Germany), Haixi Fan (FZI Research Center for Information Technology, Germany), Konstantin Ditschuneit (Karlsruhe Institute of Technology (KIT), Germany), and J. Marius Zöllner (Karlsruhe Institute of Technology (KIT), Germany ;FZI Research Center for Information Technology, Germany)</i>	
COOkeD: Ensemble-Based OOD Detection in the Era of Zero-Shot CLIP	261
<i>Galadrielle Humblot-Renaux (Aalborg University, Denmark), Gianni Franchi (ENSTA, France), Sergio Escalera (Aalborg University, Denmark ;University of Barcelona, Spain), and Thomas B. Moeslund (Aalborg University, Denmark)</i>	
Are X-ray Landmark Detection Models Fair? A Preliminary Assessment and Mitigation Strategy ...	272
<i>Roberto Di Via (MaLGA-DIBRIS), Massimiliano Ciranni (MaLGA-DIBRIS), Davide Marinelli (MaLGA-DIBRIS), Allison Clement (Oxford University, UK), Nikil Patel (Oxford University, UK), Julian Wyatt (University of Oxford, UK), Francesca Odone (MaLGA-DIBRIS), Matteo Santacesaria (University of Genoa, Italy), Irina Voiculescu (Oxford University, UK), and Vito Paolo Pastore (MaLGA-DIBRIS)</i>	
Is It Certainly a Deepfake? Reliability Analysis in Detection & Generation Ecosystem	279
<i>Neslihan Kose (Intel Labs), Anthony Rhodes (Intel Labs), Umur Aybars Çiftçi (Binghamton University), and Ilke Demir (Cauth AI)</i>	
Explain with Confidence: Fusing Saliency Maps for Faithful and Interpretable Weakly-Supervised Model	290
<i>Ayush Somani (UiT The Arctic University of Norway), Arif Ahmed Sekh (UiT The Arctic University of Norway), and Dilip K. Prasad (UiT The Arctic University of Norway)</i>	

Data Bias Mitigation and Evaluation Framework for Diffusion-based Generative Face Models	300
<i>Hao Yu (Boston University), Margrit Betke (Boston University), and Sarah Adel Bargal (Georgetown University)</i>	
Extracting Uncertainty Estimates from Mixtures of Experts for Semantic Segmentation	311
<i>Svetlana Pavlitska (Karlsruhe Institute of Technology (KIT), Germany ;FZI Research Center for Information Technology, Germany), Beyza Keskin (Karlsruhe Institute of Technology (KIT), Germany), Alwin Faßbender (Karlsruhe Institute of Technology (KIT), Germany), Christian Hubschneider (FZI Research Center for Information Technology, Germany), and J. Marius Zöllner (Karlsruhe Institute of Technology (KIT), Germany ;FZI Research Center for Information Technology, Germany)</i>	
GenAI Confessions: Black-box Membership Inference for Generative Image Models	321
<i>Matyas Bohacek (Stanford University) and Hany Farid (University of California, USA)</i>	
Enhancing Vision-Language Models for Zero-Shot Video Action Recognition via Visual-Textual Refinement and Improved Interpretability	331
<i>Adeel Yousaf (University of Central Florida, USA) and Mubarak Shah (University of Central Florida, USA)</i>	

The 1st International Workshop and Challenge on Disentangled Representation Learning for Controllable Generation

DiViD: Disentangled Video Diffusion for Static–Dynamic Factorization	341
<i>Marzieh Gheisari (Ecole Normale Supérieure PSL, France) and Auguste Genovesio (Ecole Normale Supérieure PSL, France)</i>	
FusionGen: Feature Fusion-Based Few-Shot EEG Data Generation	351
<i>Yuheng Chen (Huazhong University of Science and Technology, China), Dingkun Liu (Huazhong University of Science and Technology, China; Zhongguancun Academy, China), Xinyao Yang (Huazhong University of Science and Technology, China), Xinpeng Xu (Huazhong University of Science and Technology, China), Baicheng Chen (Huazhong University of Science and Technology, China), and Dongrui Wu (Huazhong University of Science and Technology, China; Zhongguancun Academy, China)</i>	
Textual Semantics Matters: Unsupervised Representation Disentanglement in Realistic Scenarios with Language Inductive Bias	361
<i>Junhao Geng (Hunan University, China), Lexiang Lv (Hunan University, China), and Jianxin Lin (Hunan University, China)</i>	
Why Compress What You Can Generate? When GPT-4o Generation Ushers in Image Compression Fields	371
<i>Yixin Gao (University of Science and Technology of China), Xiaohan Pan (University of Science and Technology of China), Xin Li (University of Science and Technology of China), and Zhibo Chen (University of Science and Technology of China)</i>	
A Guided Fine-tuning Framework for Diffusion Models with Disentangled Semantic Priors for Multi-Factor Image Editing	382
<i>Baorui Peng (Georgia Institute of Technology) and Zhongming Chen (Chongqing University)</i>	

Controllable Generation with Disentangled Representative Learning of Multiple Perspectives in Autonomous Driving	388
<i>Haoran Jin (KTH Royal Institute of Technology)</i>	

The 2nd AI for Visual Arts Workshop

WP-CLIP: Leveraging CLIP to Predict Wölfflin's Principles in Visual Art	396
<i>Abhijay Ghildyal (Portland State University), Li-Yun Wang (Portland State University), and Feng Liu (Portland State University)</i>	
VQArt-Bench: A semantically rich VQA Benchmark for Art and Cultural Heritage	406
<i>Andrea Alfarano (University of Zurich, Switzerland), Lorenzo Venturoli (University of Zurich, Switzerland), and Dario Negueruela Del Castillo (University of Zurich, Switzerland)</i>	
Leveraging Diffusion Models for Stylization using Multiple Style Images	417
<i>Dan Ruta (Disney Research), Abdelaziz Djelouah (Disney Research), Raphael Ortiz (Disney Research), and Christopher Schroers (Disney Research)</i>	
Composite Reflections	427
<i>Abhishek Dangeti (Tata Consultancy Services Research, INDIA), Pavan Gajula (Tata Consultancy Services Research, INDIA), and Vikram Jamwal (Tata Consultancy Services Research, INDIA)</i>	
Controllable Single-shot Animation Blending with Temporal Conditioning	436
<i>Eleni Tselepi (University of Thessaly, Greece), Spyridon Thermos (Moverse, Greece), and Gerasimos Potamianos (University of Thessaly, Greece)</i>	
Enhancing Artwork Style Clustering via Neural Representation Re-Alignment	446
<i>Abhishek Dangeti (Tata Consultancy Services Research, INDIA), Pavan Gajula (Tata Consultancy Services Research, INDIA), Vikram Jamwal (Tata Consultancy Services Research, INDIA), and Vivek Srivastava (Tata Consultancy Services Research, INDIA)</i>	

The 4th DataCV Workshop and Challenge

LoRA-Loop: Closing the Synthetic Replay Cycle for Continual VLM Learning	456
<i>Kaihong Wang (Boston University), Donghyun Kim (Korea University), and Margrit Betke (Boston University)</i>	
Synthetic Hands Meet Legacy Data: A Synthetic Dataset for Structured, Controllable, and Multimodal Evaluation	466
<i>Menghe Zhang (Samsung Semiconductor Inc., USA), Haley M. So (Stanford University, USA), Mohammad Asadi (Stanford University, USA; Samsung Semiconductor Inc., USA), Dongfang Zhao (Samsung Semiconductor Inc., USA), Yangwen Liang (Samsung Semiconductor Inc., USA), Shuangquan Wang (Samsung Semiconductor Inc., USA), Gordon Wetzstein (Stanford University, USA), Kee-Bong Song (Samsung Semiconductor Inc., USA), and Donghoon Kim (Samsung Electronics, South Korea)</i>	

On the Generalization of Optical Flow: Quantifying Robustness to Dataset Shifts	478
<i>Katrin Bauer (University of Stuttgart), Andres Bruhn (University of Stuttgart), and Jenny Schmalfuss (University of Stuttgart)</i>	
Enhancing Domain Diversity in Synthetic Data Face Recognition with Dataset Fusion	489
<i>Anjith George (Idiap Research Institute, Switzerland) and Sébastien Marcel (Idiap Research Institute, Switzerland)</i>	
Express4D: Expressive, Friendly, and Extensible 4D Facial Motion Generation Benchmark	496
<i>Yaron Aloni (Tel Aviv University), Rotem Shalev-Arkushin (Tel Aviv University), Yonatan Shafir (Tel Aviv University), Guy Tevet (Tel Aviv University), Ohad Fried (Reichman University), and Amit Haim Bermano (Tel Aviv University)</i>	
Hybrid Generative Fusion for Efficient and Privacy-Preserving Face Recognition Dataset Generation	506
<i>Feiran Li (Institute of Information Engineering; University of Chinese Academy of Sciences), Qianqian Xu (Institute of Computing Technology), Shilong Bao (University of Chinese Academy of Sciences), Boyu Han (Institute of Computing Technology; University of Chinese Academy of Sciences), Zhiyong Yang (University of Chinese Academy of Sciences), and Qingming Huang (Institute of Computing Technology; University of Chinese Academy of Sciences)</i>	
Mapillary Vistas Validation for Fine-Grained Traffic Signs: A Benchmark Revealing Vision-Language Model Limitations	513
<i>Sparsh Garg (NEC Laboratories, America) and Abhishek Aich (NEC Laboratories, America)</i>	

The Challenge of Detecting Synthetic Manipulations in ID Documents

DeepID Challenge of Detecting Synthetic Manipulations in ID Documents	521
<i>Pavel Korshunov (Idiap Research Institute), Vidit Vidit (Idiap Research Institute), Amir Mohammadi (Idiap Research Institute), Christophe Ecabert (Idiap Research Institute), Nevena Shamoska (PXL Vision), Sébastien Marcel (Idiap Research Institute), Zeqin Yu (Sun Yat-sen University), Ye Tian (Sun Yat-sen University), Jiangqun Ni (Sun Yat-sen University), Lazar Lazarevic (Incode Technologies Inc.), Renat Khizbullin (Incode Technologies Inc.), Anastasiia Evteeva (Incode Technologies Inc.), Alexey Tochin (Incode Technologies Inc.), Aleksei Grishin (Incode Technologies Inc.), Anjith George (Idiap research institute), Daniel Deacala (Universidad Autonoma de Madrid), Tamás Endrei (Universidad Autonoma de Madrid; Pázmány Péter Catholic University), Javier Muñoz-Haro (Universidad Autonoma de Madrid), Ruben Tolosana (Universidad Autonoma de Madrid), Ruben Vera-Rodriguez (Universidad Autonoma de Madrid), Aythami Morales (Universidad Autonoma de Madrid), Julian Fierrez (Universidad Autonoma de Madrid), György Cserey (Pázmány Péter Catholic University), Hardik Sharma (Reagvis Labs Pvt. Ltd. ;IIT Mandi), Sachin Chaudhary (Reagvis Labs Pvt. Ltd. ;UPES Dehradun), Akshay Dudhane (MBZUAI), Praful Hambarde (IIT Mandi), Amit Shukla (IIT Mandi), Prateek Shaily (Reagvis Labs Pvt. Ltd. ;UPES Dehradun), Jayant Kumar (Reagvis Labs Pvt. Ltd. ;UPES Dehradun), Ajinkya Hase (IIT Mandi), Satish Maurya (IIT Mandi), Mridul Sharma (IIT Mandi), and Pallav Dwivedi (IIT Mandi)</i>	

International Workshop on Computer Vision for Physiological Measurement

BeatFormer: Efficient motion-robust remote heart rate estimation through unsupervised spectral zoomed attention filters	531
<i>Joaquim Comas (Pompeu Fabra University, Spain) and Federico Sukno (Pompeu Fabra University, Spain)</i>	
A High-Throughput Platform to Bench Test Smartphone-Based Heart Rate Measurements Derived From Video	542
<i>Ming-Zher Poh (Google), Jonathan Wang (Google), Jonathan Hsu (Google), Lawrence Cai (Google), Eric Teasley (Google), James Taylor (Google), Jameson Rogers (Google), Anupam Pathak (Google), and Shwetak Patel (Google)</i>	
Remote Heart Rate Measurement Based on Near Infrared Prior Information	549
<i>Hang Shao (Qingdao University, China) and Chuanfei Hu (Southeast University, China)</i>	
VitalVideos-Worldwide: A large and diverse rPPG dataset with rich ground truths	557
<i>Pieter-Jan Toye (VitalVideos)</i>	

Biomedical Image and Signal Computing for Unbiasedness, Interpretability, and Trustworthiness

AdeptHEQ-FL: Adaptive Homomorphic Encryption for Federated Learning of Hybrid Classical-Quantum Models with Dynamic Layer Sparing	563
<i>Md Abrar Jahin (University of Southern California), Taufikur Rahman Fuad (Islamic University of Technology), M. F. Mridha (American International University-Bangladesh), Nafiz Fahad (Multimedia University), and Md. Jakir Hossen (Multimedia University)</i>	
Evaluating the Trustworthiness of Foundation Models for Skin Lesion Segmentation	573
<i>Yusung Chu (Yonsei University, Republic of Korea), Byungho Oh (Yonsei University, Republic of Korea), and Sejung Yang (Yonsei University, Republic of Korea)</i>	
Understanding Dataset Bias in Medical Imaging: A Case Study on Chest X-rays	584
<i>Ethan Dack (University of Bern) and Chengliang Dai (UCB Biopharma UK ;Imperial College London)</i>	
Dino2-DR: A Trustworthy and Explainable Vision Transformer for Cross-Domain Diabetic Retinopathy Grading	595
<i>Lucia Cascone (University of Salerno, Italy), Luigi Di Biasi (University of Salerno, Italy), Giuseppe Genito (University of Salerno, Italy), and Michele Nappi (University of Salerno, Italy)</i>	
TrustMatch: Mitigating Pseudo-Label Bias in Semi-Supervised Learning with Trust-Aware Refinement	605
<i>Hongyang He (University of Warwick) and Yundi Hong (University of Warwick)</i>	

WaveDamp: Enhancing Natural Robustness in Endoscopy Through Wavelet-Based Frequency Damping	615
<i>Haiko Middeljans (Eindhoven University of Technology, The Netherlands), Carolus H.J. Kusters (Eindhoven University of Technology, The Netherlands), Tim J.M. Jaspers (Eindhoven University of Technology, The Netherlands), Martijn R. Jong (University of Amsterdam, The Netherlands), Rixta A.H. Van Eijck Van Heslinga (University of Amsterdam, The Netherlands), Florance C. Slooter (University of Amsterdam, The Netherlands), Albert J. De Groof (University of Amsterdam, The Netherlands), Jacques J. Bergman (University of Amsterdam, The Netherlands), Peter H.N. de With (Eindhoven University of Technology, The Netherlands), and Fons Van Der Sommen (Eindhoven University of Technology, The Netherlands)</i>	
SAM-SPJunc: Self-Prompting for Junction Detection in Retinal Images via Radius-Based Representations	626
<i>Minasadat Attari (University of Missouri-Columbia, USA), Kannappan Palaniappan (University of Missouri-Columbia, USA), and Filiz Bunyak (University of Missouri-Columbia, USA)</i>	
Measuring and Addressing Information Leakage in Concept Bottleneck Models	635
<i>Raffael Schon (Université Paris-Saclay), Baptiste Abeloos (Université Paris-Saclay, France), and Stéphane Herbin (Université Paris-Saclay, France)</i>	
Are Medical Image Generative Models Biologically Trustworthy?	644
<i>Suhyun Ahn (Korea Advanced Institute of Science and Technology (KAIST)), Wonjung Park (Korea Advanced Institute of Science and Technology (KAIST)), and Jinah Park (Korea Advanced Institute of Science and Technology (KAIST))</i>	

2nd Beyond Euclidean: Hyperbolic and Hyperspherical Learning for Computer Vision

Geometric Inductive Priors in Diffusion-Based Optical Flow Estimation	655
<i>Alberto Pepe (University of Cambridge, UK), Joan Lasenby (University of Cambridge, UK), and Paulo dos Santos Mendonca (Applied Sciences Group, Microsoft One Microsoft Way, WA)</i>	
HIVE: A Hyperbolic Interactive Visualization Explorer for Representation Learning	666
<i>Thijmen Nijdam (University of Amsterdam), Derck Prinzhorn (University of Amsterdam), Jurgen de Heus (University of Amsterdam), and Thomas Brouwer (University of Amsterdam)</i>	
Flatland and Beyond: Mutual Information Across Geometries	672
<i>Youssef Wally (UiT The Arctic University of Norway, Norway), Johan Mylius-Kroken (UiT The Arctic University of Norway, Norway), Michael Kampffmeyer (UiT The Arctic University of Norway, Norway ;Norwegian Computing Center, Norway), Rezvan Ehsani (University of Bergen, Norway), Vladan Milosevic (University of Bergen, Norway), and Elisabeth Wetzer (UiT The Arctic University of Norway, Norway)</i>	

HierVision: Standardized and Reproducible Hierarchical Sources for Vision Datasets	682
<i>Tejaswi Kasarla (University of Amsterdam, The Netherlands), Ruthu Hulikal Rooparaghunath (University of Amsterdam, The Netherlands), Stefano D'arrigo (Sapenzia University of Rome, Italy), Gowreesh Mago (University of Amsterdam, The Netherlands), Abhishek Jha (KU Leuven, Belgium), Melika Ayoughi (University of Amsterdam, The Netherlands), Swasti Shreya Mishra (University of Amsterdam, The Netherlands), Ana Manzano Rodríguez (University of Amsterdam), Teng Long (University of Amsterdam, The Netherlands), Mina Ghadimi Atigh (University of Amsterdam, The Netherlands), Max Van Spengler (University of Amsterdam, The Netherlands), and Pascal Mettes (University of Amsterdam, The Netherlands)</i>	
HAPPI: Hyperbolic Hierarchical Part Prototypes for Image Recognition	696
<i>Hooman Vaseli (The University of British Columbia, Canada), Victoria Wu (The University of British Columbia, Canada), Nima Kondori (The University of British Columbia, Canada), Nguyen Nhat Minh To (The University of British Columbia, Canada), Andrea Fung (The University of British Columbia, Canada), Ang Nan Gu (The University of British Columbia, Canada), and Purang Abolmaesumi (The University of British Columbia, Canada)</i>	
Sparse Hyperbolic Convolutional Networks with Enhanced Object Localization via GradCAM Analysis	706
<i>Vijayavallabh Jayamanikandan (Indian Institute of Technology Madras, India), Jithamanyu Settur (Indian Institute of Technology Madras, India), Lokesh Rajulapati (Indian Institute of Technology Madras, India), and Raghunathan Rengaswamy (Indian Institute of Technology Madras, India)</i>	

Explainable Computer Vision: Quo Vadis?

Do VLMs Have Bad Eyes? Diagnosing Compositional Failures via Mechanistic Interpretability	715
<i>Ashwath Vaithinathan Aravindan (University of Southern California), Abha Jha (University of Southern California), and Mihir Kulkarni (University of Southern California)</i>	
Explaining Object Detection Through Difference Map	724
<i>Shujun Xia (City University of Hong Kong), Chenyang Zhao (City University of Hong Kong), and Antoni Chan (City University of Hong Kong)</i>	
Interpretable Open-Vocabulary Referring Object Detection with Reverse Contrast Attention	734
<i>Drandreb Earl O. Juanico (University of the Philippines, Quezon City), Rowel O. Atienza (University of the Philippines, Quezon City; EEEL, University of the Philippines, Quezon City), and Jeffrey Kenneth Go (Samsung R&D Institute Philippines)</i>	
GFR-CAM: Gram-Schmidt Feature Reduction for Hierarchical Class Activation Maps	744
<i>Kaveh Safavigerdini (University of Missouri-Columbia, USA), Bahram Yaghooti (Washington University in St. Louis, USA), Amir Erfan Zareei Shams Abadi (University of Missouri-Columbia, USA), and Kannappan Palaniappan (University of Missouri-Columbia, USA)</i>	

Rethinking Explainer Trust: A Position on the Inconsistencies of Visual Explanations in Weakly Supervised Segmentation	754
<i>Ayush Somani (UiT The Arctic University of Norway) and Dilip K. Prasad (UiT The Arctic University of Norway)</i>	

2nd Workshop on the Challenge Of Out Of Label Hazards In Autonomous Driving

2COOOL: 2nd Workshop on the Challenge Of Out-Of-Label Hazards in Autonomous Driving	764
<i>Ali K. AlShami (University of Colorado Colorado Springs), Ryan Rabinowitz (University of Colorado Colorado Springs), Magèd Shoman (University of Tennessee), Jianwu Fang (Xi'an Jiaotong University), Lukas Picek (University of West Bohemia), Shao-Yuan Lo (Honda Research Institute USA), Steve Cruz (University of Notre Dame), Khang Nhut Lam (Can Tho University), Nachiket Kamod (University of Colorado Colorado Springs), Lei-Lei Li (Xi'an Jiaotong University), Jugal Kalita (University of Colorado Colorado Springs), and Terrance E. Boulton (University of Colorado Colorado Springs)</i>	
VRU-Accident: A Vision-Language Benchmark for Video Question Answering and Dense Captioning for Accident Scene Understanding	772
<i>Younggun Kim (University of Central Florida, United States), Ahmed S. Abdelrahman (University of Central Florida, United States), and Mohamed Abdel-Aty (University of Central Florida, United States)</i>	
Uncertainty-Aware Likelihood Ratio Estimation for Pixel-Wise Out-of-Distribution Detection.....	783
<i>Marc Hölle (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Walter Kellermann (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), and Vasileios Belagiannis (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)</i>	
Adapt, But Don't Forget: Fine-Tuning and Contrastive Routing for Lane Detection under Distribution Shift	794
<i>Mohammed Abdul Hafeez Khan (Florida Institute of Technology, USA), Parth Ganeriwala (Florida Institute of Technology, USA), Sarah M. Lehman (NASA Langley Research Center, USA), Siddhartha Bhattacharyya (Florida Institute of Technology, USA), Amy Alvarez (Florida Institute of Technology, USA), and Natasha Neogi (NASA Langley Research Center, USA)</i>	
Interpretable Decision-Making for End-to-End Autonomous Driving	805
<i>Mona Mirzaie (Leibniz University Hannover) and Bodo Rosenhahn (Leibniz University Hannover)</i>	
Fourier Domain Adaptation for Traffic Light Detection in Adverse Weather	816
<i>Ishaan Gakhar (Manipal Institute of Technology, India), Aryaman Gupta (Manipal Institute of Technology, India), Aryesh Guha (Dept. of Electrical and Electronics Engineering, MIT, MAHE, India), Amit Agarwal (Enterprise Analytics & Data Science Artificial Intelligence, Center of Excellence Wells Fargo International Solutions Private Limited), and Ujjwal Verma (Dept. of Electronics and Communication Engineering, MIT, MAHE, India)</i>	
DRAMA-X: A Fine-grained Intent Prediction and Risk Reasoning Benchmark For Driving	826
<i>Mihir Godbole (Texas A&M University), Xiangbo Gao (Texas A&M University), and Zhengzhong Tu (Texas A&M University)</i>	

SADWA: Fine-Grained Weather Awareness with Vision-Language Models for Seamless Autonomous Driving in Real Time	832
<i>Jinwoo Kim (Electronics and Telecommunications Research Institute (ETRI), South Korea; Korea Advanced Institute of Science and Technology (KAIST), South Korea), O. Hayeon (Electronics and Telecommunications Research Institute (ETRI), South Korea), Youngmin Oh (Electronics and Telecommunications Research Institute (ETRI), South Korea), Kyoungwan An (Electronics and Telecommunications Research Institute (ETRI), South Korea), and Donghwan Lee (Korea Advanced Institute of Science and Technology (KAIST), South Korea)</i>	
FlareGS: 4D Flare Removal using Gaussian Splatting for Urban Scenes	842
<i>Mayank Chandak (Indian Institute of Technology Madras, India), Sai Sri Teja Kuppa (Indian Institute of Technology Madras, India), Rahul . (Indian Institute of Technology, Madras, India), Gopi Raju Matta (Indian Institute of Technology Madras, India), Vinayak Gupta (Indian Institute of Technology Madras, India), and Kaushik Mitra (Indian Institute of Technology Madras, India)</i>	
Towards Vision Zero! The TUM Traffic Accid3nD Dataset	852
<i>Walter Zimmer (Technical University of Munich), Ross Greer (University of California Merced), Xingcheng Zhou (Technical University of Munich), Rui Song (Technical University of Munich; Fraunhofer IVI), Hu Cao (Technical University of Munich), Daniel Lehmborg (Technical University of Munich), Marc Pavel (Technical University of Munich), Ahmed Ghita (SETLabs Research GmbH), Akshay Gopalkrishnan (University of California, San Diego), Holger Caesar (Delft University of Technology), Mohan Trivedi (University of California, San Diego), and Alois C. Knoll (Technical University of Munich)</i>	
Simplifying Traffic Anomaly Detection with Video Foundation Models	863
<i>Svetlana Orlova (Eindhoven University of Technology), Tommie Kersters (Eindhoven University of Technology), Bruno B. Englert (Eindhoven University of Technology), and Gijs Dubbelman (Eindhoven University of Technology)</i>	
LaViPlan : Language-Guided Visual Path Planning with RLVR	874
<i>Hayeon Oh (Electronics and Telecommunications Research Institute)</i>	

Computer Vision for Automated Medical Diagnosis

Efficient Self-Supervised Adaptation for Medical Image Analysis	884
<i>Moein Sorkhei (KTH Royal Institute of Technology, Sweden; Science for Life Laboratory, Sweden), Emir Konuk (KTH Royal Institute of Technology, Sweden; Science for Life Laboratory, Sweden), Jingyu Guo (KTH Royal Institute of Technology, Sweden; Science for Life Laboratory, Sweden), Chanjuan Meng (KTH Royal Institute of Technology, Sweden; Science for Life Laboratory, Sweden), Christos Matsoukas (KTH Royal Institute of Technology, Sweden; Science for Life Laboratory, Sweden), and Kevin Smith (KTH Royal Institute of Technology, Sweden; Science for Life Laboratory, Sweden)</i>	

Comparison of Digital Histology AI Models with Low-Dimensional Genomic and Clinical Models in Survival Modeling for Prostate Cancer	892
<i>Aidan McLoughlin (UC Berkeley), Ho Yin Ho (Veracyte Inc.), Xin Zhao (Veracyte Inc.), Alexander Karl Hakansson (Veracyte Inc.), Alireza Moradi (Veracyte Inc.), Qi Joslove Xu (Veracyte Inc.), and Yang Liu (Veracyte Inc.)</i>	
GADA: Graph Attention-based Detection Aggregation for Ultrasound Video Classification	903
<i>Li Chen (Philips), Naveen Balaraju (Philips), Jochen Kruecker (Philips), Balasundar Raju (Philips), and Alvin Chen (Philips)</i>	
Latent Gene Diffusion for Spatial Transcriptomics Completion	912
<i>Paula Cárdenas (Universidad de los Andes, Colombia), Leonardo Manrique (Universidad de los Andes, Colombia), Daniela Vega (Universidad de los Andes, Colombia), Daniela Ruiz (Universidad de los Andes, Colombia), and Pablo Arbeláez (Universidad de los Andes, Colombia)</i>	
From Cosmos to Clinic: Interpretable Spatial Statistics for Histopathology Prognosis	922
<i>Attila Barna (Eötvös Loránd University, Hungary), Oz Kilim (Eötvös Loránd University, Hungary; Semmelweis University, Hungary), and Istvan Csabai (Eötvös Loránd University, Hungary)</i>	
MPromer: A Unified Diffusion-Based Framework for Scalable and Generalizable Multi-Modal Medical Image Segmentation	930
<i>Wafa Al Ghallabi (Mohamed bin Zayed University of Artificial Intelligence, UAE), Akshay Dudhane (Mohamed bin Zayed University of Artificial Intelligence, UAE), Syed Waqas Zamir (Inception Institute of Artificial Intelligence, UAE), Salman Khan (Mohamed bin Zayed University of Artificial Intelligence, UAE), and Fahad Shahbaz Khan (Mohamed bin Zayed University of Artificial Intelligence, UAE; Linköping University, Sweden)</i>	
Automated Assessment of Aesthetic Outcomes in Facial Plastic Surgery	939
<i>Pegah Varghaei (Michigan State University), Kiran Abraham-Aggarwal (Cornell University), Manoj T. Abraham (New York Medical College), and Arun Ross (Michigan State University)</i>	
Weakly Supervised Intracranial Aneurysm Detection and Segmentation in MR angiography via Multi-task UNet with Vesselness Prior	949
<i>Erin Rainville (Concordia University, Canada), Amirhossein Rasoulilian (NeuroRx Research, Canada), Hassan Rivaz (Concordia University, Canada), and Yiming Xiao (Concordia University, Canada)</i>	
TextSAM-EUS: Text Prompt Learning for SAM to Accurately Segment Pancreatic Tumor in Endoscopic Ultrasound	959
<i>Pascal Spiegler (Concordia University, Canada), Taha Koleilat (Concordia University, Canada), Arash Harirpoush (Concordia University, Canada), Corey S. Miller (Jewish General Hospital, Canada; McGill University Faculty of Medicine, Canada; Lady Davis Institute for Medical Research, Canada), Hassan Rivaz (Concordia University, Canada), Marta Kersten-Oertel (Concordia University, Canada), and Yiming Xiao (Concordia University, Canada)</i>	

UGPL: Uncertainty-Guided Progressive Learning for Evidence-Based Classification in Computed Tomography	969
<i>Shravan Venkatraman (Vellore Institute of Technology, India), S. Pavan Kumar (Vellore Institute of Technology, India), Rakesh Raj Madavan (Shiv Nadar University, India), and S. Chandrakala (Shiv Nadar University, India)</i>	
QPolypNet: A Quantum-Inspired Deep Learning Model for Polyp Segmentation	980
<i>Md Majedul Islam (Kennesaw State University), Rashik Shahriar Akash (Daffodil International University), Sayed Mehedi Azim (Rutgers University, USA), and Selena He (Kennesaw State University)</i>	
EchoNet-Quality: Denoising Echocardiograms via Deep Generative Modeling of Ultrasound Noise	990
<i>David Choi (Cedars-Sinai Medical Center, United States of America), Milos Vukadinovic (Cedars-Sinai Medical Center, United States of America; University of California Los Angeles, United States of America; Division of Research, Kaiser Permanente, United States of America), Bryan He (Stanford University, United States of America; Division of Research, Kaiser Permanente, United States of America), Christina Binder (Medical University of Vienna, Austria), Yuki Sahashi (Cedars-Sinai Medical Center, United States of America), and David Ouyang (Cedars-Sinai Medical Center, United States of America; Division of Research, Kaiser Permanente, United States of America)</i>	
Breast Cancer Detection with Topological Deep Learning	998
<i>Brighton Nuwagira (Dept. of Math. Sciences, UT Dallas), Adrian Rodriguez (Dept. of Computer Science, UT Dallas), Qiwei Li (Dept. of Math. Sciences, UT Dallas), and Baris Coskunuzer (Dept. of Math. Sciences, UT Dallas)</i>	
Endonama: Graph Neural Networks based In-Silico Transcriptomics from Histology Whole Slide Images for Fertility Diagnostics	1009
<i>George Wright (University of Warwick, UK), Paul Brighton (University of Warwick, UK), Joanne Muter (University of Warwick, UK), Jan Brosens (University Hospitals Coventry and Warwickshire National Health Service Trust, UK; University of Warwick, UK), and Fayyaz Minhas (University of Warwick, UK)</i>	
Artifact Correction in Panoramic Radiographs using Deep De-shadowing	1019
<i>Omri Dan (Tel-Aviv University, Israel), Samuel Lilek (Tel-Aviv University, Israel), Ariel Hirschhorn (Sheba Medical Center, Israel), Lazar Kats (Tel Aviv University, Israel), Nahum Kiryati (Tel Aviv University, Israel), and Arnaldo Mayer (Tel Aviv University, Israel)</i>	
Unsupervised Abnormality Segmentation in Chest CT with Anatomy-Guided Latent Diffusion Model and Adaptive Thresholding	1028
<i>Xueqi Guo (Siemens Healthineers, PA), Yoshihisa Shinagawa (Siemens Healthineers, PA), Sepehr Farhand (Siemens Healthineers, PA), Halid Yerebakan (Siemens Healthineers, PA), Kritika Iyer (Siemens Healthineers, PA), Matthias Wolf (Siemens Healthineers,PA), and Gerardo Hermosillo Valadez (Siemens Healthineers, PA)</i>	

Unsupervised Nuclei Segmentation by Improving Pseudo Labels from Segment Anything Model	1036
<i>Ryota Nakai (Meijo University, Japan) and Kazuhiro Hotta (Meijo University, Japan)</i>	
MedSAM-Guided Curriculum Learning for White Matter Tract Segmentation in Block Face Imaging of Fetal Brain	1045
<i>Athira Kalladayil Shibu (Indian Institute of Technology Madras (IITM), India), Sriprabha Ramanarayanan (Healthcare Technology Innovation Centre (HTIC), India), Vinoth Kanna (Healthcare Technology Innovation Centre (HTIC), India), Jaikishan Jayakumar (Sudha Gopalakrishnan Brain Centre (SGBC), India), Keerthi Ram (Healthcare Technology Innovation Centre (HTIC), India), and Mohanasankar Sivaprakasam (Indian Institute of Technology Madras (IITM), India; 2Healthcare Technology Innovation Centre (HTIC), India; Sudha Gopalakrishnan Brain Centre (SGBC), India)</i>	
MK-UNet: Multi-kernel Lightweight CNN for Medical Image Segmentation	1053
<i>Md Mostafijur Rahman (The University of Texas at Austin, Texas) and Radu Marculescu (The University of Texas at Austin, Texas)</i>	
Render2Seg: Landmark-Guided Patch-wise Segmentation for Robust 3D Dental Mesh Analysis ...	1063
<i>Sena Lee (Yonsei University Wonju College of Medicine), Guenhye Kim (University Wonju College of Medicine), YongKyu Jin (Diorco), and Sejung Yang (Yonsei University Wonju College of Medicine)</i>	
DCEtriformer: A Hybrid Attention Transformer for DCE-MRI Synthesis in Prostate Imaging	1073
<i>S Sadhana (Indian Institute of Technology, India), Sriprabha Ramanarayanan (Healthcare Technology Innovation Centre, India), Kishore Kumar (Indian Institute of Technology, India), Keerthi Ram (Healthcare Technology Innovation Centre, India), Harsh Agarwal (GE HealthCare, India), Ramesh Venkatesan (GE HealthCare, India), and Mohanasankar Sivaprakasam (Indian Institute of Technology, India; Healthcare Technology Innovation Centre, India)</i>	
Increasing the Classification Rates of the Trained Models using Invariant Dataset Augmentation	1083
<i>Piotr Milczarski (University of Lodz, Poland)</i>	
BAMPolyp: Bi-Axial Mamba Bottleneck for Gastrointestinal Polyp Segmentation	1093
<i>Md. Farhadul Islam (BRAC University, Bangladesh; United International University, Bangladesh), Tashik Ahmed (BRAC University, Bangladesh), Partho Chanda (Daffodil International University, Bangladesh), Joyanta Jyoti Mondal (University of Delaware, USA; United International University, Bangladesh), Meem Arafat Manab (Universidad de León, Spain; Brac University, Bangladesh), Sarah Zabeen (University of Kentucky, USA), and Jannatun Noor (United International University, Bangladesh; Brac University, Bangladesh)</i>	
Unsupervised Domain Adaptation via Content Alignment for Hippocampus Segmentation	1104
<i>Hoda Kalabizadeh (University of Oxford, United Kingdom), Ludovica Griffanti (University of Oxford, United Kingdom), Pak-Hei Yeung (Nanyang Technological University, Singapore), Ana I. L. Namburete (University of Oxford, United Kingdom), Nicola K. Dinsdale (University of Oxford, United Kingdom), and Konstantinos Kamnitsas (University of Oxford, United Kingdom)</i>	

Taming Modern Point Tracking for Speckle Tracking Echocardiography via Impartial Motion	1115
<i>Md Abulkalam Azad (Norwegian University of Science and Technology, Norway; St. Olavs Hospital), John Nyberg (Norwegian University of Science and Technology, Norway), Havard Dalen (Norwegian University of Science and Technology, Norway; St. Olavs Hospital, Norway), Bjørnar Grenne (Norwegian University of Science and Technology, Norway; St. Olavs Hospital, Norway), Lasse Lovstakken (Norwegian University of Science and Technology, Norway), and Andreas Østvik (Norwegian University of Science and Technology, Norway; St. Olavs Hospital; SINTEF Digital, Norway)</i>	
MedVisionLlama: Leveraging Pre-Trained Large Language Model Layers to Enhance Medical Image Segmentation	1125
<i>Gurucharan Marthi Krishna Kumar (McGill University), Aman Chadha (Amazon), Janine Mendola (McGill University), and Amir Shmuel (McGill University)</i>	
CytoDiff: AI-Driven Cytomorphology Image Synthesis for Medical Diagnostics	1136
<i>Jan Carreras Boada (Institute of AI for Health, Helmholtz Zentrum München - German Research Center for Environmental Health, Germany; Universitat Pompeu Fabra (ESCI-UPF), Spain), Rao Muhammad Umer (Institute of AI for Health, Helmholtz Zentrum München - German Research Center for Environmental Health, Germany), and Carsten Marr (Institute of AI for Health, Helmholtz Zentrum München - German Research Center for Environmental Health, Germany; Ludwig-Maximilian-University Hospital, Germany; DKTK, German Cancer Consortium, Germany.)</i>	
Impact of Black-Box Adversarial Attacks on Deep Neural Networks for Skin Imaging	1145
<i>Bartłomiej Moniak (AGH University of Krakow) and Joanna Jaworek-Korjakowska (AGH University of Krakow)</i>	
A Dynamic Agent Framework for Large Language Model Reasoning for Medical and Visual Question Answering	1154
<i>Ziyan Xiao (University of Hong Kong, Hong Kong), Ruiyang Zhang (University of Hong Kong, Hong Kong), Yushi Feng (University of Hong Kong, Hong Kong), Lingting Zhu (University of Hong Kong, Hong Kong), Liang Peng (University of Hong Kong, Hong Kong), and Lequan Yu (University of Hong Kong, Hong Kong)</i>	
Generative Counterfactual Augmentation for Bias Mitigation	1164
<i>Jason Uwaeze (Rice University), Pranav Kulkarni (University of Maryland), Vladimir Braverman (Johns Hopkins University), Michael A. Jacobs (UTHealth Houston), and Vishwa S. Parekh (UTHealth Houston)</i>	
CECT-Mamba: a Hierarchical Contrast-enhanced-aware Model for Pancreatic Tumor Subtyping from Multi-phase CECT	1172
<i>Zhifang Gong (Southeast University), Shuo Gao (Southeast University), Ben Zhao (Southeast University; Medical School of Southeast University), Yingjing Xu (Zhejiang University), Yijun Yang (The Hong Kong University of Science and Technology (Guangzhou)), Shenghong Ju (Southeast University), and Guangquan Zhou (Southeast University)</i>	
Dual-LVT: A Dual Attention Language-Vision Transformer for Tumor Segmentation	1183
<i>Shiman Zhang (Yeshiva University, USA), Songzhu Zheng (Morgan Stanley, USA), and Ming Ma (Yeshiva University, USA)</i>	

A Large-Scale Benchmark of Cross-Modal Learning for Histology and Gene Expression in Spatial Transcriptomics	1193
<i>Rushin H. Gindra (1Helmholtz Munich, Germany; Technical University Munich, Germany), Giovanni Palla (Helmholtz Munich, Germany), Mathias Nguyen (Helmholtz Munich, Germany; Technical University Munich, Germany), Sophia J. Wagner (Helmholtz Munich, Germany; 2Technical University Munich, Germany), Manuel Tran (Helmholtz Munich, Germany; Technical University Munich, Germany), Fabian J Theis (1Helmholtz Munich, Germany; Technical University Munich, Germany), Dieter Saur (Technical University Munich, Germany), Lorin Crawford (Microsoft Research, USA), and Tingying Peng (Helmholtz Munich, Germany; Technical University Munich, Germany)</i>	
CARDIUM: Congenital Anomaly Recognition with Diagnostic Images and Unified Medical records....	1204
<i>Daniela Vega (Universidad de Los Andes, Colombia), Hannah V. Ceballos (Universidad de Los Andes, Colombia), Javier S. Vera (Universidad de Los Andes, Colombia), Santiago Rodriguez (Universidad de Los Andes, Colombia), Alejandra Perez (Universidad de los Andes, Colombia), Angela Castillo (Universidad de Los Andes), Maria Escobar (Universidad de Los Andes, Colombia), Dario Londoño (Universidad de los Andes, Colombia; Fundacion Santa Fe de Bogot ´ a, Colombia), Luis A. Sarmiento (Universidad de los Andes, Colombia; Fundacion Santa Fe de Bogot ´ a, Colombia), Camila I. Castro (Universidad de Los Andes, Colombia), Nadiezhda Rodriguez (Universidad de los Andes, Colombia; Fundacion Santa Fe de Bogot ´ a, Colombia), Juan C. Briceño (Universidad de Los Andes, Colombia), and Pablo Arbelaez (Universidad de los Andes, Colombia)</i>	
Towards Robust Multimodal Representation: A Unified Approach with Adaptive Experts and Alignment	1214
<i>Nazanin Moradinab (University of Virginia, USA), Saurav Sengupta (University of Virginia, USA), Jiebei Liu (University of Virginia, USA), Sana Syed (Duke University, USA), and Donald E. Brown (University of Virginia, USA)</i>	
MV-MLM: Bridging Multi-View Mammography and Language for Breast Cancer Diagnosis and Risk Prediction	1224
<i>Shunjie-Fabian Zheng (LMU University Hospital, LMU Munich, Germany; Lunit Inc., South Korea), Hyeonjun Lee (Lunit Inc., South Korea), Thijs Kooi (Lunit Inc., South Korea), and Ali Diba (Lunit Inc., South Korea)</i>	
MedBLINK: Probing Basic Perception in Multimodal Language Models for Medicine	1234
<i>Mahtab Bigverdi (University of Washington), Wisdom O. Ikezogwo (University of Washington), Kevin Zhang (University of Washington), Hyewon Jeong (Massachusetts Institute of Technology), Mingyu Lu (University of Washington), Sungjae Cho (Seoul National University), Linda G. Shapiro (University of Washington), and Ranjay Krishna (University of Washington)</i>	
Knowledge-Driven Query Network with Adaptive Cross-View Attention for Structured Radiology Report Generation	1245
<i>Xuege Hou (Tsinghua University), Yali Li (Tsinghua University), and Shengjin Wang (Tsinghua University)</i>	

