

2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2024)

**Seattle, Washington, USA
16-22 June 2024**

Pages 1-676



**IEEE Catalog Number: CFP2488A-POD
ISBN: 979-8-3503-6548-1**

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

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

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

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

IEEE Catalog Number:	CFP2488A-POD
ISBN (Print-On-Demand):	979-8-3503-6548-1
ISBN (Online):	979-8-3503-6547-4
ISSN:	2160-7508

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

2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) **CVPRW 2024**

Table of Contents

(3rd) Monocular Depth Estimation Challenge

The Third Monocular Depth Estimation Challenge	1
<i>Jaime Spencer (Independent Researcher), Fabio Tosi (University of Bologna, Italy), Matteo Poggi (University of Bologna, Italy), Ripudaman Singh Arora (Blue River Technology, USA), Chris Russell (Oxford Internet Institute, UK), Simon Hadfield (University of Surrey, UK), Richard Bowden (University of Surrey, UK), GuangYuan Zhou (ByteDance, China), ZhengXin Li (University of Chinese Academy of Sciences, China), Qiang Rao (ByteDance, China), YiPing Bao (ByteDance, China), Xiao Liu (ByteDance, China), Dohyeong Kim (RGA Inc., Republic of Korea), Jinseong Kim (RGA Inc., Republic of Korea), Myunghyun Kim (RGA Inc., Republic of Korea), Mykola Lavreniuk (Space Research Institute NASU-SSAU, Ukraine), Rui Li (Northwestern Polytechnical University, China), Qing Mao (Northwestern Polytechnical University, China), Jiang Wu (Northwestern Polytechnical University, China), Yu Zhu (Northwestern Polytechnical University, China), Jinqiu Sun (Northwestern Polytechnical University, China), Yanning Zhang (Northwestern Polytechnical University, China), Suraj Patni (Indian Institute of Technology, India), Aradhye Agarwal (Indian Institute of Technology, India), Chetan Arora (Indian Institute of Technology, India), Pihai Sun (Harbin Institute of Technology, China), Kui Jiang (Harbin Institute of Technology, China), Gang Wu (Harbin Institute of Technology, China), Jian Liu (Harbin Institute of Technology, China), Xianming Liu (Harbin Institute of Technology, China), Junjun Jiang (Harbin Institute of Technology, China), Xidan Zhang (Fujitsu, China), Jianing Wei (Fujitsu, China), Fangjun Wang (Fujitsu, China), Zhiming Tan (Fujitsu, China), Jiabao Wang (GuangXi University, China), Albert Luginov (University of Reading, UK), Muhammad Shahzad (University of Reading, UK), Seyed Hosseini (York University, UK), Aleksander Trajcevski (University of York, UK), and James H. Elder (University of York, UK)</i>	

The 7th Workshop and Challenge Bridging the Gap Between Computational Photography and Visual Recognition (UG2+)

UAV-Rain1k: A Benchmark for Raindrop Removal from UAV Aerial Imagery	15
<i>Wenhui Chang (Shenyang Aerospace University, China), Hongming Chen (Shenyang Aerospace University, China), Xin He (Naval Aviation University, China), Xiang Chen (Nanjing University of Science and Technology, China), and Liangduo Shen (Zhejiang Ocean University, China)</i>	

Feature Corrective Transfer Learning: End-to-End Solutions to Object Detection in Non-Ideal Visual Conditions	23
<i>Chuheng Wei (University of California Riverside), Guoyuan Wu (University of California Riverside), and Matthew J. Barth (University of California Riverside)</i>	
Source-Free Domain Adaptation of Weakly-Supervised Object Localization Models for Histology	33
<i>Alexis Guichemerre (ETS Montreal), Soufiane Belharbi (ETS Montreal), Tsiry Mayet (INSA Rouen), Shakeeb Murtaza (ETS Montreal), Pourya Shamsolmoali (East China Normal University), Luke McCaffrey (McGill University), and Eric Granger (ETS Montreal)</i>	
Mobile Aware Denoiser Network (MADNet) for Quad Bayer Images	44
<i>Pavan C. Madhusudana (Samsung Research America), Jing Li (Samsung Research America), Zeeshan Nadir (Samsung Research America), Hamid R. Sheikh (Samsung Research America), and Seok-Jun Lee (Samsung Research America)</i>	

Computer Vision for Materials Science Workshop

VolRAFT: Volumetric Optical Flow Network for Digital Volume Correlation of Synchrotron Radiation-based Micro-CT Images of Bone-Implant Interfaces	53
<i>Tak Ming Wong (Helmholtz-Zentrum Hereon, Germany), Julian Moosmann (Helmholtz-Zentrum Hereon, Germany), and Berit Zeller-Plumhoff (Helmholtz-Zentrum Hereon, Germany)</i>	
Damage Detection and Localization by Learning Deep Features of Elastic Waves in Piezoelectric Ceramic Using Point Contact Method	63
<i>Pragyan Banerjee (Indian Institute of Technology, Guwahati), Pranjali Saxena (Indian Institute of Technology, Guwahati), Nur M M Kalimullah (Indian Institute of Technology, Guwahati), Amit Shelke (Indian Institute of Technology, Guwahati), and Anowarul Habib (UiT The Arctic University of Norway)</i>	
Self-Supervised Learning with Generative Adversarial Networks for Electron Microscopy	71
<i>Bashir Kazimi (Forschungszentrum Jülich (FZJ), Germany), Karina Ruzaeva (Forschungszentrum Jülich (FZJ), Germany), and Stefan Sandfeld (Forschungszentrum Jülich (FZJ), Germany; Rheinisch Westfälische Technische Hochschule (RWTH) Aachen, Germany)</i>	

The Fifth Workshop on Fair, Data-efficient, and Trusted Computer Vision

Towards Explainable Visual Vessel Recognition Using Fine-Grained Classification and Image Retrieval	82
<i>Heiko Karus (Hensoldt Optronics GmbH, Germany), Friedhelm Schwenker (Ulm University, Germany), Michael Munz (Ulm University of Applied Sciences, Germany), and Michael Teutsch (Hensoldt Optronics GmbH, Germany)</i>	

Towards Efficient Machine Unlearning with Data Augmentation: Guided Loss-Increasing (GLI) to Prevent the Catastrophic Model Utility Drop	93
<i>Dasol Choi (Yonsei University), Soora Choi (Chung Ang University), Eunsun Lee (Kyung Hee University), Jinwoo Seo (Millennialsworks Inc.), and Dongbin Na (Pohang University of Science and Technology)</i>	
Enforcing Conditional Independence for Fair Representation Learning and Causal Image Generation	103
<i>Jensen Hwa (Stanford University), Qingyu Zhao (Weill Cornell Medicine), Aditya Lahiri (University of California San Diego), Adnan Masood (UST), Babak Salimi (University of California San Diego), and Ehsan Adeli (Stanford University)</i>	
Improving the Robustness of 3D Human Pose Estimation: A Benchmark and Learning from Noisy Input	113
<i>Trung-Hieu Hoang (University of Illinois at Urbana-Champaign, USA), Mona Zehni (University of Illinois at Urbana-Champaign, USA), Huy Phan (VinUniversity, Ha Noi, Vietnam), Duc Minh Vo (The University of Tokyo, Japan), and Minh N. Do (University of Illinois at Urbana-Champaign, USA)</i>	
DIA: Diffusion based Inverse Network Attack on Collaborative Inference	124
<i>Dake Chen (University of Southern California), Shiduo Li (Tsinghua University), Yuke Zhang (University of Southern California), Chenghao Li (University of Southern California), Souvik Kundu (Intel Labs), and Peter A. Beerel (University of Southern California)</i>	
ReweightOOD: Loss Reweighting for Distance-based OOD Detection	131
<i>Sudarshan Regmi (NepAl Applied Mathematics and Informatics Institute for Research, Nepal), Bibek Panthi (NepAl Applied Mathematics and Informatics Institute for Research, Nepal), Yifei Ming (University of Wisconsin-Madison, USA), Prashnna K Gyawali (West Virginia University, USA), Danail Stoyanov (University College London, UK), and Binod Bhattarai (University of Aberdeen, UK)</i>	
Our Deep CNN Face Matchers Have Developed Achromatopsia	142
<i>Aman Bhatta (University of Notre Dame), Domingo Mery (Pontificia Universidad Católica de Chile), Haiyu Wu (University of Notre Dame), Joyce Annan (Florida Institute of Technology), Michael C. King (Florida Institute of Technology), and Kevin W. Bowyer (University of Notre Dame)</i>	
T2FNorm: Train-time Feature Normalization for OOD Detection in Image Classification	153
<i>Sudarshan Regmi (NepAl Applied Mathematics and Informatics Institute for Research, Nepal), Bibek Panthi (NepAl Applied Mathematics and Informatics Institute for Research, Nepal), Sakar Dotel (Tribhuvan University), Prashnna K Gyawali (West Virginia University, USA), Danail Stoyanov (University College London, UK), and Binod Bhattarai (University of Aberdeen)</i>	
Fractals as Pre-training Datasets for Anomaly Detection and Localization	163
<i>Cynthia I. Ugwu (Free University of Bozen-Bolzano, Italy), Sofia Casarin (Free University of Bozen-Bolzano, Italy), and Oswald Lanz (Free University of Bozen-Bolzano, Italy)</i>	

Test-time Assessment of a Model’s Performance on Unseen Domains via Optimal Transport	173
<i>Akshay Mehra (Tulane University), Yunbei Zhang (Tulane University), and Jihun Hamm (Tulane University)</i>	
Robust and Explainable Fine-Grained Visual Classification with Transfer Learning: A Dual-Carriageway Framework	183
<i>Zheming Zuo (Newcastle University), Joseph Smith (Newcastle University), Jonathan Stonehouse (Procter and Gamble), and Boguslaw Obara (Newcastle University)</i>	
Practical Region-level Attack against Segment Anything Models	194
<i>Yifan Shen (University of Illinois Urbana-Champaign), Zhengyuan Li (University of Illinois Urbana-Champaign), and Gang Wang (University of Illinois Urbana-Champaign)</i>	
SkipPLUS: Skip the First Few Layers to Better Explain Vision Transformers	204
<i>Faridoun Mehri (Sharif University of Technology, Iran), Mohsen Fayyaz (University of Tehran, Iran), Mahdieh Soleymani Baghshah (Sharif University of Technology, Iran), and Mohammad Taher Pilehvar (Cardiff University, UK)</i>	
AR-CP: Uncertainty-Aware Perception in Adverse Conditions with Conformal Prediction and Augmented Reality For Assisted Driving	216
<i>Achref Doula (Technical University of Darmstadt, Germany), Max Mühlhäuser (Technical University of Darmstadt, Germany), and Alejandro Sanchez Guinea (Technical University of Darmstadt, Germany)</i>	
Fast-NTK: Parameter-Efficient Unlearning for Large-Scale Models	227
<i>Guihong Li (The University of Texas at Austin), Hsiang Hsu (Global Technology Applied Research, JPMorgan Chase), Chun-Fu Chen (Global Technology Applied Research, JPMorgan Chase), and Radu Marculescu (The University of Texas at Austin)</i>	
Mitigating Bias Using Model-Agnostic Data Attribution	235
<i>Sander De Coninck (Ghent University - imec, Belgium), Sam Leroux (Ghent University - imec, Belgium), and Pieter Simoens (Ghent University - imec, Belgium)</i>	
RLNet: Robust Linearized Networks for Efficient Private Inference	244
<i>Sreetama Sarkar (University of Southern California, Los Angeles, USA), Souvik Kundu (Intel Labs, San Diego, USA), and Peter A. Bearel (University of Southern California, Los Angeles, USA)</i>	
Data-free Defense of Black Box Models Against Adversarial Attacks	254
<i>Gaurav Kumar Nayak (University of Central Florida, USA), Inder Khatri (New York University, USA), Ruchit Rawal (Indian Institute of Science, India), and Anirban Chakraborty (Indian Institute of Science, India)</i>	

LatinX in Computer Vision Research Workshop

An End-to-End Approach for Handwriting Recognition: From Handwritten Text Lines to Complete Pages	264
<i>Dayvid Castro (Universidade Federal de Pernambuco), Byron Leite Dantas Bezerra (Universidade de Pernambuco), and Cleber Zanchettin (Universidade Federal de Pernambuco)</i>	

Enhancing Image Classification Robustness through Adversarial Sampling with Delta Data Augmentation (DDA)	274
<i>Ivan Reyes-Amezcuca (Center for Research and Advanced Studies of the National Polytechnic Institute, Mexico), Gilberto Ochoa-Ruiz (Tecnologico de Monterrey, Mexico), and Andres Mendez-Vazquez (Center for Research and Advanced Studies of the National Polytechnic Institute, Mexico)</i>	
High-Resolution Detection of Earth Structural Heterogeneities from Seismic Amplitudes using Convolutional Neural Networks with Attention layers	284
<i>Luiz Schirmer (Unisinos), Guilherme Schardong (University of Coimbra), Vinícius da Silva (PUC-Rio), Rogério Santos (UFF), and Hélio Lopes (PUC-Rio)</i>	
Beyond Appearances: Material Segmentation with Embedded Spectral Information from RGB-D imagery	293
<i>Fabian Perez (Universidad Industrial de Santander, Colombia) and Hoover Rueda-Chacón (Universidad Industrial de Santander, Colombia)</i>	
ST-Gait++: Leveraging Spatio-temporal Convolutions for Gait-based Emotion Recognition on Videos	302
<i>Maria Luísa Lima (Universidade Federal de Pernambuco, Brazil), Williams de Lima Costa (Voxar Labs, Brazil), Estefania Talavera Martínez (University of Twente, The Netherlands), and Veronica Teichrieb (Universidade Federal de Pernambuco, Brazil)</i>	
The Myth of the Pyramid	311
<i>Ramon Izquierdo-Cordova (University of Bristol) and Walterio Mayol-Cuevas (University of Bristol)</i>	

The Seventh International Workshop on Computer Vision for Physiological Measurement (CVPM)

GPT as Psychologist? Preliminary Evaluations for GPT-4V on Visual Affective Computing	322
<i>Hao Lu (The Hong Kong University of Science & Technology (Guangzhou)), Xuesong Niu (Beijing Institute for General Artificial Intelligence), Jiyao Wang (The Hong Kong University of Science & Technology (Guangzhou)), Yin Wang (Zhejiang University), Qingyong Hu (The Hong Kong University of Science & Technology), Jiaqi Tang (The Hong Kong University of Science & Technology (Guangzhou)), Yuting Zhang (The Hong Kong University of Science & Technology (Guangzhou)), Kaishen Yuan (Great Bay University), Bin Huang (Hangzhou Research Institute, Beihang University), Zitong Yu (Great Bay University), Dengbo He (The Hong Kong University of Science & Technology (Guangzhou)), Shuiguang Deng (Zhejiang University), Hao Chen (The Hong Kong University of Science & Technology), Yingcong Chen (The Hong Kong University of Science & Technology (Guangzhou)), and Shiguang Shan (Zhejiang University)</i>	

NurtureNet: A Multi-task Video-based Approach for Newborn Anthropometry	332
<i>Yash Khandelwal (Wadhwani Institute for Artificial Intelligence), Mayur Arvind (Wadhwani Institute for Artificial Intelligence), Sriram Kumar (Wadhwani Institute for Artificial Intelligence), Ashish Gupta (Wadhwani Institute for Artificial Intelligence), Sachin Kumar Danisetty (Wadhwani Institute for Artificial Intelligence), Piyush Bagad (Wadhwani Institute for Artificial Intelligence), Anish Madan (Wadhwani Institute for Artificial Intelligence), Mayank Lunayach (Wadhwani Institute for Artificial Intelligence), Aditya Annavajjala (Wadhwani Institute for Artificial Intelligence), Abhishek Maiti (Wadhwani Institute for Artificial Intelligence), Sansiddh Jain (Wadhwani Institute for Artificial Intelligence), Aman Dalmia (Wadhwani Institute for Artificial Intelligence), Namrata Deka (Wadhwani Institute for Artificial Intelligence), Jerome White (Wadhwani Institute for Artificial Intelligence), Jigar Doshi (Wadhwani Institute for Artificial Intelligence), Angjoo Kanazawa (University of California Berkeley), Rahul Panicker (Wadhwani Institute for Artificial Intelligence), Alpan Raval (Wadhwani Institute for Artificial Intelligence), Srinivas Rana (Wadhwani Institute for Artificial Intelligence), and Makarand Tapaswi (Wadhwani Institute for Artificial Intelligence)</i>	
Vision-language Models for Decoding Provider Attention During Neonatal Resuscitation	343
<i>Felipe Parodi (University of Pennsylvania, USA), Jordan K. Matelsky (University of Pennsylvania, USA), Alejandra Regla-Vargas (University of Pennsylvania, USA), Elizabeth E. Foglia (University of Pennsylvania Perelman School of Medicine, USA), Charis Lim (University of Pennsylvania Perelman School of Medicine, USA), Danielle Weinberg (University of Pennsylvania Perelman School of Medicine, USA), Konrad P. Kording (University of Pennsylvania, USA), Heidi M. Herrick (University of Pennsylvania Perelman School of Medicine, USA), and Michael L. Platt (University of Pennsylvania, USA)</i>	
Orientation-conditioned Facial Texture Mapping for Video-based Facial Remote Photoplethysmography Estimation	354
<i>Sam Cantrill (The Australian National University, Australia; Data61/CSIRO, Australia), David Ahmedt-Aristizabal (Data61/CSIRO, Australia), Lars Petersson (Data61/CSIRO, Australia), Hanna Suominen (The Australian National University, Australia; Data61/CSIRO, Australia; University of Turku, Finland), and Mohammad Ali Armin (Data61/CSIRO, Australia)</i>	
Paediatric Pulse Rate Measurements: a Comparison of Methods using Remote Photoplethysmography	364
<i>Simon Wegerif (Xim Ltd, United Kingdom), Ivan Veleslavov (Xim Ltd, United Kingdom), Lieke Dorine van Putten (Xim Ltd, United Kingdom), Kate Emily Bamford (Xim Ltd, United Kingdom), Gauri Misra (Mind over Matter Medtech Ltd, United Kingdom), and Niall Mullen (South Tyneside and Sunderland, NHS Foundation Trust, United Kingdom)</i>	

DECNet: A Non-Contacting Dual-Modality Emotion Classification Network for Driver Health Monitoring	371
<i>Zhekang Dong (Hangzhou Dianzi University), Chenhao Hu (Hangzhou Dianzi University), Shiqi Zhou (Hangzhou Dianzi University), Liyan Zhu (Hangzhou Dianzi University), Junfan Wang (Hangzhou Dianzi University), Yi Chen (Zhejiang University), Xudong Lv (Hangzhou Dianzi University), and Xiaoyue Ji (Tsinghua University)</i>	
Refining Remote Photoplethysmography Architectures using CKA and Empirical Methods	380
<i>Nathan Vance (University of Notre Dame) and Patrick Flynn (University of Notre Dame)</i>	
Analyzing Participants' Engagement during Online Meetings Using Unsupervised Remote Photoplethysmography with Behavioral Features	389
<i>Alexander Vedernikov (University of Oulu, Finland), Zhaodong Sun (University of Oulu, Finland), Virpi-Liisa Kykyri (University of Jyväskylä, Finland), Mikko Pohjola (University of Jyväskylä, Finland), Miriam Nokia (University of Jyväskylä, Finland), and Xiaobai Li (Zhejiang University, China; University of Oulu, Finland)</i>	
Video Based Computational Coding of Movement Anomalies in ASD Children	400
<i>Priya Singh (TCS Research India), Abhishek Pathak (Airbase Labs India Pvt Ltd), Umer Jon Ganai (Indian Institute of Technology Kanpur), Braj Bhushan (Indian Institute of Technology Kanpur), and Venkatesh K. Subramanian (Indian Institute of Technology Kanpur)</i>	
How Suboptimal is Training rPPG Models with Videos and Targets from Different Body Sites?	410
<i>Björn Braun (ETH Zürich, Switzerland), Daniel McDuff (University of Washington), and Christian Holz (ETH Zürich, Switzerland)</i>	

EarthVision: Large Scale Computer Vision for Remote Sensing Imagery

UrbanSARFloods: Sentinel-1 SLC-Based Benchmark Dataset for Urban and Open-Area Flood Mapping	419
<i>Jie Zhao (Technical University of Munich, Germany), Zhitong Xiong (Technical University of Munich, Germany), and Xiao Xiang Zhu (Technical University of Munich, Germany)</i>	
Exploring Robust Features for Few-Shot Object Detection in Satellite Imagery	430
<i>Xavier Bou (ENS Paris-Saclay), Gabriele Facciolo (ENS Paris-Saclay), Rafael Grompone von Gioi (ENS Paris-Saclay), Jean-Michel Morel (City University of Hong Kong), and Thibaud Ehret (ENS Paris-Saclay)</i>	
Efficient Local Correlation Volume for Unsupervised Optical Flow Estimation on Small Moving Objects in Large Satellite Images	440
<i>Sarra Khairi (Inria, France), Etienne Meunier (Inria, France), Renaud Fraisse (Airbus, France), and Patrick Bouthemy (Inria, France)</i>	
Deep Generative Data Assimilation in Multimodal Setting	449
<i>Yongquan Qu (Columbia University), Juan Nathaniel (Columbia University), Shuolin Li (Columbia University), and Pierre Gentine (Columbia University)</i>	

GeoSynth: Contextually-Aware High-Resolution Satellite Image Synthesis	460
<i>Srikumar Sastry (Washington University in St. Louis), Subash Khanal (Washington University in St. Louis), Aayush Dhakal (Washington University in St. Louis), and Nathan Jacobs (Washington University in St. Louis)</i>	
Implicit Assimilation of Sparse In Situ Data for Dense & Global Storm Surge Forecasting	471
<i>Patrick Ebel (ESA, Italy), Brandon Victor (La Trobe University, Australia), Peter Naylor (ESA, Italy), Gabriele Meoni (ESA, Italy), Federico Serva (Consiglio Nazionale delle Ricerche, Italy), and Rochelle Schneider (ESA, Italy)</i>	
Detecting Out-Of-Distribution Earth Observation Images with Diffusion Models	481
<i>Georges Le Bellier (CNAM, France) and Nicolas Audebert (CNAM, Gustave Eiffel University, ENSG, IGN, France)</i>	
(Street) Lights Will Guide You: Georeferencing Nighttime Astronaut Photography of Earth	492
<i>Alex Stoken (Jacobs, NASA Johnson Space Center), Peter Ilhardt (Jacobs, NASA Johnson Space Center), Mark Lambert (Jacobs, NASA Johnson Space Center), and Kenton Fisher (NASA Johnson Space Center)</i>	
Cross-sensor super-resolution of irregularly sampled Sentinel-2 time series	502
<i>Aimi Okabayashi (Université Bretagne Sud, IRISA, France), Nicolas Audebert (Université Gustave Eiffel, ENSG, IGN, France; Conservatoire national des arts et métiers, CEDRIC, France), Simon Donike (Image Processing Laboratory (IPL), Universitat de València, Spain), and Charlotte Pelletier (Université Bretagne Sud, IRISA, France)</i>	
SyntStereo2Real: Edge-Aware GAN for Remote Sensing Image-to-Image Translation while Maintaining Stereo Constraint	512
<i>Vasudha Venkatesan (University of Freiburg), Daniel Panangian (German Aerospace Center (DLR)), Mario Fuentes Reyes (German Aerospace Center (DLR)), and Ksenia Bittner (German Aerospace Center (DLR))</i>	
SUNDIAL: 3D Satellite Understanding through Direct, Ambient, and Complex Lighting Decomposition	522
<i>Nikhil Behari (MIT), Akshat Dave (MIT), Kushagra Tiwary (MIT), William Yang (MIT), and Ramesh Raskar (MIT)</i>	
Sat2Cap: Mapping Fine-Grained Textual Descriptions from Satellite Images	533
<i>Aayush Dhakal (Washington University in St. Louis, USA), Adeel Ahmad (Taylor Geospatial Institute, USA), Subash Khanal (Washington University in St. Louis, USA), Srikumar Sastry (Washington University in St. Louis, USA), Hannah Kerner (Arizona State University, USA), and Nathan Jacobs (Washington University in St. Louis, USA)</i>	
Unsupervised Domain Adaptation Architecture Search with Self-Training for Land Cover Mapping	543
<i>Clifford Broni-Bediako (RIKEN Center for Advanced Intelligence Project, Japan), Junshi Xia (RIKEN Center for Advanced Intelligence Project, Japan), and Naoto Yokoya (The University of Tokyo, Japan and RIKEN Center for Advanced Intelligence Project, Japan)</i>	

Charting New Territories: Exploring the Geographic and Geospatial Capabilities of Multimodal LLMs	554
<i>Jonathan Roberts (University of Cambridge), Timo Lüddecke (University of Göttingen), Rehan Sheikh (University of Cambridge), Kai Han (The University of Hong Kong), and Samuel Albanie (University of Cambridge)</i>	
Radar Fields: An Extension of Radiance Fields to SAR	564
<i>Thibaud Ehret (ENS Paris-Saclay, France), Roger Mari (ENS Paris-Saclay, France), Dawu Derksen (CNES, France), Nicolas Gasnier (CNES, France), and Gabriele Facciolo (ENS Paris-Saclay, France)</i>	
Contrastive Pretraining for Visual Concept Explanations of Socioeconomic Outcomes	575
<i>Ivica Obadic (Technical University of Munich, Germany), Alex Levering (Vrije Universiteit Amsterdam, Netherlands), Lars Pennig (Technical University of Munich, Germany), Dario Oliveira (Getulio Vargas Foundation, Brazil), Diego Marcos (Inria, University of Montpellier, France), and Xiaoxiang Zhu (Technical University of Munich, Germany)</i>	
GeoLLM-Engine: A Realistic Environment for Building Geospatial Copilots	585
<i>Simranjit Singh (Microsoft Corporation, USA), Michael Fore (Microsoft Corporation, USA), and Dimitrios Stamoulis (Microsoft Corporation, USA)</i>	
Let Me Show You How It's Done - Cross-modal Knowledge Distillation as Pretext Task for Semantic Segmentation	595
<i>Rudhishna Narayanan Nair (Technische Universität Berlin, Germany) and Ronny Hänsch (German Aerospace Center (DLR), Germany)</i>	

GAZE 2024: The 6th International Workshop on Gaze Estimation and Prediction in the Wild

Spatio-Temporal Attention and Gaussian Processes for Personalized Video Gaze Estimation	604
<i>Swati Jindal (University of California Santa Cruz, USA), Mohit Yadav (University of Massachusetts Amherst, USA), and Roberto Manduchi (University of California Santa Cruz, USA)</i>	
Exploring the Zero-Shot Capabilities of Vision-Language Models for Improving Gaze Following	615
<i>Anshul Gupta (Idiap Research Institute, Switzerland and EPFL, Switzerland), Pierre Vuillecard (Idiap Research Institute, Switzerland and EPFL, Switzerland), Arya Farkhondeh (Idiap Research Institute, Switzerland and EPFL, Switzerland), and Jean-Marc Odobez (Idiap Research Institute, Switzerland and EPFL, Switzerland)</i>	
Gaze Scanpath Transformer: Predicting Visual Search Target by Spatiotemporal Semantic Modeling of Gaze Scanpath	625
<i>Takumi Nishiyasu (The University of Tokyo, Japan) and Yoichi Sato (The University of Tokyo, Japan)</i>	
GESCAM : A Dataset and Method on Gaze Estimation for Classroom Attention Measurement	636
<i>Athul M. Mathew (Elm Company), Arshad Ali Khan (Elm Company), Thariq Khalid (Elm Company), and Riad Souissi (Elm Company)</i>	

Second Workshop for Learning 3D with Multi-View Supervision

Semi-Stereo: A Universal Stereo Matching Framework for Imperfect Data via Semi-supervised Learning	646
<i>Xin Yue (Tsinghua University, China), Zongqing Lu (Tsinghua University, China), Xiangru Lin (Prometheus Vision Technology Co., Ltd., China), Wenjia Ren (Tsinghua University, China), Zhijing Shao (Prometheus Vision Technology Co., Ltd., China), Haonan Hu (Tsinghua University, China), Yu Zhang (Prometheus Vision Technology Co., Ltd., China), and Qingmin Liao (Tsinghua University, China)</i>	
MonoSelfRecon: Purely Self-Supervised Explicit Generalizable 3D Reconstruction of Indoor Scenes from Monocular RGB Views	656
<i>Runfa Li (UC San Diego), Upal Mahbub (Qualcomm), Vasudev Bhaskaran (Qualcomm), and Truong Nguyen (UC San Diego)</i>	
Lifting Multi-View Detection and Tracking to the Bird's Eye View	667
<i>Torben Teepe (TUM), Philipp Wolters (TUM), Johannes Gilg (TUM), Fabian Herzog (TUM), and Gerhard Rigoll (TUM)</i>	
3D Clothed Human Reconstruction from Sparse Multi-view Images	677
<i>Jin Gyu Hong (Kwangwoon University, South Korea), Seung Young Noh (Kwangwoon University, South Korea), Hee Kyung Lee (Electronics and Telecommunications Research Institute, South Korea), Won Sik Cheong (Electronics and Telecommunications Research Institute, South Korea), and Ju Yong Chang (Kwangwoon University, South Korea)</i>	
SACReg: Scene-Agnostic Coordinate Regression for Visual Localization	688
<i>Jerome Revaud (Naver Labs Europe), Yohann Cabon (Naver Labs Europe), Romain Brégier (Naver Labs Europe), JongMin Lee (Seoul University), and Philippe Weinzaepfel (Naver Labs Europe)</i>	
DepthVoting: A Few-Shot Point Cloud Classification Model Incorporating a Projection-Based Voting Mechanism	699
<i>Yunhui Zhu (Syracuse University, USA), Jiajing Chen (Syracuse University, USA), and Senem Velipasalar (Syracuse University, USA)</i>	
Cross-Modal Self-Training: Aligning Images and Pointclouds to learn Classification without Labels	708
<i>Amaya Dharmasiri (Princeton University, USA), Muzammal Naseer (MBZUAI, UAE), Salman Khan (MBZUAI, UAE), and Fahad Shahbaz Khan (MBZUAI, UAE)</i>	
MIMIC: Masked Image Modeling with Image Correspondences	718
<i>Kalyani Marathe (University of Washington, Seattle), Mahtab Bigverdi (University of Washington, Seattle), Nishat Khan (University of Washington, Seattle), Tuhin Kundu (Independent Researcher), Patrick Howe (Independent Researcher), Sharan Ranjit S (University of Washington, Seattle), Anand Bhattad (Toyota Technological Institute at Chicago), Aniruddha Kembhavi (Allen Institute for AI), Linda G. Shapiro (University of Washington, Seattle), and Ranjay Krishna (University of Washington, Seattle)</i>	
Selective Multi-View Deep Model for 3D Object Classification	728
<i>Mona Alzahrani (King Fahd University of Petroleum and Minerals, Saudi Arabia), Muhammad Usman (King Fahd University of Petroleum and Minerals, Saudi Arabia), Saeed Anwar (King Fahd University of Petroleum and Minerals, Saudi Arabia), and Tarek Helmy (King Fahd University of Petroleum and Minerals, Saudi Arabia)</i>	

From 2D Portraits to 3D Realities: Advancing GAN Inversion for Enhanced Image Synthesis	737
<i>Wonseok Oh (University of Michigan) and Youngjoo Jo (Electronics and Telecommunications Research Institute)</i>	
DGBD: Depth Guided Branched Diffusion for Comprehensive Controllability in Multi-View Generation	747
<i>Hovhannes Margaryan (Picsart AI Research (PAIR)), Daniil Hayrapetyan (Picsart AI Research (PAIR)), Wenyan Cong (UT Austin), Zhangyang Wang (Picsart AI Research (PAIR), UT Austin), and Humphrey Shi (Picsart AI Research (PAIR), SHI Labs @ Georgia Tech, Oregon & UIUC)</i>	
2T-UNET: A Two-Tower UNet with Depth Clues for Robust Stereo Depth Estimation	757
<i>Mansi Sharma (Thapar Institute of Engineering and Technology, Patiala, Punjab, India; Indian Institute of Technology Madras, India), Rohit Choudhary (Indian Institute of Technology Madras, India), and Rithvik Anil (Indian Institute of Technology Madras, India)</i>	
AgileGAN3D: Few-Shot 3D Portrait Stylization by Augmented Transfer Learning	765
<i>Guoxian Song (ByteDance Inc), Hongyi Xu (ByteDance Inc), Jing Liu (ByteDance Inc), Tiancheng Zhi (ByteDance Inc), Yichun Shi (ByteDance inc), Jianfeng Zhang (ByteDance Inc), Zihang Jiang (ByteDance Inc), Jiashi Feng (ByteDance Inc), Shen Sang (ByteDance Inc), and Linjie Luo (ByteDance Inc)</i>	
Color-cued Efficient Densification Method for 3D Gaussian Splatting	775
<i>Sieun Kim (Seoul National University, Republic of Korea), Kyungjin Lee (Seoul National University, Republic of Korea), and Youngki Lee (Seoul National University, Republic of Korea)</i>	
PointOfView: A Multi-modal Network for Few-shot 3D Point Cloud Classification Fusing Point and Multi-view Image Features	784
<i>Huantao Ren (Syracuse University), Jiyang Wang (Syracuse University), Minmin Yang (Syracuse University), and Senem Velipasalar (Syracuse University)</i>	
OGRMPI: An Efficient Multiview Integrated Multiplane Image based on Occlusion Guided Residuals	794
<i>Dae Yeol Lee (Dolby Laboratories), Guan-Ming Su (Dolby Laboratories), and Peng Yin (Dolby Laboratories)</i>	
Sparse Multi-view Hand-object Reconstruction for Unseen Environments	803
<i>Yik Lung Pang (Queen Mary University of London), Changjae Oh (Queen Mary University of London), and Andrea Cavallaro (Idiap Research Institute)</i>	
Depth-Regularized Optimization for 3D Gaussian Splatting in Few-Shot Images	811
<i>Jaeyoung Chung (Seoul National University, South Korea), Jeongtaek Oh (Seoul National University, South Korea), and Kyoung Mu Lee (Seoul National University, South Korea)</i>	

First Workshop on Efficient and On-Device Generation (EDGE)

LD-Pruner: Efficient Pruning of Latent Diffusion Models using Task-Agnostic Insights	821
<i>Thibault Castells (Nota Inc.), Hyoung-Kyu Song (Nota Inc.), Bo-Kyeong Kim (Nota Inc.), and Shinkook Choi (Nota Inc.)</i>	

EdgeRelight360: Text-Conditioned 360-Degree HDR Image Generation for Real-Time On-Device Video Portrait Relighting	831
<i>Min-Hui Lin (Qualcomm Technologies, Inc.), Mahesh Reddy (Qualcomm Technologies, Inc.), Guillaume Berger (Qualcomm Technologies, Inc.), Michel Sarkis (Qualcomm Technologies, Inc.), Fatih Porikli (Qualcomm Technologies, Inc.), and Ning Bi (Qualcomm Technologies, Inc.)</i>	

7th International Workshop on Visual Odometry and Computer Vision Applications Based on Location Clues

Camera Motion Estimation from RGB-D-Inertial Scene Flow	841
<i>Samuel Cerezo (Universidad de Zaragoza, Spain) and Javier Civera (Universidad de Zaragoza, Spain)</i>	
BAA-NGP: Bundle-Adjusting Accelerated Neural Graphics Primitives	850
<i>Sainan Liu (Intel Labs, USA), Shan Lin (University of California San Diego, USA), Jingpei Lu (University of California San Diego, USA), Alexey Supikov (Intel Labs), and Michael Yip (University of California San Diego, USA)</i>	
Weakly Supervised End2End Deep Visual Odometry	858
<i>Amin Abouee (Spleenlab GmbH), Ashwanth Ravi (Spleenlab GmbH), Lars Hinneburg (Spleenlab GmbH), Mateusz Dziwulski (Spleenlab GmbH), Florian Ölsner (Spleenlab GmbH), Jürgen Hess (Spleenlab GmbH), Stefan Milz (Spleenlab GmbH), and Patrik Mäder (TU Ilmenau)</i>	

Implicit Neural Representation for Vision

Connecting NeRFs, Images, and Text	866
<i>Francesco Ballerini (University of Bologna, Italy), Pierluigi Zama Ramirez (University of Bologna, Italy), Roberto Mirabella (University of Bologna, Italy), Samuele Salti (University of Bologna, Italy), and Luigi Di Stefano (University of Bologna, Italy)</i>	
Contextualising Implicit Representations for Semantic Tasks	877
<i>Theo W. Costain (University of Oxford, England), Kejie Li (University of Oxford, England), and Victor A. Prisacariu (University of Oxford, England)</i>	
StegaNeRV: Video Steganography using Implicit Neural Representation	888
<i>Monsij Biswal (University of California, Santa Barbara), Tong Shao (Dolby Laboratories, Inc.), Kenneth Rose (University of California, Santa Barbara), Peng Yin (Dolby Laboratories, Inc.), and Sean McCarthy (Dolby Laboratories, Inc.)</i>	
ImplicitTerrain: a Continuous Surface Model for Terrain Data Analysis	899
<i>Haoan Feng (University of Maryland, College Park, USA), Xin Xu (University of Maryland, College Park, USA), and Leila De Floriani (University of Maryland, College Park, USA)</i>	

Workshop on Graphic Design Understanding and Generation (GDUG)

Reference-based GAN Evaluation by Adaptive Inversion	910
<i>Jianbo Wang (University of Tokyo), Heliang Zheng (University of Science and Technology of China), and Toshihiko Yamasaki (University of Tokyo)</i>	

The 5th Face Anti-Spoofing Workshop

Unified Physical-Digital Attack Detection Challenge	919
<i>Haocheng Yuan (Macau University of Science and Technology, Macau), Ajian Liu (Chinese Academy of Sciences, China), Junze Zheng (Macau University of Science and Technology, Macau), Jun Wan (Chinese Academy of Sciences, China), Jiankang Deng (Imperial College London, UK), Sergio Escalera (Computer Vision Center, Spain), Hugo Jair Escalante (INAOE, Mexico), Isabelle Guyon (ChaLearn, USA), and Zhen Lei (Chinese Academy of Sciences, China)</i>	

Multi-angle Consistent Generative NeRF with Additive Angular Margin Momentum Contrastive Learning	930
<i>Hang Zou (China Telecom Corporation Limited Beijing Research Institute, China), Hui Zhang (Tianjin University of Science & Technology, China), Yuan Zhang (China Telecom Corporation Limited Shanghai Research Institute, China), Hui Ma (Macau University of Science and Technology, China), Dexin Zhao (China Telecom Corporation Limited Beijing Research Institute, China), Qi Zhang (China Telecom Corporation Limited Beijing Research Institute, China), and Qi Li (Chinese Academy of Sciences, China)</i>	

Rethinking the Domain Gap in Near-infrared Face Recognition	940
<i>Michail Tarasiou (Imperial College London), Jiankang Deng (Imperial College London), and Stefanos Zafeiriou (Imperial College London)</i>	

IDAdapter: Learning Mixed Features for Tuning-Free Personalization of Text-to-Image Models.....	950
<i>Siying Cui (Peking University), Jia Guo (InsightFace), Xiang An (DeepGlint), Jiankang Deng (InsightFace), Yongle Zhao (DeepGlint), Xinyu Wei (Peking University), and Ziyong Feng (DeepGlint)</i>	

Unified Face Attack Detection with Micro Disturbance and a Two-Stage Training Strategy	960
<i>Jiaruo Yu (IntSig Information Co. Ltd, China), Dagong Lu (IntSig Information Co. Ltd, China), Xingyue Shi (IntSig Information Co. Ltd, China), Chenfan Qu (IntSig Information Co. Ltd, China), and Fengjun Guo (IntSig Information Co. Ltd, China)</i>	

Advancing Cross-Domain Generalizability in Face Anti-Spoofing: Insights, Design, and Metrics	970
<i>Hyojin Kim (Naver Cloud, Republic of Korea), Jiyoung Lee (Korea University, Republic of Korea), Yonghyun Jeong (Naver Cloud, Republic of Korea), Haneol Jang (Hanbat National University, Republic of Korea), and Youngjoon Yoo (Chung-Ang University, Republic of Korea)</i>	

Supervised Contrastive Learning for Snapshot Spectral Imaging Face Anti-Spoofing	980
<i>Chuanbiao Song (Ant Group), Yan Hong (Ant Group), Jun Lan (Ant Group), Huijia Zhu (Ant Group), Weiqiang Wang (Ant Group), and Jianfu Zhang (Shanghai Jiao Tong University)</i>	

A Visualization Method for Data Domain Changes in CNN Networks and the Optimization Method for Selecting Thresholds in Classification Tasks	986
<i>Minzhe Huang (Akuvox), Changwei Nie (Akuvox), and Weihong Zhong (Akuvox)</i>	
Joint Physical-Digital Facial Attack Detection Via Simulating Spoofing Clues	995
<i>Xianhua He (Vision AI Department, Meituan), Dashuang Liang (Vision AI Department, Meituan), Song Yang (Vision AI Department, Meituan), Zhanlong Hao (Vision AI Department, Meituan), Hui Ma (M.U.S.T, Macau), Binjie Mao (Vision AI Department, Meituan), Xi Li (Vision AI Department, Meituan), Yao Wang (Vision AI Department, Meituan), Pengfei Yan (Vision AI Department, Meituan), and Ajian Liu (MAIS, CASIA, China)</i>	
Snapshot Spectral Imaging for Face Anti-Spoofing: Addressing Data Challenges with Advanced Processing and Training	1005
<i>Hui Li (China Telecom Artificial Intelligence Technology Co. Ltd, China), Yaowen Xu (China Telecom Artificial Intelligence Technology Co. Ltd, China), Zhaofan Zou (China Telecom Artificial Intelligence Technology Co. Ltd, China), and Zhixiang He (China Telecom Artificial Intelligence Technology Co. Ltd, China)</i>	
Multiattention-Net: A Novel Approach to Face Anti-Spoofing with Modified Squeezed Residual Blocks	1013
<i>Sabari Nathan (Couger Inc), M.Parisa Beham (Sethu Institute of Technology), A Nagaraj (Sethu Institute of Technology), and S. Mohamed Mansoor Roomi (Thiagarajar College of Engineering)</i>	
Assessing the Performance of Efficient Face Anti-Spoofing Detection Against Physical and Digital Presentation Attacks	1021
<i>Luis S. Luevano (Inria, France), Yoanna Martínez-Díaz (Advanced Technologies Application Center (CENATAV), Cuba), Heydi Méndez-Vázquez (Advanced Technologies Application Center (CENATAV), Cuba), Miguel González-Mendoza (Tecnologico de Monterrey, Mexico), and Davide Frey (Inria, France)</i>	