L ³ Seg: Lean Linear Layers for Language-Guided Vision Transformer in Medical Image Segmentation	1255
<i>Rahul Bhardwaj (Indian Institute of Technology Guwahati, India), Utkarsh Yashwant Tambe (SRM Institute of Science & Technology, India), and Debanga Raj Neog (Indian Institute of Technology Guwahati, India)</i>	
PMC-Vid: A Large-Scale Biomedical Video Captioning Dataset	1265
<i>Yosuke Yamagishi (The University of Tokyo, Japan), Kuniaki Saito (OMRON SINIC X Corporation), Atsushi Hashimoto (OMRON SINIC X Corporation), and Yoshita Ushiku (OMRON SINIC X Corporation)</i>	
Trustworthy Medical Imaging with Large Language Models: A Study of Hallucinations Across Modalities	1276
<i>Anindya Bijoy Das (The University of Akron), Shahnewaz Karim Sakib (University of Tennessee at Chattanooga), and Shibbir Ahmed (Texas State University)</i>	
Multiplicative Loss for Enhancing Semantic Segmentation in Medical and Cellular Images	1284
<i>Yuto Yokoi (Meijo University, Japan) and Kazuhiro Hotta (Meijo University, Japan)</i>	
Enhancing the Reliability of Auto-Prompting SAM for Medical Image Segmentation with Uncertainty Estimation and Rectification	1293
<i>Yichi Zhang (Fudan University), Shiyao Hu (Xi'an Jiaotong University), Le Xue (Fudan University; Shanghai Academy of Artificial Intelligence for Science), Sijie Ren (Fudan University), Zixin Hu (Fudan University; Shanghai Academy of Artificial Intelligence for Science), Yuan Cheng (Fudan University; Shanghai Academy of Artificial Intelligence for Science), and Yuan Qi (Fudan University; Shanghai Academy of Artificial Intelligence for Science)</i>	
SurGen-Net: A Generative Approach for Surgical VQA with Structured Text Generation	1303
<i>Yongjun Jeon (Sungkyunkwan University), Seonmin Park (Samsung Medical Center), Jongmin Shin (Samsung Medical Center), Kanggil Park (Samsung Medical Center), Bogeun Kim (Samsung Medical Center), Namkee Oh (Samsung Medical Center), and Kyu-Hwan Jung (Sungkyunkwan University; Samsung Medical Center)</i>	
Not Only Grey Matter: OmniBrain for Robust Multimodal Classification of Alzheimer's Disease	1311
<i>Ahmed Sharshar (Mohamed bin Zayed University of Artificial Intelligence, UAE), Yasser Ashraf (Mohamed bin Zayed University of Artificial Intelligence, UAE), Tameem Bakr (Mohamed bin Zayed University of Artificial Intelligence, UAE), Salma Hassan (Mohamed bin Zayed University of Artificial Intelligence, UAE), and Hosam Elgendy (Mohamed bin Zayed University of Artificial Intelligence, UAE)</i>	
Building a General SimCLR Self-Supervised Foundation Model Across Neurological Diseases to Advance 3D Brain MRI Diagnoses	1321
<i>Emily Kaczmarek (McGill University, Canada; Mila - Quebec Artificial Intelligence Institute, Canada), Justin Szeto (McGill University, Canada; Mila - Quebec Artificial Intelligence Institute, Canada), Brennan Nichyporuk (McGill University, Canada; Mila - Quebec Artificial Intelligence Institute, Canada), and Tal Arbel (McGill University, Canada; Mila - Quebec Artificial Intelligence Institute, Canada)</i>	

Prompt4Trust: A Reinforcement Learning Prompt Augmentation Framework for Clinically-Aligned Confidence Calibration in Multimodal Large Language Models	1331
<i>Anita Kriz (McGill University, Canada), Elizabeth Laura Janes (McGill University, Canada), Xing Shen (McGill University, Canada; Mila - Quebec Artificial Intelligence Institute, Canada), and Tal Arbel (McGill University, Canada; Mila - Quebec Artificial Intelligence Institute, Canada)</i>	
Diffusion-Based Data Augmentation for Medical Image Segmentation	1341
<i>Maham Nazir (Beihang University, China), Muhammad Aqeel (University of Verona, Italy), and Francesco Setti (University of Verona, Italy)</i>	
Autoregressive-Conditioned Diffusion for Semi-Supervised Thyroid Ultrasound Segmentation with Optical Flow-Based Pseudo Labels	1351
<i>Xiaorui Liu (The Hong Kong University of Science and Technology), Yijun Yang (The Hong Kong University of Science and Technology), Yingjing Xu (Zhejiang University), and Lei Zhu (The Hong Kong University of Science and Technology)</i>	

3rd Workshop on Vision-based Industrial InspectiON

PCB-SAID: A Low-Cost Camera-Based Dataset for Few-Shot SMD Assembly Inspection	1362
<i>Raffaele Mineo (Deepsensing SRL, Italy; University of Catania, Italy), Amelia Sorrenti (Deepsensing SRL, Italy; University of Catania, Italy), Robin Faro (Deepsensing SRL, Italy), Gabriele Mineo (Deepsensing SRL, Italy), Francesco Cancelliere (Deepsensing SRL, Italy; University of Catania, Italy), and Alberto Faro (Deepsensing SRL, Italy)</i>	
Development and Implementation of a Digital Twin for Power Plants: Enhancing Maintenance and Operational Efficiency	1369
<i>Guneet Bhatia (Siemens Energy, USA), Nagarjun Nagarjun (Siemens Energy, USA), and Jason Voelker (Siemens Energy, USA)</i>	
Deep Learning-based Rail Surface Condition Evaluation	1377
<i>Shilin Hu (Stony Brook University, USA), Ke Ma (Snap Inc.), Sagnik Das (Stony Brook University, USA), Dichang Zhang (Stony Brook University, USA), and Dimitris Samaras (Stony Brook University, USA)</i>	
Two-stage Text-Guided Diffusion Models for Disentangled Industrial Defect Generation	1387
<i>Jing Wei (Chinese Academy of Sciences; University of Chinese Academy of Sciences), Qingfeng Shi (Chinese Academy of Sciences; University of Chinese Academy of Sciences), Fei Shen (Chinese Academy of Sciences; University of Chinese Academy of Sciences; CASI Vision Technology CO), and Zhengtao Zhang (Chinese Academy of Sciences; University of Chinese Academy of Sciences; CASI Vision Technology CO)</i>	
SupConWI-RL: wafer inspection with reinforcement learning enhanced by supervised contrastive learning	1396
<i>Aleksandr Dekhovich (Delft University of Technology) and Oleg Soloviev (Delft University of Technology; Flexible Optical B.V.)</i>	

Towards Automated Assembly Quality Inspection with Synthetic Data and Domain Randomization...	1406
<i>Xiaomeng Zhu (Scania CV AB and KTH Royal Institute of Technology, Sweden), Jacob Henningsson (Scania CV AB, Sweden), Duruo Li (Scania CV AB, Sweden), Pär Mårtensson (Scania CV AB, Sweden), Lars Hanson (Skövde University, Sweden), Mårten Björkman (KTH Royal Institute of Technology, Sweden), and Atsuto Maki (KTH Royal Institute of Technology, Sweden)</i>	
WS2: Weakly Supervised Segmentation using Before-After Supervision in Waste Sorting	1415
<i>Andrea Marelli (Politecnico di Milano), Alberto Foresti (EURECOM, France), Leonardo Pesce (Politecnico di Milano, Italy), Giacomo Boracchi (Politecnico di Milano, Italy), and Mario Grosso (Politecnico di Milano, Italy)</i>	
SynSpill: Improved Industrial Spill Detection With Synthetic Data	1425
<i>Aaditya Baranwal (University of Central Florida, USA), Abdul Mueez (University of Central Florida, USA), Jason Voelker (Siemens Energy, USA), Guneet Bhatia (Siemens Energy, USA), and Shruti Vyas (University of Central Florida, USA)</i>	
InspectVLM: Unified in Theory, Unreliable in Practice	1435
<i>Conor Wallace (Zeitview), Isaac Corley (Zeitview), and Jonathan Lwowski (Zeitview)</i>	
iSafetyBench: A video-language benchmark for safety in industrial environment	1444
<i>Raiyaan Abdullah (University of Central Florida), Yogesh Singh Rawat (University of Central Florida), and Shruti Vyas (University of Central Florida)</i>	
Robust Anomaly Detection in Industrial Environments via Meta-Learning	1454
<i>Muhammad Aqeel (University of Verona, Italy), Shakiba Sharifi (University of Verona, Italy), Marco Cristani (University of Verona, Italy; Qualyco S.r.l, Italy), and Francesco Setti (University of Verona, Italy; Qualyco S.r.l, Italy)</i>	
A Contrastive Learning-Guided Confident Meta-learning for Zero Shot Anomaly Detection	1463
<i>Muhammad Aqeel (University of Verona, Italy), Danijel Skočaj (University of Ljubljana, Slovenia), Marco Cristani (University of Verona, Italy; Qualyco S.r.l, Italy), and Francesco Setti (University of Verona, Italy; Qualyco S.r.l, Italy)</i>	
RDDPM: Robust Denoising Diffusion Probabilistic Model for Unsupervised Anomaly Segmentation	1473
<i>Mehrdad Moradi (Georgia Institute of Technology, Georgia) and Kamran Paynabar (Georgia Institute of Technology, Georgia)</i>	
UniLEAD: A Unified and LightwEight model for Anomaly Detection	1483
<i>Shih-Chih Lin (International Intercollegiate Ph.D. Program) and Shang-Hong Lai (National Tsing Hua University, Taiwan)</i>	

Artificial Social Intelligence Workshop

RealTalk: Realistic Emotion-Aware Lifelike Talking-Head Synthesis	1492
<i>Wenqing Wang (Northeastern University, Boston) and Yun Fu (Northeastern University, Boston)</i>	

Toward Socially Aware Vision-Language Models: Evaluating Cultural Competence Through Multimodal Story Generation	1502
<i>Arka Mukherjee (KIIT Deemed University, India) and Shreya Ghosh (IIT Bhubaneswar, India)</i>	

Authenticity & Provenance in the age of Generative AI

SHIELD: A Secure and Highly Enhanced Integrated Learning for Robust Deepfake Detection against Adversarial Attacks	1513
<i>Kutub Uddin (University of Michigan-Flint, USA), Awaiz Khan (University of Michigan-Flint, USA), Muhammad Umar Farooq (University of Michigan-Flint, USA), and Khalid Mahmood Malik (University of Michigan-Flint, USA)</i>	
CustomMark: Customization of Diffusion Models for Proactive Attribution	1523
<i>Vishal Asnani (Adobe Research ;Michigan State University), John Collomosse (Adobe Research ;University of Surrey), Xiaoming Liu (Michigan State University), and Shruti Agarwal (Adobe Research)</i>	
MultiNeRF: Multiple Watermark Embedding for Neural Radiance Fields	1534
<i>Yash Kulthe (University of Surrey, UK), Andrew Gilbert (University of Surrey, UK), and John Collomosse (University of Surrey, UK ;Adobe Research)</i>	
FreqPure: a High-frequency Preservation Diffusion-based Purification Method for Protective Perturbation Removal	1544
<i>Yan Ju (University at Buffalo, State University of New York), Hongfei Xue (University at North Carolina at Charlotte), and Siwei Lyu (University at Buffalo, State University of New York)</i>	
Adaptive Test-Time Semantic Debiasing for AI-Generated Image Detection	1554
<i>Yu Cai (State University of New York at Buffalo, USA), Jiahe Tian (Chinese Academy of Sciences, China), Xiaomeng Fu (Chinese Academy of Sciences, China), Jiao Dai (Chinese Academy of Sciences, China), Jizhong Han (Chinese Academy of Sciences, China), and Siwei Lyu (State University of New York at Buffalo, USA)</i>	
CatAID: Category-Guided AI-Generated Image Detection via Vision-Language Model Adaptation	1564
<i>Yu Cai (State University of New York at Buffalo, USA), Shan Jia (State University of New York at Buffalo, USA), Jiahe Tian (Institute of Information Engineering, Chinese Academy of Sciences, China), Jiao Dai (Institute of Information Engineering, Chinese Academy of Sciences, China), Jizhong Han (Institute of Information Engineering, Chinese Academy of Sciences, China), and Siwei Lyu (State University of New York at Buffalo, USA)</i>	
Is JPEG AI going to change image forensics?	1575
<i>Edoardo Daniele Cannas (Politecnico di Milano), Sara Mandelli (Politecnico di Milano), Nataša Popović (Politecnico di Milano), Ayman Alkhateeb (Universita degli Studi di Brescia), Alessandro Gnutti (Universita degli Studi di Brescia), Paolo Bestagini (Politecnico di Milano), and Stefano Tubaro (Politecnico di Milano)</i>	

Explainable AI-Generated Image Forensics: A Low-Resolution Perspective with Novel Artifact Taxonomy	1587
<i>Kaustubh Sharma (Indian Institute of Technology Roorkee, India)</i>	
Phoneme-Level Analysis for Person-of-Interest Speech Deepfake Detection	1597
<i>Davide Salvi (Politecnico di Milano, Italy), Viola Negroni (Politecnico di Milano, Italy), Sara Mandelli (Politecnico di Milano, Italy), Paolo Bestagini (Politecnico di Milano, Italy), and Stefano Tubaro (Politecnico di Milano, Italy)</i>	
PIA: Deepfake Detection Using Phoneme-Temporal and Identity-Dynamic Analysis	1607
<i>Soumya Kanti Datta (University at Buffalo, USA), Tanvi Ranga (University at Buffalo, USA), Chengzhe Sun (University at Buffalo, USA), and Siwei Lyu (University at Buffalo, USA)</i>	
Adversarial Reality for Evading Deepfake Image Detectors	1618
<i>Umur Aybars Çiftçi (Binghamton University), Nicholas Solar (Microsoft), Emily Greene (Binghamton University), Sophie Riley Saremsky (Brigham and Women's Hospital), and Ilke Demir (Cauth AI)</i>	
Witness Sensing for Verifying the Human Origin of Digital Media	1630
<i>Isabella Lenz (Arizona State University), Yu Rong (Arizona State University), Daniel Bliss (Arizona State University), Julie Liss (Arizona State University), and Visar Berisha (Arizona State University)</i>	
The Effect of Semantically Aligned Images for Deepfake Detection	1640
<i>Aritra Bandyopadhyay (University of Amsterdam), Tyme Chatupanyachotikul (University of Amsterdam), Jose Garcia (University of Amsterdam), Arijit Ghosh (University of Amsterdam), Karolína Hájková (University of Amsterdam), Carlos Patiño (University of Amsterdam), and Ivona Najdenkoska (University of Amsterdam)</i>	

Computer Vision for Developing Countries

WavePaint: Resource-efficient Token-mixer for Self-supervised Inpainting	1650
<i>Pranav Jeevan (Indian Institute of Technology Bombay, India), Dharshan Sampath Kumar (Indian Institute of Technology Bombay, India), and Amit Sethi (Indian Institute of Technology Bombay, India)</i>	
FLD+: Data-efficient Evaluation Metric for Generative Models	1660
<i>Pranav Jeevan (Indian Institute of Technology Bombay, India), Neeraj Nixon (Indian Institute of Technology Bombay, India), and Amit Sethi (Indian Institute of Technology Bombay, India)</i>	
V-RoAst: Visual Road Assessment Can VLM be a Road Safety Assessor Using the iRAP Standard?.....	1669
<i>Natchapon Jongwiriyanurak (University College London, UK), Zichao Zeng (University College London, UK), June Moh Goo (University College London, UK), Xinglei Wang (University College London, UK), Ilya Ilyankou (University College London, UK), Kerkrutt Sriroongvickrai (Chulalongkorn University, Thailand), Nicola Christie (University College London, UK), Meihui Wang (University College London, UK), Huanfa Chen (University College London, UK), and James Haworth (University College London, UK)</i>	

Simulating Refractive Distortions and Weather-Induced Artifacts for Resource-Constrained Autonomous Perception	1679
<i>Moseli Mots'oehli (University of Hawai'i at Mānoa ;MindForge AI, South Africa), Feimei Chen (University of Hawai'i at Mānoa), Hok Wai Chan (University of Hawai'i at Mānoa), Itumeleng Tlali (MindForge AI, South Africa), Thulani Babeli (MindForge AI, South Africa), Kyungim Baek (University of Hawai'i at Mānoa), and Huaijin Chen (University of Hawai'i at Mānoa)</i>	
Nayana: A Foundation for Document-Centric Vision-Language Models via Multi-Task, Multimodal, and Multilingual Data Synthesis	1689
<i>Adithya S Kolavi (CognitiveLab, India), P Samarth (CognitiveLab, India), and Vyoman Jain (CognitiveLab, India)</i>	

Embedded Vision Workshop

LightPrune: Latency-Aware Structured Pruning for Efficient Deep Inference on Embedded Devices	1699
<i>Asma Belhadi (OsloMet University, Norway), Youcef Djenouri (NORCE, Norway), and Ahmed Nabil Belbachir (NORCE, Norway)</i>	
Rethinking Backbone Design for Lightweight 3D Object Detection in LiDAR	1709
<i>Adwait Chandorkar (University of Wuppertal, Germany), Hasan Tercan (University of Wuppertal, Germany), and Tobias Meisen (University of Wuppertal, Germany)</i>	
Contextual Convolutions for Scalable Forward-Only Learning on Tiny Devices	1718
<i>Mehdi Abbassi (Fondazione Bruno Kessler), Alberto Ancilotto (Fondazione Bruno Kessler), and Elsabetta Farella (Fondazione Bruno Kessler)</i>	
Implementation of Extremely Low Power Visual Perception Algorithm on Programmable Vision Accelerator: Examples of Challenges and Solutions	1726
<i>Sanmati Kamath (Nvidia Corporation), Ching Hung (Nvidia Corporation), Jagadeesh Sankaran (Nvidia Corporation), Zoran Nikolic (Nvidia Corporation), Eric Viscito (Nvidia Corporation), and Branislav Kisanin (NVIDIA Corp, Institute for AI R&D of Serbia, and Faculty of Technical Sciences, University of Novi Sad)</i>	
Scalable Optical Convolutional Neural Networks For Edge Applications	1735
<i>Venkata Anirudh Puligandla (Institut Ruder Boskovic, Croatia), Vladimir Ceperic (University of Zagreb, Croatia), and Tihomir Knezevic (Institut Ruder Boskovic, Croatia)</i>	
Inside Knowledge: Graph-based Path Generation with Explainable Data Augmentation and Curriculum Learning for Visual Indoor Navigation	1745
<i>Daniel Airinei (National University of Science and Technology POLITEHNICA Bucharest, Romania), Elena Burceanu (Bitdefender), and Marius Leordeanu (National University of Science and Technology POLITEHNICA Bucharest, Romania; Bitdefender; Institute of Mathematics "Simion Stoilow" of the Romanian Academy, Romania)</i>	

Efficient Self-Supervised Neuro-Analytic Visual Servoing for Real-time Quadrotor Control 1755
Sebastian Mocanu (National University of Science and Technology POLITEHNICA Bucharest, Romania), Sebastian-Ion Nae (National University of Science and Technology POLITEHNICA Bucharest, Romania), Mihai-Eugen Barbu (National University of Science and Technology POLITEHNICA Bucharest, Romania), and Marius Leordeanu (National University of Science and Technology POLITEHNICA Bucharest, Romania; Institute of Mathematics "Simion Stoilov" of the Romanian Academy, Romania; NORCE Norwegian Research Center, Norway)

Foundation Models for V2X-Based Cooperative Autonomous Driving

Contextual-Personalized Adaptive Cruise Control via Fine-Tuned Large Language Models 1765
Ziye Qin (Southwest Jiaotong University, China), Xue Yao (Delft University of Technology, The Netherlands), Chuheng Wei (University of California at Riverside, USA), Ang Ji (Southwest Jiaotong University, China), Guoyuan Wu (University of California at Riverside, USA), and Zhanbo Sun (Southwest Jiaotong University, China)

RG-Attn: Radian Glue Attention for Multi-modal Multi-agent Cooperative Perception 1774
Lantao Li (Sony (China) Limited), Kang Yang (Renmin University of China ;Sony (China) Limited), Wenqi Zhang (Sony (China) Limited), Xiaoxue Wang (Sony (China) Limited), and Chen Sun (Sony (China) Limited)

D3FNet: A Differential Attention Fusion Network for Fine-Grained Road Structure Extraction in Remote Perception Systems 1784
Chang Liu (Budapest University of Technology and Economics, Hungary), Yan Xu (Budapest University of Technology and Economics, Hungary), and Tamas Sziranyi (Budapest University of Technology and Economics, Hungary)

Understanding What Vision-Language Models See in Traffic: PixelSHAP for Object-Level Attribution in Autonomous Driving 1794
Roni Goldshmidt (Nexar, Tel Aviv)

SlimComm: Doppler-Guided Sparse Queries for Bandwidth-Efficient Cooperative 3-D Perception..... 1803
Melih Yazgan (FZI Research Center for Information Technology), Qiyuan Wu (FZI Research Center for Information Technology ;Karlsruhe Institute of Technology), Iramm Hamdard (FZI Research Center for Information Technology ;Karlsruhe Institute of Technology), Shiqi Li (FZI Research Center for Information Technology ;Karlsruhe Institute of Technology), and J. Marius Zoellner (FZI Research Center for Information Technology ;Karlsruhe Institute of Technology)

Robust Scenario Mining Assisted by Multimodal Semantics 1813
Yifei Chen (Xi'an University of Technology) and Ross Greer (University of California, Merced)

MAP: End-to-End Autonomous Driving with Map-Assisted Planning 1823
Huilin Yin (Tongji University), Yiming Kan (Tongji University), and Daniel Watzenig (Graz University of Technology)

SpaRC-AD: A Baseline for Radar-Camera Fusion in End-to-End Autonomous Driving	1831
<i>Philipp Wolters (Technical University of Munich), Johannes Gilg (Technical University of Munich), Torben Teepe (Technical University of Munich), and Gerhard Rigoll (Technical University of Munich)</i>	
Improving Event-Phase Captions in Multi-View Urban Traffic Videos via Prompt-Aware LoRA Tuning of Vision Language Models	1842
<i>Ricardo Ornelas (University of California, Merced), Alton Chao (University of California, Merced), Shyam Gupta (University of California, Merced), Edmund Chao (University of California, Merced), and Ross Greer (University of California, Merced)</i>	
Research Challenges and Progress in the End-to-End V2X Cooperative Autonomous Driving Competition	1849
<i>Ruiyang Hao (Tsinghua University), Haibao Yu (Hong Kong University ;Tsinghua University), Jiaru Zhong (Tsinghua University), Chuanye Wang (Tsinghua University), Jiahao Wang (Tsinghua University), Yiming Kan (Tongji University), Wenxian Yang (Tsinghua University), Siqi Fan (Tsinghua University), Huilin Yin (Tongji University), Jianing Qiu (Chinese University of Hong Kong), Yao Mu (Hong Kong University ;Shanghai Jiao Tong University), Jiankai Sun (Stanford University), Li Chen (Hong Kong University ;OpenDriveLab), Walter Zimmer (Technical University of Munich), Dandan Zhang (Imperial College London), Shanghang Zhang (Peking University), Mac Schwager (Stanford University), Ping Luo (Hong Kong University), and Zaiqing Nie (Tsinghua University)</i>	

Human-Interactive Generation and Editing

Controllable Pedestrian Video Editing for Multi-View Driving Scenarios via Motion Sequence.....	1861
<i>Danzhen Fu (Xiaomi Inc.), Jiagao Hu (Xiaomi Inc.), Daiguo Zhou (Xiaomi Inc.), Fei Wang (Xiaomi Inc.), Zepeng Wang (Xiaomi Inc.), and Wenhua Liao (Xiaomi Inc.)</i>	
SEED-Story: Multimodal Long Story Generation with Large Language Model	1871
<i>Shuai Yang (Hong Kong University of Science and Technology (Guangzhou)), Yuying Ge (Tencent ARC Lab), Yang Li (Hong Kong University of Science and Technology (Guangzhou)), Yukang Chen (CUHK), Yixiao Ge (ARC Lab, Tencent PCG), Ying Shan (Tencent AI Lab Center of Visual Computing), and YingCong Chen (Hong Kong University of Science and Technology (Guangzhou))</i>	
UniPaint: Unified Space-time Video Inpainting via Mixture-of-Experts	1882
<i>Zhen Wan (Fudan University), Chenyang Qi (Hong Kong University of Science and Technology), Zhiheng Liu (University of Hong Kong), Tao Gui (Fudan University), and Yue Ma (The Hong Kong University of Science and Technology)</i>	

Sketch-to-Layout: Sketch-Guided Multimodal Layout Generation	1893
<i>Riccardo Brioschi (EPFL - EPF Lausanne, Switzerland), Aleksandr Alekseev (EPFL - EPF Lausanne, Switzerland), Emanuele Nevali (EPFL - EPF Lausanne, Switzerland), Berkay Döner (EPFL - EPF Lausanne, Switzerland), Omar El Malki (EPFL - EPF Lausanne, Switzerland), Blagoj Mitrevski (Google DeepMind, Switzerland), Leandro Kieliger (Google DeepMind, Switzerland), Mark Collier (Google DeepMind, Switzerland), Andrii Maksai (Google DeepMind, Switzerland), Jesse Berent (Google DeepMind, Switzerland), Claudiu Cristian Musat (Google DeepMind, Switzerland), and Efi Kokiopoulou (Google DeepMind, Switzerland)</i>	
Null text-guided interactive image editing for diffusion models	1906
<i>Jing Wang (Alibaba Group) and Hao Luo (Alibaba Group)</i>	
DreamHOI: Subject-Driven Generation of 3D Human-Object Interactions with Diffusion Priors ...	1916
<i>Hanwen Zhu (University of Oxford; Carnegie Mellon University), Ruining Li (University of Oxford), and Tomas Jakab (University of Oxford)</i>	
Concat-ID: Towards Universal Identity-Preserving Video Synthesis	1927
<i>Yong Zhong (Renmin University of China), Zhuoyi Yang (Tsinghua University), Jiayan Teng (Tsinghua University), Xiaotao Gu (Zhipu AI), and Chongxuan Li (Renmin University of China)</i>	
Beyond Flat Text: Dual Self-inherited Guidance for Visual Text Generation	1937
<i>Minxing Luo (Nankai University), Zixun Xia (Nankai University), Liaojun Chen (Nankai University), Zhenhang Li (Institute of Information Engineering, Chinese Academy of Sciences), Weichao Zeng (Institute of Information Engineering, Chinese Academy of Sciences), Jianye Wang (Nankai University), Wentao Cheng (Nankai University), Yaxing Wang (Nankai University), Yu Zhou (Nankai University), and Jian Yang (Nankai University)</i>	
DepthDance: Complex-pose Human Image Animation with Appearance-agnostic Depth Guidance	1947
<i>Yingjie Xi (Bournemouth University, UK), Zhengze Xu (Alibaba, China), Zhao Wang (Alibaba, China), Xiaosong Yang (Bournemouth University, UK), Jinsong Lan (Alibaba, China), Jian Jun Zhang (Bournemouth University, UK), and Mengting Chen (Alibaba, China)</i>	
Model as a Game: On Numerical and Spatial Consistency for Generative Games	1958
<i>Jingye Chen (Hong Kong University of Science and Technology), Yuzhong Zhao (University of the Chinese Academy of Sciences), Yupan Huang (Microsoft Research), Lei Cui (Microsoft Research Asia), Li Dong (Microsoft Research), Tengchao Lv (Microsoft Research), Qifeng Chen (Hong Kong University of Science and Technology), and Furu Wei (Microsoft Research)</i>	
StyleBooth: Image Style Editing with Multimodal Instruction	1968
<i>Zhen Han (Alibaba Group), Chaojie Mao (Alibaba Group), Zeyinzi Jiang (Alibaba Group), Yulin Pan (Alibaba Group), and Jingfeng Zhang (Alibaba Group)</i>	
ACE++: Instruction-Based Image Creation and Editing via Context-Aware Content Filling	1979
<i>Chaojie Mao (Tongyi Lab), Jingfeng Zhang (Tongyi Lab), Yulin Pan (Tongyi Lab), Zeyinzi Jiang (Tongyi Lab), Zhen Han (Tongyi Lab), Yu Liu (Tongyi Lab), and Jingren Zhou (Tongyi Lab)</i>	

2D Instance Editing in 3D Space	1988
<i>Yuhuan Xie (University of Hong Kong), Aoxuan Pan (Independent), Ming-Xian Lin (University of Hong Kong), Wei Huang (University of Hong Kong), Yi-Hua Huang (University of Hong Kong), and Xiaojuan Qi (University of Hong Kong)</i>	
Axes-and-Tags: LLM-Driven Design Galleries for Generative Content	1996
<i>Asanshay Gupta (Stanford University), Vishnu Sarukkai (Stanford University), and Kayvon Fatahalian (Stanford University)</i>	
MFT-VITON: High-Fidelity Virtual Try-On with Minimal Input via a Mask-Free Transformer-Diffusion Model	2006
<i>Zhenchen Wan (University of Melbourne ;Mohamed bin Zayed University of AI), Yanwu Xu (Google), Dongting Hu (University of Melbourne), Weilun Cheng (University of Melbourne), Tianxi Chen (University of Melbourne), Zhaoqing Wang (The University of Sydney), Feng Liu (University of Melbourne), Tongliang Liu (The University of Sydney ;Mohamed bin Zayed University of AI), and Mingming Gong (University of Melbourne ;Mohamed bin Zayed University of AI)</i>	
Enhancing Identity Preservation in Portrait Generation via Reward Optimization	2016
<i>Yang Liu (King’s College London), Hongyu Zang (Beijing Institute of Technology), Chao Xu (Sany Group), Baigui Sun (Sany Group), and Shan Luo (King’s College London)</i>	
GenEscape: Hierarchical Multi-Agent Generation of Escape Room Puzzles	2026
<i>Mengyi Shan (University of Washington, USA), Brian Curless (University of Washington, USA), Ira Kemelmacher-Shlizerman (University of Washington, USA), and Steve Seitz (University of Washington, USA)</i>	

Interactive Human-Centric Foundation Models

Superior and pragmatic talking face generation with teacher-student framework	2035
<i>Chao Liang (ByteDance), Jianwen Jiang (ByteDance), Tianyun Zhong (ByteDance), Gaojie Lin (ByteDance), Zhengkun Rong (ByteDance), Yongming Zhu (ByteDance), Jiaqi Yang (ByteDance), and Xin Chen (ByteDance)</i>	
Causal Motion Tokenizer for Streaming Motion Generation	2045
<i>Biao Jiang (Fudan University; 2StepFun), Xin Chen (ByteDance), Ailing Zeng (Independent Researcher), Xinru Sun (Fudan University), Fukun Yin (Fudan University), Xianfang Zeng (StepFun), Xuanyang Zhang (StepFun), Gang Yu (StepFun), and Tao Chen (Fudan University)</i>	
ID-Consistent, Precise Expression Generation with Blendshape-Guided Diffusion	2056
<i>Foivos Paraperas Papantoniou (Imperial College London, UK) and Stefanos Zafeiriou (Imperial College London, UK)</i>	

Computer Vision for Analysis of Underwater Imagery & Automated Analysis of Marine Visual Data for Environmental Monitoring

Automated Detection of Antarctic Benthic Organisms in High-Resolution In Situ Imagery to Aid Biodiversity Monitoring	2066
<i>Cameron Trotter (British Antarctic Survey), Huw Griffiths (British Antarctic Survey), Tasnuva Ming Khan (British Antarctic Survey; University of Cambridge), and Rowan Whittle (British Antarctic Survey)</i>	
The point is the mask: scaling coral reef segmentation with weak supervision	2077
<i>Matteo Contini (IFREMER), Victor Illien (IFREMER), Sylvain Poulain (IRD), Serge Bernard (CNRS), Julien Barde (IRD), Sylvain Bonhommeau (IFREMER), and Alexis Joly (INRIA)</i>	
A Calibration Tool for Refractive Underwater Vision	2087
<i>Felix Seegräber (Kiel University), Mengkun She (Kiel University), Felix Woelk (University of Applied Sciences Kiel), and Kevin Köser (Kiel University)</i>	
Sea-ing Through Scattered Rays: Revisiting the Image Formation Model for Realistic Underwater Image Generation	2096
<i>Vasiliki Ismiroglou (Aalborg University, Denmark), Malte Pedersen (Aalborg University, Denmark), Stefan H. Bengtson (Aalborg University, Denmark), Andreas Aakerberg (Aalborg University, Denmark), and Thomas B. Moeslund (Aalborg University, Denmark)</i>	
Uncovering Anomalous Events for Marine Environmental Monitoring via Visual Anomaly Detection	2106
<i>Laura Wehl (IT University of Copenhagen, Denmark), Nejc Novak (Anemo Robotics ApS, Denmark), Stefan H. Bengtson (Aalborg University, Denmark), and Malte Pedersen (Aalborg University, Denmark)</i>	
Refining Naive Annotations with Limited Expert Guidance for Semantic Segmentation: A Case Study on Underwater Echograms	2116
<i>Melissa Cote (University of Victoria, Canada), Amanda Dash (ASL Environmental Sciences, Canada), Julek Chawarski (ASL Environmental Sciences, Canada), Alexandra Branzan Albu (University of Victoria, Canada), Femina Bharatkumar Senjaliya (Waterworth, Canada), and Andrea Niemi (Fisheries and Oceans Canada)</i>	
DebrisVision: Bridging the Synthetic-to-Real Gap for Enhanced Underwater Debris Analysis	2126
<i>Sivaji Retta (RGUKT IIIT Nuzvid, India), Sai Manikanta Eswar Machara (RGUKT IIIT Nuzvid, India), Iyyakutti Iyappan Ganapathi (Khalifa University of Science and Technology, UAE), Divya Velayudhan (Khalifa University of Science and Technology, UAE), and Naoufel Werghi (Khalifa University of Science and Technology, UAE)</i>	
The Coralscapes Dataset: Semantic Scene Understanding in Coral Reefs	2136
<i>Jonathan Sauder (École Polytechnique Fédérale de Lausanne, Switzerland), Viktor Domazetoski (University of Exeter, United Kingdom; The University of Queensland, Australia), Guilhem Banc-Prandi (École Polytechnique Fédérale de Lausanne, Switzerland), Gabriela Perna (École Polytechnique Fédérale de Lausanne, Switzerland), Anders Meibom (École Polytechnique Fédérale de Lausanne, Switzerland; University of Lausanne, Switzerland), and Devis Tuia (EPFL, Switzerland)</i>	

Weakly Supervised MaxN Estimation in Baited Remote Underwater Video	2144
<i>Amelia Sorrenti (University of Catania, Italy), Leonardo Giuseppe Russo (University of Catania, Italy), Sarinda Samarasinghe (University of Central Florida, USA), Simone Palazzo (University of Catania, Italy), and Concetto Spampinato (University of Catania, Italy)</i>	
A Multi-purpose Tracking Framework for Salmon Welfare Monitoring in Challenging Environments	2153
<i>Espen Uri Høgstedt (Norwegian University of Science and Technology, Norway), Christian Schellewald (SINTEF Ocean, Norway), Annette Stahl (Norwegian University of Science and Technology, Norway), and Rudolf Mester (Norwegian University of Science and Technology, Norway)</i>	
Benchmarking Out-of-Distribution Detection for Plankton Recognition: A Systematic Evaluation of Advanced Methods in Marine Ecological Monitoring	2163
<i>Yingzi Han (Beijing Normal University, China), Jiakai He (Beijing Normal University, China), Chuanlong Xie (Beijing Normal University, China), and Jianping Li (Shenzhen Institutes of Advanced Technology, China)</i>	
Stellar Sea Lion Counting Across Multiple Aerial Cameras	2174
<i>Matthew Dawkins (Kitware Inc., USA), Jon Crall (Kitware Inc., USA), Katie Sweeney (NOAA AKFSC Marine Mammals Lab, USA), Burlyn Birkemeier (University of Washington, USA), Neal Siekierski (Kitware Inc., USA), and Dawei Du (Kitware Inc., USA)</i>	
KAMERA: Enhancing Aerial Surveys of Ice-associated Seals in Arctic Environments	2183
<i>Adam Romlein (Kitware, Inc., USA), Benjamin X. Hou (NOAA NMFS AFSC MML, USA), Yuval Boss (CICOES, University of Washington, USA), Cynthia L. Christman (CICOES, University of Washington, USA), Stacie Koslovsky (NOAA NMFS AFSC MML, USA), Erin E. Moreland (NOAA NMFS AFSC MML, USA), Jason Parham (Kitware, Inc., USA), and Anthony Hoogs (Kitware, Inc., USA)</i>	
CoralVOS: Dataset and Benchmark for Dense Coral Video Segmentation	2193
<i>Ziqiang Zheng (The Hong Kong University of Science and Technology), Haixin Liang (The Hong Kong University of Science and Technology), Yaofeng Xie (Ocean University of China, China), Zhibin Yu (Ocean University of China, China), and Sai-Kit Yeung (The Hong Kong University of Science and Technology)</i>	

Multispectral Imaging for Robotics and Automation

Task-Driven Neural Adaptive Gain Control: 16-bit-to-8-bit Thermal Tone Mapping for Superior Object Detection	2204
<i>Hossein Javidnia (Trinity College Dublin, Ireland)</i>	
Towards a Generalizable Fusion Architecture for Multimodal Object Detection	2213
<i>Jad Berjawi (Université Grenoble Alpes, France), Yoann Dupas (Université Grenoble Alpes, France), and Christophe Cérin (University of Paris XIII, France)</i>	

TY-RIST: Tactical YOLO Tricks for Real-time Infrared Small Target Detection	2222
<i>Abdulkarim Atrash (Middle East Technical University, Turkey), Seyda Ertekin (Middle East Technical University, Turkey), Omur Ugur (Middle East Technical University, Turkey), Omar Moured (Karlsruhe Institute of Technology, Germany), Yufan Chen (Karlsruhe Institute of Technology, Germany), and Jiaming Zhang (Karlsruhe Institute of Technology, Germany)</i>	
Would SWIR modality help for detection and segmentation in harsh weather conditions? An experimental study.	2232
<i>Rohan Mehra (Université Clermont Auvergne, France ;Indian Institute of Science Education and Research (IISER), India), Alexandre Riffard (Université Clermont Auvergne, France), Mathieu Labussière (Université Clermont Auvergne, France), Pierre Duthon (Cerema, Research Team “Intelligent Transport Systems”, France), and Romuald Aufrère (Université Clermont Auvergne, France)</i>	
IVIFormer: Illumination-Aware Infrared-Visible Image Fusion via Adaptive Domain-Switching Cross Attention	2241
<i>Incheol Park (Yonsei University), Youngwan Jin (Yonsei University), Yagiz Nalcakan (Yonsei University), Hyeongjin Ju (Yonsei University), Sanghyeop Yeo (Yonsei University), and Shiho Kim (Yonsei University)</i>	

Vision-Based AI for Digital Health: From Pixels to Practice

Video Clinical Outcome Assessments (vCOAs): A Scalable Workflow for Remote Trials	2251
<i>Rabia Aziza (Eli Lilly and Company), Clare Matthews (Eli Lilly and Company), Elisa Ferrer (Eli Lilly and Company), Fereshteh Poursaeed (Eli Lilly and Company), Alejandro Mendoza Garcia (Eli Lilly and Company), and Elin Haf Davies (Eli Lilly and Company)</i>	
Positioning a Video to Mesh Pipeline for 3D Foot Reconstruction as a Monitoring Tool for Peripheral Edema	2256
<i>Wei Yan Peh (Eli Lilly and Company, Singapore), Hui Zhang (Eli Lilly and Company, USA), and Jian Yang (Eli Lilly and Company, USA)</i>	
Magnol.AI Copilot: Multimodal LLMs for Natural, Accessible Data Interaction	2264
<i>Guangchen Ruan (Eli Lilly and Company, USA), Hui Zhang (Eli Lilly and Company, USA), Hui Xiao (Eli Lilly and Company, USA), and Jian Yang (Eli Lilly and Company, USA)</i>	
Patient-Centric Statistical Multi-Modal Fusion for Medical Diagnosis: Integrating DICOM, Radiomics, and Patient Attributes	2273
<i>Seo-Yeon Choi (Jeonbuk National University, South Korea) and Kyungsu Lee (Jeonbuk National University, South Korea)</i>	
Multi-Modal Deep Clustering Survival Machines for Alzheimer's Disease Subtype Discovery	2285
<i>Zixuan Wen (University of Pennsylvania, PA), Bojian Hou (University of Pennsylvania, PA), Weiqing He (University of Pennsylvania, PA), Shu Yang (University of Pennsylvania, PA), Jason H. Moore (Cedars-Sinai Medical Center, CA), Andrew J. Saykin (Indiana University, IN), Heng Huang (University of Maryland, MD), Paul M. Thompson (University of Southern California, CA), Marylyn D. Ritchie (University of Pennsylvania, PA), Christos Davatzikos (University of Pennsylvania, PA), and Li Shen (University of Pennsylvania, PA)</i>	