1st Workshop on Test-Time Adaptation: Model, Adapt Thyself! (MAT)

MixStyle-Based Contrastive Test-Time Adaptation: Pathway to Domain Generalization	1029
<i>Kota Yamashita (Meijo University, Japan) and Kazuhiro Hotta (Meijo University, Japan)</i>	
Fully Test-time Adaptation for Object Detection	1038
<i>Xiaoqian Ruan (University of Illinois Chicago, USA) and Wei Tang (University of Illinois Chicago, USA)</i>	
Test-time Specialization of Dynamic Neural Networks	1048
<i>Sam Leroux (Ghent University - imec), Dewant Katare (Delft University of Technology), Aaron Yi Ding (Delft University of Technology), and Pieter Simoens (Ghent University - imec)</i>	
ST2ST: Self-Supervised Test-time Adaptation for Video Action Recognition	1057
<i>Masud An-Nur Islam Fahim (University of Vaasa), Mohammed Innat (Khulna University of Engineering & Technology), and Jani Boutellier (University of Vaasa)</i>	

Unknown Sample Discovery for Source Free Open Set Domain Adaptation	1067
<i>Chowdhury Sadman Jahan (Rochester Institute of Technology, USA) and Andreas Savakis (Rochester Institute of Technology, USA)</i>	

Third Workshop of Mobile Intelligent Photography & Imaging

UDAC: Under-Display Array Cameras	1077
<i>Chengyu Wang (Samsung Research America, USA), Jing Li (Samsung Research America, USA), Pavan C. Madhusudanarao (Samsung Research America, USA), Jinhan Hu (Samsung Research America, USA), Jitesh K. Singh (Samsung Research India-Bangalore, India), WooJhon Choi (Samsung Electronics), Seok-Jun Lee (Samsung Research America, USA), and Hamid R. Sheikh (Samsung Research America, USA)</i>	
EL2NM: Extremely Low-light Noise Modeling Through Diffusion Iteration	1085
<i>Jiahao Qin (North University of China, China), Pinle Qin (North University of China, China), Rui Chai (North University of China, China), Jia Qin (North University of China, China), and Zanzia Jin (North University of China, China)</i>	
Event Camera Demosaicing via Swin Transformer and Pixel-focus Loss	1095
<i>Yunfan Lu (Hong Kong University of Science and Technology (Guangzhou)), Yijie Xu (Hong Kong University of Science and Technology (Guangzhou)), Wenzong Ma (Hong Kong University of Science and Technology (Guangzhou)), Weiyu Guo (Hong Kong University of Science and Technology (Guangzhou)), and Hui Xiong (Hong Kong University of Science and Technology (Guangzhou))</i>	
From Synthetic to Real: A Calibration-free Pipeline for Few-shot Raw Image Denoising	1106
<i>Ruoqi Li (Xiaomi Inc., China), Chang Liu (Xiaomi Inc., China), Ziyi Wang (Xiaomi Inc., China), Yao Du (Xiaomi Inc., China), Jingjing Yang (Xiaomi Inc., China), Long Bao (Xiaomi Inc., China), and Heng Sun (Xiaomi Inc., China)</i>	
LaDiffGAN: Training GANs with Diffusion Supervision in Latent Spaces	1115
<i>Xuhui Liu (Beihang University, China), Bohan Zeng (Beihang University, China), Sicheng Gao (Beihang University, China), Shanglin Li (Beihang University, China), Yutang Feng (Beihang University, China), Hong Li (Beihang University, China), Boyu Liu (Beihang University, China), Jianzhuang Liu (Shenzhen Institute of Advanced Technology, China), and Baoshang Zhang (Beihang University, China)</i>	
DemosaicFormer: Coarse-to-Fine Demosaicing Network for HybridEVS Camera	1126
<i>Senyan Xu (University of Science and Technology of China, China), Zhijing Sun (University of Science and Technology of China, China), Jiaying Zhu (University of Science and Technology of China, China), Yurui Zhu (University of Science and Technology of China, China), Xueyang Fu (University of Science and Technology of China, China), and Zheng-Jun Zha (University of Science and Technology of China, China)</i>	

MIPI 2024 Challenge on Demosaic for Hybridevs Camera: Methods and Results	1136
<p>Yaqi Wu (SenseTime Research), Zhihao Fan (SenseTime Research), Xiaofeng Chu (SenseTime Research), Jimmy S. Ren (SenseTime Research), Xiaoming Li (Nanyang Technological University), Zongsheng Yue (Nanyang Technological University), Chongyi Li (Nankai University), Shangcheng Zhou (Nanyang Technological University), Ruicheng Feng (Nanyang Technological University), Yuekun Dai (Nanyang Technological University), Peiqing Yang (Nanyang Technological University), Chen Change Loy (Nanyang Technological University), 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), Yurui Zhu (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), Jun Cao (Xiaomi Inc., China), Cheng Li (Xiaomi Inc., China), Shu Chen (Xiaomi Inc., China), Liang Ma (Xiaomi Inc., China), Shiyang Zhou (Harbin Institute of Technology (Shenzhen)), Haijin Zeng (IMEC-UGent), Kai Feng (Northwestern Polytechnical University), Yongyong Chen (Harbin Institute of Technology (Shenzhen)), Jingyong Su (Harbin Institute of Technology (Shenzhen)), Xianyu Guan (Multimedia Department, Xiaomi Inc.), Hongyuan Yu (Multimedia Department, Xiaomi Inc.), Cheng Wan (Georgia Institute of Technology), Jiamin Lin (Multimedia Department, Xiaomi Inc.), Binnan Han (Multimedia Department, Xiaomi Inc.), Yajun Zou (Multimedia Department, Xiaomi Inc.), Zhuoyuan Wu (Multimedia Department, Xiaomi Inc.), Yuan Huang (Multimedia Department, Xiaomi Inc.), Yongsheng Yu (University of Rochester), Daoan Zhang (University of Rochester), Jizhe Li (Multimedia Department, Xiaomi Inc.), Xuanwu Yin (Multimedia Department, Xiaomi Inc.), Kunlong Zuo (Multimedia Department, Xiaomi Inc.), Yunfan Lu (AI Thrust, The Hong Kong University of Science and Technology (Guangzhou)), Yijie Xu (AI Thrust, The Hong Kong University of Science and Technology (Guangzhou)), Wenzong Ma (AI Thrust, The Hong Kong University of Science and Technology (Guangzhou)), Weiyu Guo (AI Thrust, The Hong Kong University of Science and Technology (Guangzhou)), Hui Xiong (AI Thrust, The Hong Kong University of Science and Technology (Guangzhou)), Wei Yu (Harbin Institute of Technology), Bingchun Luo (Harbin Institute of Technology), Sabari Nathan (Couger Inc, Japan), and Priya Kansal (Couger Inc, Japan)</p>	
MIPI 2024 Challenge on Nighttime Flare Removal: Methods and Results	1144
<p>Yuekun Dai (Nanyang Technological University), Dafeng Zhang (Samsung Research China), Xiaoming Li (Nanyang Technological University), Zongsheng Yue (Nanyang Technological University), Chongyi Li (Nankai University), Shangchen Zhou (Nanyang Technological University), Ruicheng Feng (Nanyang Technological University), Peiqing Yang (Nanyang Technological University), Zhezhu Jin (Samsung Research China), Guanqun Liu (Samsung Research China), and Chen Change Loy (Nanyang Technological University)</p>	

MIPI 2024 Challenge on Few-shot RAW Image Denoising: Methods and Results	1153
<i>Xin Jin (Nankai University), Chunle Guo (Nankai University), Xiaoming Li (Nanyang Technological University), Zongsheng Yue (Nanyang Technological University), Chongyi Li (Nankai University), Shangchen Zhou (Nanyang Technological University), Ruicheng Feng (Nanyang Technological University), Yuekun Dai (Nanyang Technological University), Peiqing Yang (Nanyang Technological University), Chen Change Loy (Nanyang Technological University), Ruqi Li (Xiaomi Inc.), Chang Liu (Xiaomi Inc.), Ziyi Wang (Xiaomi Inc.), Yao Du (Xiaomi Inc.), Jingjing Yang (Xiaomi Inc.), Long Bao (Xiaomi Inc.), Heng Sun (Xiaomi Inc.), Xiangyu Kong (Samsung Research China – Beijing (SRCB)), Xiaoxia Xing (Samsung Research China – Beijing (SRCB)), Jinlong Wu (Samsung Research China – Beijing (SRCB)), Yuanyang Xue (Samsung Research China – Beijing (SRCB)), Hyunhee Park (Samsung Electronics), Sejun Song (Samsung Electronics), Changho Kim (Samsung Electronics), Jingfan Tan (Sun Yat-sen University), Wenhan Luo (Sun Yat-sen University), Zikun Liu (Samsung Research China – Beijing (SRCB)), Mingde Qiao (Harbin Institute of Technology, China), Junjun Jiang (Harbin Institute of Technology, China), Kui Jiang (Harbin Institute of Technology, China), Yao Xiao (Harbin Institute of Technology, China), Chuyang Sun (Harbin Institute of Technology, China), Jinhui Hu (Smart City Research Institute of China Electronics Technology Group Corporation), Weijian Ruan (Smart City Research Institute of China Electronics Technology Group Corporation), Yubo Dong (Xidian University), Kai Chen (University of Electronic Science and Technology of China), Hyejeong Jo (National Hanbat University), Jiahao Qin (North University of China), Bingjie Han (North University of China), Pinle Qin (North University of China), Rui Chai (North University of China), and Pengyuan Wang (North University of China)</i>	

2nd Workshop on Foundation Models

How to Benchmark Vision Foundation Models for Semantic Segmentation?	1162
<i>Tommie Kerssies (Eindhoven University of Technology, Netherlands), Daan de Geus (Eindhoven University of Technology, Netherlands), and Gijs Dubbelman (Eindhoven University of Technology, Netherlands)</i>	
Exploring the Benefits of Vision Foundation Models for Unsupervised Domain Adaptation	1172
<i>Brunó B. Englert (Eindhoven University of Technology, Netherlands), Fabrizio J. Piva (Eindhoven University of Technology, Netherlands), Tommie Kerssies (Eindhoven University of Technology, Netherlands), Daan de Geus (Eindhoven University of Technology, Netherlands), and Gijs Dubbelman (Eindhoven University of Technology, Netherlands)</i>	

FGVC11: 11th Workshop on Fine-grained Visual Categorization

Towards Learning Image Similarity from General Triplet Labels	1181
<i>Radu Dondera (Greenfield Vision, USA)</i>	

Coarse or Fine? Recognising Action End States without Labels	1191
<i>Davide Moltisanti (University of Bath), Hakan Bilen (The University of Edinburgh), Laura Sevilla-Lara (The University of Edinburgh), and Frank Keller (The University of Edinburgh)</i>	
Leveraging Large Language Models for Multimodal Search	1201
<i>Oriol Barbany (CSIC-UPC), Michael Huang (Amazon), Xinliang Zhu (Amazon), and Arnab Dhua (Amazon)</i>	
ConceptHash: Interpretable Fine-Grained Hashing via Concept Discovery	1211
<i>Kam Woh Ng (University of Surrey), Xiatian Zhu (University of Surrey), Yi-Zhe Song (University of Surrey), and Tao Xiang (University of Surrey)</i>	
Making Use of Unlabeled Data: Comparing Strategies for Marine Animal Detection in Long-tailed Datasets Using Self-supervised and Semi-supervised Pre-training	1224
<i>Tarun Sharma (California Institute of Technology, USA), Danelle E. Cline (MBARI, USA), and Duane Edgington (MBARI, USA)</i>	
HyperLeaf2024 - A Hyperspectral Imaging Dataset for Classification and Regression of Wheat Leaves	1234
<i>William Michael Laprade (Technical University of Denmark, Denmark), Pawel Pieta (Technical University of Denmark, Denmark), Svetlana Kutuzova (University of Copenhagen, Denmark), Jesper Cairo Westergaard (University of Copenhagen, Denmark), Mads Nielsen (University of Copenhagen, Denmark), Svend Christensen (University of Copenhagen, Denmark), and Anders Bjorholm Dahl (Technical University of Denmark, Denmark)</i>	
Monitoring Social Insect Activity with Minimal Human Supervision	1244
<i>Tarun Sharma (California Institute of Technology, USA), Julian M. Wagner (California Institute of Technology, USA), Sara Beery (MIT, USA), William B. Dickson (California Institute of Technology, USA), Michael H. Dickinson (California Institute of Technology, USA), and Joseph Parker (California Institute of Technology, USA)</i>	

The 5th Omnidirectional Computer Vision Workshop

Sensor Equivariance: A Framework for Semantic Segmentation with Diverse Camera Models	1254
<i>Hannes Reichert (University of Applied Sciences Aschaffenburg), Manuel Hetzel (University of Applied Sciences Aschaffenburg), Andreas Hubert (University of Applied Sciences Aschaffenburg), Konrad Doll (University of Applied Sciences Aschaffenburg), and Bernhard Sick (University of Kassel)</i>	
Estimating Depth of Monocular Panoramic Image with Teacher-Student Model Fusing Equirectangular and Spherical Representations	1262
<i>Jingguo Liu (Southwest University, China), Yijun Xu (Southwest University, China), Shigang Li (Hiroshima City University, Japan), and Jianfeng Li (Southwest University, China)</i>	
BGDNet: Background-guided Indoor Panorama Depth Estimation	1272
<i>Jiajing Chen (Syracuse University), Zhiqiang Wan (Zillow Group), Manjunath Narayana (Zillow Group), Yuguang Li (Zillow Group), Will Hutchcroft (Zillow Group), Senem Velipasalar (Syracuse University), and Sing Bing Kang (Zillow Group)</i>	

DQ-HorizonNet: Enhancing Door Detection Accuracy in Panoramic Images via Dynamic Quantization	1282
<i>Cing-Jia Lin (National Tsing Hua University, Taiwan), Jheng-Wei Su (National Tsing Hua University, Taiwan), Kai-Wen Hsiao (National Tsing Hua University, Taiwan), Ting-Yu Yen (National Tsing Hua University, Taiwan), Chih-Yuan Yao (National Taiwan University of Science and Technology, Taiwan), and Hung-Kuo Chu (National Tsing Hua University, Taiwan)</i>	
Cross-Domain Synthetic-to-Real In-the-Wild Depth and Normal Estimation for 3D Scene Understanding	1290
<i>Jay Bhanushali (Indian Institute of Technology Madras), Manivannan Muniyandi (Indian Institute of Technology Madras), and Praneeth Chakravarthula (UNC Chapel Hill)</i>	
Impact of Video Compression Artifacts on Fisheye Camera Visual Perception Tasks	1301
<i>Madhumitha Sakthi (Qualcomm Technologies Inc., USA), Louis Kerofsky (Qualcomm Technologies Inc., USA), Varun Ravi Kumar (Qualcomm Technologies Inc., USA), and Senthil Yogamani (Automated Driving, QT Technologies Ireland Limited)</i>	
MultiPanoWise: Holistic Deep Architecture for Multi-task Dense Prediction from a Single Panoramic Image	1311
<i>Uzair Shah (Hamad Bin Khalifa University, Qatar), Muhammad Tukur (Hamad Bin Khalifa University, Qatar), Mahmood Alzubaidi (Hamad Bin Khalifa University, Qatar), Giovanni Pintore (CRS4, Italy), Enrico Gobbetti (CRS4, Italy), Mowafa Househ (Hamad Bin Khalifa University, Qatar), Jens Schneider (Hamad Bin Khalifa University, Qatar), and Marco Agus (Hamad Bin Khalifa University, Qatar)</i>	
Multi-scale Attention-Based Inclination Angles Estimation for Panoramic Camera	1322
<i>Yuhao Shan (Southwest University), Heyu Chen (Southwest University), Jiaying Zhang (Southwest University), Shigang Li (Hiroshima City University), and Jianfeng Li (Southwest University)</i>	
FisheyeBEVSeg: Surround View Fisheye Cameras based Bird's-Eye View Segmentation for Autonomous Driving	1331
<i>Senthil Yogamani (Qualcomm Inc, Ireland), David Unger (Torc Robotics, Germany), Venkatraman Narayanan (University of Michigan - Dearborn, USA), and Varun Ravi Kumar (Qualcomm Inc, USA)</i>	
Exploring the Limits: Applying State-of-the-Art Stereo Matching Algorithms to Rectified Ultra-Wide Stereo	1335
<i>Filip Slezák (AGCO A/S, Aalborg University, DK), Morten S. Laursen (AGCO A/S, DK), and Thomas B. Moeslund (Aalborg University, DK)</i>	

4th Workshop on Physics Based Vision meets Deep Learning (PBDL2024)

Gain-first or Exposure-first: Benchmark for Better Low-light Video Photography and Enhancement	1345
<i>Haiyang Jiang (The University of Tokyo, Japan), Zhihang Zhong (Shanghai Artificial Intelligence Laboratory, China), and Yinqiang Zheng (The University of Tokyo, Japan)</i>	

Point-Supervised Semantic Segmentation of Natural Scenes via Hyperspectral Imaging	1357
<i>Tianqi Ren (Nanjing University, China), Qiu Shen (Nanjing University, China), Ying Fu (Beijing Institute of Technology, China), and Shaodi You (University of Amsterdam, Netherlands)</i>	
Computational Spectral Imaging with Unified Encoding Model and Beyond	1368
<i>Xinyuan Liu (Beijing Institute of Technology), Lingen Li (Beijing Institute of Technology), Lin Zhu (Beijing Institute of Technology), and Lizhi Wang (Beijing Institute of Technology)</i>	
ViTKD: Feature-based Knowledge Distillation for Vision Transformers	1379
<i>Zhendong Yang (Tsinghua Shenzhen International Graduate School), Zhe Li (Institute of Automation, Chinese Academy of Sciences), Ailing Zeng (International Digital Economy Academy), Zexian Li (Beihang University), Chun Yuan (Tsinghua Shenzhen International Graduate School), and Yu Li (International Digital Economy Academy)</i>	
Generalized Foggy-Scene Semantic Segmentation by Frequency Decoupling	1389
<i>Qi Bi (University of Amsterdam), Shaodi You (University of Amsterdam), and Theo Gevers (University of Amsterdam)</i>	
Generating Material-Aware 3D Models from Sparse Views	1400
<i>Shi Mao (King Abdullah University of Science and Technology, Saudi Arabia), Chenming Wu (Baidu Research, China), Ran Yi (Shanghai Jiao Tong University, China), Zhelun Shen (Baidu Research, China), Liangjun Zhang (Baidu Research, USA), and Wolfgang Heidrich (King Abdullah University of Science and Technology, Saudi Arabia)</i>	
Physics Based Camera Privacy: Lens and Network Co-Design to the Rescue	1410
<i>Marius Dufraisse (ONERA - DTIS), Marcela Carvalho (Upciti), Pauline Trouvé-Peloux (ONERA - DTIS), and Frédéric Champagnat (ONERA - DTIS)</i>	
Imaging Signal Recovery Using Neural Network Priors Under Uncertain Forward Model Parameters	1420
<i>Xiwen Chen (Clemson University, USA), Wenhui Zhu (Arizona State University, USA), Peijie Qiu (Washington University in St. Louis, USA), and Abolfazl Razi (Clemson University, USA)</i>	
GPT4Motion: Scripting Physical Motions in Text-to-Video Generation via Blender-Oriented GPT Planning	1430
<i>Jiaxi Lv (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), Yi Huang (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), Mingfu Yan (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), Jiancheng Huang (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), Jianzhuang Liu (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), Yifan Liu (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), Yafei Wen (VIVO AI Lab), Xiaoxin Chen (VIVO AI Lab), and Shifeng Chen (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences)</i>	

CVPR 2024 Biometrics Workshop

3D Kinematics Estimation from Video with a Biomechanical Model and Synthetic Training Data..	1441
<i>Zhi-Yi Lin (Delft University of Technology, the Netherlands), Bofan Lyu (Delft University of Technology, the Netherlands), Judith Cueto Fernandez (Delft University of Technology, the Netherlands), Eline van der Kruk (Delft University of Technology, the Netherlands), Ajay Seth (Delft University of Technology, the Netherlands), and Xucong Zhang (Delft University of Technology, the Netherlands)</i>	
Outsmarting Biometric Imposters: Enhancing Iris-Recognition System Security through Physical Adversarial Example Generation and PAD Fine-Tuning	1451
<i>Yuka Ogino (NEC Corporation, Japan), Kazuya Kakizaki (NEC Corporation, Japan), Takahiro Toizumi (NEC Corporation), and Atsushi Ito (NEC Corporation)</i>	
FIQA-FAS: Face Image Quality Assessment Based Face Anti-Spoofing	1462
<i>Ya-Chi Liang (National Tsing Hua University, Taiwan), Min-Xuan Qiu (National Tsing Hua University, Taiwan), and Shang-Hong Lai (National Tsing Hua University, Taiwan)</i>	
Adversarial Identity Injection for Semantic Face Image Synthesis	1471
<i>Giuseppe Tarollo (University of Parma, Italy), Tomaso Fontanini (University of Parma, Italy), Claudio Ferrari (University of Parma, Italy), Guido Borghi (University of Bologna, Italy), and Andrea Prati (University of Parma, Italy)</i>	
Confidence-Aware RGB-D Face Recognition via Virtual Depth Synthesis	1481
<i>Zijian Chen (Inspur (Beijing) Electronic Information Industry Co., Ltd), Mei Wang (Inspur (Beijing) Electronic Information Industry Co., Ltd), Weihong Deng (Inspur (Beijing) Electronic Information Industry Co., Ltd), Hongzhi Shi (Inspur (Beijing) Electronic Information Industry Co., Ltd), Dongchao Wen (Inspur (Beijing) Electronic Information Industry Co., Ltd), Yingjie Zhang (Inspur (Beijing) Electronic Information Industry Co., Ltd), Xingchen Cui (Inspur (Beijing) Electronic Information Industry Co., Ltd), and Jian Zhao (Inspur (Beijing) Electronic Information Industry Co., Ltd)</i>	
GraFIQs: Face Image Quality Assessment Using Gradient Magnitudes	1490
<i>Jan Niklas Kolf (Fraunhofer Institute for Computer Graphics Research IGD, Germany; Technical University of Darmstadt, Germany), Naser Damer (Fraunhofer Institute for Computer Graphics Research IGD, Germany; Technical University of Darmstadt, Germany), and Fadi Boutros (Fraunhofer Institute for Computer Graphics Research IGD, Germany)</i>	
One Embedding to Predict Them All: Visible and Thermal Universal Face Representations for Soft Biometric Estimation via Vision Transformers	1500
<i>Nelida Mirabet-Herranz (EURECOM, France), Chiara Galdi (EURECOM, France), and Jean-Luc Dugelay (EURECOM, France)</i>	

Generalized Single-Image-Based Morphing Attack Detection Using Deep Representations from Vision Transformer	1510
<i>Haoyu Zhang (Norwegian University of Science and Technology, Norway), Raghavendra Ramachandra (Norwegian University of Science and Technology, Norway), Kiran Raja (Norwegian University of Science and Technology, Norway), and Christoph Busch (Norwegian University of Science and Technology, Norway; Darmstadt University of Applied Sciences, Germany)</i>	
Can the Accuracy Bias by Facial Hairstyle be Reduced Through Balancing the Training Data?	1519
<i>Kagan Ozturk (University of Notre Dame, USA), Haiyu Wu (University of Notre Dame, USA), and Kevin W. Bowyer (University of Notre Dame, USA)</i>	
TattTRN: Template Reconstruction Network for Tattoo Retrieval	1529
<i>Lazaro Janier Gonzalez-Soler (Hochschule Darmstadt), Maciej Salwowski (Technical University of Denmark), Christian Rathgeb (Hochschule Darmstadt), and Daniel Fischer (Hochschule Darmstadt)</i>	

Prompting in Vision

What Makes Multimodal In-Context Learning Work?	1539
<i>Folco Bertini Baldassini (Sorbonne Université, France), Mustafa Shukor (Sorbonne Université, France), Matthieu Cord (Sorbonne Université, France; Valeo.ai, France), Laure Soulier (Sorbonne Université, France), and Benjamin Piwowarski (Sorbonne Université, France)</i>	
Conv-Adapter: Exploring Parameter Efficient Transfer Learning for ConvNets	1551
<i>Hao Chen (Carnegie Mellon University), Ran Tao (Carnegie Mellon University), Han Zhang (Carnegie Mellon University), Yidong Wang (Peking University), Xiang Li (Carnegie Mellon University), Wei Ye (Peking University), Jindong Wang (Microsoft Research Asia), Guosheng Hu (Oosto), and Marios Savvides (Carnegie Mellon University)</i>	
Enhancing Visual Question Answering through Question-Driven Image Captions as Prompts	1562
<i>Övgü Özdemir (Middle East Technical University) and Erdem Akagündüz (Middle East Technical University)</i>	
AAPL: Adding Attributes to Prompt Learning for Vision-Language Models	1572
<i>Gahyeon Kim (Korea Institute of Energy Technology(KENTECH), South Korea), Sohee Kim (Korea Institute of Energy Technology(KENTECH), South Korea), and Seokju Lee (Korea Institute of Energy Technology(KENTECH), South Korea)</i>	
Prompting Foundational Models for Omni-supervised Instance Segmentation	1583
<i>Arnav M. Das (University of Washington, Seattle), Ritwick Chaudhry (AWS AI Labs), Kaustav Kundu (AWS AI Labs), and Davide Modolo (AWS AI Labs)</i>	
Low-Rank Few-Shot Adaptation of Vision-Language Models	1593
<i>Maxime Zanella (UCLouvain, Belgium) and Ismail Ben Ayed (ETS, Canada)</i>	
PointPrompt: A Multi-modal Prompting Dataset for Segment Anything Model	1604
<i>Jorge Quesada (Georgia Institute of Technology), Mohammad Alotaibi (Georgia Institute of Technology), Mohit Prabhushankar (Georgia Institute of Technology), and Ghassan AlRegib (Georgia Institute of Technology)</i>	

Uncovering the Hidden Cost of Model Compression	1611
<i>Diganta Misra (Carnegie Mellon University, Landskape AI), Muawiz Chaudhary (Mila, Concordia), Agam Goyal (University of Wisconsin-Madison), Bharat Runwal (Mila - Quebec AI Institute), and Pin Yu Chen (IBM)</i>	

Rhobin 2024: The second Rhobin challenge on Reconstruction of Human-Object Interaction

MoCap-to-Visual Domain Adaptation for Efficient Human Mesh Estimation from 2D Keypoints .	1622
<i>Bedirhan Uguz (Middle East Technical University, Turkey), Ozhan Suat (Middle East Technical University, Turkey), Batuhan Karagoz (Middle East Technical University, Turkey), and Emre Akbas (Middle East Technical University, Turkey; METU ROMER Robotics Center, Turkey)</i>	
V-VIPE: Variational View Invariant Pose Embedding	1633
<i>Mara Levy (University of Maryland, College Park) and Abhinav Shrivastava (University of Maryland, College Park)</i>	
A Survey on 3D Egocentric Human Pose Estimation	1643
<i>Md Mushfiqur Azam (The University of Texas at San Antonio) and Kevin Desai (The University of Texas at San Antonio)</i>	

Fifth Workshop on Neural Architecture Search

CycleGANAS: Differentiable Neural Architecture Search for CycleGAN	1655
<i>Taegun An (Korea University) and Changhee Joo (Korea University)</i>	
The Devil is in Discretization Discrepancy. Robustifying Differentiable NAS with Single-Stage Searching Protocol	1665
<i>Konstanty Subbotko (University of Warsaw), Wojciech Jablonski (University of Warsaw), and Piotr Bilinski (University of Warsaw)</i>	
UP-NAS: Unified Proxy for Neural Architecture Search	1675
<i>Yi-Cheng Huang (National Taiwan University, Taiwan), Wei-Hua Li (National Taiwan University, Taiwan), Chih-Han Tsou (National Taiwan University, Taiwan), Jun-Cheng Chen (Academia Sinica, Taiwan), and Chu-Song Chen (National Taiwan University, Taiwan)</i>	
CSCO: Connectivity Search of Convolutional Operators	1685
<i>Tunhou Zhang (Duke University), Shiyu Li (Duke University), Hsin-Pai Cheng (Qualcomm AI Research), Feng Yan (University of Houston), Hai Li (Duke University), and Yiran Chen (Duke University)</i>	
GRASP-GCN: Graph-Shape Prioritization for Neural Architecture Search Under Distribution Shifts	1695
<i>Sofia Casarin (Free University of Bozen-Bolzano, Italy), Oswald Lanz (Free University of Bozen-Bolzano, Italy), and Sergio Escalera (Computer Vision Center, Barcelona, Spain; Universitat de Barcelona, Barcelona, Spain)</i>	

QuantNAS: Quantization-aware Neural Architecture Search For Efficient Deployment On Mobile Device	1704
<i>Tianxiao Gao (Ant Group, China), Li Guo (Ant Group, China), Shanwei Zhao (Ant Group, China), Peihan Xu (Ant Group, China), Yukun Yang (Ant Group, China), Xionghao Liu (Ant Group, China), Shihao Wang (Ant Group, China), Shiai Zhu (Ant Group, China), and Dajiang Zhou (Ant Group, China)</i>	

2nd Workshop on ``What is Next in Multimodal Foundation Models?"

Strategies to Leverage Foundational Model Knowledge in Object Affordance Grounding	1714
<i>Arushi Rai (University of Pittsburgh, USA), Kyle Buettner (University of Pittsburgh, USA), and Adriana Kovashka (University of Pittsburgh, USA)</i>	
Recognize Anything: A Strong Image Tagging Model	1724
<i>Youcai Zhang (OPPO Research Institute, China), Xinyu Huang (OPPO Research Institute, China), Jinyu Ma (OPPO Research Institute, China), Zhaoyang Li (OPPO Research Institute, China), Zhaochuan Luo (OPPO Research Institute, China), Yanchun Xie (OPPO Research Institute, China), Yuzhuo Qin (OPPO Research Institute, China), Tong Luo (OPPO Research Institute, China), Yaqian Li (OPPO Research Institute, China), Shilong Liu (International Digital Economy Academy, China), Yandong Guo (AI² Robotics, China), and Lei Zhang (International Digital Economy Academy, China)</i>	
ICSVR: Investigating Compositional and Syntactic Understanding in Video Retrieval Models	1733
<i>Avinash Madasu (Intel Labs) and Vasudev Lal (Intel Labs)</i>	
Continual Diffusion with STAMINA: STag-And-Mask INcremental Adapters	1744
<i>James Seale Smith (Samsung Research America, USA), Yen-Chang Hsu (Samsung Research America, USA), Zsolt Kira (Georgia Institute of Technology, USA), Yilin Shen (Samsung Research America, USA), and Hongxia Jin (Samsung Research America, USA)</i>	
Forget-Me-Not: Learning to Forget in Text-to-Image Diffusion Models	1755
<i>Gong Zhang (Georgia Tech & UIUC), Kai Wang (Georgia Tech & UIUC), Xingqian Xu (Georgia Tech & UIUC; Picsart AI Research), Zhangyang Wang (UT Austin; Picsart AI Research), and Humphrey Shi (Georgia Tech & UIUC; Picsart AI Research)</i>	
LLM-Seg: Bridging Image Segmentation and Large Language Model Reasoning	1765
<i>Junchi Wang (ETH Zurich, Switzerland) and Lei Ke (ETH Zurich, Switzerland)</i>	
Matting Anything	1775
<i>Jiachen Li (Georgia Tech & Oregon & UIUC), Jitesh Jain (Georgia Tech & Oregon & UIUC), and Humphrey Shi (Georgia Tech & Oregon & UIUC; Picsart AI Research)</i>	

Robustness Analysis on Foundational Segmentation Models	1786
<i>Madeline Chantry Schiappa (University of Central Florida, USA), Shehreen Azad (University of Central Florida, USA), Sachidanand VS (Indian Institute of Technology, Madras, India), Yunhao Ge (University of Southern California, USA), Ondrej Miksik (Microsoft Research, USA), Yogesh S Rawat (University of Central Florida, USA), and Vibhav Vineet (Microsoft Research, USA)</i>	
Probing Conceptual Understanding of Large Visual-Language Models	1797
<i>Madeline Schiappa (University Of Central Florida, USA), Raiyaan Abdullah (University of Central Florida, USA), Shehreen Azad (University of Central Florida, USA), Jared Claypoole (SRI International, USA), Michael Cogswell (SRI International, USA), Ajay Divakaran (SRI International, USA), and Yogesh Rawat (University of Central Florida, USA)</i>	
Show, Think, and Tell: Thought-Augmented Fine-Tuning of Large Language Models for Video Captioning	1808
<i>Byoungjip Kim (LG AI Research, South Korea), Dasol Hwang (LG AI Research, South Korea), Sungjun Cho (LG AI Research, South Korea), Youngsoo Jang (LG AI Research, South Korea), Honglak Lee (LG AI Research, South Korea), and Moontae Lee (LG AI Research, South Korea)</i>	
Wiki-LLaVA: Hierarchical Retrieval-Augmented Generation for Multimodal LLMs	1818
<i>Davide Caffagni (University of Modena and Reggio Emilia, Italy), Federico Cocchi (University of Modena and Reggio Emilia, Italy), Nicholas Moratelli (University of Modena and Reggio Emilia, Italy), Sara Sarto (University of Modena and Reggio Emilia, Italy), Marcella Cornia (University of Modena and Reggio Emilia, Italy), Lorenzo Baraldi (University of Modena and Reggio Emilia, Italy), and Rita Cucchiara (University of Modena and Reggio Emilia, Italy)</i>	
Benchmarking Zero-Shot Recognition with Vision-Language Models: Challenges on Granularity and Specificity	1827
<i>Zhenlin Xu (Amazon), Yi Zhu (Boson AI), Siqi Deng (Amazon), Abhay Mittal (Meta), Yanbei Chen (Amazon), Manchen Wang (Meta), Paolo Favaro (University of Bern), Joseph Tighe (Meta), and Davide Modolo (Amazon)</i>	
Towards Efficient Audio-Visual Learners via Empowering Pre-trained Vision Transformers with Cross-Modal Adaptation	1837
<i>Kai Wang (University of Toronto, Canada), Yapeng Tian (University of Texas at Dallas, USA), and Dimitrios Hatzinikos (University of Toronto, Canada)</i>	

Pixel-level Video Understanding in the Wild Challenge

ChatVTG: Video Temporal Grounding via Chat with Video Dialogue Large Language Models	1847
<i>Mengxue Qu (Beijing Jiaotong University), Xiaodong Chen (University of Science and Technology of China), Wu Liu (University of Science and Technology of China), Alicia Li (Horace Mann School), and Yao Zhao (Beijing Jiaotong University)</i>	

SAM-PM: Enhancing Video Camouflaged Object Detection using Spatio-Temporal Attention	1857
<i>Muhammad Nawfal Meeran (National Institute of Technology, Tiruchirappalli), Gokul Adethya T (National Institute of Technology, Tiruchirappalli), and Bhanu Pratyush Mantha (National Institute of Technology, Tiruchirappalli)</i>	

Workshop on Human Motion Generation

T2LM: Long-Term 3D Human Motion Generation from Multiple Sentences	1867
<i>Taeryung Lee (Seoul National University), Fabien Baradel (Naver Labs Europe), Thomas Lucas (Naver Labs Europe), Kyoung Mu Lee (Seoul National University), and Grégory Rogez (Naver Labs Europe)</i>	
Speech2UnifiedExpressions: Synchronous Synthesis of Co-Speech Affective Face and Body Expressions from Affordable Inputs	1877
<i>Uttaran Bhattacharya (Adobe Inc.), Aniket Bera (Purdue University), and Dinesh Manocha (University of Maryland)</i>	
Exploring Text-to-Motion Generation with Human Preference	1888
<i>Jenny Sheng (Tsinghua University), Matthieu Lin (Tsinghua University), Andrew Zhao (Tsinghua University), Kevin Pruvost (Tsinghua University), Yu-Hui Wen (Beijing Jiaotong University), Yangguang Li (Shanghai AI Lab), Gao Huang (Tsinghua University), and Yong-Jin Liu (Tsinghua University)</i>	
Two-Person Interaction Augmentation with Skeleton Priors	1900
<i>Baiyi Li (University of Leeds, United Kingdom), Edmond S. L. Ho (University of Glasgow, United Kingdom), Hubert P. H. Shum (Durham University, United Kingdom), and He Wang (University College London, United Kingdom)</i>	
Multi-Track Timeline Control for Text-Driven 3D Human Motion Generation	1911
<i>Mathis Petrovich (Univ Gustave Eiffel, France; Max Planck Institute for Intelligent Systems, Germany), Or Litany (NVIDIA; Technion, Israel), Umar Iqbal (NVIDIA, USA), Michael J. Black (Max Planck Institute for Intelligent Systems, Germany), Gul Varol (Univ Gustave Eiffel, France), Xue Bin Peng (NVIDIA; Simon Fraser University, Canada), and Davis Rempe (NVIDIA, USA)</i>	
DiffTED: One-shot Audio-driven TED Talk Video Generation with Diffusion-based Co-speech Gestures	1922
<i>Steven Hogue (University of Texas at Dallas, USA), Chenxu Zhang (University of Texas at Dallas, USA), Hamza Daruger (University of Texas at Dallas, USA), Yapeng Tian (University of Texas at Dallas, USA), and Xiaohu Guo (University of Texas at Dallas, USA)</i>	
A Cross-Dataset Study for Text-based 3D Human Motion Retrieval	1932
<i>Léore Bensabath (Univ Gustave Eiffel, France), Mathis Petrovich (Univ Gustave Eiffel, France; Max Planck Institute for Intelligent Systems, Germany), and Gul Varol (Univ Gustave Eiffel, France)</i>	

in2IN: Leveraging Individual Information to Generate Human Interactions	1941
<i>Pablo Ruiz-Ponce (Universidad de Alicante, Spain), German Barquero (Universitat de Barcelona, Spain), Cristina Palmero (Universitat de Barcelona, Spain), Sergio Escalera (Universitat de Barcelona, Spain), and José García-Rodríguez (Universidad de Alicante, Spain)</i>	
Fake It to Make It: Using Synthetic Data to Remedy the Data Shortage in Joint Multimodal Speech-and-Gesture Synthesis	1952
<i>Shivam Mehta (KTH Royal Institute of Technology), Anna Deichler (KTH Royal Institute of Technology), Jim O'Regan (KTH Royal Institute of Technology), Birger Moëll (KTH Royal Institute of Technology), Jonas Beskow (KTH Royal Institute of Technology), Gustav Eje Henter (KTH Royal Institute of Technology), and Simon Alexanderson (KTH Royal Institute of Technology)</i>	

7th MUltimodal Learning and Applications

Cross-Modal Fusion and Attention Mechanism for Weakly Supervised Video Anomaly Detection	1965
<i>Ayush Ghadiya (Sony Research India, Bangalore, India), Purbayan Kar (Sony Research India, Bangalore, India), Vishal Chudasama (Sony Research India, Bangalore, India), and Pankaj Wasnik (Sony Research India, Bangalore, India)</i>	
Leveraging Generative Language Models for Weakly Supervised Sentence Component Analysis in Video-Language Joint Learning	1975
<i>Zaber Ibn Abdul Hakim (Bangladesh University of Engineering and Technology, Bangladesh), Najibul Haque Sarker (Bangladesh University of Engineering and Technology, Bangladesh), Rahul Pratap Singh (Netaji Subhas University of Technology, India), Bishmoy Paul (Bangladesh University of Engineering and Technology, Bangladesh), Ali Dabouei (Carnegie Mellon University, USA), and Min Xu (Carnegie Mellon University, USA)</i>	
De-noised Vision-language Fusion Guided by Visual Cues for E-commerce Product Search	1986
<i>Zhizhang Hu (University of California, Merced), Shasha Li (Amazon Visual Search & AR), Ming Du (Amazon Visual Search & AR), Arnab Dhua (Amazon Visual Search & AR), and Douglas Gray (Amazon Visual Search & AR)</i>	
RGB-D Cube R-CNN: 3D Object Detection with Selective Modality Dropout	1997
<i>Jens Piekenbrinck (RWTH Aachen University, Germany), Alexander Hermans (RWTH Aachen University, Germany), Narunas Vaskevicius (Robert Bosch GmbH, Germany), Timm Linder (Robert Bosch GmbH, Germany), and Bastian Leibe (RWTH Aachen University, Germany)</i>	
Multimodal Understanding of Memes with Fair Explanations	2007
<i>Yang Zhong (University of Pittsburgh) and Bhiman Kumar Baghel (University of Pittsburgh)</i>	
Listen Then See: Video Alignment with Speaker Attention	2018
<i>Aviral Agrawal (Carnegie Mellon University, USA), Carlos Mateo Samudio Lezcano (Carnegie Mellon University, USA), Iqui Balam Heredia-Marin (Carnegie Mellon University, USA), and Prabhdeep Singh Sethi (Carnegie Mellon University, USA)</i>	