Adapting Vision–Language Models for 3D CT/MRI Understanding on PMBB via Slice Selection and Explanation Analysis	2294
<i>Hongzhuo Chen (University of Pennsylvania), Rahul Shukla (University of Pennsylvania), Ruiming Wu (University of Pennsylvania), Shu Yang (University of Pennsylvania), Duy Duong-Tran (The United States Naval Academy), Duy Minh Ho Nguyen (Max Planck Research School for Intelligent Systems (IMPRS-IS)), Mathias Niepert (Max Planck Research School for Intelligent Systems (IMPRS-IS)), Cameron Beeche (University of Pennsylvania), James Gee (University of Pennsylvania), Jeffrey Duda (University of Pennsylvania), Rakesh Sharma (University of Pennsylvania), Christos Davatzikos (University of Pennsylvania), Walter Witschey (University of Pennsylvania), Bojian Hou (University of Pennsylvania), and Li Shen (University of Pennsylvania)</i>	

Safe and Trustworthy Multimodal AI Systems

Privacy Preservation Using Superimposed 3D-Models for Self-Supervised Training in Action Recognition	2304
<i>Asfandyar Azhar (Carnegie Mellon University), Nidhish Shah (Eindhoven University of Technology), Shaurjya Mandal (Harvard University), and Yongjie Jessica Zhang (Carnegie Mellon University)</i>	
Pulling Back the Curtain: Unsupervised Adversarial Detection via Contrastive Auxiliary Networks	2315
<i>Eylon Mizrahi (Ben-Gurion University, Israel), Raz Lapid (DeepKeep Tel-Aviv, Israel), and Moshe Sipper (Ben-Gurion University, Israel)</i>	
Safety Without Semantic Disruptions: Editing-free Safe Image Generation via Context-preserving Dual Latent Reconstruction	2327
<i>Jordan Vice (University of Western Australia, Australia), Naveed Akhtar (University of Melbourne, Australia), Mubarak Shah (University of Central Florida, USA), Richard Hartley (Australian National University, Australia), and Ajmal Mian (University of Western Australia, Australia)</i>	
On the Importance of Conditioning for Privacy-Preserving Data Augmentation	2338
<i>Julian Lorenz (University of Augsburg, Germany), Katja Ludwig (University of Augsburg, Germany), Valentin Haug (University of Augsburg, Germany), and Rainer Lienhart (University of Augsburg, Germany)</i>	
FrEVL: Leveraging Frozen Pretrained Embeddings for Efficient Vision-Language Understanding	2348
<i>Emmanuelle Bourigault (University of Oxford) and Pauline Bourigault (Imperial College London)</i>	

Revolutionizing the World of Retail

Recommendation By Generation: Generation Augmented Complementary Fashion Item Retrieval Using Incomplete Outfit	2358
<i>Gaurab Bhattacharya (TCS Research, India), B S Vivek (TCS Research, India), P. Rajith Bhargav (TCS Research, India), Jayavardhana Gubbi (TCS Research, India), Bagya Lakshmi (TCS Research, India), and Arpan Pal (TCS Research, India)</i>	

HyperVLM: Hyperbolic Space Guided Vision Language Modeling for Hierarchical Multi-Modal Understanding	2368
<i>Sarthak Srivastava (Amazon) and Kathy Wu (Amazon)</i>	
SGBD: Sharpness-Aware Mirror Gradient with BLIP-Based Denoising for Robust Multimodal Product Recommendation	2380
<i>Sarthak Srivastava (Amazon) and Kathy Wu (Amazon)</i>	
Sari Sandbox: A Virtual Retail Store Environment for Embodied AI Agents	2390
<i>Janika Deborah Gajo (University of the Philippines, Quezon City), Gerarld Paul Merales (University of the Philippines, Quezon City), Jerome Escarcha (University of the Philippines, Quezon City), Brenden Ashley Molina (University of the Philippines, Quezon City), Gian Nartea (University of the Philippines, Quezon City), Emmanuel G. Maminta (University of the Philippines, Quezon City), Juan Carlos Roldan (University of the Philippines, Quezon City), and Rowel O. Atienza (University of the Philippines, Quezon City)</i>	
RetailAction: Dataset for Multi-View Spatio-Temporal Localization of Human-Object Interactions in Retail	2400
<i>Davide Mazzini (Standard AI), Alberto Raimondi (Standard AI), Bruno Abbate (Standard AI), Daniel Fischetti (Standard AI), and David M Woollard (Standard AI)</i>	
Relative Pose Regression with Pose Auto-Encoders: Enhancing Accuracy and Data Efficiency for Retail Applications	2409
<i>Yoli Shavit (Bar Ilan University, Israel) and Yosi Keller (Bar Ilan University, Israel)</i>	
DualFit: A Two-Stage Virtual Try-On via Warping and Synthesis	2418
<i>Minh Tran (University of Arkansas), Johnmark Clements (University of Arkansas), Annie Prasanna (University of Arkansas), Tri Nguyen (Coupang Inc.), and Ngan Le (University of Arkansas)</i>	

10th International Workshop on Recovering 6D Object Pose

Lessons and Winning Solutions in Industrial Object Detection and Pose Estimation from the 2025 Bin-Picking Perception Challenge	2429
<i>Ziqin Huang (Tsinghua University), Chengxi Li (Tsinghua University), Yingyue Li (Tsinghua University), Xingyu Liu (Tsinghua University), Chenyangguang Zhang (Tsinghua University), Ruida Zhang (Tsinghua University), Bowen Fu (Tsinghua University), Xinggang Hu (Dalian University of Technology), Yun Qu (Tsinghua University), Mengge Liu (Tsinghua University), Yixiu Mao (Tsinghua University), Wendong Huang (Tsinghua University), Gu Wang (Tsinghua University), and Xiangyang Ji (Tsinghua University)</i>	
Learning Point Cloud Representations with Pose Continuity for Depth-Based Category-Level 6D Object Pose Estimation	2436
<i>Zhujun Li (Graduate Center, CUNY), Shuo Zhang (Hunter College, CUNY; Weill Cornell Medicine), and Ioannis Stamos (Graduate Center, CUNY; Hunter College, CUNY)</i>	

MR6D: Benchmarking 6D Pose Estimation for Mobile Robots	2447
<i>Anas Gouda (TU Dortmund; LAMARR Institute for Machine Learning), Shrutarv Awasthi (TU Dortmund), Christian Blesing (Fraunhofer IML), Lokeshwaran Manohar (TU Dortmund), Frank Hoffmann (TU Dortmund), and Alice Kirchheim (TU Dortmund; Fraunhofer IML; LAMARR Institute for Machine Learning)</i>	

13th International Workshop on Assistive Computer Vision and Robotics

Design Practices and Lessons from Deploying On-device Vision-Language Interaction in Robotic Guide Dogs	2456
<i>Jinse Kwon (Electronics and Telecommunications Research Institute (ETRI), Daejeon), Jemin Lee (Electronics and Telecommunications Research Institute (ETRI), Daejeon), and Yongin Kwon (Electronics and Telecommunications Research Institute (ETRI), Daejeon)</i>	
Understanding touch through latent spaces: can images and haptic maps reflect human perception?	2466
<i>Antonio Luigi Stefani (University of Trento), Sara Baldoni (University of Padova), Niccolò Bisagno (University of Trento), Federica Battisti (University of Padova), Nicola Conci (University of Trento), and Francesco De Natale (University of Trento)</i>	
LLMs as NAO Robot 3D Motion Planners	2476
<i>Riccardo Catalini (University of Modena and Reggio Emilia, Italy), Giacomo Salici (University of Modena and Reggio Emilia, Italy), Federico Biagi (University of Modena and Reggio Emilia, Italy), Guido Borghi (University of Modena and Reggio Emilia, Italy), Luigi Biagiotti (University of Modena and Reggio Emilia, Italy), and Roberto Vezzani (University of Modena and Reggio Emilia, Italy)</i>	
CrossCompanion: An Empathetic Real-Time Assistant Supporting Street Crossing for Low-Vision Users	2487
<i>Jiazhao Liang (New York University, USA) and Yi Fang (New York University, UAE)</i>	
TriPlanNet: Triangle Path Planning Network for A Variable Truss Robot with Deep Learning	2497
<i>Choonghan Lee (Electronics and Telecommunications Research Institute, Korea), Leah Harris (University of Nevada, Las Vegas, USA), Sehyun Oh (University of Science and Technology, Korea), JooHyoung Cha (University of Science and Technology, Korea), Jemin Lee (Electronics and Telecommunications Research Institute, Korea), Yongin Kwon (Electronics and Telecommunications Research Institute, Korea), and Andrew Jang-Ho Bae (University of Nevada, Las Vegas, USA)</i>	
Flow-Guided Policies: Overcoming Diffusion Limitations for Robust Robot Imitation Learning....	2507
<i>Chanhyuk Jung (Keimyung University), Sangwon Kim (ETRI), Kwang-Ju Kim (ETRI), Dasom Ahn (Keimyung University), Joonki Baek (Keimyung University), Sungkeun Yoo (Keimyung University), and Byoung Chul Ko (Keimyung University)</i>	

Towards Proactive Social Robots: Distilling Visual Knowledge from Large Vision-Language Models	2513
<i>Giuseppe De Simone (University of Salerno, SA), Loris Roveda (SUPSI), Alessia Saggese (University of Salerno, SA), and Mario Vento (University of Salerno, SA)</i>	
Visual Language Model-based Food Safety Support for Persons with Blindness and Low Vision ..	2524
<i>Ryan Hansen (New York University, USA), Hardik Setia (New York University, USA), Giles Hamilton-Fletcher (NYU Langone Health, USA), Aryan Jain (New York University, USA), Zirui Liu (New York University, USA), Mariam Zoair (New York University, USA), Reem Aboutaleb (New York University, USA), Qing Wen (New York University, USA), Yu Li (Columbia University, USA), and John Ross Rizzo (NYU Langone Health, USA)</i>	
Towards Effective Human-in-the-Loop Assistive AI Agents	2534
<i>Filippos Bellos (University of Michigan), Yayuan Li (University of Michigan), Cary Shu (University of Michigan), Ruey Day (University of Michigan), Jeffrey Siskind (Purdue University), and Jason Corso (University of Michigan)</i>	
NaVIP: A Low-Cost, Infrastructure-free Indoor Navigation Solution for Visually Impaired Persons	2544
<i>Jun Yu (Lehigh University, USA), Yitian Yang (Lehigh University, USA), and Vinod Namboodiri (Lehigh University, USA)</i>	

Unlearning and Model Editing Workshop and Challenge

Gen?: The Generative Machine Unlearning Challenge	2554
<i>Kartik Thakral (IIT Jodhpur), Shreyansh Pathak (IIT Jodhpur), Tamar Glaser (Harman International), Tal Hassner (WEIR AI), Diego Garcia-Olano (Meta AI), Iacopo Masi (Sapienza University of Rome), Richa Singh (IIT Jodhpur), and Mayank Vatsa (IIT Jodhpur)</i>	
QAQ: Quality Adaptive Quantization for LLM KV Cache	2563
<i>Wen Cheng (Nanjing University), Shichen Dong (Nanjing University), Jiayu Qin (Nanjing University), and Wei Wang (Nanjing University)</i>	
Guardians of Generation: Dynamic Inference-Time Copyright Shielding with Adaptive Guidance for AI Image Generation	2572
<i>Soham Roy (Kalinga Institute of Industrial Technology (KIIT), Bhubaneswar), Abhishek Mishra (Kalinga Institute of Industrial Technology (KIIT), Bhubaneswar), Aakash Sen Sharma (KIIT, Bhubaneswar; TCS Research), Shirish Karande (TCS Research), and Murari Mandal (Kalinga Institute of Industrial Technology (KIIT), Bhubaneswar)</i>	
Segmentation by Merging Models Specialized to Each Class	2583
<i>Kazuya Shibata (Meijo University, Japan) and Kazuhiro Hotta (Meijo University, Japan)</i>	
Bias-Aware Machine Unlearning: Towards Fairer Vision Models via Controllable Forgetting	2592
<i>Sai Siddhartha Chary Aylapuram (BITS Pilani Dubai Campus, UAE), Veeraraju Elluru (IIT Jodhpur, India), and Shivang Agarwal (BITS Pilani Dubai Campus, UAE)</i>	

5th Workshop on Open-World 3D Scene Understanding with Foundation Models

Open-Ended 3D Point Cloud Instance Segmentation	2601
<i>Phuc Nguyen (Movian AI), Minh Luu (MovianAI), Anh Tran (Movian AI), Cuong Pham (Posts & Telecommunications Inst. of Tech), and Khoi Nguyen (Movian AI)</i>	
Sparse Multiview Open-Vocabulary 3D Detection	2612
<i>Olivier Moliner (Centre for Mathematical Sciences, Lund University, Lund, Sweden), Viktor Larsson (Centre for Mathematical Sciences, Lund University, Lund, Sweden), and Kalle Åström (Centre for Mathematical Sciences, Lund University, Lund, Sweden)</i>	

End-to-End 3D Learning

Total-Editing: Head Avatar with Editable Appearance, Motion, and Lighting	2622
<i>Yizhou Zhao (Carnegie Mellon University), Chunjiang Liu (Carnegie Mellon University), Haoyu Chen (Carnegie Mellon University), Bhiksha Raj (Carnegie Mellon University), Min Xu (Carnegie Mellon University), Tadas Baltrusaitis (Microsoft), Mitch Rundle (Microsoft), Hsiangtao Wu (Microsoft), and Kamran Ghasedi (Microsoft)</i>	
EVA-Gaussian: 3D Gaussian-based Real-time Human Novel View Synthesis under Diverse Multi-view Camera Settings	2634
<i>Yingdong Hu (the Hong Kong University of Science and Technology), Zhening Liu (the Hong Kong University of Science and Technology), Jiawei Shao (Institute of Artificial Intelligence (TeleAI), China Telecom), Zehong Lin (the Hong Kong University of Science and Technology), and Jun Zhang (the Hong Kong University of Science and Technology)</i>	
PU-Gaussian: Point Cloud Upsampling using 3D Gaussian Representation	2644
<i>Mahmoud Khater (University of Freiburg, Germany; Fraunhofer IPM, Germany), Mona Strauss (University of Freiburg, Germany; Fraunhofer IPM, Germany), Philipp Von Olshausen (Fraunhofer IPM, Germany), and Alexander Reiterer (University of Freiburg, Germany; Fraunhofer IPM, Germany)</i>	
CSG-Fusion: Consistent Sparse-View Gaussian Splatting via Matching-based Fusion	2653
<i>Yan Xia (University of Science and Technology of China), Wenbo Ji (Technical University of Munich; Munich Center for Machine Learning (MCML)), Weirong Chen (Technical University of Munich; Munich Center for Machine Learning (MCML)), and Daniel Cremers (Technical University of Munich; Munich Center for Machine Learning (MCML))</i>	
Artist-Created Mesh Generation from Raw Observation	2663
<i>Yao He (Stanford University, USA), Youngjoong Kwon (Stanford University, USA), Wenxiao Cai (Stanford University, USA), and Ehsan Adeli (Stanford University, USA)</i>	

DAPS-AGF: Depth-Aware Perceptual Similarity with Adaptive Gradient Filtering for Enhanced Outdoor Scene Reconstruction	2669
<i>Aqsa Yousaf (The University of Texas at Arlington), Arkajyoti Mitra (The University of Texas at Arlington), Paul Agbaje (The University of Texas at Arlington), Afia Anjum (The University of Texas at Arlington), and Habeeb Olufowobi (The University of Texas at Arlington)</i>	
PointSeg: A Training-Free Paradigm for 3D Scene Segmentation via Foundation Models	2678
<i>Qingdong He (Tencent Youtu Lab), Jinlong Peng (Tencent Youtu Lab), Zhengkai Jiang (Tencent Youtu Lab), Xiaobin Hu (Tencent Youtu Lab), and Jiangning Zhang (Tencent Youtu Lab)</i>	
DeepCollide: Scalable Data-Driven High DoF Configuration Space Modeling using Implicit Neural Representations	2689
<i>Gabriel Guo (Stanford University), Judah Goldfeder (Columbia University), Aniv Ray (Columbia University), Tony Dear (Columbia University), and Hod Lipson (Columbia University)</i>	

Workshop on AI-driven Skilled Activity Understanding, Assessment & Feedback Generation

Learning Robust Aligned Representations Across Multiple Visual Modalities in Human Action Recognition	2700
<i>David J. Lerch (Fraunhofer IOSB, Germany; 3 Karlsruhe Institute of Technology (KIT), Germany), Bastian Rothenburger (Goethe University Frankfurt, Germany), Zeyun Zhong (Fraunhofer IOSB, Germany; Karlsruhe Institute of Technology (KIT), Germany), Manuel Martin (Fraunhofer IOSB, Germany), Frederik Diederichs (Fraunhofer IOSB), and Rainer Stiefelhagen (Karlsruhe Institute of Technology, Germany)</i>	
EgoOops: A Dataset for Mistake Action Detection from Egocentric Videos referring to Procedural Texts	2711
<i>Yuto Haneji (Kyoto University), Taichi Nishimura (Sony Interactive Entertainment), Hiroataka Kameko (Kyoto University), Keisuke Shirai (Kyoto University), Tomoya Yoshida (Kyoto University), Keiya Kajimura (Kyoto University), Koki Yamamoto (Kyoto University), Taiyu Cui (Kyoto University), Tomohiro Nishimoto (Kyoto University), and Shinsuke Mori (Kyoto University)</i>	
EHWGesture - A dataset for multimodal understanding of clinical gestures	2722
<i>Gianluca Amprimo (Politecnico di Torino, Italy), Alberto Ancilotto (Fondazione Bruno Kessler, Italy), Alessandro Savino (Politecnico di Torino, Italy), Fabio Quazzolo (Politecnico di Torino, Italy), Claudia Ferraris (CNR-IEIIT), Gabriella Olmo (Politecnico di Torino, Italy), Elisabetta Farella (Fondazione Bruno Kessler, Italy), and Stefano Di Carlo (Politecnico di Torino, Italy)</i>	
EASG-Bench: Video Q&A Benchmark with Egocentric Action Scene Graphs	2732
<i>Ioan Rodin (University of Catania), Tz-Ying Wu (Intel), Kyle Min (Intel), Sharath Nittur Sridhar (Intel Labs), Antonino Furnari (University of Catania), Subarna Tripathi (Intel Corporation), and Giovanni Maria Farinella (University of Catania)</i>	

CricTAL: Introducing Temporal Activity Localisation using pose estimation to identify key phases in cricket batting for downstream Action Quality Assessment	2738
<i>Tevin Moodley (University of Johannesburg, South Africa) and Dustin van der Haar (University of Johannesburg, South Africa)</i>	
InstaPose: Scene-Aware Pose Recommendation via Vision Transformers and Diversity-Optimized Reranking	2747
<i>Yu Ji (Santa Clara University, USA) and I-Han Hsiao (Santa Clara University, USA)</i>	
Assessing the Quality of Soccer Shots from Single-Camera Video with Vision-Language Models and Motion Features	2754
<i>Filip Noworolnik (AGH University of Krakow, Center of Excellence in Artificial Intelligence AGH) and Joanna Jaworek-Korjakowska (AGH University of Krakow, Center of Excellence in Artificial Intelligence AGH)</i>	

Multimodal Continual Learning

Forecasting and Visualizing Air Quality from Sky Images with Vision-Language Models	2762
<i>Mohammad Saleh Vahdatpour (Georgia State University, USA), Maryam Eyvazi (Savannah College of Art and Design, USA), and Yanqing Zhang (Georgia State University, USA)</i>	
PerFairX: Is There a Balance Between Fairness and Personality in Large Language Model Recommendations?	2771
<i>Chandan Kumar Sah (Beihang University, China)</i>	
MCL for MLLMs: Benchmarking Forgetting in Task-Incremental Multimodal Learning	2781
<i>Zichao Li (Canoakbit Alliance)</i>	

Neural SLAM

DROID-Splat Combining end-to-end SLAM with 3D Gaussian Splatting	2788
<i>Christian Homeyer (Heidelberg University, Germany), Leon Begiristain (Heidelberg University, Germany), and Christoph Schnörr (Heidelberg University, Germany)</i>	

Sustainability with Earth Observation & AI

Towards High-Resolution Alignment and Super-Resolution of Multi-Sensor Satellite Imagery	2799
<i>Philip Wootack Shin (The Pennsylvania State University ;NASA Marshall Space Flight Center, Huntsville), Vishal Gaur (University of Alabama at Huntsville), Rahul Ramachandran (NASA Marshall Space Flight Center, Huntsville), Manil Maskey (NASA Marshall Space Flight Center, Huntsville), Jack Sampson (The Pennsylvania State University), Vijaykrishnan Narayanan (The Pennsylvania State University), and Sujit Roy (NASA Marshall Space Flight Center, Huntsville ;University of Alabama in Huntsville)</i>	

GasTwinFormer: A Hybrid Vision Transformer for Livestock Methane Emission Segmentation and Dietary Classification in Optical Gas Imaging	2808
<i>Toqi Tahamid Sarker (Southern Illinois University Carbondale, USA), Mohamed Embaby (University of California, USA), Taminul Islam (Southern Illinois University Carbondale, USA), Amer AbuGhazaleh (Southern Illinois University Carbondale, USA), and Khaled Ahmed (Southern Illinois University Carbondale, USA)</i>	
ViewDelta: Scaling Scene Change Detection through Text-Conditioning	2818
<i>Subin Varghese (University of Houston, TX), Joshua Gao (University of Houston, TX), and Vedhus Hoskere (University of Houston, TX)</i>	
Tree Mapping with Limited Data: Fine-Tuning Foundation Models for Multimodal Fusion	2829
<i>Xiaoyan Lu (Hong Kong Polytechnic University, Hong Kong) and Qihao Weng (Hong Kong Polytechnic University, Hong Kong)</i>	
ViT-Koop: Vision-Transformer-Koopman Operators for Efficient Time-Series Forecasting of Earth-Observation Data	2835
<i>Takayuki Shinohara (National Institute of Advanced Industrial Science and Technology, Japan) and Hidetaka Saomoto (National Institute of Advanced Industrial Science and Technology, Japan)</i>	
Resolution Revolution: A Physics-Guided Deep Learning Framework for Spatiotemporal Temperature Reconstruction	2845
<i>Shengjie Liu (University of Southern California), Lu Zhang (University of Southern California), and Siqin Wang (University of Southern California)</i>	
PyroFocus: A Deep Learning Approach to Real-Time Wildfire Detection in Multispectral Remote Sensing Imagery	2855
<i>Mark Moussa (NASA Goddard Space Flight Center), Andre Williams (NASA Goddard Space Flight Center), Seth Roffe (NASA Goddard Space Flight Center), and Douglas Morton (NASA Goddard Space Flight Center)</i>	
LEGNet: A Lightweight Edge-Gaussian Network for Low-Quality Remote Sensing Image Object Detection	2865
<i>Wei Lu (Anhui University), Si-Bao Chen (Anhui University), Hui-Dong Li (Anhui University), Qing-Ling Shu (Anhui University), Chris H. Q. Ding (The Chinese University of Hong Kong (Shenzhen)), Jin Tang (Anhui University), and Bin Luo (Anhui University)</i>	
Inland Excess Water (IEW) Monitoring Using Sentinel-1/2: A SplitClass Segmentation and Temporal Gap-Filling Approach	2875
<i>Yahya Ibrahim (HUN-REN Institute for Computer Science and Control (SZTAKI)), Márta Belényesi (Lechner Knowledge Center), Chang Liu (HUN-REN Institute for Computer Science and Control (SZTAKI)), Mátyás Richter-Cserey (Lechner Knowledge Center), Máté Simon (Lechner Knowledge Center), Tamás Szirányi (HUN-REN Institute for Computer Science and Control (SZTAKI)), and Csaba Benedek (HUN-REN Institute for Computer Science and Control (SZTAKI))</i>	

Towards Large Scale Geostatistical Methane Monitoring with Part-based Object Detection	2886
<i>Adhemar de Senneville (Université Paris-Saclay, France), Xavier Bou (Université Paris-Saclay, France), Jean-Louis Bonne (Université de Reims Champagne-Ardenne, France), Nicolas Dumelie (Université de Reims Champagne-Ardenne, France), Rafael Grompone (Université Paris-Saclay, France), Thibaud Ehret (AMIAD, Pole Recherche, France), and Gabriele Facciolo (Université Paris-Saclay, France)</i>	
IRLTrees3D: A 3D Reconstruction Dataset of Trees	2897
<i>Joseph Chai (University College Cork, Ireland), Barry O' Sullivan (University College Cork, Ireland), and Hoang D. Nguyen (University College Cork, Ireland)</i>	
Evaluating Performance of Reinforcement Learning Agents to Control Buildings Efficiently	2903
<i>Judah Goldfeder (Columbia University), Gabriel Guerra Trigo (Columbia University), Philippe Martin Wyder (University of Washington), Neil Kachappilly (Columbia University), and Hod Lipson (Columbia University)</i>	
HA-RDet: Hybrid Anchor Rotation Detector for Oriented Object Detection	2910
<i>Phuc Nguyen (University of Information Technology, Vietnam)</i>	
Multi-Scale Hybrid CNN-Transformer for Smoke Detection in Satellite Images	2920
<i>Tony Zhang (University of Michigan, MI) and Robert Dick (University of Michigan, MI)</i>	
Context-Aware Masking and Learnable Diffusion-Guided Patch Refinement in Transformers via Sparse Supervision for Hyperspectral Image Classification	2927
<i>Abhiroop Chatterjee (Jadavpur University), Susmita Ghosh (Jadavpur University), and Ashish Ghosh (IIIT Bhubaneswar)</i>	
Combining Transformers and CNNs for Efficient Object Detection in High-Resolution Satellite Imagery	2937
<i>Nicolas Drapier (Université Sorbonne Paris Nord), Aladine Chetouani (Université Sorbonne Paris Nord), and Aurélien Chateigner (Impact SAS)</i>	
AQUAH: Automatic Quantification and Unified Agent in Hydrology	2947
<i>Songkun Yan (University of Oklahoma, USA), Zhi Li (Stanford University, USA), Siyu Zhu (University of Oklahoma, USA), Yixin Wen (University of Florida, USA), Mofan Zhang (Stanford University, USA), Mengye Chen (University of Oklahoma, USA), Jie Cao (University of Oklahoma, USA), and Yang Hong (University of Oklahoma, USA)</i>	
BiodiverseNet: Multitask Learning on Fused Multispectral and Radar Data for Scalable Ecosystem Monitoring	2957
<i>Prasanth Yadla (Independent Researcher, USA)</i>	
PlantationBench: A Multiscale, Multimodal Remote Sensing Benchmark for Plantation Mapping Under Distribution Shift	2965
<i>Angela Tsao (Stanford University, CA) and David Lobell (Stanford University, CA)</i>	

Short-Form Video Understanding: The Next Frontier in Video Intelligence

Watch, Listen, Understand, Mislead: Tri-modal Adversarial Attacks on Short Videos for Content Appropriateness Evaluation	2976
<i>Sahid Hossain Mustakim (United International University), S M Jishanul Islam (United International University), Ummay Maria Muna (BRAC University), Montasir Chowdhury (United International University), Mohammad Jawwadul Islam (United International University), Sadia Ahmmed (University of British Columbia), Tashfia Sikder (Bangladesh University of Professionals), Syed Tasdid Azam Dhrubo (University of Alberta), and Swakhar Shatabda (BRAC University)</i>	
Hashtag2Action: Data Engineering and Self-Supervised Pre-Training for Action Recognition in Short-Form Videos	2986
<i>Yang Qian (University of Hawai'i at Mānoa, USA), Ali Kargarandehkordi (University of Hawai'i at Mānoa, USA), Yinan Sun (University of Hawai'i at Mānoa, USA), Parnian Azizian (Stanford University), Onur Cezmi Mutlu (Stanford University), Saimourya Surabhi (Stanford University), Zain Jabbar (University of Hawai'i at Mānoa, USA), Dennis Paul Wall (Stanford University), Peter Washington (University of California), and Huaijin Chen (University of Hawai'i at Mānoa, USA)</i>	
End-to-End Action Segmentation Transformer	2997
<i>Tieqiao Wang (Oregon State University, US) and Sinisa Todorovic (Oregon State University, US)</i>	
Difformer for Action Segmentation	3007
<i>Nicolas Aziere (Oregon State University), Tieqiao Wang (Oregon State University), and Sinisa Todorovic (Oregon State University)</i>	

The 2nd Workshop on Efficient Computing under Limited Resources: Visual Computing

Low-bit FlashAttention Accelerated Operator Design Based on Triton	3017
<i>Jinyang Du (Beihang University, China), Jinyang Guo (Beihang University, China), and Yifu Ding (Beihang University, China)</i>	
Pruning by Block Benefit: Exploring the Properties of Vision Transformer Blocks during Domain Adaptation	3027
<i>Patrick Glandorf (L3S - Leibniz University Hannover, Germany) and Bodo Rosenhahn (L3S - Leibniz University Hannover, Germany)</i>	
Kernel-based Motion Free B-frame Coding for Neural Video Compression	3038
<i>Thang Van Nguyen (FPT Software, AI Center, Vietnam)</i>	
Efficient Depth- and Spatially-Varying Image Simulation for Defocus Deblur	3047
<i>Xinge Yang (KAUST, Saudi Arabia), Chuong Nguyen (Meta Reality Labs, USA), Wenbin Wang (Meta Reality Labs, USA), Kaizhang Kang (KAUST, Saudi Arabia), Wolfgang Heidrich (KAUST, Saudi Arabia), and Xiaoxing Li (Meta Reality Labs, USA)</i>	

VCMamba: Bridging Convolutions with Multi-Directional Mamba for Efficient Visual Representation	3058
<i>Mustafa Munir (The University of Texas at Austin), Alex Zhang (The University of Texas at Austin), and Radu Marculescu (The University of Texas at Austin)</i>	
Leveraging Learned Image Prior for 3D Gaussian Compression	3068
<i>Seungjoo Shin (POSTECH), Jaesik Park (Seoul National University), and Sunghyun Cho (POSTECH)</i>	
L-GGSC: Learnable Graph-based Gaussian Splatting Compression	3078
<i>Akihiro Kuwabara (The University of Osaka, Japan), Hinata Kiriwara (The University of Osaka, Japan), Sorachi Kato (The University of Osaka, Japan), Toshiaki Koike-Akino (Mitsubishi Electric Research Laboratories (MERL), USA), and Takuya Fujihashi (The University of Osaka, Japan)</i>	
Fisheye image augmentation for overcoming domain gaps with the limited dataset	3085
<i>Hyeseong Lee (Sogang University, Republic of Korea), Sunmin Park (Sogang University, Republic of Korea), and Miyoung Lee (Sogang University, Republic of Korea)</i>	
From Binary to Semantic: Utilizing Large-Scale Binary Occupancy Data for 3D Semantic Occupancy Prediction	3094
<i>Chihiro Noguchi (Toyota Motor Corporation) and Takaki Yamamoto (Toyota Motor Corporation)</i>	
Relevance-Guided Activation Sparsification for Bandwidth-Efficient Collaborative Inference.....	3104
<i>Maximilian Andreas Hoefler (Fraunhofer Heinrich Hertz Institute, Germany), Daniel Becking (Fraunhofer Heinrich Hertz Institute, Germany), Karsten Mueller (Fraunhofer Heinrich Hertz Institute, Germany), Detlev Marpe (Fraunhofer Heinrich Hertz Institute, Germany), and Wojciech Samek (Fraunhofer Heinrich Hertz Institute, Germany)</i>	
Tiny-vGamba: Distilling Large Vision-(Language) Knowledge from CLIP into a Lightweight vGamba Network	3114
<i>Yunusa Haruna (Beihang University, China ;NewraLab, China), Adamu Lawan (Beihang University, China), Ismail Abdulrashid (The University of Tulsa, USA), and Abdulganiyu Abdu Yusuf (Beijing Institute of Technology, China)</i>	
Your Super Resolution Model is not Enough for Tackling Real-World Scenarios	3123
<i>Dongsik Yoon (HDC LABS, Republic of Korea) and Jongeun Kim (HDC LABS, Republic of Korea)</i>	
Linear Attention with Global Context: A Multipole Attention Mechanism for Vision and Physics	3130
<i>Alex Colagrande (Université Paris Dauphine-PSL, France), Paul Caillon (Université Paris Dauphine-PSL, France), Eva Feillet (Université Paris Dauphine-PSL, France), and Alexandre Allauzen (Université Paris Dauphine-PSL, France ;ESPCI PSL, France)</i>	
Adaptive Compression of Large Vision Models for Efficient Image Quality Assessment of AI-Generated Content	3140
<i>Shivam Bhardwaj (Indian Institute of Technology Madras Zanzibar) and Tushar Shinde (Indian Institute of Technology Madras Zanzibar)</i>	

Decay Pruning Method: Smooth Pruning With a Self-Rectifying Procedure	3148
<i>Minghao Yang (Ningbo University), Pengyuan Li (IBM Research), Linlin Gao (Ningbo University), Zhiying Cui (Ningbo University), Wenbo Li (Zhejiang Lab), and Jiangjian Xiao (Chinese Academy of Sciences)</i>	
Latent Representation of Microstructures using Variational Autoencoders with Spatial Statistics-Space Loss	3155
<i>Andy Cai (University of Toronto, Canada), Sayed Sajad Hashemi (Georgia Institute of Technology, Canada), Noah Paulson (Argonne National Laboratory, USA), and Michael Guerzhoy (University of Toronto, Canada)</i>	
FDAL: Leveraging Feature Distillation for Efficient and Task-Aware Active Learning	3162
<i>Rebati Gaire (University of Illinois Chicago, Illinois) and Arman Roohi (University of Illinois Chicago, Illinois)</i>	
From Coarse to Fine: Learnable Discrete Wavelet Transforms for Efficient 3D Gaussian Splatting	3170
<i>Hung Nguyen (Video Processing Lab, UC San Diego), An Le (Video Processing Lab, UC San Diego), Blark Runfa Li (Video Processing Lab, UC San Diego), and Truong Nguyen (Video Processing Lab, UC San Diego)</i>	
Compressed Diffusion: Pruning with Knowledge Distillation for Efficient Text-to-Image Generation	3180
<i>Nasrin Kalanat (University of Pittsburgh), Ohiremen Dibua (Adobe Research), Yan Kang (Adobe Research), Yifang Gong (Adobe Research), and Xiaowei Jia (University of Pittsburgh)</i>	
ECLR'25: 2nd Workshop on Efficient Computing under Limited Resources: Visual Computing ...	3189
<i>Jinyang Guo (Beihang University), Zhenghao Chen (The University of Newcastle), Yuqing Ma (Beihang University), Yifu Ding (Beihang University), Xianglong Liu (Beihang University), Jinman Kim (The University of Sydney), Wanli Ouyang (The Chinese University of Hong Kong), and Dacheng Tao (Nanyang Technological University)</i>	

The 6th Face Anti-Spoofing Workshop: Unified Physical-Digital Attacks Detection

MAYA: Multi-Attack Yielding Augmentation for Unified Face Attack Detection	3193
<i>Taehoon Kim (Chung-Ang Univ, Korea), Jongwook Choi (Chung-Ang Univ, Korea), Seungjin Jung (Chung-Ang Univ, Korea), and Jongwon Choi (Chung-Ang Univ, Korea; Chung-Ang Univ, Korea)</i>	
UniAttackData+: Unified Physical-Digital Attack Detection+ Challenge	3202
<i>Hui Ma (M.U.S.T., China), Ajian Liu (CASIA, China), Yongze Li (M.U.S.T., China), Chuanbiao Song (Ant Group, China), Xiao Guo (MSU, USA), Changtao Miao (USTC, China), Wanyi Zhuang (USTC, China), Junze Zheng (M.U.S.T., China), Shunxin Chen (NJUPT, China), Yan Hong (Ant Group, China), Jiabao Guo (HFUT, China), Jiankang Deng (ICL, UK), Jun Lan (Ant Group, China), Weiqiang Wang (Ant Group, China), Tao Gong (USTC, China), Qi Chu (USTC, China), Sergio Escalera (UB, Spain), Hugo Jair Escalante (INAOE, Mexico), Xiaoming Liu (MSU, USA), Zhen Lei (CASIA, China), Isabelle Guyon (UPSaclay, France), Yanyan Liang (M.U.S.T., China), and Jun Wan (CASIA, China)</i>	

Oculus: Hierarchical Face Spoof Detection via Frequency-Enhanced Vision Transformers with Group-Aware Classification and Post-Fusion Attention ICCV Proceedings	3212
<i>Vincent Whannou de Dravo (Open SI, Benin), Kiran Raja (NTNU, Norway), Mohammed Bouzidi (Fieldmade AS, Norway), Limamou Gueye (NOVITOM, France), and Jon Yngve Hardeberg (NTNU, Norway)</i>	
A Generalizable Face Security Detection Method via Unified Texture and Semantic Feature Framework	3221
<i>Ke-Yue Zhang (YouTu Lab, Tencent), Ruoxin Chen (YouTu Lab, Tencent), Jiamu Sun (YouTu Lab, Tencent), Jiangming Wang (YouTu Lab, Tencent), Hao Yang (Shanghai Jiao Tong University), Taiping Yao (YouTu Lab, Tencent), and Shouhong Ding (YouTu Lab, Tencent)</i>	
Applying Semantic Anchor in Face Anti-Spoofing Detection for Unified Physical-Digital Attacks	3230
<i>Xu Yang (Institute of Artificial Intelligence (TeleAI), China Telecom; Xi'an Jiaotong University), Qi Zhang (Institute of Artificial Intelligence (TeleAI), China Telecom), Yaowen Xu (Institute of Artificial Intelligence (TeleAI), China Telecom), Hui Ma (Macau University of Science and Technology), Zhaofan Zou (Institute of Artificial Intelligence (TeleAI), China Telecom), and Hao Sun (Institute of Artificial Intelligence (TeleAI), China Telecom)</i>	
CrossGAP: Unified Face Anti-Spoofing via Cross-Modal Global-Aware Prompting	3239
<i>Xihong Chen (Beijing Jiaotong University, China), Yutong Xie (Beijing Jiaotong University, China), Hui Ma (Macau University of Science and Technology, China), Ning Li (Macau University of Science and Technology, China), Yanyan Liang (Macau University of Science and Technology, Macau, China), and Jiabao Guo (Hefei University of Technology, China)</i>	
Paired-Sampling Contrastive Framework for Joint Physical-Digital Face Attack Detection	3247
<i>Andrei Balykin (IDRND), Anvar Ganiev (IDRND), Denis Kondranin (IDRND), Kirill Polevoda (IDRND), Nikolai Liudkevich (IDRND), and Artem Petrov (IDRND)</i>	
Adaptive Face Anti-Spoofing Through Enhanced Data Manipulation and Feature Contrast Techniques	3255
<i>Jielong Wang (Akuvox, China), Kaifeng Huang (Akuvox, China), Zehua Lan (Akuvox, China), Senxu He (Akuvox, China), Minzhe Huang (Akuvox, China), and Ruyi Cai (Akuvox, China)</i>	
Center-Guided Feature Selection with Dimensionality Reduction for Face Anti-Spoofing	3262
<i>Senxu He (Akuvox, China), Minzhe Huang (Akuvox, China), Ruyi Cai (Akuvox, China), Jielong Wang (Akuvox, China), Kaifeng Huang (Akuvox, China), and Zehua Lan (Akuvox, China)</i>	
Iterative Binary Training	3270
<i>Vitor da Silva (University of Lleida, Spain), Rosana Tomàs (University of Lleida, Spain), Fernando Cores (University of Lleida, Spain), and Francesc Giné (University of Lleida, Spain)</i>	
One Model, All Attacks: ACER-Optimized Vision Transformer for Robust Face Anti-Spoofing	3279
<i>Jie Jin (Shizuoka University), Mahiro Tokumasu (Shizuoka University), Yu Makino (Shizuoka University), Shumpei Nishikawa (Shizuoka University), and Tetsushi Ohki (Shizuoka University; RIKEN AIP)</i>	

Optimizing DINOv2 with Registers for Face Anti-Spoofing	3287
<i>Mika Feng (Tohoku University, Japan), Pierre Gallin-Martel (Tohoku University, Japan), Koichi Ito (Tohoku University, Japan), and Takafumi Aoki (Tohoku University, Japan)</i>	
Spoofing-aware Prompt Learning for Unified Physical-Digital Facial Attack Detection	3294
<i>Jiabao Guo (HFUT, China), Yadian Wang (HFUT, China), Hui Ma (MUST, China), Yuhao Fu (HFUT, China), Ju Jia (SEU, China), Hui Liu (CCNU, China), Shengeng Tang (HFUT, China), Lechao Cheng (HFUT, China), Yunfeng Diao (HFUT, China), and Ajian Liu (CASIA, China)</i>	
TSANet: Unified Face Attack Detection via Triplet Embedding and Soft ACER Loss	3303
<i>Gabrial Zencha Ashungafac (Carnegie Mellon University Africa, Kigali Rwanda), Miracle James Oloutuche (Carnegie Mellon University Africa, Kigali Rwanda), Mukar Wepngong Eleanor (Carnegie Mellon University Africa, Kigali Rwanda), Blessed Guda (Carnegie Mellon University Africa, Kigali Rwanda), Assane Gueye (Carnegie Mellon University Africa, Kigali Rwanda), and Moise Busogi (Carnegie Mellon University Africa, Kigali Rwanda)</i>	