InVERGe: Intelligent Visual Encoder for Bridging Modalities in Report Generation	2028
<i>Ankan Deria (Jio Institute, India), Komal Kumar (Jio Institute, India), Snehashis Chakraborty (Jio Institute, India), Dwarikanath Mahapatra (Inception Institute of Artificial Intelligence, UAE), and Sudipta Roy (Jio Institute, India)</i>	
LAformer: Trajectory Prediction for Autonomous Driving with Lane-Aware Scene Constraints ...	2039
<i>Mengmeng Liu (University of Twente), Hao Cheng (University of Twente), Lin Chen (VISCODA GmbH), Hellward Broszio (VISCODA GmbH), Jiangtao Li (PhiGent Robotics), Runjiang Zhao (PhiGent Robotics), Monika Sester (Leibniz University Hannover), and Michael Ying Yang (University of Bath)</i>	
ZInD-Tell: Towards Translating Indoor Panoramas into Descriptions	2050
<i>Tonmoay Deb (Northwestern University), Lichen Wang (Zillow Group), Zachary Bessinger (Zillow Group), Naji Khosravan (Zillow Group), Eric Penner (Zillow Group), and Sing Bing Kang (Zillow Group)</i>	
VMCML: Video and Music Matching via Cross-Modality Lifting	2060
<i>Yi-Shan Lee (National Tsing Hua University, Taiwan), Wei-Cheng Tseng (University of Toronto, Canada; Vector Institute, Canada), Fu-En Wang (National Tsing Hua University, Taiwan), and Min Sun (National Tsing Hua University, Taiwan)</i>	
AIGeN: An Adversarial Approach for Instruction Generation in VLN	2070
<i>Niyati Rawal (University of Modena and Reggio Emilia), Roberto Bigazzi (University of Modena and Reggio Emilia), Lorenzo Baraldi (University of Modena and Reggio Emilia), and Rita Cucchiara (University of Modena and Reggio Emilia)</i>	
Multi-Modal Fusion of Event and RGB for Monocular Depth Estimation Using a Unified Transformer-based Architecture	2081
<i>Anusha Devulapally (The Pennsylvania State University, USA), Md Fahim Faysal Khan (The Pennsylvania State University, USA), Siddharth Advani (Samsung Electronics America, USA), and Vijaykrishnan Narayanan (The Pennsylvania State University, USA)</i>	
Exploring the Role of Audio in Video Captioning	2090
<i>Yuhan Shen (Northeastern University), Linjie Yang (ByteDance), Longyin Wen (ByteDance), Haichao Yu (Adobe), Ehsan Elhamifar (Northeastern University), and Heng Wang (TikTok)</i>	

Embedded Vision Workshop

Dedicated Inference Engine and Binary-Weight Neural Networks for Lightweight Instance Segmentation	2101
<i>Tse-Wei Chen (Canon Inc.), Wei Tao (Canon Innovative Solution (Beijing) Co., Ltd.), Dongyue Zhao (Canon Innovative Solution (Beijing) Co., Ltd.), Kazuhiro Mima (Canon Inc.), Tadayuki Ito (Canon Inc.), Kinya Osa (Canon Inc.), and Masami Kato (Canon Inc.)</i>	
Lightweight Maize Disease Detection through Post-Training Quantization with Similarity Preservation	2111
<i>Carlos Victorino Padeiro (Nagoya University), Tse-Wei Chen (Nagoya University), Takahiro Komamizu (Nagoya University), and Ichiro Ide (Nagoya University)</i>	

Multi-bit, Black-box Watermarking of Deep Neural Networks in Embedded Applications	2121
<i>Sam Leroux (Ghent University - imec), Stijn Vanassche (Ghent University - imec), and Pieter Simoens (Ghent University - imec)</i>	
Pruning as a Binarization Technique	2131
<i>Lukas Frickenstein (BMW Group, Germany), Pierpaolo Mori (Politecnico Di Torino, Italy), Shambhavi Balamuthu Sampath (BMW Group, Germany), Moritz Thoma (BMW Group, Germany), Nael Fasfous (BMW Group, Germany), Manoj Rohit Vemparala (BMW Group, Germany), Alexander Frickenstein (BMW Group, Germany), Christian Unger (BMW Group, Germany), Claudio Passerone (Politecnico Di Torino, Italy), and Walter Stechele (Technical University of Munich, Germany)</i>	
Neuromorphic Lip-Reading with Signed Spiking Gated Recurrent Units	2141
<i>Manon Dampfhoffer (Univ. Grenoble Alpes, France) and Thomas Mesquida (Univ. Grenoble Alpes, France)</i>	
Efficient Video Stabilization via Partial Block Phase Correlation on Edge GPUs	2152
<i>Cevahir Çığla (Aselsan Inc.)</i>	
SciFlow: Empowering Lightweight Optical Flow Models with Self-Cleaning Iterations	2162
<i>Jamie Menjay Lin (Qualcomm Technologies, Inc.), Jisoo Jeong (Qualcomm AI Research), Hong Cai (Qualcomm AI Research), Risheek Garrepalli (Qualcomm AI Research), Kai Wang (Qualcomm Technologies, Inc.), and Fatih Porikli (Qualcomm Technologies, Inc.)</i>	
Structured Sparse Back-propagation for Lightweight On-Device Continual Learning on Microcontroller Units	2172
<i>Francesco Paissan (Fondazione Bruno Kessler, Italy), Davide Nadalini (Politecnico di Torino e Università di Bologna, Italy), Manuele Rusci (KU Leuven, Belgium), Alberto Ancilotto (Fondazione Bruno Kessler, Italy), Francesco Conti (Università di Bologna, Italy), Luca Benini (ETHZ Zurich, Switzerland and Università di Bologna, Italy), and Elisabetta Farella (Fondazione Bruno Kessler, Italy)</i>	
Multi-resolution Rescored ByteTrack for Video Object Detection on Ultra-low-power Embedded Systems	2182
<i>Luca Bompani (University of Bologna, Italy), Manuele Rusci (KU Leuven, Belgium), Daniele Palossi (USI-SUPSI, Switzerland), Francesco Conti (University of Bologna, Italy), and Luca Benini (University of Bologna, Italy)</i>	
ED-DCFNet: An Unsupervised Encoder-decoder Neural Model for Event-driven Feature Extraction and Object Tracking	2191
<i>Raz Ramon (The Open University of Israel), Hadar Cohen-Duwek (The Open University of Israel), and Elishai Ezra Tsur (The Open University of Israel)</i>	
RAVN: Reinforcement Aided Adaptive Vector Quantization of Deep Neural Networks	2200
<i>Anamika Jha (Mercedes-Benz Research & Development India), Aratrik Chattopadhyay (Mercedes-Benz Research & Development India), Mrinal Banerji (Mercedes-Benz Research & Development India), and Disha Jain (Mercedes-Benz Research & Development India)</i>	

Prune Efficiently by Soft Pruning	2210
<i>Parakh Agarwal (Texas Instruments), Manu Mathew (Texas Instruments), Kunal Ranjan Patel (Texas Instruments), Varun Tripathi (Texas Instruments), and Pramod Swami (Texas Instruments)</i>	
Content-aware Input Scaling and Deep Learning Computation Offloading for Low-Latency Embedded Vision	2218
<i>Omkar Prabhune (Purdue University), Tianen Chen (University of Wisconsin-Madison), and Younghyun Kim (Purdue University)</i>	

Multimodal Algorithmic Reasoning Workshop

Using Language-Aligned Gesture Embeddings for Understanding Gestures Accompanying Math Terms	2227
<i>Tristan Maidment (University of Pittsburgh), Purav J Patel (University of Maryland - College Park), Erin Walker (University of Pittsburgh), and Adriana Kovashka (University of Pittsburgh)</i>	
What Does CLIP Know About Peeling a Banana?	2238
<i>Claudia Cuttano (Politecnico di Torino), Gabriele Rosi (Politecnico di Torino), Gabriele Trivigno (Politecnico di Torino), and Giuseppe Averta (Politecnico di Torino)</i>	
Task Navigator: Decomposing Complex Tasks for Multimodal Large Language Models	2248
<i>Feipeng Ma (University of Science and Technology of China, China), Yizhou Zhou (Tencent, China), Yueyi Zhang (University of Science and Technology of China, China), Siying Wu (Hefei Comprehensive National Science Center, China), Zheyu Zhang (University of Science and Technology of China, China), Zilong He (University of Science and Technology of China, China), Fengyun Rao (Tencent, China), and Xiaoyan Sun (University of Science and Technology of China, China)</i>	
Multi-Explainable TemporalNet: An Interpretable Multimodal Approach using Temporal Convolutional Network for User-level Depression Detection	2258
<i>Anas Zafar (National University of Computer and Emerging Sciences, Karachi, Pakistan), Danyal Aftab (National University of Computer and Emerging Sciences, Karachi, Pakistan), Rizwan Qureshi (Center for Regenerative Medicine and Health, Hong Kong Institute of Science and Innovation, Chinese Academy of Sciences, Hong Kong Science Park, Hong Kong), Yaofeng Wang (Center for Regenerative Medicine and Health, Hong Kong Institute of Science and Innovation, Chinese Academy of Sciences, Hong Kong Science Park, Hong Kong), and Hong Yan (City University of Hong Kong)</i>	
ViTA: An Efficient Video-to-Text Algorithm using VLM for RAG-based Video Analysis System .	2266
<i>Md Adnan Arefeen (University of Missouri-Kansas City, USA), Biplob Debnath (NEC Labs America, USA), Md Yusuf Sarwar Uddin (University of Missouri-Kansas City, USA), and Srimat Chakradhar (NEC Labs America, USA)</i>	

Data Curation and Augmentation in Enhancing Medical Imaging Applications

Strategies to Improve Real-World Applicability of Laparoscopic Anatomy Segmentation Models..	2275
<i>Fiona R. Kolbinger (Purdue University, USA), Jiangpeng He (Purdue University, USA), Jinge Ma (Purdue University, USA), and Fengqing Zhu (Purdue University, USA)</i>	
nnMobileNet: Rethinking CNN for Retinopathy Research	2285
<i>Wenhui Zhu (Arizona State University, USA), Peijie Qiu (Washington University in St.Louis, USA), Xiwen Chen (Clemson University, USA), Xin Li (Arizona State University, USA), Natasha Lepore (Children's Hospital Los Angeles, USA), Oana M. Dumitrascu (Mayo Clinic, USA), and Yalin Wang (Arizona State University, USA)</i>	
Distribution-Aware Multi-Label FixMatch for Semi-Supervised Learning on CheXpert.	2295
<i>Sontje Ihler (Leibniz Universität Hannover, Germany), Felix Kuhnke (Independent Researcher), Timo Kuhlitz (Leibniz Universität Hannover, Germany), and Thomas Seel (Leibniz Universität Hannover, Germany)</i>	
Repurposing the Image Generative Potential: Exploiting GANs to Grade Diabetic Retinopathy	2305
<i>Isabella Poles (Politecnico di Milano, Italy), Eleonora D'Arnese (Politecnico di Milano, Italy), Luca G. Cellamare (Politecnico di Milano, Italy), Marco D. Santambrogio (Politecnico di Milano, Italy), and Darvin Yi (University of Illinois at Chicago, USA)</i>	
Repeat and Concatenate: 2D to 3D Image Translation with 3D to 3D Generative Modeling	2315
<i>Abril Corona-Figueroa (Durham University, United Kingdom), Hubert P. H. Shum (Durham University, United Kingdom), and Chris G. Willcocks (Durham University, United Kingdom)</i>	
ControlPolypNet: Towards Controlled Colon Polyp Synthesis for Improved Polyp Segmentation	2325
<i>Vanshali Sharma (Indian Institute of Technology Guwahati, India), Abhishek Kumar (Eli Lilly and Company, India), Debesh Jha (Northwestern University, USA), M.K. Bhuyan (Indian Institute of Technology Guwahati, India), Pradip K. Das (Indian Institute of Technology Guwahati, India), and Ulas Bagci (Northwestern University, USA)</i>	
Generation of Structurally Realistic Retinal Fundus Images with Diffusion Models	2335
<i>Sojung Go (Seoul National University Bundang Hospital, Rep. of Korea), Younghoon Ji (Vuno Inc., Rep. of Korea), Sang Jun Park (Seoul National University Bundang Hospital, Rep. of Korea.), and Soochahn Lee (Kookmin University, Rep. of Korea)</i>	
A Comparative Analysis of Implicit Augmentation Techniques for Breast Cancer Diagnosis Using Multiple Views	2345
<i>Yumnah Hasan (University of Limerick), Talhat Khan (University of Limerick), Darian Reyes Fernandez de Bulnes (University of Limerick), Juan F H Albarracin (University of Limerick), and Conor Ryan (University of Limerick)</i>	

Creating a Digital Twin of Spinal Surgery: A Proof of Concept	2355
<i>Jonas Hein (ETH Zurich, Switzerland), Frédéric Giraud (University of Zurich, Switzerland), Lilian Calvet (University of Zurich, Switzerland), Alexander Schwarz (ETH Zurich, Switzerland), Nicola Alessandro Cavalcanti (University of Zurich, Switzerland), Sergey Prokudin (ETH Zurich, Switzerland), Mazda Farshad (University of Zurich, Switzerland), Siyu Tang (ETH Zurich, Switzerland), Marc Pollefeys (ETH Zurich, Switzerland), Fabio Carrillo (University of Zurich, Switzerland), and Philipp Färnstahl (University of Zurich, Switzerland)</i>	
Codebook VQ-VAE Approach for Prostate Cancer Diagnosis using Multiparametric MRI	2365
<i>Ekaterina Redekop (University of California, Los Angeles, USA), Mara Pleasure (University of California, Los Angeles, USA), Zichen Wang (University of California, Los Angeles, USA), Karthik V Sarma (University of California, Los Angeles, USA), Adam Kinnaird (University of Alberta, Canada), William Speier (University of California, Los Angeles, USA), and Corey W Arnold (University of California, Los Angeles, USA)</i>	
Advancing Brain Tumor Analysis: Curating a High-Quality MRI Dataset for Deep Learning-Based Molecular Marker Profiling	2373
<i>Divya D. Reddy (University of Texas Southwestern Medical Center), Niloufar Saadat (University of Texas Southwestern Medical Center), James M. Holcomb (University of Texas Southwestern Medical Center), Benjamin C. Wagner (University of Texas Southwestern Medical Center), Nghi C. Truong (University of Texas Southwestern Medical Center), Jason Bowerman (University of Texas Southwestern Medical Center), Kimmo J. Hatanpää (University of Texas Southwestern Medical Center), Toral R Patel (University of Texas Southwestern Medical Center), Marco C. Pinho (University of Texas Southwestern Medical Center), Ananth J Madhuranthakam (University of Texas Southwestern Medical Center), Chandan Ganesh Bangalore Yogananda (University of Texas Southwestern Medical Center), and Joseph A. Maldjian (University of Texas Southwestern Medical Center)</i>	
Privacy-Preserving Collaboration for Multi-Organ Segmentation via Federated Learning from Sites with Partial Labels	2380
<i>Adway Kanhere (University of Maryland School of Medicine, USA), Pranav Kulkarni (University of Maryland School of Medicine, USA), Paul H. Yi (University of Maryland School of Medicine, USA), and Vishwa S. Parekh (University of Maryland School of Medicine, USA)</i>	
GSAM+Cutie: Text-Promptable Tool Mask Annotation for Endoscopic Video	2388
<i>Roger D. Soberanis-Mukul (Johns Hopkins University, USA), Jiahuan Cheng (Johns Hopkins University, USA), Jan Emily Mangulabnan (Johns Hopkins University, USA), S. Swaroop Vedula (Johns Hopkins University, USA), Masaru Ishii (Johns Hopkins Medical Institutions, USA), Gregory Hager (Johns Hopkins University, USA), Russell H. Taylor (Johns Hopkins University, USA), and Mathias Unberath (Johns Hopkins University, USA)</i>	

MMIST-ccRCC: A Real World Medical Dataset for the Development of Multi-Modal Systems	2395
<i>Tiago Mota (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), M. Rita Verdelho (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), Diogo J. Araújo (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), Alceu Bissoto (Institute of Computing, University of Campinas, Brazil), Carlos Santiago (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), and Catarina Barata (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal)</i>	
Hairy Ground Truth Enhancement for Semantic Segmentation	2404
<i>Sophie Fischer (University of Oxford, UK) and Irina Voiculescu (University of Oxford, UK)</i>	
Beyond Respiratory Models: A Physics-enhanced Synthetic Data Generation Method for 2D-3D Deformable Registration	2413
<i>François Lecomte (INRIA, France), Pablo Alvarez (INRIA, France), Stéphane Cotin (INRIA, France), and Jean-Louis Dillenseger (University of Rennes, France)</i>	
UltraAugment: Fan-shape and Artifact-based Data Augmentation for 2D Ultrasound Images	2422
<i>Florian Ramakers (Katholieke Universiteit Leuven), Tom Vercauteren (King's College London), Jan Deprest (Katholieke Universiteit Leuven), and Helena Williams (Katholieke Universiteit Leuven)</i>	

Women in Computer Vision

PARASOL: Parametric Style Control for Diffusion Image Synthesis	2432
<i>Gemma Canet Tarrés (University of Surrey, United Kingdom), Dan Ruta (University of Surrey, United Kingdom), Tu Bui (University of Surrey, United Kingdom), and John Collomosse (University of Surrey, United Kingdom)</i>	
Extending Global-local View Alignment for Self-supervised Learning with Remote Sensing Imagery	2443
<i>Xinye Wanyan (University of Melbourne, Australia), Sachith Seneviratne (University of Melbourne, Australia), Shuchang Shen (University of Melbourne, Australia), and Michael Kirley (University of Melbourne, Australia)</i>	
RetinaLiteNet: A Lightweight Transformer based CNN for Retinal Feature Segmentation	2454
<i>Mehwish Mehmood (Queen's University Belfast, United Kingdom), Majed Alsharari (Queen's University Belfast), Shahzaib Iqbal (Abasyn University Islamabad), Ivor Spence (Queen's University Belfast), and Muhammad Fahim (Queen's University Belfast)</i>	
ABC-CapsNet: Attention based Cascaded Capsule Network for Audio Deepfake Detection	2464
<i>Taiba Majid Wani (Sapienza University of Rome, Italy), Reeve Gulzar (Sapienza University of Rome, Italy), and Irene Amerini (Sapienza University of Rome, Italy)</i>	

GestFormer: Multiscale Wavelet Pooling Transformer Network for Dynamic Hand Gesture Recognition	2473
<i>Mallika Garg (Indian Institute of Technology, Roorkee, India), Debashis Ghosh (Indian Institute of Technology, Roorkee, India), and Pyari Mohan Pradhan (Indian Institute of Technology, Roorkee, India)</i>	
Unsupervised Domain Adaptation for Weed Segmentation Using Greedy Pseudo-labelling	2484
<i>Yingchao Huang (University of Regina) and Abdul Bais (Professor in University of Regina)</i>	

SyntaGen: Harnessing Generative Models for Synthetic Visual Datasets

RePoseDM: Recurrent Pose Alignment and Gradient Guidance for Pose Guided Image Synthesis	2495
<i>Anant Khandelwal (Indian Institute of Technology Delhi)</i>	
Is Synthetic Data all We Need? Benchmarking the Robustness of Models Trained with Synthetic Images	2505
<i>Krishnakant Singh (TU Darmstadt), Thanush Navaratnam (TU Darmstadt), Jannik Holmer (TU Darmstadt), Simone Schaub-Meyer (TU Darmstadt), and Stefan Roth (TU Darmstadt)</i>	

2nd Workshop on Scene Graphs and Graph Representation Learning

FloCoDe: Unbiased Dynamic Scene Graph Generation with Temporal Consistency and Correlation Debiasing	2516
<i>Anant Khandelwal (Indian Institute of Technology Delhi)</i>	
VideoSAGE: Video Summarization with Graph Representation Learning	2527
<i>Jose M. Rojas Chaves (Intel Corporation) and Subarna Tripathi (Intel Labs)</i>	
EgoSG: Learning 3D Scene Graphs from Egocentric RGB-D Sequences	2535
<i>Chaoyi Zhang (University of Sydney, Australia), Xitong Yang (Meta, USA), Ji Hou (Meta, USA), Kris Kitani (Meta, USA), Weidong Cai (University of Sydney, Australia), and Fu-Jen Chu (Meta, USA)</i>	
Efflex: Efficient and Flexible Pipeline for Spatio-Temporal Trajectory Graph Modeling and Representation Learning	2546
<i>Ming Cheng (Dartmouth College), Ziyi Zhou (Dartmouth College), Bowen Zhang (Shanghai Jiao Tong University), Ziyu Wang (University of California, Irvine), Jiaqi Gan (Dartmouth College), Ziang Ren (Dartmouth College), Weiqi Feng (Harvard University), Yi Lyu (Independent Researcher), Hefan Zhang (Dartmouth College), and Xingjian Diao (Dartmouth College)</i>	
Segment Anything Model for Road Network Graph Extraction	2556
<i>Congrui Hetang (Carnegie Mellon University), Haoru Xue (Carnegie Mellon University), Cindy Le (Columbia University), Tianwei Yue (Carnegie Mellon University), Wenping Wang (Carnegie Mellon University), and Yihui He (Carnegie Mellon University)</i>	

A Review and Efficient Implementation of Scene Graph Generation Metrics	2567
<i>Julian Lorenz (University of Augsburg, Germany), Robin Schön (University of Augsburg, Germany), Katja Ludwig (University of Augsburg, Germany), and Rainer Lienhart (University of Augsburg, Germany)</i>	