Transparent & Reflective objects In the wild Challenges

TRICKY 2025 Challenge on Monocular Depth from Images of Specular and Transparent Surfaces	3311
<i>Pierluigi Zama Ramirez (CVLab, University of Bologna), Alex Costanzino (CVLab, University of Bologna), Fabio Tosi (CVLab, University of Bologna), Matteo Poggi (CVLab, University of Bologna), Luigi Di Stefano (CVLab, University of Bologna), Jean-Baptiste Weibel (Automation and Control Institute, TU Wien), Doris Antensteiner (Austrian Institute of Technology), Markus Vincze (Automation and Control Institute, TU Wien), Benjamin Busam (Technical University of Munich, Germany), Guangyao Zhai (Technical University of Munich, Germany), Weihang Li (Technical University of Munich, Germany), Junwen Huang (Technical University of Munich, Germany), Hyunjun Jung (Technical University of Munich), Mykola Lavreniuk (Space Research Institute NASU-SSAU, Kyiv, Ukraine), Pihai Sun (Harbin Institute of Technology), Yijun Luo (Harbin Institute of Technology), Hongtao Wang (China Mobile Financial Technology Co., Ltd.), Manying Gao (China Mobile Financial Technology Co., Ltd.), Kui Jiang (Harbin Institute of Technology), and Junjun Jiang (Harbin Institute of Technology)</i>	

TRICKY 2025 HouseCat6D Object Pose Estimation Challenge with Specular and Transparent Surfaces	3323
--	------

Weihang Li (Technical University of Munich; Munich Center for Machine Learning), Junwen Huang (Technical University of Munich; Munich Center for Machine Learning), HyunJun Jung (Technical University of Munich; Munich Center for Machine Learning), Guangyao Zhai (Technical University of Munich; Munich Center for Machine Learning), Pierluigi Zama Ramirez (University of Bologna), Alex Costanzino (University of Bologna), Fabio Tosi (University of Bologna), Matteo Poggi (University of Bologna), Luigi Di Stefano (University of Bologna), Jean-Baptiste Weibel (TU Wien), Doris Antensteiner (Austrian Institute of Technology), Markus Vincze (TU Wien), Jing He (Xidian University, China), Yiqing Wang (Xidian University, China), Kexin Zhang (Xidian University, China), Licheng Jiao (Xidian University, China), Lingling Li (Xidian University, China), Fang Liu (Xidian University, China), Wenping Ma (Xidian University, China), and Benjamin Busam (Technical University of Munich; Munich Center for Machine Learning)

Visual Quality Assessment Competition

Better Supervised Fine-tuning for VQA: Integer-Only Loss	3334
--	------

Baihong Qian (ByteDance Inc., China), Haotian Fan (ByteDance Inc., China), Wenjie Liao (ByteDance Inc., China), Yunqiu Wang (ByteDance Inc., China), Tao Li (ByteDance Inc., China), and Junhui Cui (ByteDance Inc., China)

VQualA 2025 Document Image Quality Assessment Challenge	3344
<i>Fan Huang (Shanghai Jiaotong University), Xionghuo Min (Shanghai Jiao Tong University), Zhichao Ma (n/a), Xiaohong Liu (Shanghai Jiao Tong University), Chris Wei Zhou (Cardiff University), Guangtao Zhai (Shanghai Jiao Tong University), Junjie Gao (Mohamed bin Zayed University of Artificial Intelligence), Runze Liu (n/a), Yingzhe Peng (Southeast University), Shujian Yang (Shanghai Jiaotong University), Jin Zhang (Alibaba Group), Kai Yang (Ant Group), Zhiyuan You (The Chinese University of Hong Kong), Michael Ao (Southern University of Science and Technology), Yicheng Wu (n/a), Weixia Zhang (Shanghai Jiao Tong University), Junlin Chen (n/a), Wei Sun (Shanghai Jiaotong University), Zhihua Wang (City University of Hong Kong), Zhe Zhang (Nanjing University of Science and Technology), Yang Yang (Nanjing University of Science and Technology), Mingying Bai (n/a), Jiawang Du (Nanjing University of Posts and Telecommunications), Zilong Lu (Nanjing University of Posts and Telecommunications), Zhenyu Jiang (Nanjing University of Posts and Telecommunications), Ziguan Cui (Nanjing University of Posts and Telecommunications), Zongliang Gan (Nanjing University of Posts and Telecommunications), Guijin Tang (Nanjing University of Posts and Telecommunications), Fan Yang (n/a), Hang Ouyang (Chengdu University of Technology), Zhuohang Shi (Hebei University of Science and Technology), Tianxin Xiao (n/a), Zhizun Luo (Chengdu University of Technology), Zhaowang Wu (Sichuan University), Kaixin Deng (Hokkaido University), Ruikun Zhang (Beijing Institute of Technology), Hao Yang (Beijing Institute of Technology), and Liyuan Pan (Beijing Institute of Technology)</i>	
Hybrid Vision Transformer and Convolutional Neural Network for Super-Resolution Image Quality Assessment	3354
<i>Xinyu Li (University of Science and Technology of China ;Ant Group), Chuanbiao Song (AntGroup), Chenqi Zhang (AntGroup), Jun Lan (AntGroup), Huijia Zhu (AntGroup), Weiqiang Wang (AntGroup), and Xiaoyan Sun (University of Science and Technology of China)</i>	
PKU-AIGIQA-4K: A Perceptual Quality Assessment Database for Both Text-to-Image and Image-to-Image AI-Generated Images	3362
<i>Jiquan Yuan (Peking University, China), Jihe Li (Peking University, China), Fanyi Yang (Peking University, China), Xinyan Cao (Peking University, China), Jinming Che (Peking University, China), Jinlong Lin (Peking University, China), and Xixin Cao (Peking University, China)</i>	

VQualA 2025 Challenge on Image Super-Resolution Generated Content Quality Assessment: Methods and Results	3372
<p><i>Yixiao Li (Beihang University, China), Xin Li (University of Science and Technology of China, China), Chris Wei Zhou (Cardiff University, United Kingdom), Shuo Xing (Texas A&M University, USA), Hadi Amirpour (Alpen-Adria Universität Klagenfurt, Austria), Xiaoshuai Hao (Beijing Academy of Artificial Intelligence, China), Guanghui Yue (Shenzhen University, China), Baoquan Zhao (SUN YAT-SEN UNIVERSITY, China), Weide Liu (Harvard University, USA), Xiaoyuan Yang (Beihang University), Zhengzhong Tu (Texas A&M University, USA), Xinyu Li (University of Science and Technology of China), Chuanbiao Song (Alibaba Group), Chenqi Zhang (Ant Group), Jun Lan (AntGroup), Huijia Zhu (AntGroup), Weiqiang Wang (AntGroup), Xiaoyan Sun (University of Science and Technology of China), Shishun Tian (College of Electronics and Information Engineering, Shenzhen University), Dongyang Yan (Shenzhen University), Weixia Zhang (Shanghai Jiao Tong University), Junlin Chen (Shanghai Jiao Tong University), Wei Sun (East China Normal University), Zihua Wang (City University of Hong Kong), Zhuohang Shi (Hebei University of Technology), Zhizun Luo (Chengdu University of Technology), Hang Ouyang (Chengdu University of Technology), Tianxin Xiao (Chengdu University of Technology), Fan Yang (Chengdu University of Technology), Zhaowang Wu (Sichuan University), and Kaixin Deng (Hokkaido University)</i></p>	
VQualA 2025 Challenge on Visual Quality Comparison for Large Multimodal Models: Methods and Results	3383
<p><i>Hanwei Zhu (Nanyang Technological University), Haoning Wu (Moonshot AI), Zicheng Zhang (Shanghai Jiaotong University), Lingyu Zhu (City University of Hong Kong), Yixuan Li (City University of Hong Kong), Peilin Chen (City University of Hong Kong), Shiqi Wang (City University of Hong Kong), Chris Wei Zhou (Cardiff University), Linhan Cao (Shanghai Jiaotong University), Wei Sun (Shanghai Jiaotong University), Xiangyang Zhu (Shanghai Artificial Intelligence Laboratory), Weixia Zhang (Shanghai Jiao Tong University), Yucheng Zhu (Shanghai Jiaotong University), Jing Liu (Tianjin University), Dandan Zhu (East China Normal University), Guangtao Zhai (Shanghai Jiao Tong University), Xiongkuo Min (Shanghai Jiao Tong University), Zhichao Zhang (Shanghai Jiaotong University), Xinyue Li (Shanghai Jiaotong University), Shubo Xu (n/a), Anh Dao (Michigan State University), Yifan Li (Michigan State University), Hongyuan Yu (Not current), Jiaojiao Yi (n/a), Yiding Tian (n/a), Yupeng Wu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Feiran Sun (n/a), Lijuan Jiao (Samsung), and Song Jiang (Samsung)</i></p>	
Efficient Face Image Quality Assessment via Self-training and Knowledge Distillation	3394
<p><i>Wei Sun (East China Normal University), Weixia Zhang (Shanghai Jiao Tong University), Linhan Cao (Shanghai Jiao Tong University), Jun Jia (Shanghai Jiao Tong University), Xiangyang Zhu (Shanghai Artificial Intelligence Laboratory), Dandan Zhu (East China Normal University), Xiongkuo Min (Shanghai Jiao Tong University), and Guangtao Zhai (Shanghai Jiao Tong University)</i></p>	

Perceptual Classifiers: Detecting Generative Images using Perceptual Features	3403
<i>Krishna Srikar Durbha (The University of Texas at Austin), Asvin Kumar Venkataramanan (The University of Texas at Austin), Rajesh Sureddi (The University of Texas at Austin), and Alan C. Bovik (University of Colorado Boulder)</i>	
AIGVQA: A Unified Framework for Multi-Dimensional Quality Assessment of AI-Generated Video....	3414
<i>Jiarui Wang (Shanghai Jiao Tong University, China), Juntong Wang (Shanghai Jiao Tong University, China), Xiaorong Zhu (Shanghai Jiao Tong University, China), Huiyu Duan (Shanghai Jiao Tong University, China), Guangtao Zhai (Shanghai Jiao Tong University, China), and Xionghuo Min (Shanghai Jiao Tong University, China)</i>	
VQualA 2025 Challenge on GenAI-Bench AIGC Video Quality Assessment: Methods and Results	3422
<i>Ying Cheng (TAOBAO & TMALL Group, China), Huasheng Wang (TAOBAO & TMALL Group, China), Pengxiang Xiao (TAOBAO & TMALL Group, China), Yukang Ding (TAOBAO & TMALL Group, China), Enpeng Liu (TAOBAO & TMALL Group, China), Chris Wei Zhou (Cardiff University, UK), Baojun Li (Independent Researcher), Jiamian Huang (Independent Researcher), Jiarui Wang (Shanghai Jiao Tong University, China), Xiaorong Zhu (Shanghai Jiao Tong University, China), Juntong Wang (Shanghai Jiao Tong University, China), Huiyu Duan (Shanghai Jiao Tong University, China), Xionghuo Min (Shanghai Jiao Tong University, China), Qiang Hu (Shanghai Jiao Tong University, China), Chunlei Cai (Bilibili Inc, China), Guangao Zhai (Shanghai Jiao Tong University, China), Baihong Qian (ByteDance, China), Haotian Fan (ByteDance, China), Wenjie Liao (ByteDance, China), Yunqiu Wang (ByteDance, China), Tao Li (ByteDance, China), Junuhi Cui (ByteDance, China), Zhichao Zhang (Shanghai Jiao Tong University, China), Xinyue Li (Shanghai Jiao Tong University, China), Yunhao Li (Shanghai Jiao Tong University, China), Xiaohong Liu (Shanghai Jiao Tong University, China), Weixia Zhang (Shanghai Jiao Tong University, China), Bingkun Zheng (Shanghai Jiao Tong University, China), Wei Sun (East China Normal University, China), Zhihua Wang (City University of Hong Kong), Longwei Li (Shanghai University of Electric Power, China), Jinyu Zhao (Shanghai Administration Institute, China), Xincheng Lv (Shanghai University of Electric Power, China), Yang Cai (Shanghai University of Electric Power, China), Fangfang Lu (Shanghai University of Electric Power, China), Ritik Bompilwar (Northeastern University, USA), Saurabh Koshatwar (University of Southern California), Weifeng Cai (Guizhou University, China), Guangqian Kong (Guizhou University, China), Junfeng Yang (Hunan University of Technology and Business, China), Jing Fu (Hunan University of Technology and Business, China), Wei Zhang (Hunan University of Technology and Business, China), Wenzhi Cao (Hunan University of Technology and Business, China), Limei Liu (Hunan University of Technology and Business, China), Qin Li (Hunan University of Technology and Business, China), Wanli Ma (University of Cambridge, UK), Yixiao Li (Beihang University, China), and Xiaoshuai Hao (Beijing Academy of Artificial Intelligence, China)</i>	

VQualA 2025 Challenge on Engagement Prediction for Short Videos: Methods and Results	3432
<p><i>Dasong Li (The Chinese University of Hong Kong), Sizhuo Ma (Snap Inc.), Hang Hua (University of Rochester), Wenjie Li (Not current), Jian Wang (Snap Inc.), Chris Wei Zhou (Cardiff University), Fengbin Guan (University of Science and Technology of China), Xin Li (University of Science and Technology of China), Zihao Yu (University of Science and Technology of China), Yiting Lu (University of Science and Technology of China), Ru-Ling Liao (Alibaba Group), Yan Ye (n/a), Zhibo Chen (University of Science and Technology of China), Wei Sun (Shanghai Jiaotong University), Linhan Cao (Shanghai Jiaotong University), Yuqin Cao (Shanghai Jiaotong University), Weixia Zhang (Shanghai Jiao Tong University), Wen Wen (City University of Hong Kong), Kaiwei Zhang (Shanghai Jiaotong University), Zijian Chen (Shanghai Jiao Tong University), Fangfang Lu (n/a), Xiongkuo Min (Shanghai Jiao Tong University), Guangtao Zhai (Shanghai Jiao Tong University), Erjia Xiao (Hong Kong University of Science and Technology), Lingfeng Zhang (Beijing Academy of Artificial Intelligence), Zhenjie Su (n/a), Hao Cheng (Hong Kong University of Science and Technology), Yu Liu (Hefei University of Technology), Renjing Xu (The Hong Kong University of Science and Technology), Long Chen (Xiaomi), Xiaoshuai Hao (Beijing Academy of Artificial Intelligence), Zhenpeng Zeng (n/a), Jianqin Wu (Communication University of China), Xuxu Wang (Communication University of China), Qian Yu (n/a), Bo Hu (Chongqing University of Post and Telecommunications), Weiwei Wang (n/a), Pinxin Liu (University of Rochester), Yunlong Tang (University of Rochester), Luchuan Song (University of Rochester), Jinxi He (University of Rochester), Jiaru Wu (Carnegie Mellon University), and Hanjia Lyu (University of Rochester)</i></p>	
Engagement Prediction of Short Videos with Large Multimodal Models	3443
<p><i>Wei Sun (East China Normal University), Linhan Cao (Shanghai Jiao Tong University), Yuqin Cao (Shanghai Jiao Tong University), Weixia Zhang (East China Normal University), Wen Wen (City University of Hong Kong), Kaiwei Zhang (Shanghai Jiao Tong University), Zijian Chen (Shanghai Jiao Tong University), Fangfang Lu (Shanghai University of Electric Power), Xiongkuo Min (Shanghai Jiao Tong University), and Guangtao Zhai (Shanghai Jiao Tong University)</i></p>	
Understanding Perceptual Quality in CCTV Images: A Benchmark Dataset and Entropy-based Insights	3453
<p><i>Yujin Han (Dongduk Women's University, South Korea), Jitwoo Kang (Sookmyung Women's University, South Korea), Sanghoon Lee (Yonsei University, South Korea), and Taewan Kim (Dongduk Women's University, South Korea)</i></p>	
A Lightweight Ensemble-Based Face Image Quality Assessment Method with Correlation-Aware Loss	3461
<p><i>Mohammadali Hamidi (University of Cagliari, Italy), Hadi Amirpour (Alpen-Adria Universität Klagenfurt, Austria), Luigi Atzori (University of Cagliari, Italy), and Christian Timmerer (Alpen-Adria Universität Klagenfurt, Austria)</i></p>	

DeQA-Doc: Adapting DeQA-Score to Document Image Quality Assessment	3469
<i>Junjie Gao (Ant Group ;Mohamed bin Zayed University of Artificial Intelligence), Runze Liu (Southeast University), Yingzhe Peng (Southeast University), Shujian Yang (Shanghai Jiaotong University ;Ant Group), Jin Zhang (Alibaba Group), Kai Yang (Ant Group), and Zhiyuan You (The Chinese University of Hong Kong)</i>	
QualiVision: Multi-Modal Video Quality Assessment with Quality-Aware Fusion and Discriminative Learning Strategies	3479
<i>Ritik Bompilwar (Northeastern University) and Saurabh Koshatwar (University of Southern California)</i>	
VQualA 2025 Challenge on Face Image Quality Assessment: Methods and Results	3489
<i>Sizhuo Ma (Snap Inc.), Wei-Ting Chen (Microsoft), Qiang Gao (Snap Inc.), Jian Wang (Snap Inc.), Chris Wei Zhou (Cardiff University), Wei Sun (East China Normal University), Weixia Zhang (Shanghai Jiao Tong University), Linhan Cao (Shanghai Jiao Tong University), Jun Jia (Shanghai Jiao Tong University), Xiangyang Zhu (Shanghai Artificial Intelligence Laboratory), Dandan Zhu (East China Normal University), Xiongkuo Min (Shanghai Jiao Tong University), Guangtao Zhai (Shanghai Jiao Tong University), Baoying Chen (Alibaba Group), Xiongwei Xiao (Hong Kong Polytechnic University), Jishen Zeng (Alibaba Group), Wei Wu (Donghai Laboratory), Tiexuan Lou (Zhejiang University), Yuchen Tan (Zhejiang University), Chunyi Song (Zhejiang University), Zhiwei Xu (Zhejiang University), MohammadAli Hamidi (University of Cagliari), Hadi Amirpour (University of Klagenfurt), Mingyin Bai (Nanjing University of Posts and Telecommunications), Jiawang Du (Nanjing University of Posts and Telecommunications), Zhenyu Jiang (Nanjing University of Posts and Telecommunications), Zilong Lu (Nanjing University of Posts and Telecommunications), Ziguan Cui (Nanjing University of Posts and Telecommunications), Zongliang Gan (Nanjing University of Posts and Telecommunications), Xinpeng Li (East China Normal University), Shiqi Jiang (East China Normal University), Chenhui Li (East China Normal University), Changbo Wang (East China Normal University), Weijun Yuan (Jinan University), Zhan Li (Jinan University), Yihang Chen (Jinan University), Yifan Deng (Jinan University), Ruting Deng (Jinan University), Zhanglu Chen (Jinan University), Boyang Yao (Jinan University), Shuling Zheng (Jinan University), Feng Zhang (Jinan University), Zhiheng Fu (Jinan University), Abhishek Joshi (Aftershoot), Aman Agarwal (Aftershoot), Rakhil Immidiseti (Blue River Technology), and Ajay Narasimha Mopidevi (Lendbuzz)</i>	
MMF-QE: Advanced Multi-Modal Fusion for Quality Assessment and Engagement Prediction in User-Generated Short Videos	3499
<i>Fengbin Guan (University of Science and Technology of China), Zihao Yu (University of Science and Technology of China), Yiting Lu (University of Science and Technology of China), Wei Luo (University of Science and Technology of China), Yixin Gao (University of Science and Technology of China), Xin Li (University of Science and Technology of China), and Zhibo Chen (University of Science and Technology of China)</i>	

RankCORE: Self-Supervised Ranking-Aware Correlation Optimized Regression for Face Image Quality Assessment	3507
<i>Abhishek Joshi (Aftershoot Pvt. Ltd., India) and Aman Agarwal (Aftershoot Pvt. Ltd., India)</i>	

Exploring MLLM in Fine-Grained Visual Quality Comparison with Quality Token	3516
<i>Song Jiang (Samsung R&D Institute China Xi'an (SRCX)), Lijuan Jiao (Samsung R&D Institute China Xi'an (SRCX)), Yuguang Li (Samsung R&D Institute China Xi'an (SRCX)), Pei You (Samsung R&D Institute China Xi'an (SRCX)), Hui Li (Samsung R&D Institute China Xi'an (SRCX)), Ran Yang (Samsung R&D Institute China Xi'an (SRCX)), Feng Zhu (Samsung R&D Institute China Xi'an (SRCX)), Sehwan Ki (Samsung Electronics Co., LTD., South Korea), and Hyong-Euk Lee (Samsung Electronics Co., LTD., South Korea)</i>	

Woman in Computer Vision Workshop

Self-Supervised Representation Learning with Diffusion-Based Refinement for Image Rectification	3527
<i>Pooja Kumari (IIT Madras, India) and Sukhendu Das (IIT Madras)</i>	

Efficient Additive Attention for Transformer-based Semi-supervised Document Layout Analysis	3536
<i>Tahira Shehzadi (University of Kaiserslautern, Germany; University of Kaiserslautern, Germany; German Research Center for Artificial Intelligence (DFKI), Germany), Ifza Ifza (University of Kaiserslautern, Germany; German Research Center for Artificial Intelligence (DFKI), Germany), Didier Stricker (University of Kaiserslautern, Germany; University of Kaiserslautern, Germany; German Research Center for Artificial Intelligence (DFKI), Germany), and Muhammad Zeshan Afzal (University of Kaiserslautern, Germany; University of Kaiserslautern, Germany; German Research Center for Artificial Intelligence (DFKI), Germany)</i>	

Pose to Protect: Federated Skeleton-Based Anomaly Detection for Privacy-Conscious Video Surveillance	3545
<i>Dina Famouri (Southern Illinois University Carbondale, USA), Md Zarif Hossain (Southern Illinois University Carbondale, USA), and Ahmed Imteaj (Florida Atlantic University, USA)</i>	

Workshop on Graphic Design Understanding and Generation

MG-Gen: Single Image to Motion Graphics Generation	3554
<i>Takahiro Shirakawa (CyberAgent), Tomoyuki Suzuki (CyberAgent), Takuto Narumoto (CyberAgent), and Daichi Haraguchi (CyberAgent)</i>	
Embedding Font Impression Word Tags Based on Co-occurrence	3564
<i>Yugo Kubota (Kyushu University, Japan) and Seiichi Uchida (Kyushu University, Japan)</i>	

Anomaly Detection with Foundation Models

A Comprehensive Library for Benchmarking Multi-class Visual Anomaly Detection	3574
<i>Jiangning Zhang (Zhejiang University), Haoyang He (Zhejiang University), Zhenye Gan (Tencent), Qingdong He (Tencent), Yuxuan Cai (Huazhong University of Science and Technology), Zhucun Xue (Zhejiang University), Yabiao Wang (Tencent), Chengjie Wang (Tencent), Lei Xie (Zhejiang University), and Yong Liu (Zhejiang University)</i>	
Bridge Feature Matching and Cross-Modal Alignment with Mutual-filtering for Zero-shot Anomaly Detection	3584
<i>Yuhu Bai (Zhejiang university), Jiangning Zhang (Tencent), Yunkang Cao (Hunan University), Guangyuan Lu (Zhejiang University), Qingdong He (Tencent), Xiangtai Li (Peking University), and Guanzhong Tian (Zhejiang University)</i>	
DinoAtten3D: Slice-Level Attention Aggregation of DinoV2 for 3D Brain MRI Anomaly Classification	3594
<i>Fazle Rafsani (Arizona State University), Jay Shah (Arizona State University), Catherine D. Chong (Mayo Clinic, Arizona), Todd J. Schwedt (Mayo Clinic, Arizona), and Teresa Wu (Arizona State University)</i>	
Zero-Shot Image Anomaly Detection Using Generative Foundation Models	3604
<i>Lemar Abdi (Eindhoven University of Technology, Netherlands), Amaan Valiuddin (Eindhoven University of Technology, Netherlands), Francisco Caetano (Eindhoven University of Technology, Netherlands), Christiaan Viviers (Eindhoven University of Technology, Netherlands), and Fons Van Der Sommen (Eindhoven University of Technology, Netherlands)</i>	

Computer Vision for Materials Science

Automated Feature Tracking for Real-Time Kinematic Analysis and Shape Estimation of Carbon Nanotube Growth	3614
<i>Kaveh Safaavigerdini (University of Missouri-Columbia, USA), Ramakrishna Surya (University of Missouri-Columbia, USA), Jaired Collins (University of Missouri-Columbia, USA), Prasad Calyam (University of Missouri-Columbia, USA), Filiz Bunyak (University of Missouri-Columbia, USA), Matthew R. Maschmann (University of Missouri-Columbia, USA), and Kannappan Palaniappan (University of Missouri-Columbia, USA)</i>	
Neural Object Detection for 4D-STEM: High-Throughput Sub-Pixel Electron Diffraction Pattern Recognition	3624
<i>Arda Genc (University of California Santa Barbara, USA) and Ravit Silverstein (University of California Santa Barbara, USA; University of Florida, USA)</i>	
MatSSL: Robust Self-Supervised Representation Learning for Metallographic Image Segmentation	3635
<i>Hoang Hai Nam Nguyen (Korea Institute of Materials Science; University of Science and Technology), Phan Nguyen Duc Hieu (Korea Institute of Materials Science; University of Science and Technology), and Ho Won Lee (Korea Institute of Materials Science; University of Science and Technology)</i>	

Application of Semantic Segmentation on SEM Images of Rubber Composites and Verification of Correlation with Material Properties	3643
<i>Sho Sakamoto (Sumitomo Rubber Industries), Wakana Miyoshi (Sumitomo Rubber Industries), Kentaro Oishi (Sumitomo Rubber Industries), Ryuta Asami (NEC), Akira Kobayashi (NEC), Hiroki Iwana (NEC), Taichi Tokunaga (NEC), and Yosuke Motohashi (NEC)</i>	
Unveiling Process–Structure Mapping with a Deep Variational Autoencoder	3652
<i>Avanish Mishra (Los Alamos National Laboratory, USA), Brenden William Hamilton (Los Alamos National Laboratory, USA), Mashroor Shafat Nitol (Los Alamos National Laboratory, USA), Nithin Mathew (Los Alamos National Laboratory, USA), Kipton Marcos Barros (Los Alamos National Laboratory, USA), Timothy Clark Germann (Los Alamos National Laboratory, USA), and Saryu Jindal Fensin (Los Alamos National Laboratory, USA)</i>	
Improving Multislice Electron Ptychography with a Generative Prior	3662
<i>Christian K. Belardi (Cornell University), Chia-Hao Lee (Cornell University), Yingheng Wang (Cornell University), Justin Lovelace (Cornell University), Kilian Q. Weinberger (Cornell University), David A. Muller (Cornell University), and Carla P. Gomes (Cornell University)</i>	
SAM- and μ SAM- Based Inference of Nuclear Materials Processing History from SEM Imagery ...	3673
<i>Mayolo Valencia Mendoza (Los Alamos National Laboratory), Alexei N. Skurikhin (Los Alamos National Laboratory), Judith D. Cohn (Los Alamos National Laboratory), Luther McDonald (Oregon State University), and Kari Sentz (Los Alamos National Laboratory)</i>	
Improving U-Net Confidence on TEM Image Data with L2-Regularization, Transfer Learning, and Deep Fine-Tuning	3681
<i>Aiden Ochoa (Penn State University), Xinyuan Xu (Penn State University), and Xing Wang (Penn State University)</i>	
AtomDiffuser: Time-Aware Degradation Modeling for Drift and Beam Damage in STEM Imaging	3690
<i>Hao Wang (Clemson University), Hongkui Zheng (University of California, Irvine), Kai He (University of California, Irvine), and Abolfazl Razi (Clemson University)</i>	
AI-Driven Analysis of Fracture Types in Metallic Materials: Classification, Segmentation, and Quantification with Deep Neural Networks	3700
<i>Anna Wójcicka (AGH University of Krakow, Poland), Mateusz Wojtulewicz (AGH University of Krakow, Poland), Andrzej Brodzicki (AGH University of Krakow, Poland), Magdalena Leonkiewicz (AGH University of Krakow, Poland), Patryk Maciejewski (AGH University of Science and Technology, Center of Excellence in Artificial Intelligence), Joanna Jaworek-Korjakowska (AGH University of Krakow, Poland)], and Krzysztof Mroczka (Cracow University of Technology, Poland)</i>	

A Transformer-Based GAN Architecture for Simulating Carbon Nanotube Structures in the Frequency Domain	3709
<i>Trevor Bohl (Department of Electrical Engineering and Computer Science), Minasadat Attari (Department of Electrical Engineering and Computer Science), Matthew R. Maschmann (Department of Mechanical and Aerospace Engineering; University of Missouri-Columbia, USA), and Filiz Bunyak (Department of Electrical Engineering and Computer Science; University of Missouri-Columbia, USA)</i>	

Foundation and Generative Models in Biometrics

Foundation versus Domain-specific Models: Performance Comparison, Fusion, and Explainability in Face Recognition	3715
<i>Redwan Sony (Michigan State University), Parisa Farmanifard (Michigan State University), Arun Ross (Michigan State University), and Anil K. Jain (Michigan State University)</i>	
GaitCrafter: Diffusion Model for Biometric Preserving Gait Synthesis	3726
<i>Sirshapan Mitra (University of Central Florida) and Yogesh S. Rawat (University of Central Florida)</i>	
FaceLLM: A Multimodal Large Language Model for Face Understanding	3736
<i>Hatef Otroschi Shahreza (Idiap Research Institute, Switzerland) and Sébastien Marcel (Idiap Research Institute, Switzerland)</i>	
A Multi-domain Image Translative Diffusion StyleGAN for Iris Presentation Attack Detection.....	3747
<i>Shivangi Yadav (Michigan State University, Michigan) and Arun Ross (Michigan State University, Michigan)</i>	
FLUXSynID: A Framework for Identity-Controlled Synthetic Face Generation with Document and Live Images	3757
<i>Raul Ismayilov (University of Twente, Netherlands), Dzemila Sero (University of Twente, Netherlands), and Luuk Spreeuwers (University of Twente, Netherlands)</i>	

From street to space: 3D Vision Across altitudes

Fusing Convolution and Vision Transformer Encoders for Object Height Estimation from Monocular Satellite and Aerial Images	3768
<i>Furkan Gültekin (Locdus AI), Alper Koz (Center for Image Analysis, METU), Reza Bahmanyar (German Aerospace Center (DLR), Germany), Seyed Majid Azimi (German Aerospace Center (DLR), Germany), and Mehmet Lütfi Süzen (Geological Engineering Department, METU)</i>	
Tile and Slide : A New Framework for Scaling NeRF from Local to Global 3D Earth Observation	3778
<i>Camille Billouard (Centre National d'Etudes Spatiales (CNES), France); Univ Gustave Eiffel, France), Dawa Derksen (Centre National d'Etudes Spatiales (CNES), France), Alexandre Constantin (Centre National d'Etudes Spatiales (CNES), France), and Bruno Vallet (Univ Gustave Eiffel, France)</i>	

3DGS-to-PC: 3D Gaussian Splatting to Dense Point Clouds	3789
<i>Lewis A G Stuart (University of Nottingham), Andrew Morton (University of Saskatchewan), Ian Stavness (University of Saskatchewan), and Michael P Pound (University of Nottingham)</i>	
MIRAGE: Unsupervised Single Image to Novel View Generation with Cross Attention Guidance	3799
<i>Llukman Cerkezi (University of Bern), Aram Davtyan (University of Bern), Sepehr Sameni (Independent Researcher), and Paolo Favaro (University of Bern)</i>	
Adapting Stereo Vision From Objects To 3D Lunar Surface Reconstruction with the StereoLunar Dataset	3810
<i>Clémentine Grethen (Université de Toulouse, France), Simone Gasparini (Université de Toulouse, France), Géraldine Morin (Université de Toulouse, France), Jérémy Lebreton (Airbus Defence and Space), Lucas Marti (Airbus Defence and Space), and Manuel Sanchez Gestido (ESA (European Space Agency))</i>	

Generative AI for Storytelling

Re:Verse - Can your VLM read a Manga?	3820
<i>Aaditya Baranwal (University of Central Florida, USA), Madhav Kataria (Indian Institute of Technology, IN), Naitik Agarwal (Indian Institute of Technology, IN), Yogesh Singh Rawat (University of Central Florida, USA), and Shruti Vyas (University of Central Florida, USA)</i>	
Generating Visually Consistent Images for Storytelling via Narrative Graph Prompting	3831
<i>Andrew Shin (Keio University) and Kunitake Kaneko (Keio University)</i>	
Generative AI for Cel-Animation: A Survey	3837
<i>Yunlong Tang (University of Rochester), Junjia Guo (University of Rochester), Pinxin Liu (University of Rochester), Zhiyuan Wang (UCSB), Hang Hua (University of Rochester), Jia-Xing Zhong (University of Oxford), Yunzhong Xiao (CMU), Chao Huang (University of Rochester), Luchuan Song (University of Rochester), Susan Liang (University of Rochester), Yizhi Song (Purdue University), Liu He (Purdue University), Jing Bi (University of Rochester), Mingqian Feng (University of Rochester), Xinyang Li (University of Rochester), Zeliang Zhang (University of Rochester), and Chenliang Xu (University of Rochester)</i>	
From Sound to Sight: Towards AI-authored Music Videos	3851
<i>Leo Vitasovic (IT University of Copenhagen), Stella Graßhof (IT University of Copenhagen), Agnes Mercedes Kloft (Aalto University), Ville V. Lehtola (University of Twente), Martin Cunneen (University of Limerick), Justyna Starostka (IT University of Copenhagen), Glenn McGarry (University of Nottingham), Kun Li (University of Twente), and Sami S. Brandt (IT University of Copenhagen)</i>	

Large-scale Video Object Segmentation

TransFlow: Motion Knowledge Transfer from Video Diffusion Models to Video Salient Object Detection	3862
<i>Suhwan Cho (GenGenAI), Minhyeok Lee (Yonsei University), Junggho Lee (Yonsei University), Sunghun Yang (Yonsei University), and Sangyoun Lee (Yonsei University)</i>	
Find First, Track Next: Decoupling Identification and Propagation in Referring Video Object Segmentation	3873
<i>Suhwan Cho (GenGenAI), Seunghoon Lee (Yonsei University), Minhyeok Lee (Yonsei University), Junggho Lee (Yonsei University), and Sangyoun Lee (Yonsei University)</i>	
DepthFlow: Exploiting Depth-Flow Structural Correlations for Unsupervised Video Object Segmentation	3884
<i>Suhwan Cho (GenGenAI), Minhyeok Lee (Yonsei University), Junggho Lee (Yonsei University), Donghyeong Kim (Yonsei University), and Sangyoun Lee (Yonsei University)</i>	
Local2Global query Alignment for Video Instance Segmentation	3895
<i>Rajat Koner (AWS AI Lab ;Ludwig Maximilian University of Munich), Zhipeng Wang (LinkedIn Corporation), Srinivas Parthasarathy (Amazon), and Chinghang Chen (Amazon)</i>	
Beyond Appearance: Geometric Cues for Robust Video Instance Segmentation	3905
<i>Quanzhu Niu (Wuhan University, China), Yikang Zhou (Wuhan University, China), Shihao Chen (Wuhan University, China), Tao Zhang (Wuhan University, China), and Shunping Ji (Wuhan University, China)</i>	

Mobile Intelligent Photography & Imaging

MIPI 2025 Challenge on Detailed Image Quality Assessment : Methods and Results	3916
<i>Wenjie Liao (VCIP, CS, Nankai University; Bytedance Inc.), Haotian Fan (Bytedance Inc.), Yifang Xu (Bytedance Inc.), Meijia Song (Bytedance Inc.), Qiufang Ma (Bytedance Inc.), Shuhao Han (VCIP, CS, Nankai University; Bytedance Inc.), Chunle Guo (VCIP, CS, Nankai University), Chongyi Li (VCIP, CS, Nankai University), Jianhui Sun (Tencent, WeChat), Xinli Yue (Tencent, WeChat), Yuhao Xie (Tencent, WeChat), Tao Shao (Tencent, WeChat), Zhaoran Zhao (Beijing University of Posts and Telecommunications), Xinjun Ma (Bilibili Inc., China), Lu Liu (Shanghai Jiao Tong University), Chunlei Cai (Bilibili Inc., China), Qiang Hu (Shanghai Jiao Tong University), Shaocheng Shen (Shanghai Jiao Tong University), Huiyu Duan (Shanghai Jiao Tong University), Tianxiao Ye (Bilibili Inc., China), Xiaoyun Zhang (Shanghai Jiao Tong University), Hong Yi (Ricoh Software Research Center Beijing), Yupeng Zhang (Ricoh Software Research Center Beijing), Zhao Linnan (Xidian University, China), Xinyi You (Xidian University, China), Ziang Li (Xidian University, China), Chenhao Qiu (MangoTV), Alireza Talebpour (Shahid Beheshti University, Iran), Azadeh Mansouri (Kharazmi University, Iran), Ahmad Mahmoudi-Aznavah (Shahid Beheshti University, Iran), and Hossein Motamednia (Shahid Beheshti University, Iran; Institute for Research in Fundamental Sciences, Iran)</i>	

MIPI 2025 Challenge on Deblurring for Hybrid EVS Camera: Methods and Results	3924
<i>Yaqi Wu (Pentas Vision Pte Ltd, Singapore), Zhihao Fan (AlpsenTek, China), Hirotaka Shinozaki (Pentas Vision Pte Ltd, Singapore), Frank Zhang (Pentas Vision Pte Ltd, Singapore), Xander Li (Sony Electronics Inc, America), Alexis Baudron (Sony Electronics Inc, America), Wenbin Feng (AlpsenTek, China), Shuang Zhao (AlpsenTek, China), Jin Han (HUAWEI Noah's Ark Lab), Cheng Li (Xiaomi Inc., China), Yihui Shi (HUAWEI Noah's Ark Lab), Dehua Song (HUAWEI Noah's Ark Lab), Zheng Chen (Shanghai Jiaotong University), Wenbo Li (HUAWEI Noah's Ark Lab), Fenglong Song (HUAWEI Noah's Ark Lab), Yihong Leng (Xidian University; vivo Mobile Communication Co., Ltd), Siming Zheng (vivo Mobile Communication Co., Ltd), Peng-Tao Jiang (vivo Mobile Communication Co., Ltd), Linxiao Shi (vivo Mobile Communication Co., Ltd), Jinwei Chen (vivo Mobile Communication Co., Ltd), Bo Li (vivo Mobile Communication Co., Ltd), Jiaojiao Li (Xidian University), Cheng Li (Xiaomi Inc., China), Jinao Song (Xiaomi Inc., China), Yan Chen (Xiaomi Inc., China), Yajing Wei (Xiaomi Inc., China), Yuqiang Yang (Xiaomi Inc., China), Jian Tang (Xiaomi Inc., China), Long Bao (Xiaomi Inc., China), Heng Sun (Xiaomi Inc., China), Yang Qiu (University of Science and Technology of China), Kean Liu (University of Science and Technology of China), Senyan Xu (University of Science and Technology of China), Zhijing Sun (University of Science and Technology of China), Jiaying Zhu (University of Science and Technology of China), Chengjie Ge (University of Science and Technology of China), Xingbo Wang (University of Science and Technology of China), Yidi Liu (University of Science and Technology of China), Xin Lu (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), Zheng-Jun Zha (University of Science and Technology of China), Qinglin Liu (Harbin Institute of Technology, Weihai), Wei Yu (Tsinghua University), Yunfan Lu (HKUST (GZ)), Xiangming Wang (Harbin Institute of Technology, Shenzhen), Shuqi Ren (Harbin Institute of Technology, Shenzhen), Haijin Zeng (Harvard University), Jinchao Li (Harbin Institute of Technology, Shenzhen), Shiyang Zhou (Harbin Institute of Technology, Shenzhen), Yongyong Chen (Harbin Institute of Technology, Shenzhen), Qichao Dong (University of Macau), and Junqi Lu (University of Macau)</i>	
Monochromatic Event Guided Image Deblurring with Event-triggering-aware Decomposition	3935
<i>Minggui Teng (Peking University), Boyu Li (Peking University), Yixin Yang (Peking University), Chu Zhou (National Institute of Informatics), Yan Chen (SenseTime Research), Jimmy S. Ren (SenseTime Research; Hong Kong Metropolitan University), and Boxin Shi (Peking University)</i>	
Learning Camera-Agnostic White-Balance Preferences	3945
<i>Luxi Zhao (AI Center-Toronto, Samsung Electronics), Mahmoud Afifi (AI Center-Toronto, Samsung Electronics), and Michael S. Brown (AI Center-Toronto, Samsung Electronics)</i>	
Syncalizer - Precise Automatic Timecodes for Video and Audio Devices Using Consumer Hardware	3955
<i>Oliver Zendel (AIT Austrian Institute Of Technology, Austria), Matthias Schörghuber (AIT Austrian Institute Of Technology, Austria), and Clemens Hofbauer (AIT Austrian Institute Of Technology, Austria)</i>	

AesCrop: Aesthetic-driven Cropping Guided by Composition	3965
<i>Yen-Hong Wong (Multimedia University, Malaysia) and Lai-Kuan Wong (Multimedia University)</i>	
GuideSR: Rethinking Guidance for One-Step High-Fidelity Diffusion-Based Super-Resolution	3973
<i>Aditya Arora (TU Darmstadt), Zhengzhong Tu (Texas A&M University), Yufei Wang (Snap Inc.), Ruizheng Bai (Texas A&M University), Jian Wang (Snap Inc.), and Sizhuo Ma (Snap Inc.)</i>	
EGVD: Event-Guided Video Diffusion Model for Physically Realistic Large-Motion Frame Interpolation	3982
<i>Ziran Zhang (Zhejiang University; Shanghai AI Laboratory), Xiaohui Li (Shanghai AI Laboratory; Shanghai Jiao Tong University), Yihao Liu (Shanghai AI Laboratory), Yujin Wang (Shanghai AI Laboratory), Yueting Chen (Zhejiang University), Tianfan Xue (The Chinese University of Hong Kong; Shanghai AI Laboratory), and Shi Guo (Shanghai AI Laboratory)</i>	
HiLLIE: Human-in-the-Loop Training for Low-Light Image Enhancement	3992
<i>Xiaorui Zhao (University of Electronic Science and Technology of China), Xinyue Zhou (University of Electronic Science and Technology of China), Peibei Cao (Nanjing University of Information Science and Technology), Junyu Lou (University of Electronic Science and Technology of China), and Shuhang Gu (University of Electronic Science and Technology of China)</i>	
iDETEX: Empowering MLLMs for Intelligent DETailed EXplainable IQA	4003
<i>Zhaoran Zhao (Tencent, WeChat), Xinli Yue (Tencent, WeChat), Jianhui Sun (Tencent, WeChat), Yuhao Xie (Tencent, WeChat), Tao Shao (Tencent, WeChat), Liangchao Yao (Tencent, WeChat), FAN XIA (Tencent, WeChat), and Yuetang Deng (Tencent, WeChat)</i>	
Predictive Quality Assessment for Mobile Secure Graphics	4013
<i>Cas Steigstra (Scantrust / University of Amsterdam), Sergey Milyaev (Scantrust), and Shaodi You (University of Amsterdam)</i>	
Diffusion Autoencoders are Foundation Video Compressors	4023
<i>Niccolò Niccoli (University of Florence), Leonardo Galteri (Pegaso Telematic University), and Lorenzo Seidenari (University of Florence)</i>	

The 3rd Binary and Extreme Quantization for Computer Vision

Probabilistic dynamic quantization for memory constrained devices	4032
<i>Gabriele Santini (Fondazione Bruno Kessler), Francesco Paissan (Fondazione Bruno Kessler), and Elisabetta Farella (Fondazione Bruno Kessler)</i>	
Extreme Compression of Adaptive Neural Images	4042
<i>Leo Hoshikawa (Sony Interactive Entertainment), Marcos V. Conde (Sony Interactive Entertainment), Takeshi Ohashi (Sony Group Corporation), and Atsushi Irie (Sony Group Corporation)</i>	

HC?PTQ: Poincaré?Based Hyperbolic Clustering for Data?Free Quantization of Vision Transformers	4053
<i>Raffaele Mineo (University of Catania, Italy), Simone Palazzo (University of Catania, Italy), Concetto Spampinato (University of Catania, Italy), and Francesco Rundo (University of Catania, Italy)</i>	
Enhancing Generalization in Data-free Quantization via Mixup-class Prompting	4061
<i>Jiwoong Park (HyperAccel), Chaeun Lee (KRAFTON), Yongseok Choi (SK Telecom), Sein Park (Rebellions), Deokki Hong (Rebellions), and Jungwook Choi (Hanyang University)</i>	
PTQAT: A Hybrid Parameter-Efficient Quantization Algorithm for 3D Perception Tasks	4071
<i>Xinhao Wang (Wangxuan Institute of Computer Technology, Peking University), Zhiwei Lin (Peking University), Zhongyu Xia (Peking University), and Yongtao Wang (Peking University)</i>	
MoPEQ: Mixture of Mixed Precision Quantized Experts	4082
<i>Krishna Teja Chitty-Venkata (Argonne National Laboratory), Jie Ye (Illinois Institute of Technology), and Murali Emani (Argonne National Laboratory)</i>	
Exploiting Information Redundancy in Attention Maps for Extreme Quantization of Vision Transformers	4092
<i>Lucas Maisonnave (CEA), Karim Haroun (i3S / CNRS, Université Côte d'Azur), and Tom Pegeot (CEA)</i>	
PREFILT: Prefiltering for Fully Quantized Image Restoration Neural Networks	4101
<i>Denis Makhov (Huawei), Ruslan Ostapets (Huawei), Irina Zhelavskaya (Huawei), and Dehua Song (Huawei)</i>	
Binary SqueezeNet: Enhancing Parameter Efficiency in Binary Neural Networks	4111
<i>Salih Atabey (Middle East Technical University, T"urkiye) and Erdem Akagündüz (Middle East Technical University, T"urkiye)</i>	

Multimodal Representation and Retrieval

Document Haystack: A Long Context Multimodal Image/Document Understanding Vision LLM Benchmark	4121
<i>Goeric Huybrechts (Amazon AGI), Srikanth Ronanki (Amazon AGI), Sai Muralidhar Jayanthi (Amazon AGI), Jack Fitzgerald (Amazon AGI), and Srinivasan Veeravanallur (Amazon AGI)</i>	
Refining Skewed Perceptions in Vision-Language Contrastive Models through Visual Representations	4130
<i>Haocheng Dai (University of Utah) and Sarang Joshi (University of Utah)</i>	
Med-GRIM: Enhanced Zero-Shot Medical VQA using prompt-embedded Multimodal Graph RAG	4140
<i>Rakesh Raj Madavan (Shiv Nadar University, Chennai), Akshat Kaimal (Shiv Nadar University, Chennai), Hashim Faisal (Shiv Nadar University, Chennai), and S. Chandrakala (Shiv Nadar University, Chennai)</i>	

Chrono: A Simple Blueprint for Representing Time in MLLMs	4151
<i>Boris Meinardus (TU Berlin, Germany and TU Darmstadt & Hessian.ai, Germany), Hector G. Rodriguez (TU Darmstadt & Hessian.ai, Germany), Anil Batra (University of Edinburgh, UK), Anna Rohrbach (TU Darmstadt & Hessian.ai, Germany), and Marcus Rohrbach (TU Darmstadt & Hessian.ai, Germany)</i>	
Towards Reporting Bias in Visual-Language Datasets: Bi-modal Data Augmentation by Decoupling Object-Attribute Association	4157
<i>Qiyu Wu (Sony Group Corporation), Mengjie Zhao (Sony Group Corporation), Yutong He (Carnegie Mellon University ;Sony AI), Lang Huang (The University of Tokyo), Junya Ono (Sony Group Corporation), Hiromi Wakaki (Sony Group Corporation), and Yuki Mitsufuji (Sony Group Corporation ;Sony AI)</i>	
Global-to-Local or Local-to-Global? Enhancing Image Retrieval with Efficient Local Search and Effective Global Re-ranking	4167
<i>Dror Aiger (Google), Bingyi Cao (Google DeepMind), Kaifeng Chen (Google DeepMind), and Andre Araujo (Google DeepMind)</i>	
Smart Routing for Multimodal Video Retrieval: When to Search What	4177
<i>Kevin Dela Rosa (Cloudglue)</i>	
IRR-LMM: Improving On-demand Retail Recommendation with Large Multi-Modal Models	4186
<i>Yihao Zhao (Meituan Group ;Sun Yat-sen University), Nan Lai (Meituan Group), Xiaoming Li (Meituan Group), Xu Yan (Meituan Group), Wenhao Deng (Meituan Group), Hujiang Huang (Meituan Group), Shuai Zhang (Meituan Group), and Wei Lin (Meituan Group)</i>	
Visual Adaptive Prompting for Compositional Zero-Shot Learning	4196
<i>Kyle Stein (University of West Florida, USA), Andrew Arash Mahyari (Florida Institute for Human and Machine Cognition (IHMC), USA), Guillermo Francia (University of West Florida, USA), and Eman El-Sheikh (University of West Florida, USA)</i>	
Rate-Distortion Limits for Multimodal Retrieval: Theory, Optimal Codes, and Finite-Sample Guarantees	4206
<i>Thomas Y. Chen (Columbia University, New York)</i>	
MIND-RAG: Multimodal Context-Aware and Intent-Aware Retrieval-Augmented Generation for Educational Publications	4216
<i>Jiayang Yu (Beijing University of Posts and Telecommunications, China), Yuxi Xie (Beijing University of Posts and Telecommunications, China), Guixuan Zhang (Beijing University of Posts and Telecommunications, China), Jie Liu (Beijing University of Posts and Telecommunications, China), Zhi Zeng (Beijing University of Posts and Telecommunications, China), Ying Huang (Beijing University of Posts and Telecommunications, China), and Shuwu Zhang (Beijing University of Posts and Telecommunications, China)</i>	

What is Next in Multimodal Foundation Models?

Mitigating Language Confusion for Multimodal Foundation Models via Confusion-Aware Preference Optimization Pipeline	4224
<i>Seunghyun Hwang (GenON, Republic of Korea), Sungjun Lim (GenON, Republic of Korea), Soyeon Shin (GenON, Republic of Korea), Hyun-Geun Kim (GenON, Republic of Korea), Jungwon Lim (GenON, Republic of Korea), Juncheol Kim (GenON, Republic of Korea), Byungseok Kang (GenON, Republic of Korea), and Daewoo Myoung (GenON, Republic of Korea)</i>	
Enhancing Circuit Diagram Understanding via Near Sight Correction Using VLMs	4234
<i>Shreyas Kulkarni (Research & Development Centre, Hitachi India Pvt Ltd, India), Vivek Kumar (Research & Development Centre, Hitachi India Pvt Ltd, India), Remish Leonard Minz (Research & Development Centre, Hitachi India Pvt Ltd, India), Munender Varshney (Research & Development Centre, Hitachi India Pvt Ltd, India), Thiruvengadam Samon (Research & Development Centre, Hitachi India Pvt Ltd, India), Abhishek Mitra (Research & Development Centre, Hitachi India Pvt Ltd, India), Nikhil Kulkarni (Research & Development Centre, Hitachi India Pvt Ltd, India), Nilanjan Chakravorty (Research & Development Centre, Hitachi India Pvt Ltd, India), Prateek Mital (Research & Development Centre, Hitachi India Pvt Ltd, India), and Kingshuk Banerjee (Research & Development Centre, Hitachi India Pvt Ltd, India)</i>	
Uncertainty-Aware ControlNet: Bridging Domain Gaps with Synthetic Image Generation	4243
<i>Joshua Niemeijer (German Aerospace Center (DLR)), Jan Ehrhardt (University of Lübeck; German Research Center for Artificial Intelligence (DFKI)), Heinz Handels (University of Lübeck; German Research Center for Artificial Intelligence (DFKI)), and Hristina Uzunova (German Research Center for Artificial Intelligence (DFKI))</i>	
Evaluating Variance in Visual Question Answering Benchmarks	4253
<i>Nikitha SR (Media and Data Science Research Lab, Adobe)</i>	
Low-Rank Prompt Adaptation for Open-Vocabulary Object Detection	4263
<i>Zekun Zhang (Stony Brook University), Vu Quang Truong (The University of Adelaide), and Minh Hoai (The University of Adelaide)</i>	
Vocabulary-free Fine-grained Visual Recognition via Enriched Contextually Grounded Vision-Language Model	4275
<i>Dmitry Demidov (Mohamed bin Zayed University of Artificial Intelligence, UAE), Zaigham Zaheer (Mohamed bin Zayed University of Artificial Intelligence, UAE), Omkar Thawakar (Mohamed bin Zayed University of Artificial Intelligence, UAE), Salman Khan (Mohamed bin Zayed University of Artificial Intelligence, UAE), and Fahad Shahbaz Khan (Mohamed bin Zayed University of Artificial Intelligence, UAE)</i>	
Infusing fine-grained visual knowledge to Vision-Language Models	4285
<i>Nikolaos-Antonios Ypsilantis (Czech Technical University in Prague), Kaifeng Chen (Google DeepMind), André Araujo (Google DeepMind), and Ondřej Chum (Czech Technical University in Prague)</i>	

MORFI: Mutimodal Zero-Shot Reasoning for Financial Time-Series Inference	4295
<i>Tina Khezresmaeilzadeh (University of Southern California), Parsa Razmara (University of Southern California), Mohammad Erfan Sadeghi (University of Southern California), Seyedarmin Azizi (University of Southern California), and Erfan Baghaei Potraghloo (University of Southern California)</i>	
Audio-Visual LLM for Video Understanding	4305
<i>Fangxun Shu (Alibaba Group), Lei Zhang (University of California), Hao Jiang (Alibaba Group), and Cihang Xie (University of California)</i>	
What Holds Back Open-Vocabulary Segmentation?	4315
<i>Josip Šarić (University of Ljubljana), Ivan Martinović (University of Zagreb), Matej Kristan (University of Ljubljana), and Siniša Šegvić (University of Zagreb)</i>	
TULIP: Contrastive Image-Text Learning With Richer Vision Understanding	4326
<i>Zineng Tang (University of California), Long Lian (University of California), Seun Eisaep (University of California), Xudong Wang (University of California), Roei Herzig (University of California), Adam Yala (University of California), Alane Suhr (University of California), Trevor Darrell (University of California), and David M. Chan (University of California)</i>	
LLaVA-MORE: A Comparative Study of LLMs and Visual Backbones for Enhanced Visual Instruction Tuning	4337
<i>Federico Cocchi (University of Modena and Reggio Emilia, Italy ;University of Pisa, Italy), Nicholas Moratelli (University of Modena and Reggio Emilia, Italy), Davide Caffagni (University of Modena and Reggio Emilia, Italy), Sara Sarto (University of Modena and Reggio Emilia, Italy), Lorenzo Baraldi (University of Modena and Reggio Emilia, Italy), Marcella Cornia (University of Modena and Reggio Emilia, Italy), and Rita Cucchiara (University of Modena and Reggio Emilia, Italy ;IIT-CNR, Italy)</i>	
Strefer: Empowering Video LLMs with Space-Time Referring and Reasoning via Synthetic Instruction Data	4348
<i>Honglu Zhou (Salesforce AI Research), Xiangyu Peng (Salesforce AI Research), Shrikant Kendre (Salesforce AI Research), Michael S Ryou (Salesforce AI Research), Silvio Savarese (Salesforce AI Research), Caiming Xiong (Salesforce AI Research), and Juan Carlos Niebles (Salesforce AI Research)</i>	
CoT-Pose: Chain-of-Thought Reasoning for 3D Pose Generation from Abstract Prompts	4360
<i>Junuk Cha (Korea Advanced Institute of Science and Technology, South Korea) and Jihyeon Kim (Korea Telecom, South Korea)</i>	
GLAD: Generalizable Tuning for Vision-Language Models	4369
<i>Yuqi Peng (Chinese Academy of Sciences ;Northeastern University), Pengfei Wang (The Hong Kong Polytechnic University), Jianzhuang Liu (Chinese Academy of Sciences), and Shifeng Chen (Chinese Academy of Sciences ;Shenzhen University of Advanced Technology)</i>	

A Plug-and-Play Approach for Robust Image Editing in Text-to-Image Diffusion Models	4380
<i>Hyunwook Jo (Inha University, Republic of Korea), Jiseung Maeng (Inha University, Republic of Korea), Jun Hyung Park (Inha University, Republic of Korea), Namhyuk Ahn (Inha University, Republic of Korea), and In Kyu Park (Inha University, Republic of Korea)</i>	
CobraVPS: Code Template Optimization for Better Question Reasoning Accuracy with Visual Program Synthesis	4390
<i>Jiajing Chen (Syracuse University ;KLA), Xiu Zhang (KLA), Yang Li (KLA), Renyu Zhang (KLA), Yujie Dong (KLA), Senem Velipasalar (Syracuse University), and Jing Zhang (KLA)</i>	
Learning by Taking Notes: Memory-Guided Continual Learning for Generative Multimodal Models	4400
<i>Yanhui Guo (Amazon, USA), Chenghuan Guo (Amazon, USA), Yan Gao (Amazon, USA), and Yi Sun (Amazon, USA)</i>	

Workshop on Advanced Perception for Autonomous Healthcare

TAGS: 3D Tumor-Adaptive Guidance for SAM	4411
<i>Sirui Li (Southern University of Science and Technology), Linkai Peng (Northwestern University), Zheyuan Zhang (Northwestern University), Gorkem Durak (Northwestern University), and Ulas Bagci (Northwestern University)</i>	
UltraNBA Neural Bundle Adjustment for Pose Refinement in 3D Freehand Ultrasound	4422
<i>V. Bugra Yesilkaynak (Technical University of Munich), Vanessa Gonzalez Duque (Technical University of Munich), Magdalena Wysocki (Technical University of Munich), Mohammad Farid Azampour (Technical University of Munich), Nassir Navab (Technical University of Munich), and Diana Mateus (Ecole Centrale de Nantes)</i>	
Best Foot Forward: Robust Foot Reconstruction in-the-wild	4432
<i>Kyle Fogarty (University of Cambridge), Jing Yang (University of Cambridge), Chayan Kumar Patodi (Hike Medical), Jack Foster (Hike Medical), Aadi Bhanti (Hike Medical), Steven Chacko (Hike Medical), Cengiz Oztireli (University of Cambridge), and Ujwal Bonde (Hike Medical)</i>	
MAPS: A Morphology-Aware PPE Segmentation Framework for Healthcare Settings	4442
<i>Wanzhao Yang (Rutgers University), Syed Anwar (Children's National Hospital), Beomseok Park (Rutgers University), Sifan Yuan (Rutgers University), Aleksandra Sarcevic (Drexel University), Marius G. Linguraru (Children's National Hospital), Randall S. Burd (Children's National Hospital), and Ivan Marsic (Rutgers University)</i>	
Automated C-Arm Positioning via Conformal Landmark Localization	4451
<i>Ahmad Arrabi (University of Vermont, USA), Jay Hwasung Jung (University of Vermont, USA), Jax Luo (Cleveland Clinic, USA), Nathan Franssen (University of Vermont, USA), Scott Raymond (Neurological Institute, USA), and Safwan Wshah (University of Vermont, USA)</i>	

Learning Generalizable Diabetic Retinopathy Grading by Decoupled State Space Decoding 4461
Jingjun Yi (Tencent Jarvis Lab, China), Qi Bi (Westlake University, China), Hao Zheng (Tencent Jarvis Lab, China), Haolan Zhan (Monash University, Australia), Wei Ji (Yale University, United States), Huimin Huang (Tencent Jarvis Lab, China), Yuexiang Li (Macau University, Macau), Xian Wu (Tencent Jarvis Lab, China), and Yefeng Zheng (Westlake University, China)

Evaluating and Improving the Effectiveness of Synthetic Chest X-Rays for Medical Image Analysis 4472
Eva Prakash (Stanford University, USA), Jeya Maria Jose Valanarasu (Stanford University, USA), Zhihong Chen (Stanford University, USA), Eduardo Pontes Reis (Stanford University, USA), Andrew Johnston (Stanford University, USA), Anuj Pareek (Stanford University, USA), Christian Bluethgen (Stanford University, USA), Sergios Gatidis (Stanford University, USA), Cameron Olsen (Stanford University, USA), Akshay Chaudhari (Stanford University, USA), Andrew Ng (Stanford University, USA), and Curtis Langlotz (Stanford University)

Workshop on Cultural Continuity of Artists

Preservation vs. Fabrication: An Ethical Framework of Consent, Transparency, and Integrity for Posthumous AI Art 4481
Xiantao Zhang (Beihang University, China)

Visual Design from Cultural Heritage through Inpainting-based Restoration 4491
Jungmin Lee (Chung-Ang University, Republic of Korea) and Jongwon Choi (Chung-Ang University, Republic of Korea)

Style Composition within Distinct LoRA modules for Traditional Art 4497
Jaehyun Lee (Seoul National University), Wonhark Park (Seoul National University), Wonsik Shin (Seoul National University), Hyunho Lee (Seoul National University), Hyoung Min Na (Kyunghee University), and Nojun Kwak (Seoul National University)

HaSPeR: An Image Repository for Hand Shadow Puppet Recognition 4505
Syed Rifat Raiyan (Systems and Software Lab (SSL)), Zibrán Zarif Amio (Islamic University of Technology (IUT), Bangladesh), and Sabbir Ahmed (Computer Vision Lab (CVLab))

A Proposed XR-based Digital Framework for Nakhwa-nori: Preserving Regional Traditional Festivals and Addressing Safety and Environmental Barrier 4516
Hatnim Kim (Sogang University, South Korea), Geonwoo Song (Sogang University, South Korea), Seungryeol Eom (Sogang University, South Korea), and Seungkwan Choi (Sogang University, South Korea)

K-StyleLoRA: Information-Guided Image Generation via Selective Feature Learning 4524
Soyoung Lee (Chung-Ang University), Hyoungseo Cho (Seoul National University), Myungjoo Kang (Seoul National University), and YoungJoon Yoo (Chung-Ang University)

FashionRNA: Interactive Reimagination of Fashion Heritage 4532
Haneol Kim (Sogang University, South Korea), Jonghwan Bae (Sogang University, South Korea), Seonwoo Shin (Sogang University, South Korea), and Sanghun Park (Sogang University, South Korea)

Lost in Translation: A Position Paper on Probing Cultural Bias in Vision-Language Models via Hanbok VQA	4538
<i>Hyeokjun Lee (Chung-Ang University), Jeonghyo Song (Chung-Ang University), Geonhui Jang (Chung-Ang University), and Youngjoon Yoo (Chung-Ang University)</i>	
Towards a New Copyright Paradigm for Generative AI: Bridging Human-Machine Creativity Through Legal and Policy Reform	4546
<i>Daiwon Hyun (Sogang University), Sunho Park (Sogang University), and Eutteum Kim (Sogang University)</i>	
WCCA-AK: A Multimodal Dataset of Andre Kim’s Fashion Legacy for AI-Driven Cultural Heritage Research	4554
<i>Seongyeon Oh (Sogang University), Soyoung Lee (Chung-Ang University), Hyeon Seong Jeong (Sogang University), Sangwoo Jo (Sogang University), JinYoung Kim (Chung-Ang University), Yeonsoo Choi (Chung-Ang University), YoungJoon Yoo (Chung-Ang University), and Taehoon Kim (Sogang University)</i>	

Workshop on Distillation of Foundation Models for Autonomous Driving

Distillation Improves Visual Place Recognition for Low Quality Images	4560
<i>Anbang Yang (New York University), Ge Jin (New York University), Junjie Huang (New York University), Yao Wang (New York University), John-Ross Rizzo (NYU Grossman School of Medicine), and Chen Feng (New York University)</i>	
From Static to Dynamic: a Survey of Topology-Aware Perception in Autonomous Driving	4570
<i>Yixiao Chen (Sems), Ruining Yang (Sems), Xin Chen (Shandong University), Jia He (Shandong University), Dongliang Xu (Shandong University), and Yue Yao (Shandong University)</i>	
A Survey on Vision-Language-Action Models for Autonomous Driving	4583
<i>Sicong Jiang (McGill University, Canada), Zilin Huang (University of Wisconsin–Madison, USA), Kangan Qian (Tsinghua University, China), Ziang Luo (Tsinghua University, China), Tianze Zhu (Tsinghua University, China), Yang Zhong (Xiaomi Corporation), Yihong Tang (McGill University, Canada), Menglin Kong (McGill University, Canada), Yunlong Wang (Tsinghua University, China), Sixen Jiao (Xiaomi Corporation), Hao Ye (Xiaomi Corporation), Zihao Sheng (University of Wisconsin–Madison, USA), Xin Zhao (Tsinghua University, China), Tuopu Wen (Tsinghua University, China), Zheng Fu (Tsinghua University, China), Sikai Chen (University of Wisconsin–Madison, USA), Kun Jiang (Tsinghua University, China), Diange Yang (Tsinghua University, China), Seongjin Choi (University of Minnesota–Twin Cities, USA), and Lijun Sun (McGill University, Canada)</i>	
Cross-Camera Module Training of Raw Sensor Data-Based Automotive Machine Vision: Challenges and Solutions	4596
<i>Setareh Kian (University of Dayton), Shannon Brooks-Lehnert (Ford Motor Company), and Keigo Hirakawa (University of Dayton)</i>	

SafeRoute: Enhancing Traffic Scene Understanding via a Unified Deep Learning and Multimodal LLM	4606
<i>Ankit Kumar Shaw (Tsinghua University), Chandan Kumar Sah (Beihang University), Xiaoli Lian (Beihang University), Arsalan Shahid Baig (Beihang University), Tuopu Wen (Tsinghua University), Kun Jiang (Tsinghua University), Mengmeng Yang (Tsinghua University), Diange Yang (Tsinghua University), and Li Zhang (Beihang University)</i>	
Analyzing Closed-loop Training Techniques for Realistic Traffic Agent Models in Autonomous Highway Driving Simulations	4616
<i>Matthias Bitzer (Robert Bosch GmbH, Germany), Reinis Cimurs (Robert Bosch GmbH, Germany), Benjamin Coors (Robert Bosch GmbH, Germany), Johannes Goth (Robert Bosch GmbH, Germany), Sebastian Ziesche (Robert Bosch GmbH, Germany), Philipp Geiger (Robert Bosch GmbH, Germany), and Maximilian Naumann (Robert Bosch GmbH, Germany)</i>	
NuScenes-SpatialQA: A Spatial Understanding and Reasoning Benchmark for Vision-Language Models in Autonomous Driving	4626
<i>Kexin Tian (Texas A&M University), Jingrui Mao (Texas A&M University), Yunlong Zhang (Texas A&M University), Jiwan Jiang (University of Wisconsin-Madison), Yang Zhou (Texas A&M University), and Zhengzhong Tu (Texas A&M University)</i>	
TinyBEV: Cross-Modal Knowledge Distillation for Efficient Multi-Task Bird’s-Eye-View Perception and Planning	4636
<i>Reeshad Khan (University of Arkansas, Arkansas) and John Gauch (University of Arkansas, Arkansas)</i>	
SynSHRP2: A Synthetic Multimodal Benchmark for Driving Safety-critical Events Derived from Real-world Driving Data	4645
<i>Liang Shi (Virginia Tech Transportation Institute), Boyu Jiang (Virginia Tech), Zhenyuan Yuan (Virginia Tech Transportation Institute), Miguel A. Perez (Virginia Tech Transportation Institute ;Virginia Tech), and Feng Guo (Virginia Tech Transportation Institute ;Virginia Tech)</i>	

Neuromorphic Vision: Advantages and Applications of Event Cameras

Enhancing Event-Based Optical Camera Communication Via Dynamic Timing Correction	4656
<i>Matthew Howard (University of Dayton, USA) and Keigo Hirakawa (University of Dayton, USA)</i>	
DEIO: Deep Event Inertial Odometry	4665
<i>Weipeng Guan (The University of Hong Kong), Fuling Lin (The University of Hong Kong), Peiyu Chen (The University of Hong Kong), and Peng Lu (The University of Hong Kong)</i>	
Neural Ganglion Sensors: Learning Task-specific Event Cameras Inspired by the Neural Circuit of the Human Retina	4675
<i>Haley M. So (Stanford University) and Gordon Wetzstein (Stanford University)</i>	

Comparing Representations for Event Camera-based Visual Object Tracking	4686
<i>Oussama Abdul Hay (Khalifa University of Science and Technology), Sara Alansari (Khalifa University of Science and Technology), Mohamad Alansari (Khalifa University of Science and Technology), and Yahya Zweiri (Khalifa University of Science and Technology)</i>	
Quantifying Accuracy of an Event-Based Star Tracker via Earth’s Rotation	4696
<i>Dennis Melamed (Kitware), Connor Hashemi (Kitware), and Scott McCloskey (Kitware)</i>	
Lattice-allocated Real-time Line Segment Feature Detection and Tracking Using Only an Event-based Camera	4704
<i>Mikihiro Ikura (Istituto Italiano di Tecnologia), Arren Glover (Istituto Italiano di Tecnologia), Masayoshi Mizuno (Sony Interactive Entertainment Inc.), and Chiara Bartolozzi (Istituto Italiano di Tecnologia)</i>	
Event-based Spinning Object SLAM	4714
<i>Ethan Elms (The University of Adelaide), Yasir Latif (The University of Adelaide), and Tat-Jun Chin (The University of Adelaide)</i>	
GraphEnet: Event-driven Human Pose Estimation with a Graph Neural Network	4724
<i>Gaurvi Goyal (Maastricht University Istituto Italiano di Tecnologia), Pham Cong Thuong (Istituto Italiano di Tecnologia), Arren Glover (Istituto Italiano di Tecnologia), Masayoshi Mizuno (Sony Interactive Entertainment Inc.), and Chiara Bartolozzi (Istituto Italiano di Tecnologia)</i>	
SIS-Challenge: Event-based Spatio-temporal Instance Segmentation Challenge at the CVPR 2025 Event-based Vision Workshop	4734
<i>Friedhelm Hamann (TU Berlin), Emil Mededovic (RWTH Aachen), Fabian Gülhan (RWTH Aachen), Yuli Wu (RWTH Aachen), Johannes Stegmaier (RWTH Aachen), Jing He (Xidian University), Yiqing Wang (Xidian University), Kexin Zhang (Xidian University), Lingling Li (Xidian University), Licheng Jiao (Xidian University), Mengru Ma (Xidian University), Honxiang Huang (Hong Kong University of Science and Technology), Yuhao Yan (Hong Kong University of Science and Technology), Hongwei Ren (Hong Kong University of Science and Technology), Xiaopeng Lin (Hong Kong University of Science and Technology), Yulong Huang (Hong Kong University of Science and Technology), Bojun Cheng (Hong Kong University of Science and Technology), Se Hyun Lee (Wonkwang University), Gyu Sung Ham (Wonkwang University), Kanghan Oh (Wonkwang University), Gi Hyun Lim (Wonkwang University), Boxuan Yang (Tongji University), Bowen Du (Tongji University), and Guillermo Gallego (TU Berlin)</i>	
Multimodal Neuromorphic Event-Frame Fusion in Domain-Generalized Vision Transformer for Dynamic Object Tracking	4743
<i>Taha Razzaq (Tibbling Technologies) and Asim Iqbal (Tibbling Technologies)</i>	
Event-driven Robust Fitting on Neuromorphic Hardware	4751
<i>Tam Ngoc-Bang Nguyen (The University of Adelaide), Anh-Dzung Doan (The University of Adelaide), Zhipeng Cai (Intel Labs), and Tat-Jun Chin (The University of Adelaide)</i>	

Drone Detection with Event Cameras	4762
<i>Gabriele Magrini (University of Florence), Lorenzo Berlincioni (University of Florence), Federico Becattini (University of Siena), Luca Cultrera (University of Florence), and Pietro Pala (University of Florence)</i>	
Toward Low-SWaP Cognitive Agents: Neuromorphic Intelligence and FPGA-Based Deployments of Event Neural Networks	4774
<i>Trevor Bihl (Ohio University, OH), Rajashree Majumder (Ohio University, OH), Zhewei Wang (Ohio University, OH), Avinash Karanth (Ohio University, OH), and Jundong Liu (Ohio University, OH)</i>	
Exploring spatial-temporal dynamics in event-based facial micro-expression analysis	4782
<i>Nicolas Mastropasqua (Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Argentina ;CONICET-UBA, Instituto de Ciencias de la Computacion (ICC), Argentina), Ignacio Bugueno-Cordova (University of O'Higgins, Chile ;L3S Research Center, Leibniz Universitat Hannover, Germany), Rodrigo Verschae (University of O'Higgins, Chile), Daniel Acevedo (Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Argentina ;CONICET-UBA, Instituto de Ciencias de la Computacion (ICC), Argentina), Pablo Negri (Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Argentina ;CONICET-UBA, Instituto de Ciencias de la Computacion (ICC), Argentina), and Maria Elena Buemi (Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Argentina ;CONICET-UBA, Instituto de Ciencias de la Computacion (ICC), Argentina)</i>	

Human-Inspired Computer Vision

OscNet v1.5: Energy Efficient Hopfield Network on CMOS Oscillators for Image Classification	4792
<i>Wenxiao Cai (Stanford University), Zongru Li (Stanford University), Iris Wang (Carnegie Mellon University), Yu-Neng Wang (Stanford University), and Thomas H. Lee (Stanford University)</i>	
Human Vision Constrained Super-Resolution	4801
<i>Volodymyr Karpenko (Universita della Svizzera Italiana, Switzerland), Taimoor Tariq (Università della Svizzera Italiana, Switzerland), Jorge Condor (Università della Svizzera italiana, Switzerland), and Piotr Didyk (Università della Svizzera italiana, Switzerland)</i>	
Long-Tailed Data Classification by Increasing and Decreasing Neurons During Training	4811
<i>Taigo Sakai (Meijo-University, Japan) and Kazuhiro Hotta (Meijo-University, Japan)</i>	
Towards Human-Like Invariance: Self-Supervised Learning with Feature-Level Rotation Alignment	4820
<i>Sangjun Han (Seoul National University; CRABS, South Korea), Woojin Cheong (Seoul National University), Changhoon Song (Seoul National University), and Myungjoo Kang (Seoul National University)</i>	

From Neural Activity to Computation: Biological Reservoirs for Pattern Recognition in Digit Classification	4830
<i>Ludovico Iannello (ISTI-CNR, Italy), Luca Ciampi (ISTI-CNR, Italy), Fabrizio Tonelli (Bio@SNS, Italy), Gabriele Lagani (ISTI-CNR, Italy), Lucio M. Calcagnile (IBF-CNR, Italy), Federico Cremisi (Bio@SNS, Italy), Angelo Di Di Garbo (IBF-CNR, Italy), and Giuseppe Amato (ISTI-CNR, Italy)</i>	
Using Human Perception to Regularize Transfer Learning	4840
<i>Steve Cruz (University of Notre Dame), Justin Dulay (Struction), and Walter Scheirer (University of Notre Dame)</i>	
HiSS: Human-inspired Semantic Segmentation for Vehicle Interior Scene Understanding	4851
<i>Aleksander Kostuch (AGH University of Krakow) and Joanna Jaworek-Korjakowska (AGH University of Krakow)</i>	
Co-VisiON: Co-Visibility ReasONing on Sparse Image Sets of Indoor Scenes	4861
<i>Chao Chen (New York University), Nobel Dang (New York University), Juexiao Zhang (New York University), Wenkai Sun (New York University), Pengfei Zheng (New York University), Xuhang He (New York University), Yimeng Ye (New York University), Jiasheng Zhang (New York University), Taarun Srinivas (New York University), and Chen Feng (New York University)</i>	
Human-Inspired Summarization: Cluster Scene Videos into Diverse Frames	4872
<i>Chao Chen (New York University), Mingzhi Zhu (New York University), Ankush Pratap Singh (New York University), Yu Yan (New York University), Felix Juefei-Xu (New York University), and Chen Feng (New York University)</i>	
SeeEEG: Semantic-aware EEG-based Multi-Modal Retrieval-Augmented Generation for High-Fidelity Visual Brain Decoding	4883
<i>Jun-Mo Kim (Korea University, Korea), Woohyeok Choi (Korea University, Korea), Sang-Jun Park (Korea University, Korea), Keun-Soo Heo (Korea University, Korea), Young-Han Son (Korea University, Korea), Ji-Hye Oh (Korea University, Korea), Dong-Hee Shin (Korea University, Korea), and Tae-Eui Kam (Korea University, Korea)</i>	
AttZoom: Attention Zoom for Better Visual Features	4893
<i>Daniel DeAlcala (Universidad Autónoma de Madrid, Spain), Aythami Morales (Universidad Autónoma de Madrid, Spain), Julian Fierrez (Universidad Autónoma de Madrid, Spain), and Ruben Tolosana (Universidad Autónoma de Madrid, Spain)</i>	
Who Walks With You Matters: Perceiving Social Interactions with Groups for Pedestrian Trajectory Prediction	4903
<i>Ziqian Zou (Huazhong University of Science and Technology), Conghao Wong (Huazhong University of Science and Technology), Beihao Xia (Huazhong University of Science and Technology), and Xinge You (Huazhong University of Science and Technology; Research Institute of Huazhong University of Science and Technology in Shenzhen)</i>	

Exploring Human-Model Alignment in Visual Social Attention During Help-and-Hinder Social Interaction Classification	4913
<i>Lucia Schiatti (Istituto Italiano di Tecnologia, Italy), Guido Vallarino (University of Genoa, Italy), Sabrina M. Lopez (University of Virginia, USA), Yen-Ling Kuo (University of Virginia, USA), Matteo Moro (University of Genoa, Italy), Mengmi Zhang (Nanyang Technological University, Singapore), Monica Gori (Istituto Italiano di Tecnologia, Italy), Alessio Del Bue (Istituto Italiano di Tecnologia, Italy), Boris Katz (Massachusetts Institute of Technology, USA), and Andrei Barbu (Massachusetts Institute of Technology, USA)</i>	

Multimodal Sign Language Recognition Workshop

A Comparative Study of RGB-based Continuous Sign Language Recognition Techniques	4923
<i>Sarah Alyami (King Fahd University of Petroleum and Minerals; Imam Abdulrahman Bin Faisal University) and Hamzah Luqman (King Fahd University of Petroleum and Minerals; SDAIA-KFUPM Joint Research Center for Artificial Intelligence)</i>	
Development of an Intelligent System for Recognizing Islamic Religious Visual Signs in the Arabic Language	4933
<i>Duaa Mhnaa (Tartous University, Syria), Yaroub Dayoub (Tartous University, Syria), and Jafar Salman (Tartous University, Syria)</i>	
Iterative Latent Refinement for Robust Non-Autoregressive Sign Language Production	4942
<i>Tuğçe Kızıltepe (Hacettepe University, Turkey), Sümeyye Meryem Taşyürek (Hacettepe University, Turkey), and Hacer Yalim Keles (Hacettepe University, Turkey)</i>	
Text-Aligned Radar-Based Sign Language Recognition for Healthcare Communication	4953
<i>Raffaele Mineo (University of Catania, Italy), Amelia Sorrenti (University of Catania, Italy), Gaia Caligiore (University of Modena and Reggio Emilia, Italy), Federica Proietto Salanitri (University of Catania, Italy), Giovanni Bellitto (University of Catania, Italy), Senya Polikovskiy (Max Planck Institute for Intelligent Systems, Germany), Sabina Fontana (University of Catania, Italy), Egidio Ragonese (University of Catania, Italy), Concetto Spampinato (University of Catania, Italy), and Simone Palazzo (University of Catania, Italy)</i>	
Inclusive Sign Language AI: Towards Authentic Accessibility Through Community Collaboration	4962
<i>Abraham Glasser (Gallaudet University, USA)</i>	
A Closer Look at Skeleton-based Continuous Sign Language Recognition	4968
<i>Yuecong Min (Chinese Academy of Sciences, China), Yifan Yang (Huazhong University of Science and Technology, China), Peiqi Jiao (Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Zixi Nan (University of Chinese Academy of Sciences, China), and Xilin Chen (Chinese Academy of Sciences, China; Institute of Computing Technology, China)</i>	

Point-Supervised Japanese Fingerspelling Localization via HR-Pro and Contrastive Learning	4975
<i>Ryota Murai (Tokyo Polytechnic University, Japan), Naoto Tsuta (Tokyo Polytechnic University, Japan), Duk Shin (Tokyo Polytechnic University, Japan), and Yousun Kang (Tokyo Polytechnic University, Japan)</i>	
FusionEnsemble-Net: An Attention-Based Ensemble of Spatiotemporal Networks for Multimodal Sign Language Recognition	4983
<i>Md. Milon Islam (University of Waterloo), Md Rezwanaul Haque (University of Waterloo), S M Taslim Uddin Raju (University of Waterloo), and Fakhri Karray (University of Waterloo; Zayed University of Artificial Intelligence)</i>	
A Signer-Invariant Conformer and Multi-Scale Fusion Transformer for Continuous Sign Language Recognition	4990
<i>Md Rezwanaul Haque (University of Waterloo), Md. Milon Islam (University of Waterloo), S M Taslim Uddin Raju (University of Waterloo), and Fakhri Karray (University of Waterloo; Zayed University of Artificial Intelligence)</i>	
RF-ChessSIGN: Radar-enabled Human-Computer Interaction in a Real-Time Sign Language-Controlled Game	5000
<i>Kenneth Dehaan (Gallaudet University, USA), Emre Kurtoglu (Rockefeller University, USA), Sabyasachi Biswas (Mississippi State University, MS), Caroline Kobek Pezzarossi (Gallaudet University, USA), Darrin J. Griffin (The University of Alabama, USA), Chris Crawford (The University of Alabama, USA), Ali C. Gurbuz (North Carolina State University), Evie A. Malaia (The University of Alabama, USA), Abraham Glasser (Gallaudet University, USA), Raja Kushalnagar (Gallaudet University, USA), and Sevgi. Z Gurbuz (North Carolina State University, USA)</i>	
SAGE: Segment-Aware Gloss-Free Encoding for Token-Efficient Sign Language Translation	5011
<i>Low Jian He (University of Surrey, United Kingdom), Ozge Mercanoglu Sincan (University of Surrey, United Kingdom), and Richard Bowden (University of Surrey, United Kingdom)</i>	
Region-Aware Pose Modeling and Permutation Decoding for Signer-Independent Sign Language Recognition	5021
<i>Sieu Tran (University of Information Technology, Vietnam; Vietnam National University, Vietnam; DayOne AI Lab), Duc Nguyen Minh (University of Information Technology, Vietnam; Vietnam National University, Vietnam), Truong Nguyen Thanh (University of Information Technology, Vietnam; Vietnam National University, Vietnam), Hao Vo (University of Information Technology, Vietnam; Vietnam National University, Vietnam), Thanh Duc Ngo (University of Information Technology, Vietnam; Vietnam National University, Vietnam), Tien Do (University of Information Technology, Vietnam; Vietnam National University, Vietnam), Khiem Le (University of Information Technology, Vietnam; Vietnam National University, Vietnam), and Duy-Dinh Le (University of Information Technology, Vietnam; Vietnam National University, Vietnam)</i>	