L3D-IVU: 3rd Workshop on Learning with Limited Labelled Data for Image and Video Understanding

SemiGPC: Distribution-Aware Label Refinement for Imbalanced Semi-Supervised Learning Using Gaussian Processes	2576
<i>Abdelhak Lemkhenter (AWS AI Labs), Manchen Wang (AWS AI Labs), Luca Zancato (AWS AI Labs), Gurumurthy Swaminathan (AWS AI Labs), Paolo Favaro (AWS AI Labs), and Davide Modolo (AWS AI Labs)</i>	
Uncertainty-based Forgetting Mitigation for Generalized Few-Shot Object Detection	2586
<i>Karim Guirguis (Robert Bosch GmbH), George Eskandar (University of Stuttgart), Mingyang Wang (Robert Bosch GmbH), Matthias Kayser (Robert Bosch GmbH), Eduardo Monari (Robert Bosch GmbH), Bin Yang (University of Stuttgart), and Jürgen Beyerer (Fraunhofer IOSB)</i>	
Image-caption Difficulty for Efficient Weakly-supervised Object Detection from In-the-wild Data	2596
<i>Giacomo Nebbia (University of Pittsburgh, USA) and Adriana Kovashka (University of Pittsburgh, USA)</i>	
Learning Tracking Representations from Single Point Annotations	2606
<i>Qiangqiang Wu (City University of Hong Kong) and Antoni B. Chan (City University of Hong Kong)</i>	
CDAD-Net: Bridging Domain Gaps in Generalized Category Discovery	2616
<i>Sai Bhargav Rongali (Indian Institute of Technology Bombay), Sarthak Mehrotra (Indian Institute of Technology Bombay), Ankit Jha (INRIA, Grenoble France), Mohamad Hassan N C (Indian Institute of Technology Bombay), Shirsha Bose (Technical University of Munich), Tanisha Gupta (Indian Institute of Technology Bombay), Mainak Singha (Indian Institute of Technology Bombay), and Biplab Banerjee (Indian Institute of Technology Bombay)</i>	
Audio-Visual Generalized Zero-Shot Learning using Pre-Trained Large Multi-Modal Models	2627
<i>David Kurzendörfer (Localyzer GmbH, Germany), Otniel-Bogdan Mercea (University of Tübingen, Germany; Tübingen AI Center, Germany; Helmholtz Munich, Germany), A. Sophia Koepke (University of Tübingen, Germany; Tübingen AI Center, Germany), and Zeynep Akata (Tübingen AI Center, Germany; Helmholtz Munich, Germany; Technical University of Munich, Germany)</i>	
Latent-based Diffusion Model for Long-tailed Recognition	2639
<i>Pengxiao Han (Australian National University), Changkun Ye (Australian National University, CSIRO, Australia), Jieming Zhou (Australian National University, CSIRO, Australia), Jing Zhang (Australian National University, CSIRO, Australia), Jie Hong (Australian National University, CSIRO, Australia), and Xuesong Li (Australian National University, CSIRO, Australia)</i>	

MoDA: Leveraging Motion Priors from Videos for Advancing Unsupervised Domain Adaptation in Semantic Segmentation	2649
<i>Fei Pan (University of Michigan), Xu Yin (Korea Advanced Institute of Science and Technology), Seokju Lee (Korea Institute of Energy Technology), Axi Niu (Korea Advanced Institute of Science and Technology), Sungeui Yoon (Korea Advanced Institute of Science and Technology), and In So Kweon (Korea Advanced Institute of Science and Technology)</i>	
Active Transferability Estimation	2659
<i>Tarun Ram Menta (Adobe Systems, India), Surjan Jandial (Adobe Systems, India), Akash Patil (IIT Madras, India), Saketh Bachu (IIT Hyderabad, India), Vimal K.B (IIT Hyderabad, India), Balaji Krishnamurthy (Adobe Systems, India), Vineeth N. Balasubramanian (IIT Hyderabad, India), Mausoom Sarkar (Adobe Systems, India), and Chirag Agarwal (Harvard University, USA)</i>	
What is Point Supervision Worth in Video Instance Segmentation?	2671
<i>Shuaiyi Huang (University of Maryland, College Park, USA), De-An Huang (NVIDIA, USA), Zhiding Yu (NVIDIA, USA), Shiyi Lan (NVIDIA, USA), Subhashree Radhakrishnan (NVIDIA, USA), Jose M. Alvarez (NVIDIA, USA), Abhinav Shrivastava (University of Maryland, College Park, USA), and Anima Anandkumar (Caltech, USA)</i>	
UVIS: Unsupervised Video Instance Segmentation	2682
<i>Shuaiyi Huang (University of Maryland, College Park, USA), Saksham Suri (University of Maryland, College Park, USA), Kamal Gupta (University of Maryland, College Park, USA), Sai Saketh Rambhatla (Meta, USA), Ser-nam Lim (University of Central Florida, USA), and Abhinav Shrivastava (University of Maryland, College Park, USA)</i>	
Open-world Instance Segmentation: Top-down Learning with Bottom-up Supervision	2693
<i>Tarun Kalluri (UC San Diego), Weiyao Wang (Meta AI), Heng Wang (Meta AI), Manmohan Chandraker (UC San Diego), Lorenzo Torresani (Meta AI), and Du Tran (Meta AI)</i>	
Weakly-Supervised Temporal Action Localization with Multi-Modal Plateau Transformers	2704
<i>Xin Hu (Tulane University, USA), Kai Li (NEC Labs America, USA), Deep Patel (NEC Labs America, USA), Erik Kruus (NEC Labs America, USA), Martin Renqiang Min (NEC Labs America, USA), and Zhengming Ding (Tulane University, USA)</i>	
On Accuracy and Speed of Geodesic Regression: Do Geometric Priors Improve Learning on Small Datasets?	2714
<i>Adele Myers (UC Santa Barbara) and Nina Miolane (UC Santa Barbara)</i>	
Human-in-the-Loop Segmentation of Multi-species Coral Imagery	2723
<i>Scarlett Raine (Queensland University of Technology, Australia), Ross Marchant (Image Analytics, Australia), Brano Kusy (CSIRO Data61, Australia), Frederic Maire (Queensland University of Technology, Australia), Niko Sunderhauf (Queensland University of Technology, Australia), and Tobias Fischer (Queensland University of Technology, Australia)</i>	

Zero-Shot Monocular Motion Segmentation in the Wild by Combining Deep Learning with Geometric Motion Model Fusion	2733
<i>Yuxiang Huang (University of Waterloo, Canada), Yuhao Chen (University of Waterloo, Canada), and John Zelek (University of Waterloo, Canada)</i>	
Generalized Few-Shot Meets Remote Sensing: Discovering Novel Classes in Land Cover Mapping via Hybrid Semantic Segmentation Framework	2744
<i>Zhuohong Li (Wuhan University, China), Fangxiao Lu (Wuhan University, China), Jiaqi Zou (Wuhan University, China), Lei Hu (Wuhan University, China), and Hongyan Zhang (Wuhan University, China)</i>	
Learnable Prompt for Few-Shot Semantic Segmentation in Remote Sensing Domain	2755
<i>Steve Andreas Immanuel (TelePIX) and Hagai Raja Sinulingga (TelePIX)</i>	
Class Similarity Transition: Decoupling Class Similarities and Imbalance from Generalized Few-shot Segmentation	2762
<i>Shihong Wang (Xi'an Jiaotong University), Ruixun Liu (Xi'an Jiaotong University), Kaiyu Li (Xi'an Jiaotong University), Jiawei Jiang (Sun Yat-Sen University), and Xiangyong Cao (Xi'an Jiaotong University)</i>	
Enrich, Distill and Fuse: Generalized Few-Shot Semantic Segmentation in Remote Sensing Leveraging Foundation Model's Assistance	2771
<i>Tianyi Gao (Wuhan University, China), Wei Ao (Wuhan University, China), Xing-ao Wang (Wuhan University, China), Yuanhao Zhao (Robert Gordon University, UK), Ping Ma (Robert Gordon University, UK), Mengjie Xie (Wuhan University, China), Hang Fu (Robert Gordon University, UK), Jinchang Ren (Robert Gordon University, UK), and Zhi Gao (Wuhan University, China)</i>	
Dynamic Knowledge Adapter with Probabilistic Calibration for Generalized Few-Shot Semantic Segmentation	2781
<i>Jintao Tong (Huazhong University of Science and Technology), Haichen Zhou (Huazhong University of Science and Technology), Yicong Liu (Huazhong University of Science and Technology), Yiman Hu (Huazhong University of Science and Technology), and Yixiong Zou (Huazhong University of Science and Technology)</i>	

Neural Rendering Intelligence

Localised-NeRF: Specular Highlights and Colour Gradient Localising in NeRF	2791
<i>Dharmendra Selvaratnam (University of Plymouth, United Kingdom) and Dena Bazazian (University of Plymouth, United Kingdom)</i>	
Recon3D: High Quality 3D Reconstruction from a Single Image Using Generated Back-View Explicit Priors	2802
<i>Ruiyang Chen (Beijing University of Technology, China), Mohan Yin (Beijing University of Technology, China), Jiawei Shen (Beijing University of Technology, China), and Wei Ma (Beijing University of Technology, China)</i>	

GHNeRF: Learning Generalizable Human Features with Efficient Neural Radiance Fields	2812
<i>Arnab Dey (I3S-CNRS/Université Côte d’Azur, France), Di Yang (Inria Center at Université Côte d’Azur, France), Rohith Agaram (IIIT Hyderabad, India), Antitza Dantcheva (Inria Center at Université Côte d’Azur, France), Andrew I. Comport (I3S-CNRS/Université Côte d’Azur, France), Srinath Sridhar (Brown University, USA), and Jean Martinet (I3S-CNRS/Université Côte d’Azur, France)</i>	
Analyzing the Internals of Neural Radiance Fields	2822
<i>Lukas Radl (Graz University of Technology, Austria), Andreas Kurz (Graz University of Technology, Austria), Michael Steiner (Graz University of Technology, Austria), and Markus Steinberger (Graz University of Technology, Austria)</i>	
Unveiling the Ambiguity in Neural Inverse Rendering: A Parameter Compensation Analysis	2832
<i>Georgios Kourou (KU Leuven), Minye Wu (KU Leuven), Sushruth Nagesh (Ford Motor Company), Xianling Zhang (Ford Motor Company), and Tinne Tuytelaars (KU Leuven)</i>	
SAD-GS: Shape-aligned Depth-supervised Gaussian Splatting	2842
<i>Pou-Chun Kung (University of Michigan, USA), Seth Isaacson (University of Michigan, USA), Ram Vasudevan (University of Michigan, USA), and Katherine A. Skinner (University of Michigan, USA)</i>	
CoLa-SDF: Controllable Latent StyleSDF for Disentangled 3D Face Generation	2852
<i>Rahul Dey (Michigan State University, USA), Bernhard Egger (Friedrich-Alexander-University Erlangen, Germany), Vishnu Naresb Boddeti (Michigan State University, USA), Ye Wang (Mitsubishi Electric Research Laboratories, USA), and Tim K. Marks (Mitsubishi Electric Research Laboratories, USA)</i>	
SLAIM: Robust Dense Neural SLAM for Online Tracking and Mapping	2862
<i>Vincent Cartillier (GeorgiaTech), Grant Schindler (Google Research), and Irfan Essa (Google Research)</i>	
NeRF as Pretraining at Scale: Generalizable 3D-Aware Semantic Representation Learning from View Prediction	2872
<i>Wenyan Cong (University of Texas at Austin, USA), Hanxue Liang (University of Cambridge, UK), Zhiwen Fan (University of Texas at Austin, USA), Peihao Wang (University of Texas at Austin, USA), Yifan Jiang (University of Texas at Austin, USA), Dejie Xu (University of Texas at Austin, USA), A. Cengiz Oztireli (University of Cambridge, UK), and Zhangyang Wang (University of Texas at Austin, USA)</i>	
Neural Fields for Co-Reconstructing 3D Objects from Incidental 2D Data	2883
<i>Dylan Campbell (Australian National University), Eldar Insafutdinov (University of Oxford), Joao F. Henriques (University of Oxford), and Andrea Vedaldi (University of Oxford)</i>	

The 4th Workshop of Adversarial Machine Learning on Computer Vision: Robustness of Foundation Models

Large Language Models in Wargaming: Methodology, Application, and Robustness	2894
<i>Yuwei Chen (Aviation Industry Development Research Center of China, China) and Shiyong Chu (Aviation Industry Development Research Center of China, China)</i>	
Enhancing Targeted Attack Transferability via Diversified Weight Pruning	2904
<i>Hung-Jui Wang (National Taiwan University, Taiwan), Yu-Yu Wu (National Taiwan University, Taiwan), and Shang-Tse Chen (National Taiwan University, Taiwan)</i>	
Enhancing the Transferability of Adversarial Attacks with Stealth Preservation	2915
<i>Xinwei Zhang (Beihang University, China), Tianyuan Zhang (Beihang University, China), Yitong Zhang (Beihang University, China), and Shuangcheng Liu (Beihang University, China)</i>	
Benchmarking Robustness in Neural Radiance Fields	2926
<i>Chen Wang (University of Pennsylvania), Angtian Wang (Johns Hopkins University), Junbo Li (UC Santa Cruz), Alan Yuille (Johns Hopkins University), and Cihang Xie (UC Santa Cruz)</i>	
Sharpness-Aware Optimization for Real-World Adversarial Attacks for Diverse Compute Platforms with Enhanced Transferability	2937
<i>Muchao Ye (Pennsylvania State University), Xiang Xu (Amazon AI Labs), Qin Zhang (Amazon AI Labs), and Jonathan Wu (Amazon AI Labs)</i>	
Red-Teaming Segment Anything Model	2947
<i>Krzysztof Jankowski (University of Warsaw, Poland), Bartłomiej Sobieski (Warsaw University of Technology, Poland), Mateusz Kwiatkowski (University of Warsaw, Poland), Jakub Szulc (University of Warsaw, Poland), Michał Janik (University of Warsaw, Poland), Hubert Baniecki (University of Warsaw, Poland), and Przemysław Biecek (University of Warsaw; Warsaw University of Technology, Poland)</i>	
Learning to Schedule Resistant to Adversarial Attacks in Diffusion Probabilistic Models Under the Threat of Lipschitz Singularities	2957
<i>SangHwa Hong (Seoul National University of Science and Technology, South Korea)</i>	
Multimodal Attack Detection for Action Recognition Models	2967
<i>Furkan Mumcu (University of South Florida, USA) and Yasin Yilmaz (University of South Florida, USA)</i>	

20th Workshop on Perception Beyond the Visible Spectrum

Deep Learning-Based Identification of Arctic Ocean Boundaries and Near-Surface Phenomena in Underwater Echograms	2977
<i>Femina Senjaliya (University of Victoria, Canada), Melissa Cote (University of Victoria, Canada), Amanda Dash (ASL Environmental Sciences, Canada), Alexandra Branzan Albu (University of Victoria, Canada), Andrea Niemi (Fisheries and Oceans Canada, Canada), Stéphane Gauthier (Fisheries and Oceans Canada, Canada), Julek Chawarski (ASL Environmental Sciences, Canada), Steve Pearce (ASL Environmental Sciences, Canada), Kaan Ersahin (ASL Environmental Sciences, Canada), and Keath Borg (ASL Environmental Sciences, Canada)</i>	

BiMAE - A Bimodal Masked Autoencoder Architecture for Single-Label Hyperspectral Image Classification	2987
<i>Maksim Kukushkin (Leipzig University), Martin Bogdan (Leipzig University), and Thomas Schmid (Martin Luther University Halle-Wittenberg)</i>	
DaFF: Dual Attentive Feature Fusion for Multispectral Pedestrian Detection	2997
<i>Afnan Althoupety (Portland State University), Li-Yun Wang (Portland State University), Wu-chi Feng (Portland State University), and Banafsheh Rekabdar (Portland State University)</i>	
HNN: Hierarchical Noise-Deinterlace Net Towards Image Denoising	3007
<i>Amogh Joshi (KLE Technological University, India), Nikhil Akalwadi (KLE Technological University, India), Chinmayee Mandi (KLE Technological University, India), Chaitra Desai (KLE Technological University, India), Ramesh Ashok Tabib (KLE Technological University, India), Ujwala Patil (KLE Technological University, India), and Uma Mudenagudi (KLE Technological University, India)</i>	
Seeing the Vibration from Fiber-Optic Cables: Rain Intensity Monitoring using Deep Frequency Filtering	3017
<i>Zhuocheng Jiang (NEC Laboratories America, Inc), Yangmin Ding (NEC Laboratories America, Inc), Junhui Zhao (Eversource Energy), Yue Tian (NEC Laboratories America, Inc), Shaobo Han (NEC Laboratories America, Inc), Sarper Ozharar (NEC Laboratories America, Inc), Ting Wang (NEC Laboratories America, Inc), and James M. Moore (Verizon)</i>	
SwinFuSR: An Image Fusion-inspired Model for RGB-guided Thermal Image Super-resolution ...	3027
<i>Cyprien Arnold (Polytechnique Montréal, Canada), Philippe Jovet (CHU Sainte Justine, Canada), and Lama Seoud (Polytechnique Montréal, Canada)</i>	
CAFF-DINO: Multi-spectral Object Detection Transformers with Cross-attention Features Fusion	3037
<i>Kevin Helvig (Université Paris-Saclay, France), Baptiste Abeloos (Université Paris-Saclay, France), and Pauline Trouvé-Peloux (Université Paris-Saclay, France)</i>	
Learning Surface Terrain Classifications from Ground Penetrating Radar	3047
<i>Anja Sheppard (The University of Michigan, USA), Jason Brown (The University of Michigan, USA), Nilton Renno (The University of Michigan, USA), and Katherine A. Skinner (The University of Michigan, USA)</i>	
Scattering Prompt Tuning: A Fine-tuned Foundation Model for SAR Object Recognition	3056
<i>Weilong Guo (Technology and the Engineering Center for Space Utilization, Chinese Academy of Sciences, China), Shengyang Li (Technology and the Engineering Center for Space Utilization, Chinese Academy of Sciences, China), and Jian Yang (Technology and the Engineering Center for Space Utilization, Chinese Academy of Sciences, China)</i>	

MvAV-pix2pixHD: Multi-view Aerial View Image Translation	3066
<i>Jun Yu (University of Science and Technology of China), Keda Lu (University of Science and Technology of China), Shenshen Du (University of Science and Technology of China), Lin Xu (University of Science and Technology of China), Peng Chang (PAII Inc.), Houde Liu (Tsinghua Shenzhen International Graduate School), Bin Lan (Jianghuai Advance Technology Center), and Tianyu Liu (Jianghuai Advance Technology Center)</i>	
Flexible Window-based Self-attention Transformer in Thermal Image Super-Resolution	3076
<i>Hongcheng Jiang (University of Missouri Kansas City, USA) and ZhiQiang Chen (University of Missouri Kansas City, USA)</i>	
Multi-Scale Feature Fusion using Channel Transformers for Guided Thermal Image Super Resolution	3086
<i>Raghunath Sai Puttagunta (University of Missouri - Kansas City, USA), Birendra Kathariya (University of Missouri - Kansas City, USA), Zhu Li (University of Missouri - Kansas City, USA), and George York (US Air Force Academy)</i>	
Multi-modal Aerial View Image Challenge: Sensor Domain Translation	3096
<i>Spencer Low (Brigham Young University), Oliver Nina (Air Force Research Laboratory), Dylan Bowald (Air Force Research Laboratory), Angel D. Sappa (ESPOL Polytechnic University), Nathan Inkawich (Air Force Research Laboratory), and Peter Bruns (University of Utah)</i>	
Multi-modal Aerial View Image Challenge: SAR Classification	3105
<i>Spencer Low (Brigham Young University), Oliver Nina (Air Force Research Laboratory), Dylan Bowald (Air Force Research Laboratory), Angel D. Sappa (ESPOL Polytechnic University), Nathan Inkawich (Air Force Research Laboratory), and Peter Bruns (University of Utah)</i>	
Thermal Image Super-Resolution Challenge Results - PBVS 2024	3113
<i>Rafael E. Rivadeneira (Escuela Superior Politécnica del Litoral ESPOL, Ecuador), Angel D. Sappa (Computer Vision Center, Spain), Chenyang Wang (Harbin Institute of Technology, China), Junjun Jiang (Harbin Institute of Technology, China), Zhiwei Zhong (City University of Hong Kong, China), Peilin Chen (City University of Hong Kong, China), and Shiqi Wang (City University of Hong Kong, China)</i>	
Exploring the Usage of Diffusion Models for Thermal Image Super-resolution: A Generic, Uncertainty-aware Approach for Guided and Non-guided Schemes	3123
<i>Carlos Cortés-Mendez (CIMAT) and Jean-Bernard Hayet (CIMAT)</i>	
Narrowing the Synthetic-to-Real Gap for Thermal Infrared Semantic Image Segmentation Using Diffusion-based Conditional Image Synthesis	3131
<i>Christian Mayr (Hensoldt Optronics GmbH, Germany), Christian Kubler (Hensoldt Optronics GmbH, Germany), Norbert Haala (University of Stuttgart, Germany), and Michael Teutsch (Hensoldt Optronics GmbH, Germany)</i>	
Performance Evaluation of Segment Anything Model with Variational Prompting for Application to Non-Visible Spectrum Imagery	3142
<i>Yona Falinie A. Gaus (Durham University), Neelanjan Bhowmik (Durham University), Brian K. S. Isaac-Medina (Durham University), and Toby P. Breckon (Durham University)</i>	