Generalizable Sign Language Recognition via Local Temporal Convolutions and Region-Aware Pose Encoding	5030
<i>Sieu Tran (University of Information Technology, Vietnam National University), Duc Nguyen Minh (University of Information Technology, Vietnam; Vietnam National University, Vietnam), Truong Nguyen Thanh (University of Information Technology, Vietnam; Vietnam National University, Vietnam), Hao Vo (University of Information Technology, Vietnam; Vietnam National University, Vietnam), Thanh Duc Ngo (University of Information Technology, Vietnam; Vietnam National University, Vietnam), Tien Do (University of Information Technology, Vietnam; Vietnam National University, Vietnam), Khiem Le (University of Information Technology, Vietnam; Vietnam National University, Vietnam), and Duy-Dinh Le (University of Information Technology, Vietnam; Vietnam National University, Vietnam)</i>	
Can a Lightweight Transformer Deliver a Robust Multimodal Sign Language Word Recognition?	5038
<i>Eva Odima Berepiki (University of Calgary, Canada), Philip Ciunkiewicz (University of Calgary, Canada), and Svetlana Yanushkevich (University of Calgary, Canada)</i>	
Modality-specific Benchmarks and Radar Range-Doppler Envelope Classification for Multimodal Isolated Sign Language Recognition	5046
<i>Dmitriy Sazonov (University of Alabama, USA), Kamrul Islam (North Carolina State University), Evie Malaia (University of Alabama), and Seogi Z. Gurbuz (North Carolina State University)</i>	
A Smart Glove to Convert Gestures to Speech & Text to Assist Deaf & Mute People Using Machine Learning	5054
<i>A Manisha (Sri Sivasubramaniya Nadar College of Engineering, India) and S Tamilselvi (Sri Sivasubramaniya Nadar College of Engineering, India)</i>	
A Multimodal Video and Radar Fusion Framework for High-Accuracy Isolated Sign Language Recognition	5061
<i>Sultan Mohammad Manjur (North Carolina State University), Sabyasachi Biswas (Mississippi State University), and Ali C. Gurbuz (North Carolina State University)</i>	
AutoSign: Direct Pose-to-Text Translation for Continuous Sign Language Recognition	5071
<i>Samuel Ebimobowe Johnny (Carnegie Mellon University Africa), Blessed Guda (Carnegie Mellon University Africa), Andrew Blayama Stephen (Carnegie Mellon University Africa), and Assane Gueye (Carnegie Mellon University Africa)</i>	
Multimodal Italian Sign Language Recognition with Radar-Video Late Fusion on the MultiMeDaLIS Dataset	5079
<i>Jakub F. Juranek (Polish Academy of Sciences, Poland)</i>	

The SignEval 2025 Challenge at the ICCV Multimodal Sign Language Recognition Workshop: Results and Discussion	5086
<p><i>Hamzah Luqman (King Fahd University of Petroleum & Minerals), Raffaele Mineo (University of Catania), Murtadha Aljubran (Mohamed Bin Zayed University of Artificial Intelligence), Ahmed Abul Hasanaath (King Fahd University of Petroleum & Minerals), Amelia Sorrenti (University of Catania), Sarah Alyami (Imam Abdulrahman Bin Faisal University), Sadam Al-Azani (King Fahd University of Petroleum & Minerals), Maad Alowaifeer (King Fahd University of Petroleum & Minerals), Jihwan Moon (Yonsei University), Václav Javorek (University of West Bohemia), Tomáš Železný (University of West Bohemia), Marek Hrúz (University of West Bohemia), Gaia Caligiore (University of Modena and Reggio Emilia), Silvio Giancola (King Abdullah University of Science and Technology), Senya Polikovsky (Max Planck Institute for Intelligent Systems), Motaz Alfarraj (King Fahd University of Petroleum & Minerals), Sabina Fontana (University of Catania), Mufti Mahmud (King Fahd University of Petroleum & Minerals), Muhammad Haris Khan (Mohamed Bin Zayed University of Artificial Intelligence), Kamrul Islam (North Carolina State University), Seogi Gurbuz (North Carolina State University), Egidio Ragonese (University of Catania), Giovanni Bellitto (University of Catania), Federica Proietto Salanitri (University of Catania), Concetto Spampinato (University of Catania), and Simone Palazzo (University of Catania)</i></p>	

2nd AI for Content Generation, Quality Enhancement and Streaming

VidMP3: Video Editing by Representing Motion with Pose and Position Priors	5096
<p><i>Sandeep Mishra (University of Texas at Austin), Oindrila Saha (University of Massachusetts Amherst), and Alan C. Bovik (University of Texas at Austin)</i></p>	
LSSGen: Leveraging Latent Space Scaling in Flow and Diffusion for Efficient Text to Image Generation	5107
<p><i>Jyun-Ze Tang (Inventec Corporation, Taiwan), Chih-Fan Hsu (Inventec Corporation, Taiwan), Jeng-Lin Li (Inventec Corporation, Taiwan), Ming-Ching Chang (University at Albany, State University of New York, USA), and Wei-Chao Chen (Inventec Corporation, Taiwan)</i></p>	
C2D-ISR: Optimizing Attention-based Image Super-resolution from Continuous to Discrete Scales	5117
<p><i>Yuxuan Jiang (University of Bristol, UK), Chengxi Zeng (University of Bristol, UK), Siyue Teng (University of Bristol), Fan Zhang (University of Bristol, UK), Xiaoqing Zhu (Netflix Inc., USA), Joel Sole (Netflix Inc., USA), and David Bull (University of Bristol)</i></p>	
Multi-Scale Contrastive-Adversarial Distillation for Super-Resolution	5128
<p><i>Donggeun Ko (AiM Future Inc.), Youngsang Kwak (AiM Future Inc.), San Kim (Sungkyunkwan University), Jaehwa Kwak (AiM Future Inc.), and Jaekwang Kim (Sungkyunkwan University)</i></p>	

High-Fidelity 4x Neural Reconstruction of Real-time Path Traced Videos	5138
<i>Zhiqiang Lao (Futurewei Technologies Inc., USA), Zekai Chen (Futurewei Technologies Inc., USA; Cornell University), Jiali Cui (Futurewei Technologies Inc., USA; Stevens Institute of Technology), Marcos Conde (University of Würzburg), and Heather Yu (Futurewei Technologies Inc., USA)</i>	

Computer Vision for Ecology

Species-level Detection and Tracking of Caribbean Coral Reef Fish	5147
<i>Levi Cai (Massachusetts Institute of Technology ;Woods Hole Oceanographic Institution), Austin Greene (Woods Hole Oceanographic Institution), Daniel Yang (Massachusetts Institute of Technology ;Woods Hole Oceanographic Institution), Nadège Aoki (Massachusetts Institute of Technology ;Woods Hole Oceanographic Institution), Sierra D. Jarriel (Woods Hole Oceanographic Institution), Jasper Ha (Woods Hole Oceanographic Institution ;University of California, USA), Nathan Formel (Woods Hole Oceanographic Institution), T. Aran Mooney (Woods Hole Oceanographic Institution), and Yogesh Girdhar (Woods Hole Oceanographic Institution)</i>	
Fine-Grained Beetle Taxonomy with Vision Models: A Benchmark on Long-Tailed and Domain-Adaptive Classification	5152
<i>S M Rayeed (Rensselaer Polytechnic Institute), Alyson East (The University of Maine), Samuel Stevens (The Ohio State University), Sydne Record (The University of Maine), and Charles V. Stewart (Rensselaer Polytechnic Institute)</i>	
Back Home: A Computer Vision Solution to Seashell Identification for Ecological Restoration	5159
<i>Alexander Valverde (FIFCO), Luis Solano (FIFCO), and André Montoya (FIFCO)</i>	
Bridging Domain Gaps for Fine-Grained Moth Classification Through Expert-Informed Adaptation and Foundation Model Priors	5169
<i>Ross J. Gardiner (University of Exeter), Guillaume Mougeot (Aarhus University), Sareh Rowlands (University of Exeter), Benno I. Simmons (University of Exeter), Flemming Helsing (Aarhus University), and Toke Thomas Høye (Aarhus University)</i>	
Divide and Conquer: Structured Reranking for Expert-Level Ecological Image Retrieval	5175
<i>Asmi Kumar (Massachusetts Institute of Technology), Edward Vendrow (Massachusetts Institute of Technology), and Sara Beery (Massachusetts Institute of Technology)</i>	

6th Workshop on Continual Learning in Computer Vision

MONET: Multi-Modal Online Continual Learning with Novelty Estimation	5185
<i>Evelyn Chee (National University of Singapore), Wynne Hsu (National University of Singapore), and Mong Li Lee (National University of Singapore)</i>	
Adapting curricula by tracking the gap between validation and training loss	5195
<i>Raghavendra Singh (Ashoka University, India)</i>	

Accuracy Improvement of Prompt-based Continual Learning with Past Information	5203
<i>Ikki Kato (Meijo University, Japan) and Kazuhiro Hotta (Meijo University, Japan)</i>	
RICO: Two Realistic Benchmarks and an In-Depth Analysis for Incremental Learning in Object Detection	5212
<i>Matthias Neuwirth Trapp (ETH Zurich, Switzerland ;Bosch Research), Maarten Bieshaar (Bosch Research), Danda Pani Paudel (Sofia University "St. Kliment Ohridski"), and Luc Van Gool (Sofia University "St. Kliment Ohridski")</i>	
Incremental Object Detection with Prompt-based Methods	5224
<i>Matthias Neuwirth Trapp (ETH Zurich, Switzerland ;Bosch Research), Maarten Bieshaar (Bosch Research), Danda Pani Paudel (Sofia University "St. Kliment Ohridski"), and Luc Van Gool (Sofia University "St. Kliment Ohridski")</i>	
Core Tokensets for Data-efficient Sequential Training of Transformers	5233
<i>Subarnaduti Paul (University of Bremen, Germany), Manuel Brack (Hessian Centre for Artificial Intelligence (Hessian.AI), Germany), Patrick Schramowski (German Research Center for Artificial Intelligence (DFKI), Germany), Kristian Kersting (Hessian Centre for Artificial Intelligence (Hessian.AI), Germany ;German Research Center for Artificial Intelligence (DFKI), Germany ;TU Darmstadt, Germany), and Martin Mundt (University of Bremen, Germany)</i>	
Does Prior Data Matter? Exploring Joint Training in the Context of Few-Shot Class-Incremental Learning	5244
<i>Shiwon Kim (Yonsei University), Dongjun Hwang (Sogang University), Sungwon Woo (Sogang University), and Rita Singh (Carnegie Mellon University)</i>	
Evolving from Unknown to Known: Retentive Angular Representation Learning for Incremental Open Set Recognition	5254
<i>Runqing Yang (Northwestern Polytechnical University), Yimin Fu (Hong Kong Baptist University), Changyuan Wu (Northwestern Polytechnical University), and Zhunga Liu (Northwestern Polytechnical University)</i>	
B-cos Networks as an Architectural Inductive Bias for Mitigating Catastrophic Forgetting	5264
<i>Amelia Sorrenti (University of Catania, Italy), Giovanni Bellitto (University of Catania, Italy), Salvatore Calcagno (University of Catania, Italy), Rutger Hendrix (University of Catania, Italy), Concetto Spampinato (University of Catania, Italy), and Simone Palazzo (University of Catania, Italy)</i>	
Can Synthetic Images Conquer Forgetting? Beyond Unexplored Doubts in Few-Shot Class-Incremental Learning	5273
<i>Junsu Kim (Ulsan National Institute of Science and Technology ;NVIDIA Foundation Models Lab, MODULABS), Yunhoe Ku (DeepBrain AI ;NVIDIA Foundation Models Lab, MODULABS), and Seungryul Baek (Ulsan National Institute of Science and Technology)</i>	

9th AI Ctiy Challenge

Warehouse Spatial Question Answering with LLM Agent 1st Place Solution of the 9th AI City Challenge Track 3	5283
<i>Hsiang-Wei Huang (University of Washington, USA), Pyongkun Kim (Electronics and Telecommunications Research Institute, South Korea), Jen-Hao Cheng (University of Washington, USA), Kuang-Ming Chen (University of Washington, USA), Cheng-Yen Yang (University of Washington, USA), Bahaa Alattar (University of Washington, USA), Yi-Ru Lin (University of Washington, USA), Sangwon Kim (Electronics and Telecommunications Research Institute, South Korea), Kwangju Kim (Electronics and Telecommunications Research Institute, South Korea), Chung-I Huang (National Center for High-performance Computing, Taiwan), and Jenq-Neng Hwang (University of Washington, USA)</i>	
TrafficInternVL: Understanding Traffic Scenarios with Vision–Language Models	5288
<i>Hsiu-Fu Wu (Chunghwa Telecom Laboratories), Ya-Ting Yang (Chunghwa Telecom Laboratories), Yung-Ter Chen (Chunghwa Telecom Laboratories), and I-Fan Chou (Chunghwa Telecom Laboratories)</i>	
Enhanced Fisheye Object Detection via YOLO Ensemble Learning and Weighted Box Fusion	5296
<i>Chun-Ming Tsai (University of Taipei, Taiwan), Li-Li Wu (University of Taipei, Taiwan), and Tsung-Yu Chen (University of Taipei, Taiwan)</i>	
MCBLT: Multi-Camera Multi-Object 3D Tracking in Long Videos	5304
<i>Yizhou Wang (NVIDIA), Tim Meinhardt (NVIDIA), Orcun Cetintas (NVIDIA ;Technical University of Munich), Cheng-Yen Yang (NVIDIA ;University of Washington), Sameer S. Pusegaonkar (NVIDIA), Benjamin Missaoui (NVIDIA), Sujit Biswas (NVIDIA), Zheng Tang (NVIDIA), and Laura Leal-Taixé (NVIDIA)</i>	
A Unified Detection Pipeline for Robust Object Detection in Fisheye-Based Traffic Surveillance	5314
<i>Neema Jakisa Owor (University of Missouri–Columbia), Joshua Kofi Asamoah (North Dakota State University ;SMART Lab), Tanner Wambui Muturi (University of Missouri–Columbia), Anneliese Jakisa Owor (University of Missouri–Columbia), Blessing Agyei Kyem (North Dakota State University ;SMART Lab), Andrews Danyo (North Dakota State University ;SMART Lab), Yaw Adu-Gyamfi (University of Missouri–Columbia), and Armstrong Aboah (North Dakota State University ;SMART Lab)</i>	
Augmentation, Distillation and Optimization: A Practical Pipeline for Fisheye Object Detection on Edge Devices	5322
<i>Viet Hung Duong (VNPT AI ;Hanoi University of Science and Technology), Van Duy Truong (VNPT AI), Duy Khanh Dinh (VNPT AI), Huan Vu (National Economics University), Thien Van Luong (National Economics University), and Tien Cuong Nguyen (VNPT AI)</i>	
SmolRGPT: Efficient Spatial Reasoning for Warehouse Environments with 600M Parameters	5330
<i>Abdrahamane Traore (University of Moncton, Canada), Éric Hervet (University of Moncton, Canada), and Andy Couturier (University of Moncton, Canada)</i>	

Prompt-Guided Spatial Understanding with RGB-D Transformers for Fine-Grained Object Relation Reasoning	5339
<i>Tanner W. Muturi (University of Missouri-Columbia), Blessing A. Kyem (North Dakota State University), Joshua K. Asamoah (North Dakota State University), Neema J. Owor (University of Missouri-Columbia), Richard Dyzinela (Texas A&M), Andrews Danyo (North Dakota State University), Yaw Adu-Gyamfi (University of Missouri-Columbia), and Armstrong Aboah (North Dakota State University)</i>	
DepthTrack: Cluster Meets BEV for Multi-Camera Multi-Target 3D Tracking	5348
<i>Tai Huu-Phuong Tran (Sungkyunkwan University, South Korea), Duong Nguyen-Ngoc Tran (Sungkyunkwan University, South Korea), Ngoc Doan-Minh Huynh (Sungkyunkwan University, South Korea), Chi Dai Tran (Sungkyunkwan University, South Korea), Long Hoang Pham (Sungkyunkwan University, South Korea), Quoc Pham-Nam Ho (Sungkyunkwan University, South Korea), Huy-Hung Nguyen (Sungkyunkwan University, South Korea), Duong Khac Vu (Sungkyunkwan University, South Korea), Hyung-Min Jeon (Sungkyunkwan University, South Korea), Hyung-Joon Jeon (Sungkyunkwan University, South Korea), Son Hong Phan (Sungkyunkwan University, South Korea), Trinh Le Ba Khanh (Sungkyunkwan University, South Korea), and Jae Wook Jeon (Sungkyunkwan University, South Korea)</i>	
TrafficInternVL: Spatially-Guided Fine-Tuning with Caption Refinement for Fine-Grained Traffic Safety Captioning and Visual Question Answering	5358
<i>Sasin Phimsiri (AI and Robotics Ventures), Sarut Sunpawatr (AI and Robotics Ventures), Riu Cherdchusakulchai (Carnegie Mellon University), Pornprom Kiaojak (King Mongkut's Institute of Technology Ladkrabang), Teepakorn Tosawadi (AI and Robotics Ventures), Suchat Tungjitnob (AI and Robotics Ventures), Visarut Trairattanapa (AI and Robotics Ventures), Supawit Vatathanavaro (AI and Robotics Ventures), Wasu Kudisthalert (AI and Robotics Ventures), Chaitat Utintu (AI and Robotics Ventures), Worawit Saetan (AI and Robotics Ventures), Nathamon Kongsawat (AI and Robotics Ventures), Phawat Borisuitsawat (AI and Robotics Ventures), Kasisdis Mahakijdechachai (AI and Robotics Ventures), Nitipan Su-Inn (AI and Robotics Ventures), Ek Thamwivattthana (AI and Robotics Ventures), and Vasin Suttichaya (Mahidol University)</i>	
TrafficVILA: Scaling Vision-Language Models to High-Resolution Video Understanding for Traffic Safety Analysis	5366
<i>Zaid Pervaiz Bhat (NVIDIA, USA), Seunghwan Cha (NVIDIA, USA), Rohan Gulati (NVIDIA, USA ;UC Berkeley), Monika Jhuria (NVIDIA, USA), Varun Praveen (NVIDIA, USA), Tomasz Kornuta (NVIDIA, USA), Yao Lu (NVIDIA, USA), and Vidya Murali (NVIDIA, USA)</i>	
EKI-GAN: Context-Aware Vehicle Trajectory Forecasting with Vehicle Factors and Environmental Information at Signalized Intersections	5374
<i>Chuheng Wei (University of California, Riverside), Guoyuan Wu (University of California, Riverside), and Matthew J. Barth (University of California, Riverside)</i>	

Task-Specific Dual-Model Framework for Comprehensive Traffic Safety Video Description and Analysis	5384
<i>Blessing A. Kyem (North Dakota State University), Neema J. Owor (University of Missouri–Columbia), Andrews Danyo (North Dakota State University), Joshua K. Asamoah (North Dakota State University), Eugene Denteh (North Dakota State University), Tanner Muturi (University of Missouri–Columbia), Anthony Dontoh (University of Memphis), Yaw Adu-Gyamfi (University of Missouri–Columbia), and Armstrong Aboah (North Dakota State University)</i>	
Data Augmentation Is All You Need For Robust Fisheye Object Detection	5393
<i>Long Hoang Pham (Sungkyunkwan University), Quoc Pham-Nam Ho (Sungkyunkwan University), Duong Khac Vu (Sungkyunkwan Universit), Huy-Hung Nguyen (Sungkyunkwan University), Chi Dai Tran (Sungkyunkwan University), Duong Nguyen-Ngoc Tran (Sungkyunkwan University), Tai Huu-Phuong Tran (Sungkyunkwan University), Ngoc Doan-Minh Huynh (Sungkyunkwan University), Hyung Joon Jeon (Sungkyunkwan University), Hyung-Min Jeon (Sungkyunkwan University), Son Hong Phan (Sungkyunkwan University), Trinh Le Ba Khanh (Sungkyunkwan University), and Jae Wook Jeon (Sungkyunkwan University)</i>	
Hierarchical Multi-Modal Fusion for Roadside VRU Detection: Method Complementarity Under Sparse Label Constraints	5402
<i>Chuheng Wei (University of California, Riverside), Ziyang Zhang (University of California, Riverside), Haishan Liu (University of California, Riverside), Guoyuan Wu (University of California, Riverside), and Matthew J. Barth (University of California, Riverside)</i>	
Multimodal and Multi-task Fusion for Spatial Reasoning	5410
<i>Van-Minh Dang (Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, Vietnam) and Thanh-Sach Le (Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, Vietnam)</i>	
Multi-Camera 3D Object Tracking via 3D Point Clouds and Re-Identification	5417
<i>Jaewon Lee (ZIOVISION Co., Ltd., Republic of Korea), Heecheol Kim (ZIOVISION Co., Ltd., Republic of Korea), Doohee Lee (ZIOVISION Co., Ltd., Republic of Korea ;Kangwon National University, Republic of Korea), and Kanghee Lee (ZIOVISION Co., Ltd., Republic of Korea ;Kangwon National University, Republic of Korea)</i>	
Domain-Aware Enhancements to Vision-Language Models for Urban Traffic Safety Question Answering	5425
<i>Vu Thanh Dat Ha (VNPT IT), Tuan Huy Tran (VNPT IT), Gang Thiep Dong (VNPT IT), Ngoc Chien Chu (VNPT IT), Huan Vu (National Economics University), and Tien Cuong Nguyen (VNPT IT)</i>	

VGCRTTrack: Multi-Camera 3D Tracking with View-Aware Geometric Center Refinement	5434
<i>Trieu-Huy Phan (Ho Chi Minh City University of Technology and Education, Vietnam), Duc-Duy Dinh (Ho Chi Minh City University of Technology and Education, Vietnam), Thanh-Phong Huynh (Ho Chi Minh City University of Technology and Education, Vietnam), Quoc-Thinh Le (Ho Chi Minh City University of Technology and Education, Vietnam), Huy-Hoang Dang (Ho Chi Minh City University of Technology and Education, Vietnam), Vu-Hoang Tran (Ho Chi Minh City University of Technology and Education, Vietnam), Van-Tin Luu (National Yang Ming Chiao Tung University, Taiwan), and Ching-Chun Huang (National Yang Ming Chiao Tung University, Taiwan)</i>	
A Real-time Vehicle Detection Pipeline with Data-centric Enhancements and Multi-stage DETR Distillation	5441
<i>Huy Minh Nhat Nguyen (Vietnamese German University, Vietnam), Hieu Dinh Trung Pham (Vietnamese-German University, Vietnam), Khang Minh Le (Vietnamese-German University, Vietnam), and Cuong Tuan Nguyen (Vietnamese-German University, Vietnam)</i>	
Online 3D Multi-Camera Perception through Robust 2D Tracking and Depth-based Late Aggregation	5449
<i>Vu-Minh Le (Viettel Aerospace Institute ;University of Engineering and Technology, Vietnam National University), Thao-Anh Tran (Viettel Aerospace Institute ;Hanoi University of Science and Technology), Duc Huy Do (Viettel Aerospace Institute ;University of Engineering and Technology, Vietnam National University), Xuan Canh Do (University of Engineering and Technology, Vietnam National University), Huong Ninh (Viettel Aerospace Institute ;Hanoi University of Science and Technology), and Hai Tran (Viettel Aerospace Institute)</i>	
TrafficVILA: A Multimodal Framework for Traffic Safety Description and Analysis	5460
<i>Beomseok Park (Rutgers University, USA), Wanzhao Yang (Rutgers University, USA), Sifan Yuan (Rutgers University, USA), Syed Muhammad Anwar (Childrens National Hospital, USA; George Washington University, USA), and Ivan Marsic (Rutgers University, USA)</i>	
Efficient and Distortion-Aware Fisheye Object Detection for Edge Devices	5469
<i>Bao Tran Gia (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Tuong Bui Cong Khanh (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Tam Le Thi Thanh (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Hien Ho Trong (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Tien Do (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Thanh Duc Ngo (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Duy-Dinh Le (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), and Shin'ichi Satoh (National Institute of Informatics, Japan)</i>	
Real-Time Object Detection on Edge Devices: A Fisheye Specific DFINE	5476
<i>Jaewoo Park (Xiilab, Republic of Korea), Jihae Lee (Xiilab, Republic of Korea), and Yunjeong Yong (Xiilab, Republic of Korea)</i>	

Multi-Agent Cooperation for Traffic Safety Description and Analysis	5486
<i>Ridham Kachhadiya (Santa Clara University), Dhanishtha Patil (Santa Clara University), and David C Anastasiu (Santa Clara University)</i>	
Boosting Fisheye Detection with Augmentations and Ensembles	5495
<i>Chengzhi Qian (City University of Hong Kong, China), Jing Li (City University of Hong Kong, China), Yangyang Huang (City University of Hong Kong, China), Zhixin Lin (Shandong University, China), Yue Yao (Shandong University, China), and Jianping Wang (City University of Hong Kong, China)</i>	
A Lightweight and Data-Centric Framework for Real-Time Object Detection in Fisheye Camera ..	5500
<i>An To Vinh (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Nguyen Mai Vinh (University of Information Technology, Vietnam ; Vietnam National University, Vietnam), Nguyen Luong Si (University of Information Technology, Vietnam ; Vietnam National University, Vietnam), Duy Do Quoc (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Duy Tran Khanh (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Nguyen Nguyen Hoang (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Tien Do (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Thanh Duc Ngo (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), Duy-Dinh Le (University of Information Technology, Vietnam ;Vietnam National University, Vietnam), and Shin'ichi Satoh (National Institute of Informatics, Japan)</i>	
TinyGiantVLM: A Lightweight Vision-Language Architecture for Spatial Reasoning under Resource Constraints	5508
<i>Vinh-Thuan Ly (University of Science, Vietnam ;Vietnam National University, Vietnam), Hoang M. Truong (University of Science, Vietnam ;Vietnam National University, Vietnam), and Xuan-Huong Nguyen (University of Science, Vietnam ;Vietnam National University, Vietnam)</i>	
STER-VLM: Spatio-Temporal With Enhanced Reference Vision-Language Models	5516
<i>Tinh-Anh Nguyen-Nhu (Ho Chi Minh University Of Technology, Vietnam; Vietnam National University, Vietnam), Triet Dao Hoang Minh (Vietnamese-German University, Vietnam), Dat To-Thanh (University of Science, Vietnam; Ho Chi Minh University of Science, Vietnam), Phuc Le-Gia (University of Science, Vietnam; Ho Chi Minh University of Science, Vietnam), Tuan Vo-Lan (University of Science, Vietnam; Ho Chi Minh University of Science, Vietnam), and Tien-Huy Nguyen (University of Information and Technology, Vietnam; Ho Chi Minh University of Science, Vietnam)</i>	

The 9th AI City Challenge	5526
<i>Zheng Tang (NVIDIA Corporation, USA), Shuo Wang (NVIDIA Corporation, USA), David C. Anastasiu (Santa Clara University, USA), Ming-Ching Chang (University at Albany, USA), Anuj Sharma (Iowa State University, USA), Quan Kong (Woven by Toyota, Japan), Norimasa Kobori (Woven by Toyota, Japan), Munkhjargal Gochoo (United Arab Emirates University, UAE ;Emirates Center for Mobility Research, UAE), Ganzorig Batnasan (United Arab Emirates University, UAE), Munkh-Erdene Otgonbold (United Arab Emirates University, UAE ;Emirates Center for Mobility Research, UAE), Fady Alnajjar (United Arab Emirates University, UAE), Jun-Wei Hsieh (National Yang-Ming Chiao-Tung University, Taiwan), Tomasz Kornuta (NVIDIA Corporation, USA), Xiaolong Li (NVIDIA Corporation, USA), Yilin Zhao (NVIDIA Corporation, USA), Han Zhang (NVIDIA Corporation, USA), Subhashree Radhakrishnan (NVIDIA Corporation, USA), Arihant Jain (NVIDIA Corporation, USA), Ratnesh Kumar (NVIDIA Corporation, USA), Vidya N. Murali (NVIDIA Corporation, USA), Yuxing Wang (NVIDIA Corporation, USA), Sameer Satish Pusegaonkar (NVIDIA Corporation, USA), Yizhou Wang (NVIDIA Corporation, USA), Sujit Biswas (NVIDIA Corporation, USA), Xunlei Wu (NVIDIA Corporation, USA), Zhedong Zheng (University of Macau, Macau), Pranamesh Chakraborty (Indian Institute of Technology, India), and Rama Chellappa (Johns Hopkins University, USA)</i>	

Advances in Image Manipulation Workshop

DMS: Diffusion-Based Multi-Baseline Stereo Generation for Improving Self-Supervised Depth Estimation	5536
<i>Zihua Liu (Institute of Science Tokyo, Japan), Yizhou Li (Sony Semiconductor Solutions Group, Japan), Songyan Zhang (Nanyang Technological University, Singapore), and Masatoshi Okutomi (Institute of Science Tokyo, Japan)</i>	
ReBaIR: Reference-Based Image Restoration	5548
<i>Michael Bernasconi (ETH Zürich; Disney Research Zürich), Abdelaziz Djelouah (ETH Zürich; Disney Research Zürich), Yang Zhang (Disney Research), Markus Gross (ETH Zürich; Disney Research Zürich), and Christopher Schroers (Disney Research)</i>	
MedShift: Implicit Conditional Transport for X-Ray Domain Adaptation	5558
<i>Francisco Caetano (Eindhoven University of Technology, The Netherlands), Christiaan Viviers (Eindhoven University of Technology, The Netherlands), Peter H.N. De With (Eindhoven University of Technology, The Netherlands), and Fons van der Sommen (Eindhoven University of Technology, The Netherlands)</i>	
LMLT : Low-to-high Multi-Level Vision Transformer for Lightweight Image Super-Resolution	5568
<i>Jeongsoo Kim (Sogang University), Jongho Nang (Sogang University), and Junsuk Choe (Sogang University)</i>	
DIFFRACT: Diffusion-based Restoration via Adaptive Control and Thresholding for Diffraction Imaging	5579
<i>Nikolay Falaleev (Fanis Technologies, UK) and Nikolai Orlov (AMOLF, Netherlands)</i>	

JFFRA : Joint Flow And Feature Refinement Using Attention For Video Restoration	5589
<i>Ranjith Merugu (Samsung R&D Institute, India), Mohammed Sameer Suhail (Samsung R&D Institute, India), Akshay P Sarashetti (Samsung R&D Institute, India), Venkata Bharath Reddy Reddem (Samsung R&D Institute, India), Pankaj Kumar Bajpai (Samsung R&D Institute, India), and Amit Satish Unde (Samsung R&D Institute, India)</i>	
AIM 2025 Challenge on High FPS Motion Deblurring: Methods and Results	5600
<i>George Ciubotariu (University of Würzburg, Germany), Florin-Alexandru Vasluianu (University of Würzburg, Germany), Zhuyun Zhou (University of Würzburg, Germany), Nancy Mehta (University of Würzburg, Germany), Radu Timofte (University of Würzburg, Germany), Ke Wu (Nanjing University of Science and Technology, China), Long Sun (Nanjing University of Science and Technology, China), Lingshun Kong (Nanjing University of Science and Technology, China), Zhongbao Yang (Nanjing University of Science and Technology, China), Jinshan Pan (Nanjing University of Science and Technology, China), Jiangxin Dong (Nanjing University of Science and Technology, China), Jinhui Tang (Nanjing University of Science and Technology, China), Hao Chen (Nanjing University of Science and Technology, China), Yinghui Fang (Nanjing University of Science and Technology, China), Dafeng Zhang (Samsung R&D Institute China-Beijing (SRC-B)), Yongqi Song (Samsung R&D Institute China-Beijing (SRC-B)), Jiangbo Guo (Samsung R&D Institute China-Beijing (SRC-B)), Shuhua Jin (Samsung R&D Institute China-Beijing (SRC-B)), Zeyu Xiao (National University of Singapore), Rui Zhao (Nanyang Technology University), Zhuoyuan Li (University of Science and Technology of China), Cong Zhang (The Chinese University of Hong Kong), Yufeng Peng (University of Science and Technology of China, China), Xin Lu (University of Science and Technology of China, China), Zhijing Sun (University of Science and Technology of China, China), Chengjie Ge (University of Science and Technology of China, China), Zihao Li (University of Science and Technology of China, China), Zishun Liao (University of Science and Technology of China, China), Ziang Zhou (University of Science and Technology of China, China), Qiyu Kang (University of Science and Technology of China, China), Xueyang Fu (University of Science and Technology of China, China), Zheng-Jun Zha (University of Science and Technology of China, China), Yuqian Zhang (Xiaomi), Shuai Liu (Xiaomi), Jie Liu (Nanjing University), Zhuhao Zhang (Nanjing University), Lishen Qu (Nankai University), Zhihao Liu (Nankai University), Shihao Zhou (Nankai University), Yaqi Luo (Nankai University), Juncheng Zhou (Nankai University), Jufeng Yang (Nankai University), Qianfeng Yang (Dalian Polytechnic University), Qiyuan Guan (Dalian Polytechnic University), Xiang Chen (Dalian Polytechnic University), Guiyue Jin (Nanjing University of Science and Technology), and Jiyu Jin (Dalian Polytechnic University)</i>	

AIM 2025 Rip Current Segmentation (RipSeg) Challenge Report	5610
<i>Andrei Dumitriu (University of Bucharest, Romania), Florin Miron (University of Bucharest, Romania), Florin Tatui (University of Bucharest, Romania), Radu Tudor Ionescu (University of Bucharest, Romania), Radu Timofte (University of Wurzburg, Germany), Aakash Ralhan (University of Wurzburg, Germany), Florin-Alexandru Vasluianu (University of Wurzburg, Germany), Shenyang Qian (UNSW Sydney, Australia), Mitchell Harley (UNSW Sydney, Australia), Imran Razzak (Mohamed bin Zayed University of Artificial Intelligence, UAE), Yang Song (UNSW Sydney, Australia), Pu Luo (Xidian University, China), Yumei Li (Xidian University, China), Cong Xu (Xidian University, China), Jinming Chai (Xidian University, China), Kexin Zhang (Xidian University, China), Licheng Jiao (Xidian University, China), Lingling Li (Xidian University, China), Siqi Yu (Xidian University, China), Chao Zhang (Xidian University, China), Kehuan Song (Xidian University, China), Fang Liu (Xidian University, China), Puhua Chen (Xidian University, China), Xu Liu (Xidian University, China), Jin Hu (Xidian University, China), Jinyang Xu (Xidian University, China), and Biao Liu (Xidian University, China)</i>	
Diffusion-based Compression Quality Tradeoffs without Retraining	5620
<i>Jonas Brenig (University of Würzburg, Germany) and Radu Timofte (University of Würzburg, Germany)</i>	
AIM 2025 challenge on Inverse Tone Mapping Report: Methods and Results	5630
<i>Chao Wang (MPI-Info, Germany & Peng Cheng Laboratory, China), Francesco Banterle (ISTI-CNR, Italy), Bin Ren (University of Trento, Italy), Radu Timofte (University of Wurzburg, Germany), Xin Lu (University of Science and Technology of China, Hefei, China), Yufeng Peng (University of Science and Technology of China), Chengjie Ge (University of Science and Technology of China), Zhijing Sun (University of Science and Technology of China), Ziang Zhou (University of Science and Technology of China), Zihao Li (University of Science and Technology of China), Zishun Liao (University of Science and Technology of China), Qiyu Kang (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), Zheng-Jun Zha (University of Science and Technology of China), Zhijing Sun (University of Science and Technology of China), Xingbo Wang (University of Science and Technology of China), Kean Liu (University of Science and Technology of China), Senyan Xu (University of Science and Technology of China), Yang Qiu (University of Science and Technology of China), Yifan Ding (Linköping University, Sweden), Gabriel Eilertsen (Linköping University, Sweden), Jonas Unger (Linköping University, Sweden), Zihao Wang (Nanjing University of Science and Technology, China), Ke Wu (Nanjing University of Science and Technology, China), Jinshan Pan (Nanjing University of Science and Technology, China), Zhen Liu (University of Electronic Science and Technology of China, China), Zhongyang Li (University of Electronic Science and Technology of China, China), Shuaicheng Liu (University of Electronic Science and Technology of China, China), and S. M. Nadim Uddin (Deep In Sight Co., South Korea)</i>	

Efficient High FPS Non-Uniform Motion Deblurring via Progressive Learning	5644
<i>Xin Lu (University of Science and Technology of China), Zhijing Sun (University of Science and Technology of China), Chengjie Ge (University of Science and Technology of China), Yufeng Peng (University of Science and Technology of China), Ziang Zhou (University of Science and Technology of China), Zihao Li (University of Science and Technology of China), Zishun Liao (University of Science and Technology of China), Dong Li (University of Science and Technology of China), Qiyu Kang (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), and Zheng-Jun Zha (University of Science and Technology of China)</i>	
Boosting Inverse Tone Mapping via Diffusion Regularization	5654
<i>Xin Lu (University of Science and Technology of China), Yufeng Peng (University of Science and Technology of China), Chengjie Ge (University of Science and Technology of China), Zhijing Sun (University of Science and Technology of China), Ziang Zhou (University of Science and Technology of China), Zishun Liao (University of Science and Technology of China), Zihao Li (University of Science and Technology of China), Dong Li (University of Science and Technology of China), Qiyu Kang (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), and Zheng-Jun Zha (University of Science and Technology of China)</i>	
RCENet : Recursive Concatenation and Enhancement Network for Real-Time Super-Resolution ..	5664
<i>Kihwan Yoon (BLUEDOT), Ganzorig Gankhuyag (Korea Electronics Technology Institute, South Korea), and Jinwoo Jeong (Korea Electronics Technology Institute, South Korea)</i>	
AIM 2025 Challenge on Robust Offline Video Super-Resolution: Dataset, Methods and Results	5674
<i>Nikolai Karetin (Lomonosov Moscow State University, Russia), Ivan Molodetskikh (Lomonosov Moscow State University, Russia), Dmitry Vatolin (Lomonosov Moscow State University, Russia), Radu Timofte (University of Wurzburg, Germany), Yixin Yang (Nanjing University of Science and Technology, China), Junyang Chen (Nanjing University of Science and Technology, China), Jiangxin Dong (Nanjing University of Science and Technology, China), Jinshan Pan (Nanjing University of Science and Technology, China), Zhihao Liu, Lishen Qu (Nanjing University of Science and Technology, China), Shihao Zhou (Nanjing University of Science and Technology, China), Jufeng Yang (Nanjing University of Science and Technology, China), Yuxuan Jiang (University of Bristol, UK), Siyue Teng (University of Bristol, UK), Chengxi Zeng (University of Bristol, UK), Fan Zhang (University of Bristol, UK), David Bull (University of Bristol, UK), Qi Tang (Nanjing University, China), Jie Liu (State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China), Jie Tang (State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China), and Gangshan Wu (State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China)</i>	

LiteRT-Optimized INT8 LLM for Raspberry Pi4 Deployment	5683
<i>Kihwan Yoon (BLUEDOT), Hyeon-Cheol Moon (Korea Electronics Technology Institute), Aeri Kim (Korea Electronics Technology Institute), Sungjei Kim (Korea University of Technology and Education), Sang-Seol Lee (Korea Electronics Technology Institute), Sung-Joon Jang (Korea Electronics Technology Institute), Ganzorig Gankhuyag (Korea Electronics Technology Institute), and Jinwoo Jeong (Korea Electronics Technology Institute)</i>	
Multi-Scale Tensorial Summation and Dimensional Reduction Guided Neural Network for Edge Detection	5693
<i>Lei Xu (Tampere University, Finland), Mehmet Yamac (Tampere University, Finland), Mete Ahishali (Tampere University, Finland), and Moncef Gabbouj (Tampere University, Finland)</i>	
AIM 2025 Low-light RAW Video Denoising Challenge: Dataset, Methods and Results	5703
<i>Alexander Yakovenko (Lomonosov Moscow State University, Russia; MSU Institute for Artificial Intelligence, Russia), George Chakvetadze (Lomonosov Moscow State University, Russia), Ilya Khrapov (Lomonosov Moscow State University, Russia), Maksim Zhelezov (Lomonosov Moscow State University, Russia), Dmitry Vatolin (Lomonosov Moscow State University, Russia; MSU Institute for Artificial Intelligence, Russia), Radu Timofte (University of Würzburg, Germany), Youngjin Oh (Seoul National University, Republic of Korea), Junhyeong Kwon (Seoul National University, Republic of Korea), Junyoung Park (Seoul National University, Republic of Korea), Nam Ik Cho (Seoul National University, Republic of Korea), Senyan Xu (University of Science and Technology of China, China), Ruixuan Jiang (University of Science and Technology of China, China), Long Peng (University of Science and Technology of China, China), Xueyang Fu (University of Science and Technology of China, China), Zheng-Jun Zha (University of Science and Technology of China, China), Xiaoping Peng (Axera Semiconductor Co., Ltd, China), Hansen Feng (Beijing Institute of Technology, China), Zhanyi Tie (Beijing Institute of Technology, China), Ziming Xia (Beijing Institute of Technology, China), and Lizhi Wang (Beijing Normal University, China)</i>	
AIM 2025 Challenge on Screen-Content Video Quality Assessment: Methods and Results	5714
<i>Nickolay Safonov (MSU Institute for Artificial Intelligence, Russia), Mikhail Rakhmanov (Lomonosov Moscow State University, Russia), Dmitriy Vatolin (Lomonosov Moscow State University), Radu Timofte (University of Würzburg), Chunyu Wu (Guangxi Normal University, China), Kejing Wu (Guangxi Normal University, China), Kishor Patro (IIT Hyderabad, India), Pankaj Rathour (IIT Hyderabad, India), Sumohana S. Channappayya (IIT Hyderabad, India), Pravin Pardhi (Department of Electronics and Communication Engineering, VNIT Nagpur), Vipin Kamble (Department of Electronics and Communication Engineering, VNIT Nagpur), Kishor Bhurchandi (Department of Electronics and Communication Engineering, VNIT, India), Biao Liu (Xidian University, China), Jin Hu (Xidian University, China), Jinyang Xu (Xidian University, China), Yang Dayu (Guangxi Normal University, China), and Chen Yihua (Guangxi Normal University, China)</i>	

Optuna vs Code Llama: Are LLMs a New Paradigm for Hyperparameter Tuning?	5723
<i>Roman Kochnev (University of Würzburg, Germany), Arash Torabi Goodarzi (University of Würzburg, Germany), Zofia Antonina Bentyn (University of Würzburg, Germany), Dmitry Ignatov (University of Würzburg, Germany), and Radu Timofte (University of Würzburg, Germany)</i>	
Practical Manipulation Model for Robust Deepfake Detection	5734
<i>Benedikt Hopf (University of Würzburg, Germany) and Radu Timofte (University of Würzburg, Germany)</i>	
AIM 2025 Challenge on Real-World RAW Image Denoising	5744
<i>Feiran Li (Sony Research), Jiacheng Li (Sony Research), Marcos V. Conde (University of Würzburg), Beril Besbinar (Sony Research), Vlad Hosu (Sony Research), Daisuke Iso (Sony Research), and Radu Timofte (University of Würzburg)</i>	
Efficient Perceptual Image Super Resolution: AIM 2025 Study and Benchmark	5758
<i>Bruno Longarela (Cidaut AI, Spain), Marcos V Conde (Cidaut AI, Spain), Álvaro García (Cidaut AI, Spain), and Radu Timofte (University of Würzburg)</i>	
Efficient Real-World Deblurring using Single Images: AIM 2025 Challenge Report	5772
<i>Daniel Feijoo (Cidaut AI, Spain), Paula Garrido-Mellado (Cidaut AI, Spain), Marcos V. Conde (University of Würzburg; Cidaut AI, Spain), Jaesung Rim (POSTECH, Korea), Alvaro Garcia (Cidaut AI, Spain), Sunghyun Cho (POSTECH, Korea), and Radu Timofte (University of Würzburg)</i>	

BioImage Computing

Probabilistic Domain Adaptation for Biomedical Image Segmentation	5781
<i>Anwai Archit (Georg-August-University Göttingen; Campus Institute Data Science (CIDAS), Germany) and Constantin Pape (Georg-August-University Göttingen; Campus Institute Data Science (CIDAS), Germany; CAIMed - Lower Saxony Center for AI & Causal Methods in Medicine, Göttingen; Georg-August-University Göttingen)</i>	
From Natural to Nanoscale: Training ControlNet on Scarce FIB-SEM Data for Augmenting Semantic Segmentation Data	5790
<i>Hannah Kniesel (Ulm University), Pascal Rapp (Ulm University), Pedro Hermosilla (TU Vienna), and Timo Ropinski (Ulm University)</i>	
A comparison of data-driven shape quantification methods for 2D microscopy images	5800
<i>Anna Foix Romero (EMBL-EBI, UK), Alexander Krull (University of Birmingham, UK), and Virginie Uhlmann (University of Zurich, CH)</i>	
NeuReg: Domain-Invariant Brain Registration Across Species and Modalities	5810
<i>Taha Razzaq (Tibbling Technologies) and Asim Iqbal (Tibbling Technologies)</i>	
Topology-Preserving Image Segmentation with Spatial-Aware Persistent Feature Matching	5821
<i>Bo Wen (University of California, San Diego), Haochen Zhang (University of California, San Diego), Dirk-Uwe G. Bartsch (University of California, San Diego), William R. Freeman (University of California, San Diego), Truong Q. Nguyen (University of California, San Diego), and Cheolhong An (University of California, San Diego)</i>	

cubic: CUDA-accelerated 3D BioImage Computing	5831
<i>Alexandr A. Kalinin (Broad Institute of MIT and Harvard, USA), Anne E. Carpenter (Broad Institute of MIT and Harvard, USA), Shantanu Singh (Broad Institute of MIT and Harvard, USA), and Matthew J. O’Meara (University of Michigan Medical School, USA)</i>	
Preserving instance continuity and length in segmentation through connectivity-aware loss computation	5841
<i>Karol Szustakowski (Helmholtz Munich, Germany), Luk Frank (German Center for Neurodegenerative Diseases (DZNE), Germany), Julia Esser (German Center for Neurodegenerative Diseases (DZNE), Germany), Jan Gründemann (German Center for Neurodegenerative Diseases (DZNE), Germany; University of Bonn, Germany)], and Marie Piraud (Helmholtz Munich, Germany)</i>	
An Investigation of Unsupervised Cell Tracking and Interactive Fine-Tuning	5851
<i>Manan Lalit (HHMI Janelia Research Campus, USA) and Jan Funke (HHMI Janelia Research Campus, USA)</i>	
SynapFlow: A Modular Framework Towards Large-Scale Analysis of Dendritic Spines	5860
<i>Pamela Osuna-Vargas (Frankfurt Institute for Advanced Studies, Germany; Goethe University Frankfurt, Germany), Altug Kamacioglu (University Medical Center, Germany), Dominik F. Aschauer (University Medical Center, Germany), Petros E. Vlachos (Frankfurt Institute for Advanced Studies, Germany), Sercan Alipek (Frankfurt Institute for Advanced Studies, Germany; Universität Siegen), Jochen Triesch (Frankfurt Institute for Advanced Studies, Germany; Goethe University Frankfurt, Germany), Simon Rumpel (University Medical Center, Germany), and Matthias Kaschube (Frankfurt Institute for Advanced Studies, Germany; University Frankfurt, Germany)</i>	
SelfAdapt: Unsupervised Domain Adaptation of Cell Segmentation Models	5871
<i>Fabian H. Reith (Charité - Universitätsmedizin, Germany; Humboldt-Universität zu Berlin, Germany; Helmholtz Imaging; Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Germany), Jannik Franzen (Charité - Universitätsmedizin, Germany; Universität Potsdam, Digital Engineering Faculty, Germany; Helmholtz Imaging; Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Germany), J. Lorenz Rumberger (Humboldt-Universität zu Berlin, Germany; Helmholtz Imaging; Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Germany), and Dagmar Kainmueller (Universität Potsdam, Digital Engineering Faculty, Germany; Helmholtz Imaging; Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Germany)</i>	

6th Workshop on Closing the Loop Between Vision and Language (Decade Mark)

Noise is an Efficient Learner for Zero-Shot Vision-Language Models	5879
<i>Raza Imam (Mohamed bin Zayed University of Artificial Intelligence, UAE), Asif Hanif (Mohamed bin Zayed University of Artificial Intelligence, UAE), Jian Zhang (Mohamed bin Zayed University of Artificial Intelligence, UAE), Khaled Waleed Dawoud (Mohamed bin Zayed University of Artificial Intelligence, UAE), Yova Kementchedjheva (Mohamed bin Zayed University of Artificial Intelligence, UAE), and Mohammad Yaqub (Mohamed bin Zayed University of Artificial Intelligence, UAE)</i>	
STORM: Token-Efficient Long Video Understanding for Multimodal LLMs	5889
<i>Jindong Jiang (NVIDIA; Rutgers University), Xiuyu Li (NVIDIA; UC Berkeley), Zhijian Liu (NVIDIA), Muyang Li (NVIDIA; MIT), Guo Chen (NVIDIA; Nanjing University), Zhiqi Li (NVIDIA; Nanjing University), De-An Huang (NVIDIA), Guilin Liu (NVIDIA), Zhiding Yu (NVIDIA), Kurt Keutzer (UC Berkeley), Sungjin Ahn (KAIST), Jan Kautz (NVIDIA), Hongxu Yin (NVIDIA), Yao Lu (NVIDIA), Song Han (NVIDIA, MIT), and Wonmin Byeon (NVIDIA)</i>	
Context-Aware Image Caption Editing via Hallucination-Resistant Visual Instruction Tuning	5901
<i>Yoonhyung Kim (Electronics and Telecommunications Research Institute (ETRI)), Byung Ok Kang (Electronics and Telecommunications Research Institute (ETRI)), and Hwa Jeon Song (Electronics and Telecommunications Research Institute (ETRI))</i>	
PLOT-TAL: Prompt-Learning with Optimal Transport for Few-Shot Temporal Action Localization.....	5912
<i>Edward Fish (University of Surrey) and Andrew Gilbert (University of Surrey)</i>	
AutoConcept: Unsupervised Extraction of Constituent Concepts from Single Image	5922
<i>Pranav Singh Chib (Indian Institute Of Technology, India), Kirtankumar Vijaykumar Patel (Indian Institute Of Technology, India), Mudit Gupta (Indian Institute Of Technology, India), Pise Ashutosh Kalidas (Indian Institute Of Technology, India), and Pravendra Singh (Indian Institute Of Technology, India)</i>	

Computer Vision for Biometrics, Identity & Behaviour

Demographic Differentials in Face Image Quality: Evaluation and Comparison on Real and Synthetic Data	5933
<i>Andre Dörsch (Hochschule Darmstadt, Germany), Johannes Merkle (secunet Security Networks AG, Germany), Benjamin Tams (secunet Security Networks AG, Germany), Gerardo Gutierrez Alvarez (secunet Security Networks AG, Germany), Peter Munch (Hochschule Darmstadt, Germany), Christoph Busch (Hochschule Darmstadt, Germany), and Christian Rathgeb (Hochschule Darmstadt, Germany ;secunet Security Networks AG, Germany)</i>	
Intrinsically-Interpretable Siamese Networks for Identity Recognition	5942
<i>Marco A. Rocha (Universidade do Porto, Portugal), Jaime S. Cardoso (Universidade do Porto, Portugal), and Helena Montenegro (Universidade do Porto, Portugal)</i>	

Evaluation of Human Visual Privacy Protection: A Three-Dimensional Framework and Benchmark Dataset	5952
<i>Sara Abdulaziz (Eindhoven University of Technology, The Netherlands), Giacomo D'amicantonio (Eindhoven University of Technology, The Netherlands), and Egor Bondarev (Eindhoven University of Technology, The Netherlands)</i>	
TAIGen: Training-Free Adversarial Image Generation via Diffusion Models	5962
<i>Susim Roy (University at Buffalo ;IIT Jodhpur), Anubhooti Jain (IIT Jodhpur), Mayank Vatsa (IIT Jodhpur), and Richa Singh (IIT Jodhpur)</i>	
NegFaceDiff: The Power of Negative Context in Identity-Conditioned Diffusion for Synthetic Face Generation	5973
<i>Eduarda Caldeira (Fraunhofer IGD, Germany), Naser Damer (Fraunhofer IGD, Germany ;TU Darmstadt, Germany), and Fadi Boutros (Fraunhofer IGD, Germany)</i>	
Now You See Me, Now You Don't: A Unified Framework for Expression Consistent Anonymization in Talking Head Videos	5984
<i>Anil Egin (Inria Center at Université Côte d'Azur, France), Andrea Tangherloni (Bocconi University, Italy), and Antitza Dantcheva (Inria Center at Université Côte d'Azur, France)</i>	
ViT-FIQA: Assessing Face Image Quality using Vision Transformers	5994
<i>Andrea Atzori (Fraunhofer Institute for Computer Graphics Research IGD, Germany), Fadi Boutros (Fraunhofer Institute for Computer Graphics Research IGD, Germany), and Naser Damer (Fraunhofer Institute for Computer Graphics Research IGD, Germany ;TU Darmstadt, Germany)</i>	
Are you In or Out (of gallery)? Wisdom from the Same-Identity Crowd	6005
<i>Aman Bhatta (University of Notre Dame), Maria Dhakal (University of Notre Dame), Michael C. King (Florida Institute of Technology), and Kevin W. Bowyer (University of Notre Dame)</i>	
On Adversarial Robustness of Face Presentation Attack Detection Algorithms	6015
<i>Akshay Agarwa (IISER Bhopal, India), Mayank Vatsa (IIT Jodhpur, India), and Richa Singh (IIT Jodhpur, India)</i>	

Computer Vision in Advertising and Marketing

AI vs. Human Moderators: A Comparative Evaluation of Multimodal LLMs in Content Moderation for Brand Safety	6024
<i>Adi Levi (Zefr Inc), Or Levi (Zefr Inc), Sardhendu Mishra (Zefr Inc), and Jonathan Morra (Zefr Inc)</i>	
A Survey of Human Synergy in Influencer Marketing Through Authenticity-Preserving Content Generation Approaches	6033
<i>Nafi Diallo (Amazon, USA) and Pegah Ojaghi (Amazon, USA)</i>	
CLASP: Adaptive Spectral Clustering for Unsupervised Per-Image Segmentation	6041
<i>Max Curie (Integral Ad Science, USA) and Paulo da Costa (Integral Ad Science, USA)</i>	
Seeing the Signs: A Survey of Edge-Deployable OCR Models for Billboard Visibility Analysis.....	6051
<i>Maciej Szankin (SiMa.ai, CA), Vidhyananth Venkatasamy (SiMa.ai, CA), and Lihang Ying (University of Alberta)</i>	

Privacy-Preserving Audience Analytics: Lightweight Thermal Face Recognition for Real-Time Marketing Intelligence at the Edge	6060
<i>Maciej Szankin (SiMa.ai, CA) and Jacek Ruminski (Gdansk University of Technology, Poland, Biomedical Engineering Department)</i>	
Align Before You Recommend: Parameter Efficient Personalization via Cross-Attentive Fusion of Hierarchical Language Models	6069
<i>Alicja Kwasniewska (Amazon Web Services), Marcin Bednarz (Amazon Web Services), and Chad Neal (Amazon Web Services)</i>	
Training-Free Diffusion Framework for Stylized Image Generation with Identity Preservation.....	6079
<i>Mohammad Ali Rezaei (Part AI Research Center, Iran), Helia Hajikazem (Part AI Research Center, Iran), Saeed Khanehgir (Part AI Research Center, Iran), and Mahdi Javanmardi (Amirkabir University of Technology, Iran)</i>	
A Multi-Stage Pipeline for Accurate Handwritten Information Extraction from Financial Forms	6089
<i>Guanghui Wang (AWS Generative AI Innovation Center), Jinze Yu (AWS Generative AI Innovation Center), Xing Zhang (AWS Generative AI Innovation Center), Tomal Deb (AWS Generative AI Innovation Center), Xuefeng Liu (AWS Generative AI Innovation Center), and Peiyang He (AWS Generative AI Innovation Center)</i>	
MGT: Extending Virtual Try-Off to Multi-Garment Scenarios	6098
<i>Riza Velioglu (Bielefeld University, Germany), Petra Bevandic (Bielefeld University, Germany), Robin Chan (Bielefeld University, Germany), and Barbara Hammer (Bielefeld University, Germany)</i>	
Cross-lingual Visual Text Stylization and Synthesis Incorporating Text Rendering and Diffusion Model	6108
<i>Minmin Shen (Amazon) and Caren Chen (Amazon)</i>	
ByDeWay: Boost Your multimodal LLM with DEpth prompting in a Training-Free Way	6117
<i>Rajarshi Roy (Kalyani Government Engineering College, India), Deoleena Das (Kalyani Government Engineering College, India), Ankesh Banerjee (Kalyani Government Engineering College, India), Arjya Bhattacharjee (Kalyani Government Engineering College, India), Kousik Dasgupta (Kalyani Government Engineering College, India), and Subarna Tripathi (Intel Labs, USA)</i>	
Reasoning-Enhanced Prompt Strategies for Multi-Label Classification	6124
<i>Jinze Yu (AWS Generative AI Innovation Center)</i>	
From Pixels to Context: Adapting Generative Models for Advertising at Scale	6130
<i>Hyunhee Chung (SK Telecom Republic of Korea) and Taeyoung Na (SK telecom Republic of Korea)</i>	
Toward Scalable Video Narration: A Training-free Approach Using Multimodal Large Language Models	6138
<i>Tz-Ying Wu (Intel), Tahani Trigui (Intel), Sharath Nittur Sridhar (Intel), Anand Bodas (Intel), and Subarna Tripathi (Intel)</i>	

Findings of ICCV

Similarity-Aware Selective State-Space Modeling for Semantic Correspondence	6147
<i>Seungwook Kim (Pohang University of Science and Technology (POSTECH), South Korea) and Minsu Cho (Pohang University of Science and Technology (POSTECH), South Korea)</i>	
Video-MMLU: A Massive Multi-Discipline Lecture Understanding Benchmark	6158
<i>Enxin Song (Zhejiang University), Wenhao Chai (University of Washington), Weili Xu (Zhejiang University), Jianwen Xie (Lambda, Inc.), Yuxuan Liu (Zhejiang University), and Gaoang Wang (Zhejiang University)</i>	
Zero-shot Customized Video Editing with Diffusion Feature Transfer	6173
<i>Wei Chen (Purdue University), Huidong Liu (Amazon), Yang Liu (Amazon), Chien-Chih Wang (Amazon), Moyan Li (Amazon), Hongdong Li (Australian National University), and Bryan Wang (Amazon)</i>	
BLIP-3: A Family of Open Large Multimodal Models	6183
<i>Le Xue (Salesforce), Manli Shu (Salesforce Research), Anas Awadalla (Salesforce AI Research), Jun Wang (Salesforce Research), An Yan (Salesforce AI Research), Senthil Purushwalkam (Salesforce AI Research), Honglu Zhou (Salesforce.com), Viraj Prabhu (Salesforce AI Research), Yutong Dai (Salesforce AI Research), Michael S Ryoo (Salesforce AI Research), Shrikant Kendre (Salesforce AI Research), Jieyu Zhang (University of Washington), Shaoyen Tseng (Intel), Gustavo A Lujan-Moreno (Intel Labs), Matthew L Olson (Intel Labs), Musashi Hinck (Intel Labs), David Cobbley (Intel Labs), Vasudev Lal (Intel Labs), Can Qin (Salesforce AI Research), Shu Zhang (University of Washington), Chia-Chih Chen (Salesforce AI Research), Ning Yu (Salesforce AI Research), Juntao Tan (Salesforce AI Research), Tulika Manoj Atwalgaonkar (Salesforce AI Research), Shelby Heinecke (Salesforce AI Research), Huan Wang (Salesforce AI Research), Yejin Choi (University of Washington), Ludwig Schmidt (University of Washington), Zeyuan Chen (Salesforce AI Research), Silvio Savarese (Salesforce AI Research), Juan Carlos Niebles (Salesforce AI Research), Caiming Xiong (Salesforce AI Research), and Ran Xu (Salesforce AI Research)</i>	
Zero-Shot Subject-Centric Generation for Creative Application Using Entropy Fusion	6195
<i>Kaifeng Zou (LINK-TO, China), Xiaoyi Feng (Macau University of Science and Technology, China), Tao Huang (LINK-TO, China), Zizhou Huang (LINK-TO, China), Haihang Zhang (LINK-TO, China), Yuntao Zou (Macau University of Science and Technology, China), and Dagang Li (Macau University of Science and Technology, China)</i>	
Multimodal Representation Alignment for Image Generation: Text-Image Interleaved Control Is Easier Than You Think	6205
<i>Liang Chen (Peking University), Shuai Bai (Alibaba Group), Wenhao Chai (University of Washington), Weichu Xie (Beijing Institute of Technology), Haozhe Zhao (Peking University), Leon Vinci (Bainance Labs), Junyang Lin (Alibaba Group), and Baobao Chang (Peking University)</i>	

CST Anti-UAV: A Thermal Infrared Benchmark for Tiny UAV Tracking in Complex Scenes	6216
<i>Bin Xie (Nanchang Hangkong University, China), Congxuan Zhang (Nanchang Hangkong University, China), Fagan Wang (Nanchang Hangkong university, China), Peng Liu (Beihang University, China), Feng Lu (Nanchang Hangkong University, China), Zhen Chen (Nanchang Hangkong University, China), and Weiming Hu (Chinese academy of science, China)</i>	
Probing the Representational Power of Sparse Autoencoders in Vision Models	6226
<i>Matthew L. Olson (Intel Labs, Santa Clara, CA, USA), Musashi Hinck (Intel, USA), Neale Ratzlaff (Intel, USA), Changbai Li (Oregon State University, USA), Phillip Howard (Intel Labs, Santa Clara, USA), Vasudev Lal (Intel, USA), and Shao-Yen Tseng (Intel, USA)</i>	
M3DocVQA: Multi-modal Multi-page Multi-document Understanding	6237
<i>Jaemin Cho (University of North Carolina at Chapel Hill), Debanjan Mahata (Bloomberg), Ozan Irsoy (Bloomberg), Yujie He (Bloomberg), and Mohit Bansal (University of North Carolina at Chapel Hill)</i>	
From Flat to Round: Redefining Brain Decoding with Surface-Based fMRI and Cortex Structure...	6248
<i>Sijin Yu (South China University of Technology), Zijiao Chen (Stanford University), Wenxuan Wu (South China University of Technology), Shengxian Chen (South China University of Technology), Zhongliang Liu (South China University of Technology), Jingxin Nie (South China Normal University), Xiaofen Xing (South China University of Technology), Xiangmin Xu (South China University of Technology), and Xin Zhang (South China University of Technology)</i>	
Debias your Large Multi-Modal Model at Test-Time via Non-Contrastive Visual Attribute Steering	6258
<i>Neale Ratzlaff (Intel Labs, CA), Matthew Lyle Olson (Intel Labs, CA), Musashi Hinck (Intel Labs, CA), Shao-Yen Tseng (Intel Labs, CA), Vasudev Lal (Intel Labs, CA), and Phillip Howard (Intel Labs, CA)</i>	
DEARLi: Decoupled Enhancement of Recognition and Localization for Semi-supervised Panoptic Segmentation	6268
<i>Ivan Martinović (Faculty of Electrical Engineering and Computing, University of Zagreb), Josip Šarić (Faculty of Computer and Information Science, University of Ljubljana), Marin Oršić (Faculty of Electrical Engineering and Computing, University of Zagreb), Matej Kristan (Faculty of Computer and Information Science, University of Ljubljana), and Siniša Šegvić (Faculty of Electrical Engineering and Computing, University of Zagreb)</i>	
Multi-Objective Optimization for Deep Neural Network Calibration	6280
<i>Dexter Neo (National University of Singapore) and Tsuhan Chen (National University of Singapore)</i>	
SCRAMBLE : Enhancing Multimodal LLM Compositionality with Synthetic Preference Data	6292
<i>Samarth Mishra (Boston University), Kate Saenko (Boston University), and Venkatesh Saligrama (Boston University)</i>	
BadPatch: Diffusion-Based Generation of Physical Adversarial Patches	6303
<i>Zhixiang Wang (Fudan University), Xingjun Ma (Fudan University), and Yu-Gang Jiang (Fudan University)</i>	

AdCorDA: Classifier Refinement via Adversarial Correction and Domain Adaptation	6314
<i>Lulan Shen (McGill University), Ali Edalati (Cohere), Xiangyu Li (Seyso), Brett H. Meyer (McGill University), Warren Gross (McGill University), and James J. Clark (McGill University)</i>	
ORXE: Orchestrating Experts for Dynamically Configurable Efficiency	6324
<i>Qingyuan Wang (University College Dublin, Ireland), Guoxin Wang (University College Dublin, Ireland), Barry Cardiff (University College Dublin, Ireland), and Deepu John (University College Dublin, Ireland)</i>	
SynBalance: Harnessing Synthetic Data in Long-tailed Recognition	6335
<i>Zhongyu Jiang (Apple), Jiarui Cai (Amazon), Chang Liu (Amazon), Dongsheng An (Amazon), and Jonathan Wu (Amazon)</i>	
CAFE: Unifying Representation and Generation with Contrastive-Autoregressive Finetuning	6345
<i>Hao Yu (Meta), Zhuokai Zhao (Meta), Shen Yan (Meta), Lukasz Korycki (Meta), Jianyu Wang (Meta), Baosheng He (Meta), Jiayi Liu (Meta), Lizhu Zhang (Meta), Xiangjun Fan (Meta), and Hanchao Yu (Meta)</i>	
GeoDiff: Geometry-Guided Diffusion for Metric Depth Estimation	6357
<i>Tuan Pham (University of California, Irvine), Thanh Tung Le (University of California, Irvine), Xiaohui Xie (Equal contribution), and Stephan Mandt (Equal advising)</i>	
Data Leakage in Visual Datasets	6368
<i>Patrick Ramos (The University of Osaka, Japan), Ryan Ramos (The University of Osaka, Japan), and Noa Garcia (The University of Osaka, Japan)</i>	
IDMR: Towards Instance-Driven Precise Visual Correspondence in Multimodal Retrieval	6379
<i>Bangwei Liu (East China Normal University), Yicheng Bao (East China Normal University), Shaohui Lin (East China Normal University), Xuhong Wang (Shanghai AI Laboratory), Xin Tan (East China Normal University), Yingchun Wang (Shanghai AI Laboratory), Yuan Xie (East China Normal University), and Chaochao Lu (Shanghai AI Laboratory)</i>	
MIDAS: Modeling Ground-Truth Distributions with Dark Knowledge for Domain Generalized Stereo Matching	6389
<i>Peng Xu (Zhejiang University, China), Zhiyu Xiang (Zhejiang University, China), Jingyun Fu (Zhejiang University, China), Tianyu Pu (Zhejiang University, China), Hanzhi Zhong (Zhejiang University, China), and Eryun Liu (Zhejiang University, China)</i>	
OpenInsGaussian: Open-vocabulary Instance Gaussian Segmentation with Context-aware Cross-view Fusion	6400
<i>Tianyu Huang (University of Sydney), Runnan Chen (University of Sydney), Dongting Hu (University of Melbourne), Fengming Huang (University of Sydney), Mingming Gong (University of Melbourne), and Tongliang Liu (University of Sydney)</i>	
Scaling Open-Vocabulary Action Detection	6410
<i>Zhen Hao Sia (University of Central Florida) and Yogesh Singh Rawat (University of Central Florida)</i>	

CAM-Seg: A Continuous-valued Embedding Approach for Semantic Image Generation	6421
<i>Masud Ahmed (University of Maryland, Baltimore County, USA), Zahid Hasan (University of Maryland, Baltimore County, USA), Syed Arefinul Haque (Northeastern University, USA), Abu-Zaher Faridee (Amazon, USA), Sanjay Purushotham (University of Maryland, Baltimore County, USA), Suyu You (DEVCOM Army Research Laboratory, USA), and Nirmalya Roy (University of Maryland, Baltimore County, USA)</i>	
NAS just once: Neural Architecture Search for joint Image-Video Recognition	6431
<i>Sofia Casarin (Free University of Bozen, Italy), Sergio Escalera (Universitat de Barcelona, Spain), and Oswald Lanz (Free University of Bozen-Bolzano, Italy)</i>	
DNF-Avatar: Distilling Neural Fields for Real-time Animatable Avatar Relighting	6442
<i>Zeren Jiang (University of Oxford), Shaofei Wang (Department of Computer Science, Swiss Federal Institute of Technology), and Siyu Tang (Department of Computer Science, Swiss Federal Institute of Technology)</i>	
SPIE: Semantic and Structural Post-Training of Image Editing Diffusion Models with AI feedback	6454
<i>Elior Benarous (Harvard University), Yilun Du (Harvard University), and Heng Yang (Harvard University)</i>	
LOCAL: Latent Orthonormal Contrastive Learning for Paired Image Classification	6467
<i>Fei Dou (University of Georgia), Jin Lu (University of Georgia), Tan Zhu (University of Connecticut), and Jinbo Bi (University of Connecticut)</i>	
Tuning-Free Multi-Event Long Video Generation via Synchronized Coupled Sampling	6477
<i>Subin Kim (Korea Advanced Institute of Science & Technology), Seoung Wug Oh (Adobe Research), Jui-Hsien Wang (Adobe Research), Joon-Young Lee (Adobe Research), and Jinwoo Shin (Korea Advanced Institute of Science & Technology)</i>	

Generating Digital Twins from Images and Videos

3D Gaussian Representations with Motion Trajectory Field for Monocular Dynamic Scene Reconstruction	6489
<i>Xuesong Li (CSIRO ;The Australian National University), Lars Petersson (CSIRO), and Vivien Rolland (CSIRO)</i>	
Improving Viewpoint Consistency in 3D Generation via Structure Feature and CLIP Guidance ...	6499
<i>Qing Zhang (The Australian National University), Jinguang Tong (The Australian National University ;CSIRO), Jing Zhang (The Australian National University), Jie Hong (The University of Hong Kong), and Xuesong Li (The Australian National University ;CSIRO)</i>	
Articulated Object Understanding from a Single Video Sequence	6509
<i>Arslan Artykov (École des Ponts et Chaussées, France), Clémentin Boittiaux (École des Ponts et Chaussées, France), and Vincent Lepetit (École des Ponts et Chaussées, France)</i>	

Multi-Modal Reasoning for Agentic Intelligence

- Temporal Multimodal Memory Banks for Agentic Reasoning 6519
Prasanth Yadla (Independent Researcher, USA)
- Collaborative Multimodal Agent Networks: Dynamic Specialization and Emergent Communication
for Complex Scene Understanding 6527
Prasanth Yadla (Independent Researcher, USA)

Multimodal Reasoning and Slow Thinking in Large Model Era: Towards System 2 and Beyond

- How Well Do Vision–Language Models Understand Cities? A Comparative Study on Spatial
Reasoning from Street-View Images 6535
*Juneyoung Ro (Korea Advanced Institute of Science and Technology),
Namwoo Kim (Korea Advanced Institute of Science and Technology), and
Yoonjin Yoon (Korea Advanced Institute of Science and Technology)*
- Multimodal Dual-domain Learning for Image Fusion 6545
*Heng Wang (Northwestern Polytechnical University, China), Mingxin Jin
(Northwestern Polytechnical University, China), Cong Wang
(Northwestern Polytechnical University, China), and Yuan Yuan
(Northwestern Polytechnical University, China)*
- Multimodal Monocular Geometric Clues-Informed Multi-View Stereo Network 6555
*Wanjuan Su (Interdisciplinary Artificial Intelligence Research
Institute, China) and Shili Xiong (Wuhan University of Technology,
China; Interdisciplinary Artificial Intelligence Research Institute,
China)*
- Test-time Prompt Refinement for Text-to-Image Models 6565
*Mohammad Abdul Hafeez Khan (Florida Institute of Technology, USA),
Yash Jain (Microsoft Research, USA), Siddhartha Bhattacharyya (Florida
Institute of Technology, USA), and Vibhav Vineet (Microsoft Research,
USA)*

Peng Xu (Tsinghua University), Shengwu Xiong (Wuhan University of Technology), Jiajun Zhang (Institute of Automation, Chinese Academy of Sciences), Yaxiong Chen (Wuhan University of Technology), Bowen Zhou (Tsinghua University), Chen Change Loy (Nanyang Technological University), David A. Clifton (University of Oxford), Kyoung Mu Lee (Seoul National University), Luc Van Gool (INSAIT, Sofia University St. Kliment Ohridski), Ruiming He (Wuhan University of Technology), Ruilin Yao (Wuhan University of Technology), Xinwei Long (Tsinghua University), Jirui Huang (Wuhan University of Technology), Kai Tian (Tsinghua University), Sa Yang (Peking University), Yihua Shao (Institute of Automation, Chinese Academy of Sciences), Jin Feng (KuaiShou Inc.), Yue Zhong (China University of Political Science and Law), Jiakai Zhou (Tsinghua University), Cheng Tang (Tsinghua University), Tianyu Zou (Wuhan University of Technology), Yifang Zhang (Wuhan University of Technology), Junming Liang (Wuhan University of Technology), Guoyou Li (Wuhan University of Technology), Zhaoxiang Wang (Wuhan University of Technology), Qiang Zhou (52CV Computer Vision Academic Community), Yichen Zhao (Wuhan University of Technology), Shili Xiong (Wuhan University of Technology), Hyeongjin Nam (Seoul National University), Jaerin Lee (Seoul National University), Jaeyoung Chung (Seoul National University), JoonKyu Park (Seoul National University), Junghun Oh (Seoul National University), Kanggeon Lee (Seoul National University), Wooseok Lee (Seoul National University), Juneyoung Ro (Korea Advanced Institute of Science and Technology), Turghun Osman (Xinjiang Technical Institute of Physics & Chemistry of the Chinese Academy of Sciences), Can Hu (MGTV Corporation, Changsha, China), Chaoyang Liao (Xidian University), Cheng Chen (Samsung Electronics (China) R&D Center), Chengcheng Han (Meituan-M17), Chenhao Qiu (MGTV Corporation, Changsha, China), Chong Peng (Meituan-M17), Cong Xu (Xidian University), Dailin Li (Dalian University of Technology), Feiyu Wang (Beijing Language and Culture University), Feng Gao (Institute of Automation, Chinese Academy of Sciences), Guibo Zhu (Institute of Automation, Chinese Academy of Sciences), Guopeng Tang (Tianjin University of Technology and Education), Haibo Lu (MGTV Corporation, Changsha, China), Han Fang (Institute of Artificial Intelligence, China Telecom), Han Qi (MGTV Corporation, Changsha, China), Hanxiao Wu (Institute of Automation, Chinese Academy of Sciences), Haobo Cheng (University of Science and Technology of China), Hongbo Sun (Institute of Artificial Intelligence, China Telecom), Hongyao Chen (Jiangnan University), Huayong Hu (Jilin University), Hui Li (Institute of Artificial Intelligence, China Telecom), Jiaheng Ma (Beijing Institute of Technology), Jiang Yu (Samsung Electronics (China) R&D Center), Jianing Wang (Meituan-M17), Jie Yang (MGTV Corporation, Changsha, China), Jing He (Xidian University), Jinglin Zhou (Jiangnan University), Jingxuan Li (Harbin Engineering University), Josef Kittler (University of Surrey, UK), Lihao Zheng (Hangzhou Dianzi University, China), Linnan Zhao (Xidian University), Mengxi Jia (Institute of Artificial Intelligence, China Telecom), Muyang Yan (East China Normal University), Nguyen Thanh Thien (University of

Information Technology, Ho Chi Minh City, Vietnam), Pu Luo (Xidian University), Qi Li (Jilin University), Shien Song (MGTV Corporation, Changsha, China), Shijie Dong (Beijing Institute of Technology), Shuai Shao (Hangzhou Dianzi University, China), Shutao Li (Hunan University), Taofeng Xue (Meituan-M17), Tianyang Xu (Jiangnan University), Tianyi Gao (Xi'an Jiaotong University), Tingting Li (Xidian University), Wei Zhang (Jiangnan University), Weiyang Su (Samsung Electronics (China) R&D Center), Xiaodong Dong (Institute of Artificial Intelligence, China Telecom), Xiao-Jun Wu (Jiangnan University), Xiaopeng Zhou (Xidian University), Xin Chen (Zhejiang University), Xin Wei (Institute of Artificial Intelligence, China Telecom), Xinyi You (Xidian University), Xudong Kang (Hunan University), Xujie Zhou (Jiangnan University), Xusheng Liu (MGTV Corporation, Changsha, China), Yanan Wang (Dalian University of Technology), Yanbin Huang (Huazhong University of Science and Technology), Yang Liu (China Electronics Technology Group Corporation), Yang Yang (Nanjing University of Science and Technology), Yanglin Deng (Jiangnan University), Yashu Kang (Zhejiang Supcon Information Technology), Ye Yuan (Institute of Artificial Intelligence, China Telecom), Yi Wen (Xidian University), Yicen Tian (Dalian University of Technology), Yilin Tao (University of Chinese Academy of Sciences), Yin Tang (Samsung Electronics (China) R&D Center), Yipeng Lin (Nanjing University of Science and Technology), Yiqing Wang (Xidian University), Yiting Xi (Communication University of China), Yongkang Yu (Beijing University of Posts and Telecommunications), Yumei Li (Xidian University), Yuxin Qin (Jilin University), Yuying Chen (Xidian University), Yuzhe Cen (Zhejiang Supcon Information Technology), Zhaofan Zou (Institute of Artificial Intelligence, China Telecom), Zhaohong Liu (Communication University of China), Zhehao Shen (Zhejiang Supcon Information Technology), Zhenglin Du (Xidian University), Zhengyang Li (Xidian University), Zhenni Huang (Communication University of China), Zhenwei Shao (Hangzhou Dianzi University, China), Zhilong Song (Xidian University), Zhiyong Feng (Harbin Institute of Technology), Zhiyu Wang (Hunan University), Zhou Yu (Hangzhou Dianzi University, China), Ziang Li (Xidian University), Zihan Zhai (Xidian University), Zijian Zhang (Meituan-M17), Ziyang Peng (Samsung Electronics (China) R&D Center), Ziyun Xiao (Communication University of China), and Zongshu Li (Wuhan Artificial Intelligence Research)

Reliable and Interactable World Models: Geometry, Physics, Interactivity and Real-World Generalization

A Multimodal Physics-Informed Neural Network Approach for Mean Radiant Temperature Modeling 6606
Pouya Shaeri (Arizona State University), Saud Alkhaled (Kuwait University), and Ariane Middel (Arizona State University)

The 12th IEEE International Workshop on Analysis and Modeling of Faces and Gestures

Generating Tennis Action Instruction Based on a Large Language Model	6616
<i>Zhiyuan Wang (Southeast University, China), Zhiqian Xia (Southeast University, China), Guangyao Zhao (Southeast University, China), Kaiyue Lu (Nantong Productivity Promotion Center, China), Haifeng Xia (Southeast University, China), and Siyu Xia (Southeast University, China)</i>	
A Generalized Two-stage Approach to Motion Style Transfer	6626
<i>Tong Guo (Southeast University, China), Zhiqian Xia (Southeast University, China), Feng Yu (Nanjing Xinduan Intelligent Technology Co., Ltd.), Haifeng Xia (Southeast University, China), and Siyu Xia (Southeast University, China)</i>	
FreqCross: A Multi-Modal Frequency-Spatial Fusion Network for Robust Detection of Stable Diffusion 3.5 Generated Images	6636
<i>Guang Yang (University of California, USA)</i>	
Fitting Image Diffusion Models on Video Datasets	6644
<i>Juhun Lee (Sungkyunkwan University, South Korea) and Simon S. Woo (Sungkyunkwan University, South Korea)</i>	
High-Fidelity Character Animation: Generating Coherent and Controllable Motion Videos from Static Images	6651
<i>Yongming Huang (Southeast University, China ;Ministry of Education, China), Zhuojun Xia (Southeast University, China ;Ministry of Education, China), and Can Bu (Southeast University, China ;Ministry of Education, China)</i>	
Attend and Replay: Efficient Action Understanding in Long Videos via Mechanistic Interpretability	6661
<i>Puyue Hou (Beihang University, China), Jinjin Zhang (Beihang University, China), and Di Huang (Beihang University, China)</i>	
Foundational Multi-Task Multimodal Model for Upper GI Endoscopy	6671
<i>Yuxuan He (University of Massachusetts Lowell), Qilei Chen (University of Massachusetts Lowell), Benyuan Liu (University of Massachusetts Lowell), and Yu Cao (University of Massachusetts Lowell)</i>	

Vision Foundation Models and Generative AI for Accessibility: Challenges and Opportunities

SignLLM: Sign Language Production Large Language Models	6681
<i>Sen Fang (Rutgers University), Chen Chen (University of Central Florida), Lei Wang (Australian National University; Data61/CSIRO), Ce Zheng (Carnegie Mellon University), Chunyu Sui (Columbia University), and Yapeng Tian (The University of Texas at Dallas)</i>	
The Escalator Problem: Identifying Implicit Motion Blindness in AI for Accessibility	6694
<i>Xiantao Zhang (Beihang University)</i>	

Who’s Asking? Investigating Bias Through the Lens of Disability-Framed Queries in LLMs	6703
<i>Vishnu Hari (Birla Institute of Technology and Science (BITS), India), Kalpana Panda (Birla Institute of Technology and Science (BITS), India), Srikant Panda (Oracle AI), Amit Agarwal (Oracle AI), and Hitesh Laxmichand Patel (Oracle AI)</i>	
RampNet: A Two-Stage Pipeline for Bootstrapping Curb Ramp Detection in Streetscape Images from Open Government Metadata	6715
<i>John S. O’Meara (Issaquah High School; University of Washington), Jared Hwang (University of Washington), Zeyu Wang (University of Washington), Michael Saugstad (University of Washington), and Jon E. Froehlich (University of Washington)</i>	
VisualSpeaker: Visually-Guided 3D Avatar Lip Synthesis	6725
<i>Alexandre Symeonidis-Herzig (University of Surrey, United Kingdom), Ozge Mercanoglu Sincan (University of Surrey, United Kingdom), and Richard Bowden (University of Surrey, United Kingdom)</i>	
Seeing in 2D, Thinking in 3D: 3D Hand Mesh-Guided Feature Learning for Continuous Fingerspelling	6735
<i>K. Papadimitriou (University of Thessaly, Greece; Athena Research Center, Greece), P. Filntisis (Athena Research Center, Greece), G. Retsinas (Athena Research Center, Greece), G. Potamianos (University of Thessaly, Greece; Athena Research Center, Greece), and P. Maragos (National Technical University of Athens, Greece; Athena Research Center, Greece; HERON - Hellenic Robotics Center of Excellence, Greece)</i>	
Automated Context-Aware Navigation Support for Individuals with Visual Impairment Using Multimodal Language Models in Urban Environments	6745
<i>Alton Chao (University of California, Merced), Erika Maquiling (University of California, Merced), Edmund Chao (University of California, Merced), Roshan Sanjeev (University of California, Merced), Tonko Bossen (University of California, Merced), and Ross Greer (University of California, Merced)</i>	
Introduction to the First Workshop on Vision Foundation Models and Generative AI for Accessibility	6753
<i>Yapeng Tian (University of Texas at Dallas, USA), Yuhang Zhao (University of Wisconsin - Madison, USA), Jon E. Froehlich (University of Washington, USA), Chu Li (University of Washington, USA), and Yuheng Wu (University of Wisconsin - Madison, USA)</i>	