Forward-Forward Algorithm for Hyperspectral Image Classification	3153
<i>Abel A. Reyes-Angulo (Michigan Technological University, USA) and Sidike Paheding (Fairfield University, USA)</i>	
Revisiting Pre-trained Remote Sensing Model Benchmarks: Resizing and Normalization Matters..	3162
<i>Isaac Corley (University of Texas at San Antonio, USA), Caleb Robinson (Microsoft AI for Good Research Lab, USA), Rahul Dodhia (Microsoft AI for Good Research Lab, USA), Juan M. Lavista Ferres (Microsoft AI for Good Research Lab, USA), and Peyman Najafirad (University of Texas at San Antonio, USA)</i>	

2nd Face Recognition Challenge in the Era of Synthetic Data (FRCSyn)

Second Edition FRCSyn Challenge at CVPR 2024: Face Recognition Challenge in the Era of Synthetic Data	3173
<i>Ivan DeAndres-Tame (Universidad Autonoma de Madrid, Spain), Ruben Tolosana (Universidad Autonoma de Madrid, Spain), Pietro Melzi (Universidad Autonoma de Madrid, Spain), Ruben Vera-Rodriguez (Universidad Autonoma de Madrid, Spain), Minchul Kim (Michigan State University, USA), Christian Rathgeb (Hochschule Darmstadt, Germany), Xiaoming Liu (Michigan State University, USA), Aythami Morales (Universidad Autonoma de Madrid, Spain), Julian Fierrez (Universidad Autonoma de Madrid, Spain), Javier Ortega-Garcia (Universidad Autonoma de Madrid, Spain), Zhizhou Zhong (Fudan University, China), Yuge Huang (Tencent Youtu Lab, China), Yuxi Mi (Fudan University, China), Shouhong Ding (Tencent Youtu Lab, China), Shuigeng Zhou (Fudan University, China), Shuai He (Interactive Entertainment Group of Netease Inc, China), Lingzhi Fu (Interactive Entertainment Group of Netease Inc, China), Heng Cong (Interactive Entertainment Group of Netease Inc, China), Rongyu Zhang (Interactive Entertainment Group of Netease Inc, China), Zhihong Xiao (Interactive Entertainment Group of Netease Inc, China), Evgeny Smirnov (ID R&D Inc., USA), Anton Pimenov (ID R&D Inc., USA), Aleksei Grigorev (ID R&D Inc., USA), Denis Timoshenko (ID R&D Inc., USA), Kaleb Mesfin Asfaw (Korea Advanced Institute of Science & Technology, Korea), Cheng Yaw Low (Institute for Basic Science, Korea), Hao Liu (China Telecom AI, China), Chuyi Wang (China Telecom AI, China), Qing Zuo (China Telecom AI, China), Zhixiang He (China Telecom AI, China), Hatef Otroschi Shahreza (Idiap Research Institute, Switzerland), Anjith George (Idiap Research Institute, Switzerland), Alexander Unnervik (Idiap Research Institute, Switzerland), Parsa Rahimi (Idiap Research Institute, Switzerland), Sébastien Marcel (Idiap Research Institute, Switzerland), Pedro C. Neto (INESC TEC, Portugal), Marco Huber (Fraunhofer IGD, Germany), Jan Niklas Kolf (Fraunhofer IGD, Germany), Naser Damer (Fraunhofer IGD, Germany), Fadi Boutros (Fraunhofer IGD, Germany), Jaime S. Cardoso (INESC TEC, Portugal), Ana F. Sequeira (INESC TEC, Portugal), Andrea Atzori (University of Cagliari, Italy), Gianni Fenu (University of Cagliari, Italy), Mirko Marras (University of Cagliari, Italy), Vitomir Struc (University of Ljubljana, Slovenia), Jiang Yu (Samsung Electronics (China) R&D Centre, China), Zhangjie Li (Samsung Electronics (China) R&D Centre, China), Jichun Li (Samsung Electronics (China) R&D Centre, China), Weisong Zhao (CAS, China), Zhen Lei (CASIA, China), Xiangyu Zhu (CASIA, China), Xiao-Yu Zhang (CAS, China), Bernardo Biesseck (Federal University of Paraná, Brazil), Pedro Vidal (Federal University of Paraná, Brazil), Luiz Coelho (unico - idTech, Brazil), Roger Granada (unico - idTech, Brazil), and David Menotti (Federal University of Paraná, Brazil)</i>	

10th IEEE International Workshop on Computer Vision in Sports (CVsports)

FineRehab: A Multi-modality and Multi-task Dataset for Rehabilitation Analysis	3184
<i>Jianwei Li (Beijing Sport University, China), Jun Xue (Beijing Sport University, China), Rui Cao (Beijing Sport University, China), Xiaoxia Du (Rehabilitation Research Center, China), Siyu Mo (Beijing Sport University, China), Kehao Ran (Beijing Sport University, China), and Zeyan Zhang (Rehabilitation Research Center, China)</i>	
Augmenting Pass Prediction via Imitation Learning in Soccer Simulations	3194
<i>Takeshi Kaneko (Tokyo Institute of Technology), Rei Kawakami (Tokyo Institute of Technology), Takeshi Naemura (The University of Tokyo), and Nakamasa Inoue (Tokyo Institute of Technology)</i>	
Hierarchical NeuroSymbolic Approach for Comprehensive and Explainable Action Quality Assessment	3204
<i>Lauren Okamoto (Princeton University) and Paritosh Parmar (IHPC, A*STAR, Singapore)</i>	
AutoSoccerPose: Automated 3D Posture Analysis of Soccer Shot Movements	3214
<i>Calvin Yeung (Nagoya University), Kenjiro Ide (Nagoya University), and Keisuke Fujii (Nagoya University)</i>	
Video Interaction Recognition using an Attention Augmented Relational Network and Skeleton Data	3225
<i>Farzaneh Askari (University of McGill, Canada), Cyril Yared (University of McGill, Canada), Rohit Ramaprasad (University of California San Diego, USA), Devin Garg (University of California San Diego, USA), Anjun Hu (University of Oxford, United Kingdom), and James J. Clark (University of McGill, Canada)</i>	
A General Framework for Jersey Number Recognition in Sports Video	3235
<i>Maria Koshkina (York University, Canada) and James H. Elder (York University, Canada)</i>	
MV-Soccer: Motion-Vector Augmented Instance Segmentation for Soccer Player Tracking	3245
<i>Fahad Majeed (Hamad Bin Khalifa University, Qatar), Nauman Ullah Gilal (Hamad Bin Khalifa University, Qatar), Khaled Al-Thelaya (Hamad Bin Khalifa University, Qatar), Yin Yang (Hamad Bin Khalifa University, Qatar), Marco Agus (Hamad Bin Khalifa University, Qatar), and Jens Schneider (Hamad Bin Khalifa University, Qatar)</i>	
Rugby Scene Classification Enhanced by Vision Language Model	3256
<i>Naoki Nonaka (Advanced Data Science Project, RIKEN Information R&D and Strategy Headquarters), Ryo Fujihira (Advanced Data Science Project, RIKEN Information R&D and Strategy Headquarters), Toshiki Koshiba (Advanced Data Science Project, RIKEN Information R&D and Strategy Headquarters), Akira Maeda (Hakata Knee & Sports Clinic), and Jun Seita (Advanced Data Science Project, RIKEN Information R&D and Strategy Headquarters)</i>	

X-VARS: Introducing Explainability in Football Refereeing with Multi-Modal Large Language Models	3267
<i>Jan Held (University of Liege, Belgium), Hani Itani (King Abdullah University of Science and Technology, Saudi Arabia), Anthony Cioppa (University of Liege, Belgium), Silvio Giancola (King Abdullah University of Science and Technology, Saudi Arabia), Bernard Ghanem (King Abdullah University of Science and Technology, Saudi Arabia), and Marc Van Droogenbroeck (University of Liege, Belgium)</i>	
SoccerNet-Depth: a Scalable Dataset for Monocular Depth Estimation in Sports Videos	3280
<i>Arnaud Leduc (University of Liège), Anthony Cioppa (University of Liège), Silvio Giancola (KAUST), Bernard Ghanem (KAUST), and Marc Van Droogenbroeck (University of Liège)</i>	
SoccerNet Game State Reconstruction: End-to-End Athlete Tracking and Identification on a Minimap	3293
<i>Vladimir Somers (UCLouvain, Belgium), Victor Joos (UCLouvain, Belgium), Anthony Cioppa (University of Liège, Belgium), Silvio Giancola (KAUST, Saudi Arabia), Seyed Abolfazl Ghasemzadeh (UCLouvain, Belgium), Floriane Magera (University of Liège, Belgium), Baptiste Standaert (UCLouvain, Belgium), Amir M. Mansourian (Sharif University of Technology, Iran), Xin Zhou (Baidu, USA), Shohreh Kasaei (Sharif University of Technology, Iran), Bernard Ghanem (KAUST, Saudi Arabia), Alexandre Alahi (EPFL, Switzerland), Marc Van Droogenbroeck (University of Liège, Belgium), and Christophe De Vleeschouwer (UCLouvain, Belgium)</i>	
Multi-Modal Hit Detection and Positional Analysis in Padel Competitions	3306
<i>Robbe Decorte (Ghent University - imec, Belgium), Martin Paré (Ghent University - imec, Belgium), Jelle Vanhaeverbeke (Ghent University - imec, Belgium), Joachim Taelman (Ghent University - imec, Belgium), Maarten Slembrouck (Ghent University - imec, Belgium), and Steven Verstockt (Ghent University - imec, Belgium)</i>	
Pseudo-label Based Unsupervised Fine-tuning of a Monocular 3D Pose Estimation Model for Sports Motions	3315
<i>Tomohiro Suzuki (Nagoya University, Japan), Ryota Tanaka (Nagoya University, Japan), Kazuya Takeda (Nagoya University, Japan), and Keisuke Fujii (Nagoya University, Japan)</i>	
No Bells, Just Whistles: Sports Field Registration by Leveraging Geometric Properties	3325
<i>Marc Gutiérrez-Pérez (Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Spain) and Antonio Agudo (Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Spain)</i>	
A Universal Protocol to Benchmark Camera Calibration for Sports	3335
<i>Floriane Magera (EVS Broadcast Equipment, Belgium; University of Liege, Belgium), Thomas Hoyoux (EVS Broadcast Equipment, Belgium), Olivier Barnich (EVS Broadcast Equipment, Belgium), and Marc Van Droogenbroeck (University of Liege, Belgium)</i>	
Table Tennis Ball Spin Estimation with an Event Camera	3347
<i>Thomas Gossard (University of Tübingen), Julian Krismer (University of Tübingen), Andreas Ziegler (University of Tübingen), Jonas Tebbe (University of Tübingen), and Andreas Zell (University of Tübingen)</i>	

TeamTrack: A Dataset for Multi-Sport Multi-Object Tracking in Full-pitch Videos	3357
<i>Atom Scott (Nagoya University, Japan), Ikuma Uchida (University of Tsukuba, Japan), Ning Ding (Nagoya University, Japan), Rikuhei Umemoto (Nagoya University, Japan), Rory Bunker (Nagoya University, Japan), Ren Kobayashi (Nagoya University, Japan), Takeshi Koyama (Tokai University, Japan), Masaki Onishi (National Institute of Advanced Industrial Science and Technology, Japan), Yoshinari Kameda (University of Tsukuba, Japan), and Keisuke Fujii (Nagoya University, Japan)</i>	
Event-based Ball Spin Estimation in Sports	3367
<i>Takuya Nakabayashi (Keio University, Japan), Kyota Higa (NEC Corporation, Japan), Masahiro Yamaguchi (NEC Corporation, Japan), Ryo Fujiwara (NEC Corporation, Japan), and Hideo Saito (Keio University, Japan)</i>	
A Stroke of Genius: Predicting the Next Move in Badminton	3376
<i>Magnus Ibh (IT University of Copenhagen, Denmark), Stella Graßhof (IT University of Copenhagen, Denmark), and Dan Witzner Hansen (IT University of Copenhagen, Denmark)</i>	
Beyond the Premier: Assessing Action Spotting Transfer Capability Across Diverse Domains	3386
<i>Bruno Cabado (Universidade da Coruña), Anthony Cioppa (University of Liège), Silvio Giancola (KAUST), Andrés Villa (KAUST), Bertha Guijarro-Berdiñas (Universidade da Coruña), Emilio J. Padrón (Universidade da Coruña), Bernard Ghanem (KAUST), and Marc Van Droogenbroeck (University of Liège)</i>	
Medium Scale Benchmark for Cricket Excited Actions Understanding	3399
<i>Altaf Hussain (Sejong University, Republic of Korea), Noman Khan (Sejong University, Republic of Korea), Muhammad Munsif (Sejong University, Republic of Korea), Min Je Kim (Sejong University, Republic of Korea), and Sung Wook Baik (Sejong University, Republic of Korea)</i>	
T-DEED: Temporal-Discriminability Enhancer Encoder-Decoder for Precise Event Spotting in Sports Videos	3410
<i>Artur Xarles (Universitat de Barcelona, Spain; Computer Vision Center, Spain), Sergio Escalera (Universitat de Barcelona, Spain; Computer Vision Center, Spain; Aalborg University, Denmark), Thomas B. Moeslund (Aalborg University, Denmark), and Albert Clapés (Universitat de Barcelona, Spain; Computer Vision Center, Spain)</i>	
PitcherNet: Powering the Moneyball Evolution in Baseball Video Analytics	3420
<i>Jerrin Bright (University of Waterloo, Canada), Bavesh Balaji (University of Waterloo, Canada), Yuhao Chen (University of Waterloo, Canada), David A Clausi (University of Waterloo, Canada), and John S Zelek (University of Waterloo, Canada)</i>	
ExerAIde: AI-assisted Multimodal Diagnosis for Enhanced Sports Performance and Personalised Rehabilitation	3430
<i>Ahmed Qazi (Tibbling Technologies, USA) and Asim Iqbal (Tibbling Technologies, USA)</i>	

Safe Artificial Intelligence for All Domains (SAIAD)

Look, Listen, and Attack: Backdoor Attacks Against Video Action Recognition	3439
<i>Hasan Abed Al Kader Hammoud (KAUST), Shuming Liu (KAUST), Mohammed Alkhrashi (SDAIA), Fahad AlBalawi (Taif University), and Bernard Ghanem (KAUST)</i>	
Understanding ReLU Network Robustness Through Test Set Certification Performance	3451
<i>Nicola Franco (Fraunhofer Institute for Cognitive Systems IKS, Germany), Jeanette Miriam Lorenz (Fraunhofer Institute for Cognitive Systems IKS, Germany), Karsten Roscher (Fraunhofer Institute for Cognitive Systems IKS, Germany), and Stephan Günnemann (Technical University of Munich, Germany)</i>	
Reliable Trajectory Prediction and Uncertainty Quantification with Conditioned Diffusion Models	3461
<i>Marion Neumeier (Technische Hochschule Ingolstadt, Germany), Sebastian Dorn (Technische Hochschule Augsburg, Germany), Michael Botsch (Technische Hochschule Ingolstadt, Germany), and Wolfgang Utschick (Technische Universität München, Germany)</i>	
Hinge-Wasserstein: Estimating Multimodal Aleatoric Uncertainty in Regression Tasks	3471
<i>Ziliang Xiong (Linköping University, Sweden), Arvi Jonnarth (Linköping University, Husqvarna Group, Sweden), Abdelrahman Eldesokey (King Abdullah University of Science and Technology, Saudi Arabia), Joakim Johnander (Linköping University, Zenseact, Sweden), Bastian Wandt (Linköping University, Sweden), and Per-Erik Forssén (Linköping University, Sweden)</i>	
AdvDenoise: Fast Generation Framework of Universal and Robust Adversarial Patches Using Denoise	3481
<i>Jing Li (Tsinghua University, China), Zigan Wang (Tsinghua University, China), and Jinliang Li (Tsinghua University, China)</i>	
Understanding the (Extra-)Ordinary: Validating Deep Model Decisions with Prototypical Concept-based Explanations	3491
<i>Maximilian Dreyer (Fraunhofer Heinrich Hertz Institute, Germany), Reduan Achtibat (Fraunhofer Heinrich Hertz Institute, Germany), Wojciech Samek (Fraunhofer Heinrich Hertz Institute, Germany), and Sebastian Lapuschkin (Fraunhofer Heinrich Hertz Institute, Germany)</i>	
Situation Monitor: Diversity-Driven Zero-Shot Out-of-Distribution Detection using Budding Ensemble Architecture for Object Detection	3502
<i>Syed Sha Qutub (Technical University of Munich, Germany; Intel Labs, Germany), Michael Paulitsch (Intel Labs, Germany), Kay-Ulrich Scholl (Intel Labs, Germany), Neslihan Kose Cihangir (Intel Labs, Germany), Korbinian Hagn (Intel Labs, Germany), Fabian Oboril (Intel Labs, Germany), Gereon Hinz (Technical University of Munich, Germany), and Alois Knoll (Technical University of Munich, Germany)</i>	
The Penalized Inverse Probability Measure for Conformal Classification	3512
<i>Paul Melki (IMS Bordeaux, France; EXXACT Robotics, France), Lionel Bombrun (IMS Bordeaux, France; Bordeaux Sciences Agro, France), Boubacar Diallo (EXXACT Robotics, France), Jérôme Dias (EXXACT Robotics, France), and Jean-Pierre Da Costa (IMS Bordeaux, France; Bordeaux Sciences Agro, France)</i>	

Run-time Monitoring of 3D Object Detection in Automated Driving Systems Using Early Layer Neural Activation Patterns	3522
<i>Hakan Yekta Yatbaz (University of Warwick, UK), Mehrdad Dianati (University of Warwick, UK; Queen's University of Belfast, UK), Konstantinos Koufos (University of Warwick, UK), and Roger Woodman (University of Warwick, UK)</i>	
Reactive Model Correction: Mitigating Harm to Task-Relevant Features via Conditional Bias Suppression	3532
<i>Dilyara Bareeva (Fraunhofer Heinrich Hertz Institute, Germany), Maximilian Dreyer (Fraunhofer Heinrich Hertz Institute, Germany), Frederik Pahde (Fraunhofer Heinrich Hertz Institute, Germany), Wojciech Samek (Fraunhofer Heinrich Hertz Institute, Germany), and Sebastian Lapuschkin (Fraunhofer Heinrich Hertz Institute, Germany)</i>	
Investigating Calibration and Corruption Robustness of Post-hoc Pruned Perception CNNs: An Image Classification Benchmark Study	3542
<i>Pallavi Mitra (Continental AG, Germany), Gesina Schwalbe (University of Lübeck, Germany), and Nadja Klein (TU Dortmund, Germany)</i>	
Towards Weakly-Supervised Domain Adaptation for Lane Detection	3553
<i>Jingxing Zhou (Porsche Engineering Group GmbH, Germany), Chongzhe Zhang (University of Stuttgart, Germany), and Jürgen Beyerer (Fraunhofer IOSB, Germany)</i>	
Towards Engineered Safe AI with Modular Concept Models	3564
<i>Lena Heidemann (Fraunhofer IKS, Germany), Iwo Kurzidem (Fraunhofer IKS, Germany), Maureen Monnet (Fraunhofer IKS, Germany), Karsten Roscher (Fraunhofer IKS, Germany), and Stephan Günнемann (Technical University of Munich, Germany)</i>	
Conformal Semantic Image Segmentation: Post-hoc Quantification of Predictive Uncertainty	3574
<i>Luca Mossina (IRT Saint Exupéry, France), Joseba Dalmau (IRT Saint Exupéry, France), and Léo Andéol (Institut de Mathématiques de Toulouse, France and SNCF, France)</i>	
A Comprehensive Analysis of Factors Impacting Membership Inference	3585
<i>Daniel DeAlcala (Universidad Autonoma de Madrid), Gonzalo Mancera (Universidad Autonoma de Madrid), Aythami Morales (Universidad Autonoma de Madrid), Julian Fierrez (Universidad Autonoma de Madrid), Ruben Tolosana (Universidad Autonoma de Madrid), and Javier Ortega-Garcia (Universidad Autonoma de Madrid)</i>	
Exploiting CLIP Self-Consistency to Automate Image Augmentation for Safety Critical Scenarios	3594
<i>Sujan Sai Gannamaneni (Fraunhofer IAIS, Germany; Lamarr Institute, Germany), Frederic Klein (University of Bonn, Germany), Michael Mock (Fraunhofer IAIS, Germany), and Maram Akila (Fraunhofer IAIS, Germany; Lamarr Institute, Germany)</i>	