Vision-Language Modeling in 3D Medical Imaging Workshop

T3D: Advancing 3D Medical Vision-Language Pre-training by Learning Multi-View Visual Consistency	6763
<i>Che Liu (Imperial College London, UK), Cheng Ouyang (Imperial College London, UK), Yinda Chen (University of Science and Technology of China, China), César Quilodrán-Casas (Imperial College London, UK), Lei Ma (Peking University, China), Jie Fu (Intelligence Laboratory, China), Yike Guo (University of Science and Technology, Hong Kong SAR, China), Anand Shah (Imperial College London, UK), Wenjia Bai (Imperial College London, UK), and Rossella Arcucci (Imperial College London, UK)</i>	

GTGM: Generative Text-Guided 3D Vision-Language Pretraining for Medical Image Segmentation.....	6774
<i>Yinda Chen (University of Science and Technology of China), Che Liu (Imperial College London), Wei Huang (University of Science and Technology of China), Xiaoyu Liu (University of Science and Technology of China), Haoyuan Shi (University of Science and Technology of China), Sibó Cheng (Imperial College London), Rossella Arcucci (Imperial College London), and Zhiwei Xiong (University of Science and Technology of China)</i>	
Unified Supervision For Vision-Language Modeling in 3D Computed Tomography	6784
<i>Hao-Chih Lee (BioMedical Engineering & Imaging Institute, USA), Zelong Liu (BioMedical Engineering & Imaging Institute, USA), Hamza Ahmed (BioMedical Engineering & Imaging Institute, USA), Spencer Kim (BioMedical Engineering & Imaging Institute, USA), Sean Huver (NVIDIA, USA), Vishwesh Nath (NVIDIA, USA), Zahi A. Fayad (BioMedical Engineering & Imaging Institute, USA), Timothy Deyer (East River Medical Imaging and Weill Cornell Medicine, USA), and Xueyan Mei (BioMedical Engineering & Imaging Institute, USA)</i>	
Prediction Degeneracy in Medical Vision-Language Models: Implications for Robustness and Interpretability	6793
<i>Martin Goetze (University Hospital RWTH Aachen, Germany), Dennis Eschweiler (University Hospital RWTH Aachen, Germany), Brendan Huang (University of California, San Francisco, United States), Gustav Müller-Franzes (University Hospital RWTH Aachen, Germany), Carolina Ramirez (University of California, San Francisco, United States), Madeline Hess (University of California, San Francisco, United States), Lavinia Goldermann (RWTH Aachen University, Germany), Sharmila Majumdar (University of California, San Francisco, United States), and Daniel Truhn (University Hospital RWTH Aachen, Germany)</i>	
RadZero3D: Bridging Self-Supervised Video Models and Medical Vision-Language Alignment for Zero-Shot Chest CT Interpretation	6801
<i>Jonggwon Park (DEEPNOID Inc.), Kyoyun Choi (DEEPNOID Inc.), Byungmu Yoon (DEEPNOID Inc.), Hong Geun Cho (DEEPNOID Inc.), and Bumcheol Hwang (DEEPNOID Inc.)</i>	
CTFlow: Video-Inspired Latent Flow Matching for 3D CT Synthesis	6809
<i>Jiayi Wang (Friedrich–Alexander University Erlangen Nu rnberg, DE), Hadrien Reynaud (Friedrich–Alexander University Erlangen Nu rnberg, DE), Franciskus Xaverius Erick (Friedrich–Alexander University Erlangen Nu rnberg, DE), and Bernhard Kainz (Department of Computing, Imperial College London, London, UK, Friedrich-Alexander University Erlangen Nürnberg DE,)</i>	
Foundation Models for Multimodal MRI Synthesis with Language Guidance	6818
<i>Mahmut Yurt (Stanford University), Xiaozhi Cao (Stanford University), Zihan Zhou (Stanford University), Kawin Setsompop (Stanford University), Shreyas Vasanawala (Stanford University), and John Pauly (Stanford University)</i>	
Patch-wise Intensity Mapping for Individualized Brain Abnormality Detection in Alzheimer’s Disease	6824
<i>Yangshuang Xu (Florida State University, USA), Rongjie Liu (University of Georgia, USA), and Chao Huang (University of Georgia, USA)</i>	

CT-GRAPH: Hierarchical Graph Attention Network for Anatomy-Guided CT Report Generation	6834
<i>Hamza Kalisch (Institute for AI in Medicine (IKIM), University Hospital Essen, Essen, Germany), Fabian Hörst (Institute for AI in Medicine (IKIM), University Hospital Essen, Essen, Germany), Jens Kleesiek (Institute for AI in Medicine (IKIM), University Hospital Essen, Essen, Germany), Ken Herrmann (Department of Nuclear Medicine, University Hospital Essen, Essen, Germany), and Constantin Seibold (Institute for AI in Medicine (IKIM), University Hospital Essen, Essen, Germany)</i>	
Preprocessing–Architecture Co-Design for Multi-Label Classification in 3D Chest CT Scans	6844
<i>Jung Jun Ah (Seoul National University Hospital, Republic of Korea), Juhui Lee (Seoul National University, Republic of Korea), Jihye Heo (Seoul National University Hospital, Republic of Korea), and Dongheon Lee (Seoul National University Hospital, Republic of Korea)</i>	

Curated Data for Efficient Learning

Task-Specific Generative Dataset Distillation with Difficulty-Guided Sampling	6853
<i>Mingzhuo Li (Hokkaido University), Guang Li (Hokkaido University), Jiafeng Mao (The University of Tokyo), Linfeng Ye (University of Toronto), Takahiro Ogawa (Hokkaido University), and Miki Haseyama (Hokkaido University)</i>	
LG-Traj: LLM Guided Pedestrian Trajectory Prediction	6861
<i>Pranav Singh Chib (Department of Computer Science and Engineering, India) and Pravendra Singh (Department of Computer Science and Engineering, India)</i>	
Objaverse++: Curated 3D Object Dataset with Quality Annotations	6872
<i>Chendi Lin (Carnegie Mellon University), Heshan Liu (Carnegie Mellon University), Qunshu Lin (Zhejiang University), Zachary Bright (Exascale Labs), Shitao Tang (Simon Fraser University), Yihui He (Carnegie Mellon University), Minghao Liu (2077AI), Ling Zhu (Exascale Labs), and Cindy Le (Columbia University)</i>	
How to Train your Text-to-Image Model: Evaluating Design Choices for Synthetic Training Captions	6882
<i>Manuel Brack (Adobe Applied Research; hessian.AI), Sudeep Katakol (Adobe Applied Research), Felix Friedrich (hessian.AI; TU Darmstadt), Patrick Schramowski (DFKI; TU Darmstadt; hessian.AI), Hareesh Ravi (Adobe Applied Research), Kristian Kersting (TU Darmstadt; hessian.AI; DFKI), and Ajinkya Kale (Adobe Applied Research)</i>	
VolDoGer: LLM-assisted Datasets for Domain Generalization in Vision-Language Tasks	6892
<i>Juhwan Choi (AITRICS, Republic of Korea), Junehyoung Kwon (Chung-Ang University, Republic of Korea), Jungmin Yun (Chung-Ang University, Republic of Korea), Seunguk Yu (Chung-Ang University, Republic of Korea), and Youngbin Kim (Chung-Ang University, Republic of Korea)</i>	

DISC-GAN: Disentangling Style and Content for Cluster-Specific Synthetic Underwater Image Generation	6903
<i>Sneha Varur (KLE Technological University, India), Anirudh R Hanchinamani (KLE Technological University, India), Tarun S Bagewadi (KLE Technological University, India), Uma Mudenagudi (KLE Technological University, India), Chaitra D Desai (KLE Technological University, India), C Sujata (KLE Technological University, India), Padmashree Desai (KLE Technological University, India), and Sumit Meharwade (KLE Technological University, India)</i>	
Efficient Learning for Product Attributes with Compact Multimodal Models	6912
<i>Mandar Kulkarni (Flipkart Data Science, USA)</i>	
SYM3D: Canonicalizing Triplanes via Symmetry for Single-View 3D Learning	6919
<i>Jing Yang (University of Cambridge, UK), Kyle Fogarty (University of Cambridge, UK), Fangcheng Zhong (University of Cambridge, UK), and Cengiz Oztireli (University of Cambridge, UK)</i>	
Class-Proportional Coreset Selection for Difficulty-Separable Data	6928
<i>Elisa Tsai (University of Michigan), Haizhong Zheng (Carnegie Mellon University), and Atul Prakash (University of Michigan)</i>	
Federated Active Learning for Target Domain Generalisation	6938
<i>Razvan Caramalau (University College London, UK), Binod Bhattarai (University of Aberdeen, UK), and Danail Stoyanov (University College London, UK)</i>	
Synthetic Captions for Open-Vocabulary Zero-Shot Segmentation	6948
<i>Tim Lebailly (KU Leuven; Meta), Vijay Veerabadran (Meta), Satwik Kottur (Meta), Karl Ridgeway (Meta), and Michael Louis Iuzzolino (Meta)</i>	
Few-Shot Vision-Language Reasoning for Satellite Imagery via Verifiable Rewards	6960
<i>Aybora Köksal (Middle East Technical University) and A. Aydın Alatan (Middle East Technical University)</i>	
OmViD: Omni-supervised active learning for video action detection	6970
<i>Aayush Rana (University of Central Florida), Akash Kumar (University of Central Florida), Vibhav Vineet (Microsoft), and Yogesh S. Rawat (University of Central Florida)</i>	

Personalization in Generative AI Workshop

MINDiff: Mask-Integrated Negative Attention for Controlling Overfitting in Text-to-Image Personalization	6981
<i>Seulgi Jeong (Kyungpook National University, South Korea) and Jaeh Kim (Kyungpook National University, South Korea)</i>	
Locally Controlled Face Aging with Latent Diffusion Models	6991
<i>Lais Isabelle Alves Dos Santos (L'Oréal AI Research), Julien Despois (L'Oréal AI Research), Thibaut Chauffier (L'Oréal AI Research), Siley O. Ba (L'Oréal AI Research), and Giovanni Palma (L'Oréal AI Research)</i>	

Only-Style: Stylistic Consistency in Image Generation without Content Leakage	7000
<i>Tilemachos Aravanis (National Technical University of Athens, Greece ;Czech Technical University in Prague, Czech Republic), Panagiotis Filntisis (Robotics Institute, Greece), Petros Maragos (National Technical University of Athens, Greece ;Robotics Institute, Greece ;HERON - Hellenic Robotics Center of Excellence, Greece), and George Retsinas (Robotics Institute, Greece)</i>	
Target Attribute Diffusion Models	7011
<i>William Loh (University of Waterloo), Yanting Miao (University of Waterloo), Pascal Poupart (University of Waterloo), and Suraj Kothawade (Google)</i>	
Dual Orthogonal Guidance for Robust Diffusion-based Handwritten Text Generation	7020
<i>Konstantina Nikolaidou (Luleå University of Technology), George Retsinas (National Technical University of Athens), Giorgos Sfikas (University of West Attica), Silvia Cascianelli (University of Modena and Reggio Emilia), Rita Cucchiara (University of Modena and Reggio Emilia), and Marcus Liwicki (Luleå University of Technology)</i>	
UP-VTON: A Unified Virtual Try-On Framework Supporting Mask, Mask-Free, and Prompt-Driven Guidance	7030
<i>Youngjoo Jo (ETRI), Minho Park (ETRI ;KAIST), and Dong-oh Kang (ETRI)</i>	
Tight Inversion: Image-Conditioned Inversion for Real Image Editing	7039
<i>Edo Kadosh (Tel Aviv University), Nir Goren (Tel Aviv University), Or Patashnik (Tel Aviv University ;Snap Research), Daniel Garibi (Tel Aviv University), and Daniel Cohen-Or (Tel Aviv University ;Snap Research)</i>	
MotionMatcher: Cinematic Motion Customization of Text-to-Video Diffusion Models via Motion Feature Matching	7050
<i>Yen-Siang Wu (National Taiwan University), Chi-Pin Huang (National Taiwan University), Fu-En Yang (NVIDIA), and Yu-Chiang Frank Wang (National Taiwan University; NVIDIA)</i>	
Human Preference-Aligned Concept Customization Benchmark via Decomposed Evaluation	7061
<i>Reina Ishikawa (Keio University), Ryo Fujii (Keio University), Hideo Saito (Keio University), and Ryo Hachiuma (NVIDIA)</i>	

1st Workshop on Long Multi-Scene Video Foundations

Reinforcement Learning meets Masked Video Modeling : Trajectory-Guided Adaptive Token Selection	7071
<i>Ayush K. Rai (Dublin City University), Kyle Min (Intel Labs, USA), Tarun Krishna (Dublin City University), Feiyan Hu (Dublin City University, USA), Alan F. Smeaton (Dublin City University, USA), and Noel E. O'connor (Dublin City University, USA)</i>	

A Survey on Long-Video Storytelling Generation: Architectures, Consistency, and Cinematic Quality	7082
<i>Mohamed Elmoghany (Adobe Research), Ryan Rossi (Adobe Research), Seunghyun Yoon (Adobe Research), Subhojyoti Mukherjee (Adobe Research), Eslam Bakr (King Abdullah University of Science and Technology), Puneet Mathur (Adobe Research), Gang Wu (Adobe Research), Viet Dac Lai (Adobe Research), Nedim Lipka (Adobe Research), Ruiyi Zhang (Adobe Research), Varun Manjunatha (Adobe Research), Chien Nguyen (University of Oregon), Daksh Dangi (Independent Researcher), Abel Salinas (University of Southern California), Hongjie Chen (Dolby Labs), Xiaolei Huang (University of Memphis), Joe Barrow (Pattern Data), Nesreen Ahmed (Cisco), Hoda Eldardiry (Virginia Tech), Namyong Park (Meta AI), Yu Wang (University of Oregon), Zhengzhong Tu (Texas A&M University), Thien Nguyen (Adobe Research), Dinesh Manocha (University of Maryland, College Park), Mohamed Elhoseiny (KAUST), and Franck Derroncourt (Adobe Research)</i>	

2nd Workshop on Scalable 3D Scene Generation and Geometric Scene Understanding

Forgetting-Free Incremental Panoptic Lifting by Maximum-Visibility Viewpoint Selection	7095
<i>Akira Kohjin (Shiga University; RIKEN), Motoharu Sonogashira (RIKEN), Masaaki Iiyama (Shiga University), and Yasutomo Kawanishi (RIKEN)</i>	
SELDOM: Scene Editing via Latent Diffusion with Object-centric Modifications	7105
<i>Richard E. L. Higgins (University of Michigan) and David F. Fouhey (New York University)</i>	
Globally Optimal Registration of Dense Terrestrial Laser Scans From Coarse Sampling	7118
<i>Dorian Kempf (Université de Picardie Jules Verne, France; CNRS-AIST JRL (Joint Robotics Laboratory)), Guillaume Caron (Université de Picardie Jules Verne, France; CNRS-AIST JRL (Joint Robotics Laboratory)), El Mustapha Mouaddib (Université de Picardie Jules Verne, France), and Fumio Kanehiro (National Institute of Advanced Industrial Science and Technology AIST, Japan)</i>	

Computer Vision in Plant Phenotyping and Agriculture

PlantDreamer: Achieving Realistic 3D Plant Models with Diffusion-Guided Gaussian Splatting....	7128
<i>Zane K J Hartley (University of Nottingham), Lewis A G Stuart (University of Nottingham), Andrew P French (University of Nottingham), and Michael P Pound (University of Nottingham)</i>	
AgMIC: Agricultural Masked Image Consistency for Cross-Domain Segmentation	7139
<i>Muhib Ullah (University of Regina, Canada), Nisar Ali (University of Regina, Canada), Numair Nadeem (University of Regina, Canada), and Abdul Bais (University of Regina, Canada)</i>	

Dynamic Monitoring of Crop Canopies Using Time-Series Point Clouds: Insights into Phenotypic Variation and Leaf-Level Photosynthetic Performance	7150
<i>Jiaren Zhou (China Agricultural University, China), Mengqi Zhang (China Agricultural University, China), Shulin Sun (China Agricultural University, China), Man Zhang (China Agricultural University, China), and Minjuan Wang (China Agricultural University, China)</i>	
First Place Solution to the MLCAS 2025 GWFSS Challenge: The Devil is in the Detail and Minority	7161
<i>Songliang Cao (Huazhong University of Science and Technology, China), Tianqi Hu (Huazhong University of Science and Technology, China), and Hao Lu (Huazhong University of Science and Technology, China)</i>	
High-Throughput Estimation of Photosynthetic Phenotypic Parameters Using Hyperspectral Data	7169
<i>Mengqi Zhang (China Agricultural University, China), Jiaren Zhou (China Agricultural University, China), Man Zhang (China Agricultural University, China), and Minjuan Wang (China Agricultural University, China)</i>	
Comparative Analysis of Image-Based Deep Learning and Genomic Models for Yield and Protein Content Prediction in Winter Wheat	7178
<i>Xiaoran Chen (ETH Zürich and EPFL, Swiss Data Science Center, Switzerland), Paraskevi Nousi (ETH Zürich and EPFL, Swiss Data Science Center, Switzerland), Mike Boss (ETH Zürich, Institute of Agricultural Sciences, Switzerland), Michele Volpi (ETH Zürich and EPFL, Swiss Data Science Center, Switzerland), and Lukas Roth (ETH Zürich, Institute of Agricultural Sciences, Switzerland)</i>	
A Case for the Use of Chroma Cartesian Colour Representations for Image Classification on Plant-Based Domains	7187
<i>Alexis J S Payne (University of Nottingham), Gail Hopkins (University of Nottingham), Shreyank N Gowda (University of Nottingham), Isaac Triguero (University of Granada; DaSCI Andalusian Institute in Data Science and Computational Intelligence, Spain; University of Nottingham), and Michael P Pound (University of Nottingham)</i>	
Decoder-aware Self-Supervised Continual Pretraining and Uncertainty-Guided Pseudo-Labeling for Wheat Organ Segmentation	7197
<i>Tapotosh Ghosh (University of Calgary, Canada), Md Jaber Al Nahian (University of Calgary, Canada), Farnaz Sheikhi (University of Calgary, Canada), and Farhad Maleki (University of Calgary, Canada)</i>	
Modeling Time-Lapse Trajectories to Characterize Cranberry Growth	7208
<i>Ronan John (Rutgers University - New Brunswick), Anis Chihoub (Rutgers University - New Brunswick), Ryan Meegan (Rutgers University - New Brunswick), Gina Sidelli (Rutgers University - New Brunswick), Jeffery Neyhart (Rutgers University - New Brunswick), Peter Oudemans (Rutgers University - New Brunswick), and Kristin Dana (Rutgers University - New Brunswick)</i>	
Improving Lightweight Weed Detection via Knowledge Distillation	7219
<i>Ahmet Oğuz Saltık (University of Hohenheim, Germany), Max Voigt (University of Hohenheim, Germany), Sourav Modak (University of Hohenheim, Germany), Mike Beckworth (University of Hohenheim, Germany), and Anthony Stein (University of Hohenheim, Germany)</i>	

Multimodal Fusion of X-Ray Transmission and Dark Field Imaging for Apple Internal Disorders Detection	7229
<i>Jiaqi He (KU Leuven, Belgium), Astrid Tempelaere (KU Leuven, Belgium), Janne Vignero (KU Leuven, Belgium), Pieter Verboven (KU Leuven, Belgium), and Bart Nicolai (KU Leuven, Belgium)</i>	
WeedSense: Multi-Task Learning for Weed Segmentation, Height Estimation, and Growth Stage Classification	7239
<i>Toqi Tahamid Sarker (Southern Illinois University Carbondale, USA), Khaled R Ahmed (Southern Illinois University Carbondale, USA), Taminul Islam (Southern Illinois University Carbondale, USA), Cristiana Bernardi Rankrape (Southern Illinois University Carbondale, USA), and Karla Gage (Southern Illinois University Carbondale, USA)</i>	

Observing and Understanding Hands in Action

DyTact: Capturing Dynamic Contacts in Hand-Object Manipulation	7250
<i>Xiaoyan Cong (Brown University), Angela Xing (Brown University), Chandradeep Pokhariya (Indian Institute of Technology, Madras), Rao Fu (Brown University), and Srinath Sridhar (Brown University)</i>	
Host3R: Keypoint-free Hand-Object 3D Reconstruction from RGB images	7263
<i>Anilkumar Swamy (NAVER LABS Europe), Vincent Leroy (Naver Labs Europe), Philippe Weinzaepfel (Naver Labs Europe), Jean-Sebastien Franco (Inria centre at the University Grenoble Alpes), and Grégory Rogez (Naver Labs Europe)</i>	
WACU: Multi-Modal Wristband Assistant for Contextual Understanding	7273
<i>Constantin Patsch (Technical University of Munich, Germany), Jaden Goter (Meta Reality Labs Research, Redmond, USA), Joseph Greer (Meta Reality Labs Research, Redmond, USA), Lingni Ma (Meta Reality Labs Research, Redmond, USA), and Raj Sodhi (Meta Reality Labs Research, Redmond, USA)</i>	

Large Scale Cross Device Localization

PixCuboid: Room Layout Estimation from Multi-view Featuremetric Alignment	7283
<i>Gustav Hanning (Lund University, Sweden), Kalle Åström (Lund University, Sweden), and Viktor Larsson (Lund University, Sweden)</i>	
The Overlooked Value of Test-time Reference Sets in Visual Place Recognition	7293
<i>Mubariz Zaffar (TU Delft, Netherlands.), Liangliang Nan (TU Delft, Netherlands.), Sebastian Scherer (Carnegie Mellon University, USA), and Julian F. P. Kooij (TU Delft, Netherlands)</i>	
LightGlueStick: a Fast and Robust Glue for Joint Point-Line Matching	7303
<i>Aidyn Ubingazhibov (ETH Zurich), Rémi Pautrat (Microsoft Spatial AI Lab), Iago Suárez (Qualcomm XR Labs Europe), Shaohui Liu (ETH Zurich), Marc Pollefeys (ETH Zurich/Microsoft Spatial AI Lab), and Viktor Larsson (Lund University)</i>	
Triangulation of 3D target points from radar range and bearing data	7314
<i>Magnus Oskarsson (Lund University, Sweden)</i>	

PHAROS Adaptation, Fairness, Explainability in AI Medical Imaging Workshop

PHAROS-AFE-AIMI: Multi-source & Fair Disease Diagnosis	7324
<i>Dimitrios Kollias (Queen Mary University London), Anastasios Arsenos (National Kapodistrian University Athens, Greece), and Stefanos Kollias (National Technical University of Athens, Greece)</i>	
Proactive HIV Care: AI-Based Comorbidity Prediction from Routine EHR Data	7333
<i>Solomon Russom (Queen Mary University London), Dimitrios Kollias (Queen Mary University London), and Qianni Zhang (Queen Mary University London)</i>	
Med-VLM: Enhancing Medical Image Segmentation Accuracy through Vision-Language Model ..	7342
<i>Yihao Zhao (Chinese Academy of Medical Sciences and Peking Union Medical College, China ;Sun Yat-sen University, China), Enhao Zhong (Academy of Medical Sciences and Peking Union Medical College, China ;Sun Yat-sen University, China), Cuiyun Yuan (Chinese Academy of Medical Sciences and Peking Union Medical College, China), Yang Li (Chinese Academy of Medical Sciences and Peking Union Medical College, China), Man Zhao (Chinese Academy of Medical Sciences and Peking Union Medical College, China), Chunxia Li (Macau University), Jun Hu (Sun Yat-sen University, China), Wei Liu (Mayo Clinic, U.S.A), and Chenbin Liu (Sun Yat-sen University)</i>	
Dopamine-Inspired Neuro-Modulated Radiomic Pipeline for Computed-Tomography-Driven Prediction of Immunotherapy Response in Metastatic Cancer Treatment	7353
<i>Francesco Rundo (University of Catania, Italy), Massimo Orazio Spata (University of Catania, Italy), Carmelo Pino (STMicroelectronics), and Sebastiano Battiato (University of Catania, Italy)</i>	
Multi-Source Covid-19 Detection via Variance Risk Extrapolation	7363
<i>Runtian Yuan (Fudan University), Qingqiu Li (Fudan University), Junlin Hou (The Hong Kong University of Science and Technology), Jilan Xu (Fudan University), Yuejie Zhang (Fudan University), Rui Feng (Fudan University), and Hao Chen (The Hong Kong University of Science and Technology)</i>	
In-hoc Concept Representations to Regularise Deep Learning in Medical Imaging	7371
<i>Valentina Corbetta (The Netherlands Cancer Institute, The Netherlands ;Utrecht University, The Netherlands ;Maastricht University, The Netherlands), Floris Six Dijkstra (The Netherlands Cancer Institute, The Netherlands ;University of Amsterdam, The Netherlands), Regina Beets-Tan (The Netherlands Cancer Institute, The Netherlands ;Maastricht University, The Netherlands), Hoel Kervadec (University of Amsterdam, The Netherlands ;Amsterdam UMC, The Netherlands), Kristoffer Wickstrøm (UiT The Arctic University of Norway, Norway), and Wilson Silva (The Netherlands Cancer Institute, The Netherlands ;Utrecht University, The Netherlands)</i>	
Quantifying Inter-Operator Variability and its Causes for Medical Semantic Segmentation	7381
<i>Sophie Fischer (University of Oxford, UK) and Irina Voiculescu (University of Oxford, UK)</i>	

Stable-Drift: A Patient-Aware Latent Drift Replay Method for Stabilizing Representations in Continual Learning	7390
<i>Paraskevi-Antonia Theofilou (National Technical University of Athens, Greece), Anuhyha Thota (London School of Economics and Political Science, UK), Stefanos Kollias (National Technical University of Athens, Greece), and Mamatha Thota (University of Lincoln, UK)</i>	
Fairness in Breast Cancer Diagnosis with Deep Learning	7400
<i>Laura Mata Le Bot (University of York, UK), Sze Chai Kwok (Duke Kunshan University, China), Emmanuel Amankwaa-Frempong (Sweden-Ghana Medical Center, Ghana), and Kofi Appiah (University of York, UK)</i>	
ProtoMedX: Towards Explainable Multi-Modal Prototype Learning for Bone Health Classification	7407
<i>Alvaro Lopez Pellicer (Lancaster University, United Kingdom), André Mariucci (Lancaster University, United Kingdom), Plamen Angelov (Lancaster University, United Kingdom), Marwan Bukhari (Lancaster University, United Kingdom), and Jemma G. Kerns (Lancaster University, United Kingdom)</i>	
DeepSeg-MS: A 3D network based on hierarchical multi-scale learning for MRI multiple sclerosis lesion segmentation	7417
<i>Alessia Rondinella (University of Catania, Italy), Francesco Guarnera (University of Catania, Italy), Francesco Rundo (University of Catania, Italy), and Sebastiano Battiato (University of Catania, Italy)</i>	
Advancing Lung Disease Diagnosis in 3D CT Scans	7427
<i>Qingqiu Li (Fudan University), Runtian Yuan (Fudan University), Junlin Hou (The Hong Kong University of Science and Technology), Jilan Xu (Fudan University), Yuejie Zhang (Fudan University), Rui Feng (Fudan University), and Hao Chen (The Hong Kong University of Science and Technology)</i>	

Responsible Imaging

Lighten-MST: Low-light Spectral Reconstruction via Illumination-guided Spectral-aware Transformer	7433
<i>Yanan Hu (Zhejiang University, China ;Westlake University, China), Xiaodong Wang (Zhejiang University, China ;Westlake University, China), Zijun He (Zhejiang University, China ;Westlake University, China), and Xin Yuan (Westlake University, China)</i>	
See the Invisible with SWIR: Learning to Enhance via Synthetic Noise Modeling	7440
<i>Haiyang Jiang (The University of Tokyo, Japan), Hongjun Wang (The University of Tokyo, Japan), and Yinqiang Zheng (The University of Tokyo, Japan)</i>	
Boosting Diffusion-Based Text Image Super-Resolution Model Towards Generalized Real-World Scenarios	7450
<i>Chenglu Pan (Zhejiang University), Xiaogang Xu (The Chinese University of Hong Kong), Ganggui Ding (Zhejiang University), Yunke Zhang (Zhejiang University), Wenbo Li (The Chinese University of Hong Kong), Jiarong Xu (Fudan University), and Qingbiao Wu (Zhejiang University)</i>	

Learning to See Through Flare	7460
<i>Xiaopeng Peng (Rochester Institute of Technology), Heath Gemar (U.S. Naval Research Laboratory), Erin Fleet (U.S. Naval Research Laboratory), Kyle Novak (U.S. Naval Research Laboratory), Abbie Watnik (U.S. Naval Research Laboratory), and Grover Swartzlander (Rochester Institute of Technology)</i>	

Visual Object Tracking and Segmentation Challenge Workshop

The Third Visual Object Tracking Segmentation VOTS2025 Challenge Results	7472
<i>Matej Kristan (University of Ljubljana, Slovenia), Jiří Matas (Czech Technical University, Czech Republic), Pavel Tokmakov (Toyota Research Institute, USA), Alan Lukežič (University of Ljubljana, Slovenia), Michael Felsberg (Linköping University, Sweden), Luka Čehovin Zajc (University of Ljubljana, Slovenia), Khanh-Tung Tran (Umeå University, Sweden), Xuan-Son Vu (Lund University, Sweden ;DeepTensor AB, Sweden), Johanna Björklund (Umeå University, Sweden), Michal Neoral (Czech Technical University, Czech Republic), Hyung Jin Chang (University of Birmingham, UK), Gustavo Fernández (Austrian Institute of Technology, Austria), Minasadat Attari (University of Missouri, USA), Matteo Dunnhofer (University of Udine, Italy ;York University, Canada), Wei Feng (Tianjin University, China), Zhenhua Feng (Jiangnan University, China), Jin Gao (Chinese Academy of Sciences, China), Yameng Gu (Harbin Institute of Technology, China), Ruize Han (Shenzhen University of Advanced Technology, China), Jiawei He (Chongqing University, China), Zhenyu He (Harbin Institute of technology, China), Junhui Hou (City University of Hong Kong, Hong Kong), Weiming Hu (Chinese Academy of Sciences, China), Xiantao Hu (Nanjing University of Science and Technology, China), Xingsen Huang (Harbin Institute of Technology, China), Yuqing Huang (PengCheng Laboratory, China), Gleb Kirichenko (n/a), Josef Kittler (University of Surrey, UK), Yutong Kou (Chinese Academy of Sciences, China), Simiao Lai (Dalian University of Technology, China), Bing Li (Chinese Academy of Sciences, China), Xin Li (PengCheng Laboratory, China), Shubo Lin (Chinese Academy of Sciences, China), Huchuan Lu (Dalian University of Technology, China), Deshui Miao (Harbin Institute of Technology, China), Christian Micheloni (University of Udine, Italy), Juan Mogollon (University of Missouri, USA), Moritz Nottebaum (University of Udine, Italy), Kannappan Palaniappan (University of Missouri, USA), Ziqi Pang (University of Illinois Urbana-Champaign, USA), Zekun Qian (Tianjin University, China), Gani Rahmon (University of Missouri, USA), Aleksandr Romanov (n/a), Liangtao Shi (Hefei University of Technology, China), Roman Solovyev (n/a), Elham Soltani Kazemi (University of Missouri, USA), and Imad Eddine Toubal (University of Missouri, USA)</i>	

Computer Vision Systems for Document Analysis and Recognition

Text Image Generation for Low-Resource Languages with Dual Translation Learning	7491
<i>Chihiro Noguchi (Toyota Motor Corporation), Shun Fukuda (Toyota Motor Corporation), Shoichiro Mihara (Toyota Motor Corporation), and Masao Yamanaka (Toyota Motor Corporation)</i>	
TAP-VL: Text Layout Aware Pretraining for Enriched Vision-Language Models	7502
<i>Jonathan Fhima (Technion - Israel Institute of Technology, Technion), Elad Ben Aoraham (AWS AI Labs), Oren Nuriel (AWS AI Labs), Yair Kittenplon (AWS AI Labs), Roy Ganz (AWS AI Labs), Aviad Aberdam (AWS AI Labs), and Ron Litman (AWS AI Labs)</i>	

Structure-aware Contrastive Learning for Diagram Understanding of Multimodal Models	7513
<i>Hiroshi Sasaki (The Japan Research Institute, Limited)</i>	
Quo Vadis Handwritten Text Generation for Handwritten Text Recognition?	7523
<i>Vittorio Pippi (University of Modena and Reggio Emilia), Konstantina Nikolaidou (Lulea University of Technology), Silvia Cascianelli (University of Modena and Reggio Emilia), George Retsinas (National Technical University of Athens), Giorgos Sfikas (University of West Attica), Rita Cucchiara (University of Modena and Reggio Emilia), and Marcus Liwicki (Lulea University of Technology)</i>	
Describe Anything Model for Visual Question Answering on Text-rich Images	7534
<i>Yen-Linh Vu (AI VIETNAM Lab, Vietnam), Dinh-Thang Duong (AI VIETNAM Lab, Vietnam), Truong-Binh Duong (AI VIETNAM Lab, Vietnam), Anh-Khoi Nguyen (AI VIETNAM Lab, Vietnam), Thanh-Huy Nguyen (Carnegie Mellon University, USA), Le Thien Phuc Nguyen (University of Wisconsin - Madison, USA), Jianhua Xing (University of Pittsburgh, USA), Xingjian Li (Carnegie Mellon University, USA), Tianyang Wang (University of Alabama at Birmingham, USA), Ulas Bagci (Northwestern University, USA), and Min Xu (Carnegie Mellon University, USA)</i>	
ZOD : Zero-shot and Out-of-Distribution Detection Dataset for Document Images	7545
<i>Talha Uddin Sheikh (DFKI, Germany; RPTU Kaiserslautern-Landau; Mindgarage), Sankalp Sinha (DFKI, Germany; RPTU Kaiserslautern-Landau; Mindgarage), Shino Sam (DFKI, Germany; RPTU Kaiserslautern-Landau; Mindgarage), Didier Stricker (DFKI, Germany; RPTU Kaiserslautern-Landau; Mindgarage), and Muhammad Zeshan Afzal (DFKI, Germany; RPTU Kaiserslautern-Landau; Mindgarage)</i>	
CoSMo: A Multimodal Transformer for Page Stream Segmentation in Comic Books	7556
<i>Marc Serra Ortega (Computer Vision Center, Universitat Aut`onoma de Barcelona), Emanuele Vivoli (Universidad Autonom de Barcelona), Artemis Llabres (Computer Vision Center, Universitat Autonom de Barcelona), and Dimosthenis Karatzas (Universitat Autonom de Barcelona)</i>	
Scanned documents forensics: detecting inserted characters through noise and chromatic artifacts	7565
<i>Marina Gardella (Ecole Normale Supérieure, France), Julieta Umpierrez (Ecole Normale Supérieure Paris Saclay, France), Antoine Tadros (Ecole Normale Supérieure, France), Seginus Mowlavi (Ecole Normale Supérieure, France), Natalia Bottaioli (Ecole Normale Supérieure, France), Diego Belzarena (Ecole Normale Supérieure, Hong Kong), Gabriele Facciolo (Ecole Normale Supérieure Paris-Saclay, France), Roy Y. He (City University of Hong Kong), Jean-Michel Morel (City University of Hong Kong), and Rafael Grompone Von Gioi (Ecole Normale Supérieure, France)</i>	
PRISM: Pruning for Rank-adaptive Interpretable Segmentation Model with Application to Historical Document Multiband Images	7576
<i>Kilian Declercq (École de Technologie Supérieure (ÉTS), Canada), Abderrahmane Rahiche (École de Technologie Supérieure (ÉTS), Canada), and Mohamed Cheriet (École de Technologie Supérieure (ÉTS), Canada)</i>	

DocSemi: Efficient Document Layout Analysis with Guided Queries	7586
<i>Tahira Shehzadi (Rheinland-Pfalzische Technische Universität Kaiserslautern-Landau), Ifza Ifza (German Research Center for AI), Didier Stricker (Technical University Kaiserslautern, Germany), and Muhammad Zeshan Afzal (German Research Center for AI, Germany)</i>	
DIVE-Doc: Downscaling foundational Image Visual Encoder into hierarchical architecture for DocVQA	7597
<i>Rayane Bencharef (École de Technologie Supérieure, Canada; INU Champollion, France), Abderrahmane Rahiche (École de Technologie Supérieure, Canada), and Mohamed Cheriet (École de Technologie Supérieure, Canada)</i>	
CTC Transcription Alignment of the Bullinger Letters: Automatic Improvement of Annotation Quality	7607
<i>Marco Peer (University of Fribourg, Switzerland), Anna Scius Bertrand (University of Applied Sciences and Arts Western Switserland), and Andreas Fischer (University of Fribourg, Switzerland)</i>	
Deep Learning-Based Intrusion Detection Systems for Phishing Email Detection: A Short Survey	7617
<i>Axel De Nardin (University of Udine, Italy), Silvia Zottin (University of Udine, Italy), Claudio Piciarelli (University of Udine, Italy), and Gian Luca Foresti (University of Udine, Italy)</i>	
Improved Information Extraction by Leveraging Multi-Hypothesis OCR at Inference Time	7626
<i>Arthur Hemmer (L3i La Rochelle, La Rochelle, France), Nicola Bartolo (Shift Technology, Paris, France), Mickaël Coustaty (L3i La Rochelle, La Rochelle, France), and Jean-Marc Ogier (L3i La Rochelle, La Rochelle, France)</i>	
A Survey on Reading Order, Table of Contents, and Structure Extraction in Document Analysis	7635
<i>Simone Giovannini (University of Florence) and Simone Marinai (University of Florence)</i>	
ChemMiner: A Large Language Model Agent System for Chemical Literature Data Mining	7645
<i>Kexin Chen (The Chinese University of Hong Kong), Yuyang Du (The Chinese University of Hong Kong), Junyou Li (Zhejiang Lab), Hanqun Cao (The Chinese University of Hong Kong), Menghao Guo (Zhejiang Lab), Xilin Dang (The Chinese University of Hong Kong), Lanqing Li (The Chinese University of Hong Kong ;Zhejiang Lab), Jiezhong Qiu (Hangzhou Institute of Medicine), Guangyong Chen (Hangzhou Institute of Medicine), and Pheng Ann Heng (The Chinese University of Hong Kong)</i>	

Scene Graphs and Graph Representation Learning

CoPa-SG: Dense Scene Graphs with Parametric and Proto-Relations	7654
<i>Julian Lorenz (University of Augsburg), Mrunmai Phatak (University of Augsburg), Robin Schön (University of Augsburg), Katja Ludwig (University of Augsburg), Nico Hörmann (University of Augsburg), Annemarie Friedrich (University of Augsburg), and Rainer Lienhart (University of Augsburg)</i>	

Measuring Image-Relation Alignment: Reference-Free Evaluation of VLMs and Synthetic Pre-training for Open-Vocabulary Scene Graph Generation	7664
<i>Maëlic Neau (Umeå University, Sweden), Zoe Falomir (Umeå University, Sweden), Cédric Buche (CNRS IRL 2010 CROSSING, Australia; IMT Atlantique, France), and Akihiro Sugimoto (National Institute of Informatics, Japan)</i>	
Long-form Reasoning for Keystep Recognition using Graph Neural Networks	7674
<i>Julia Lee Romero (University of Colorado Boulder), Kyle Min (Intel Labs), Subarna Tripathi (Intel Labs), and Morteza Karimzadeh (University of Colorado Boulder)</i>	

Critical Evaluation of Generative Models and Their Impact on Society

An Information-Theoretic Approach to Diversity Evaluation of Prompt-based Generative Models	7684
<i>Mohammad Jalali (The Chinese University of Hong Kong), Azim Osipov (The Chinese University of Hong Kong), Amin Gohari (The Chinese University of Hong Kong), and Farzan Farnia (The Chinese University of Hong Kong)</i>	
On the Distributed Evaluation of Generative Models	7694
<i>Zixiao Wang (The Chinese University of Hong Kong), Farzan Farnia (The Chinese University of Hong Kong), Zhenghao Lin (South China Agricultural University), Yunheng Shen (Tsinghua University), and Bei Yu (The Chinese University of Hong Kong)</i>	

Fairness and Ethics Towards Transparent AI: Facing the Challenge Through Model Debiasing

Modeling Cognitive and Implicit Biases in Multi-Agent Medical Systems for Clinical Diagnoses	7704
<i>Niyel Hassan (AlgoVerse AI Research), Benjamin Liu (AlgoVerse AI Research), Raghav Thallapragada (AlgoVerse AI Research), Ryan Bui (AlgoVerse AI Research), Nishant Chinta (AlgoVerse AI Research), Ayushman Bisht (AlgoVerse AI Research), Dev J. Chaliha (AlgoVerse AI Research), and Kevin Zhu (AlgoVerse AI Research)</i>	
Evaluating the Impact of Racial Cues on MLLMs Judgements of Politeness and Offensiveness	7713
<i>Mahammed Kamruzzaman (University of South Florida) and Gene Louis Kim (University of South Florida)</i>	
BAdd: Bias Mitigation through Bias Addition	7723
<i>Ioannis Sarridis (Information Technologies Institute, Greece), Christos Koutlis (Information Technologies Institute, Greece), Symeon Papadopoulos (Information Technologies Institute, Greece), and Christos Diou (Harokopio University of Athens, Greece)</i>	

Fairness Without Labels: Pseudo-Balancing for Bias Mitigation in Face Gender
Classification7733
*Haohua Dong (National Institute of Informatics, Japan ;University of
Lisbon, Portugal), Ana Manzano Rodríguez (National Institute of
Informatics, Japan ;University of Amsterdam, The Netherlands), Camille
Guinaudeau (Université Paris-Saclay, France), and Shin'ichi Satoh
(National Institute of Informatics, Japan)*

Author Index