Efficient Large Vision Models

Adaptive Memory Replay for Continual Learning	3605
<i>James Seale Smith (Georgia Institute of Technology, USA), Lazar Valkov (MIT-IBM Watson AI Lab, USA), Shaunak Halbe (Georgia Institute of Technology, USA), Vyshnavi Gutta (Georgia Institute of Technology, USA), Rogerio Feris (MIT-IBM Watson AI Lab, USA), Zsolt Kira (Georgia Institute of Technology, USA), and Leonid Karlinsky (MIT-IBM Watson AI Lab, USA)</i>	
Adapting the Segment Anything Model During Usage in Novel Situations	3616
<i>Robin Schön (Universität Augsburg), Julian Lorenz (Universität Augsburg), Katja Ludwig (Universität Augsburg), and Rainer Lienhart (Universität Augsburg)</i>	
PMAFusion: Projection-Based Multi-Modal Alignment for 3D Semantic Occupancy Prediction	3627
<i>Shiyao Li (Tsinghua Shenzhen International Graduate School, China), Wenming Yang (Tsinghua Shenzhen International Graduate School, China), and Qingmin Liao (Tsinghua University, China)</i>	
SAM-CLIP: Merging Vision Foundation Models Towards Semantic and Spatial Understanding ...	3635
<i>Haoxiang Wang (University of Illinois Urbana-Champaign), Pavan Kumar Anasosalu Vasu (Apple), Fartash Faghri (Apple), Raviteja Vemulapalli (Apple), Mehrdad Farajtabar (Apple), Sachin Mehta (Apple), Mohammad Rastegari (Apple), Oncel Tuzel (Apple), and Hadi Pouransari (Apple)</i>	
QAttn: Efficient GPU Kernels for Mixed-precision Vision Transformers	3648
<i>Piotr Kluska (IBM Research/Universitat Politècnica de València), Adrián Castelló (Universitat Politècnica de València), Florian Scheidegger (IBM Research), A. Cristiano I. Malossi (IBM Research), and Enrique S. Quintana-Ortí (Universitat Politècnica de València)</i>	
Efficient Transformer Adaptation with Soft Token Merging	3658
<i>Xin Yuan (University of Chicago), Hongliang Fei (Google), and Jinoo Baek (Google)</i>	
HaLViT: Half of the Weights are Enough	3669
<i>Onur Can Koyun (Istanbul Technical University, Turkey) and Behçet Uğur Töreyn (Istanbul Technical University, Turkey)</i>	
Parameter Efficient Fine-tuning of Self-supervised ViTs without Catastrophic Forgetting	3679
<i>Reza Akbarian Bafghi (University of Colorado, Boulder, USA), Nidhin Harilal (University of Colorado, Boulder, USA), Claire Monteleoni (University of Colorado, Boulder, USA), and Maziar Raissi (University of California, Riverside, USA)</i>	

MetaFood Workshop (MTF)

Automatic Recognition of Food Ingestion Environment from the AIM-2 Wearable Sensor	3685
<i>Yuning Huang (Purdue University, USA), M A Hassan (University of California, Davis, USA), Jiangpeng He (Purdue University, USA), J. Higgins (University of Colorado, USA), Megan McCrory (Boston University, USA), Heather Eicher-Miller (Purdue University, USA), J. Graham Thomas (Brown University, USA), Edward Sazonov (University of Alabama, USA), and Fengqing Zhu (Purdue University, USA)</i>	

Learning to Classify New Foods Incrementally Via Compressed Exemplars	3695
<i>Justin Yang (Purdue University, West Lafayette, Indiana, USA), Zhihao Duan (Purdue University, West Lafayette, Indiana, USA), Jiangpeng He (Purdue University, West Lafayette, Indiana, USA), and Fengqing Zhu (Purdue University, West Lafayette, Indiana, USA)</i>	
MP-PolarMask: A Faster and Finer Instance Segmentation for Concave Images	3705
<i>Ke-Lei Wang (National Yang Ming Chiao Tung University, Taiwan), Pin-Hsuan Chou (National Yang Ming Chiao Tung University, Taiwan), Young-Ching Chou (National Yang Ming Chiao Tung University, Taiwan), Chia-Jen Liu (National Yang Ming Chiao Tung University, Taiwan), Cheng-Kuan Lin (National Yang Ming Chiao Tung University, Taiwan), and Yu-Chee Tseng (National Yang Ming Chiao Tung University, Taiwan)</i>	
Segment Anything in Food Images	3715
<i>Saeed S. Alahmari (Najran University, Saudi Arabia), Michael Gardner (King Faisal University, Saudi Arabia), and Tawfiq Salem (Purdue University, USA)</i>	
Shape-Preserving Generation of Food Images for Automatic Dietary Assessment	3721
<i>Guangzong Chen (University of Pittsburgh), Zhi-Hong Mao (University of Pittsburgh), Mingui Sun (University of Pittsburgh), Kangni Liu (University of Pittsburgh), and Wenyan Jia (University of Pittsburgh)</i>	
A Generative Exploration of Cuisine Transfer	3732
<i>Philip Wootack Shin (The Pennsylvania State University, USA), Ajay Narayanan Sridhar (The Pennsylvania State University, USA), Jack Sampson (The Pennsylvania State University, USA), and Vijaykrishnan Narayanan (The Pennsylvania State University, USA)</i>	
Food Portion Estimation via 3D Object Scaling	3741
<i>Gautham Vinod (Purdue University), Jiangpeng He (Purdue University), Zeman Shao (Purdue University), and Fengqing Zhu (Purdue University)</i>	
LOFI: LOng-tailed FIne-Grained Network for Food Recognition	3750
<i>Jesús M. Rodríguez-de-Vera (Universitat de Barcelona, Spain), Imanol G. Estepa (Universitat de Barcelona, Spain), Marc Bolaños (AIGecko Technologies SL, Spain), Bhalaji Nagarajan (Universitat de Barcelona, Spain), and Petia Radeva (Universitat de Barcelona, Spain)</i>	
How Much You Ate? Food Portion Estimation on Spoons	3761
<i>Aaryam Sharma (University of Waterloo, Vision and Image Processing Lab), Chris Czarnecki (University of Waterloo, Vision and Image Processing Lab), Yuhao Chen (University of Waterloo, Vision and Image Processing Lab), Pengcheng Xi (National Research Council Canada), Linlin Xu (University of Waterloo, Vision and Image Processing Lab), and Alexander Wong (University of Waterloo, Vision and Image Processing Lab)</i>	

2nd Workshop and Challenge on DeepFake Analysis and Detection

Faster Than Lies: Real-time Deepfake Detection using Binary Neural Networks	3771
<i>Romeo Lanzino (Sapienza University of Rome), Federico Fontana (Sapienza University of Rome), Anxhelo Diko (Sapienza University of Rome), Marco Raoul Marini (Sapienza University of Rome), and Luigi Cinque (Sapienza University of Rome)</i>	

Latent Flow Diffusion for Deepfake Video Generation	3781
<i>Aashish Chandra K (BITS PILANI), Aashutosh A V (BITS PILANI), Srijan Das (UNCC), and Abhijit Das (BITS PILANI)</i>	
Deepfake Catcher: Can a Simple Fusion be Effective and Outperform Complex DNNs?	3791
<i>Akshay Agarwal (IISER Bhopal) and Nalini Ratha (University at Buffalo)</i>	
DiffSeg: Towards Detecting Diffusion-Based Inpainting Attacks Using Multi-Feature Segmentation	3802
<i>Raphael Antonius Frick (Fraunhofer SIT ATHENE Center) and Martin Steinebach (Fraunhofer SIT ATHENE Center)</i>	
PUDD: Towards Robust Multi-modal Prototype-based Deepfake Detection	3809
<i>Alvaro Lopez Pellicer (Lancaster University), Yi Li (Lancaster University), and Plamen Angelov (Lancaster University)</i>	
Demographic Bias Effects on Face Image Synthesis	3818
<i>Roberto Leyva (University of Warwick, UK), Victor Sanchez (University of Warwick, UK), Gregory Epiphaniou (University of Warwick, UK), and Carsten Maple (University of Warwick, UK)</i>	
Evaluating the Integration of Morph Attack Detection in Automated Face Recognition Systems....	3827
<i>Andrea Panzino (University of Cagliari, Italy), Simone Maurizio la Cava (University of Cagliari, Italy), Giulia Orrù (University of Cagliari, Italy), and Gian Luca Marcialis (University of Cagliari, Italy)</i>	
Temporal Surface Frame Anomalies for Deepfake Video Detection	3837
<i>Andrea Ciamarra (University of Florence; Universitas Mercatorum, Italy), Roberto Caldelli (CNIT; Universitas Mercatorum, Italy), and Alberto Del Bimbo (University of Florence, Italy)</i>	
Quality-based Artifact Modeling for Facial Deepfake Detection in Videos	3845
<i>Sara Concas (University of Cagliari), Simone Maurizio la Cava (University of Cagliari), Roberto Casula (University of Cagliari), Giulia Orrù (University of Cagliari), Giovanni Puglisi (University of Cagliari), and Gian Luca Marcialis (University of Cagliari)</i>	
MaskSim: Detection of Synthetic Images by Masked Spectrum Similarity Analysis	3855
<i>Yanhao Li (Ecole Normale Supérieure Paris-Saclay, France), Quentin Bammey (Ecole Normale Supérieure Paris-Saclay, France), Marina Gardella (Instituto de Matemática Pura e Aplicada, Brazil), Tina Nikoukhah (Ecole Normale Supérieure Paris-Saclay, France), Jean-Michel Morel (City University of Hong Kong, Hong Kong SAR, China), Miguel Colom (Ecole Normale Supérieure Paris-Saclay, France), and Rafael Grompone von Gioi (Ecole Normale Supérieure Paris-Saclay, France)</i>	

VAND 2.0: Visual Anomaly and Novelty Detection

Divide and Conquer: High-Resolution Industrial Anomaly Detection via Memory Efficient Tiled Ensemble	3866
<i>Blaž Rolih (University of Ljubljana, Slovenia), Dick Ameln (Intel, Netherlands), Ashwin Vaidya (Intel, Netherlands), and Samet Akcay (Intel, United Kingdom)</i>	

OmniCrack30k: A Benchmark for Crack Segmentation and the Reasonable Effectiveness of Transfer Learning	3876
<i>Christian Benz (Bauhaus-Universität Weimar) and Volker Rodehorst (Bauhaus-Universität Weimar)</i>	
Video Anomaly Detection via Spatio-Temporal Pseudo-Anomaly Generation : A Unified Approach....	3887
<i>Ayush K. Rai (Dublin City University), Tarun Krishna (Dublin City University), Feiyan Hu (Dublin City University), Alexandru Drimborean (Tobii Corporation), Kevin McGuinness (Dublin City University), Alan F. Smeaton (Dublin City University), and Noel E. O'Connor (Dublin City University)</i>	
Blind Localization and Clustering of Anomalies in Textures	3900
<i>Andrei-Timotei Ardelean (Friedrich-Alexander-Universität Erlangen-Nürnberg) and Tim Weyrich (Friedrich-Alexander-Universität Erlangen-Nürnberg)</i>	
Test Time Training for Industrial Anomaly Segmentation	3910
<i>Alex Costanzino (University of Bologna, Italy), Pierluigi Zama Ramirez (University of Bologna, Italy), Mirko Del Moro (University of Bologna, Italy), Agostino Aiezzo (University of Bologna, Italy), Giuseppe Lisanti (University of Bologna, Italy), Samuele Salti (University of Bologna, Italy), and Luigi Di Stefano (University of Bologna, Italy)</i>	
TAB: Text-Align Anomaly Backbone Model for Industrial Inspection Tasks	3921
<i>Ho-Weng Lee (National Tsing Hua University, Taiwan) and Shang-Hong Lai (National Tsing Hua University, Taiwan)</i>	
Tri-VAE: Triplet Variational Autoencoder for Unsupervised Anomaly Detection in Brain Tumor MRI	3930
<i>Hansen Wijanarko (National Yang Ming Chiao Tung University, Taiwan), Evelynne Calista (National Yang Ming Chiao Tung University, Taiwan), Li-Fen Chen (National Yang Ming Chiao Tung University, Taiwan), and Yong-Sheng Chen (National Yang Ming Chiao Tung University, Taiwan)</i>	
Dynamic Addition of Noise in a Diffusion Model for Anomaly Detection	3940
<i>Justin Tebbe (Otto von Guericke University Magdeburg, Germany) and Jawad Tayyub (Endress + Hauser, Germany)</i>	
SplatPose & Detect: Pose-Agnostic 3D Anomaly Detection	3950
<i>Mathis Kruse (Leibniz University Hannover, Germany), Marco Rudolph (Leibniz University Hannover, Germany), Dominik Woiwode (Leibniz University Hannover, Germany), and Bodo Rosenhahn (Leibniz University Hannover, Germany)</i>	
Dynamic Distinction Learning: Adaptive Pseudo Anomalies for Video Anomaly Detection	3961
<i>Demetris Lappas (Kingston University, UK), Vasileios Argyriou (Kingston University, UK), and Dimitrios Makris (Kingston University, UK)</i>	
COOD: Combined Out-of-distribution Detection Using Multiple Measures for Anomaly & Novel Class Detection in Large-scale Hierarchical Classification	3971
<i>Laurens E. Hogeweg (Intel Benelux BV), Rajesh Gangireddy (Intel Benelux BV), Django Brunink (Naturalis Biodiversity Center), Vincent J. Kalkman (Naturalis Biodiversity Center), Ludo Cornelissen (Intel Benelux BV), and Jacob W. Kamminga (University of Twente)</i>	

Model-guided Contrastive Fine-tuning for Industrial Anomaly Detection	3981
<i>Aitor Artola (Ens Paris-Saclay, France), Yannis Kolodziej (Visionairy, France), Jean-Michel Morel (University of Hong Kong, Hong Kong), and Thibaud Ehret (ENS Paris-Saclay, France)</i>	
Tracklet-based Explainable Video Anomaly Localization	3992
<i>Ashish Singh (University of Massachusetts Amherst, USA), Michael J. Jones (Mitsubishi Electric Research Laboratories, USA), and Erik G. Learned-Miller (University of Massachusetts Amherst, USA)</i>	
Context-aware Video Anomaly Detection in Long-Term Datasets	4002
<i>Zhengye Yang (Rensselaer Polytechnic Institute, USA) and Richard J. Radke (Rensselaer Polytechnic Institute, USA)</i>	
Manifold DivideMix: A Semi-Supervised Contrastive Learning Framework for Severe Label Noise	4012
<i>Fahimeh Fooladgar (University of British Columbia, Canada), Minh Nguyen Nhat To (University of British Columbia, Canada), Parvin Mousavi (Queen's University, Canada), and Purang Abolmaesumi (University of British Columbia, Canada)</i>	
LogicAL: Towards Logical Anomaly Synthesis for Unsupervised Anomaly Localization	4022
<i>Ying Zhao (Ricoh Software Research Center (Beijing) Co., Ltd.)</i>	
DMR: Disentangling Marginal Representations for Out-of-Distribution Detection	4032
<i>Dasol Choi (Yonsei University) and Dongbin Na (Pohang University of Science and Technology)</i>	
BMAD: Benchmarks for Medical Anomaly Detection	4042
<i>Jinan Bao (University of Alberta, Canada), Hanshi Sun (Carnegie Mellon University, USA), Hanqiu Deng (University of Alberta, Canada), Yinsheng He (University of Alberta, Canada), Zhaoxiang Zhang (University of Alberta, Canada), and Xingyu Li (University of Alberta, Canada)</i>	

5th Workshop on Continual Learning in Computer Vision (CLVISION)

DELTA: Decoupling Long-Tailed Online Continual Learning	4054
<i>Siddeshwar Raghavan (Purdue University), Jiangpeng He (Purdue University), and Fengqing Zhu (Purdue University)</i>	
Unveiling the Anomalies in an Ever-Changing World: A Benchmark for Pixel-Level Anomaly Detection in Continual Learning	4065
<i>Nikola Bugarin (University of Padova), Jovana Bugaric (University of Padova), Manuel Barusco (University of Padova), Davide Dalle Pezze (University of Padova), and Gian Antonio Susto (University of Padova)</i>	
Calibrating Higher-Order Statistics for Few-Shot Class-Incremental Learning with Pre-trained Vision Transformers	4075
<i>Dipam Goswami (Computer Vision Center, Spain; Autonomous University of Barcelona, Spain), Bartłomiej Twardowski (Computer Vision Center, Spain; Autonomous University of Barcelona, Spain; IDEAS NCBR, Poland), and Joost van de Weijer (Computer Vision Center, Spain; Autonomous University of Barcelona, Spain)</i>	

Active Data Collection and Management for Real-World Continual Learning via Pretrained Oracle	4085
<i>Vivek Chavan (Fraunhofer IPK), Paul Koch (Fraunhofer IPK), Marian Schlüter (Fraunhofer IPK), Clemens Briebe (Fraunhofer IPK), and Jörg Krüger (TU Berlin)</i>	
Class-Incremental Mixture of Gaussians for Deep Continual Learning	4097
<i>Lukasz Korycki (Virginia Commonwealth University, USA) and Bartosz Krawczyk (Rochester Institute of Technology, USA)</i>	
MultIOD: Rehearsal-free Multihead Incremental Object Detector	4107
<i>Eden Belouadah (Datakalab, France), Arnaud Dapogny (Datakalab, France), and Kevin Bailly (Datakalab, France)</i>	
Wake-Sleep Energy Based Models for Continual Learning	4118
<i>Vaibhav Singh (Concordia University(Mila), Canada), Anna Choromanska (New York University, USA), Shuang Li (University of Toronto, Canada), and Yilun Du (Massachusetts Institute of Technology, USA)</i>	
Continual-Zoo: Leveraging Zoo Models for Continual Classification of Medical Images	4128
<i>Nourhan Bayasi (University of British Columbia), Ghassan Hamarneh (Simon Fraser University), and Rafeef Garbi (University of British Columbia)</i>	
TAME: Task Agnostic Continual Learning using Multiple Experts	4139
<i>Haoran Zhu (New York University, USA), Maryam Majzoubi (Google, USA), Arihant Jain (New York University, USA), and Anna Choromanska (New York University, USA)</i>	
Tackling Domain Shifts in Person Re-Identification: A Survey and Analysis	4149
<i>Vuong D. Nguyen (University of Houston, USA), Samiha Mirza (University of Houston, USA), Abdollah Zakeri (University of Houston, USA), Ayush Gupta (Johns Hopkins University, USA), Khadija Khaldi (University of Houston, USA), Rahma Aloui (University of Houston, USA), Pranav Mantini (University of Houston, USA), Shishir K. Shah (University of Houston, USA), and Fatima Merchant (University of Houston, USA)</i>	
Calibration of Continual Learning Models	4160
<i>Lanpei Li (CNR, Italy), Elia Piccoli (University of Pisa, Italy), Andrea Cossu (University of Pisa, Italy), Davide Bacciu (University of Pisa, Italy), and Vincenzo Lomonaco (University of Pisa, Italy)</i>	
VLM-PL: Advanced Pseudo Labeling Approach for Class Incremental Object Detection via Vision-Language Model	4170
<i>Junsu Kim (UNIST; MODULABS), Yunhoe Ku (UNIST; MODULABS), Jihyeon Kim (UNIST), Junuk Cha (UNIST), and Seungryul Baek (UNIST)</i>	
The Expanding Scope of the Stability Gap: Unveiling its Presence in Joint Incremental Learning of Homogeneous Tasks	4182
<i>Sandesh Kamath (Computer Vision Center, Spain), Albin Soutif-Cormerais (Computer Vision Center, Spain), Joost van de Weijer (Computer Vision Center, Spain), and Bogdan Raducanu (Computer Vision Center, Spain)</i>	
Continual Learning with Weight Interpolation	4187
<i>Jędrzej Kozal (Wrocław University of Science and Technology, Poland), Jan Wasilewski (Rochester Institute of Technology, USA), Bartosz Krawczyk (Rochester Institute of Technology, USA), and Michał Woźniak (Wrocław University of Science and Technology, Poland)</i>	

An Analysis of Best-practice Strategies for Replay and Rehearsal in Continual Learning	4196
<i>Alexander Krawczyk (Fulda University of Applied Sciences) and Alexander Gepperth (Fulda University of Applied Sciences)</i>	

The 3rd International Workshop on Federated Learning for Computer Vision (FedVision-2024)

FedProK: Trustworthy Federated Class-Incremental Learning via Prototypical Feature Knowledge Transfer	4205
<i>Xin Gao (Southwestern University of Finance and Economics), Xin Yang (Southwestern University of Finance and Economics), Hao Yu (Southwest University of Finance and Economics), Yan Kang (Webank), and Tianrui Li (Southwest Jiaotong University)</i>	
Collaborative Visual Place Recognition through Federated Learning	4215
<i>Mattia Dutto (Politecnico di Torino), Gabriele Berton (Politecnico di Torino), Debora Caldarola (Politecnico di Torino), Eros Fani (Politecnico di Torino), Gabriele Trivigno (Politecnico di Torino), and Carlo Masone (Politecnico di Torino)</i>	
On the Efficiency of Privacy Attacks in Federated Learning	4226
<i>Nawrin Tabassum (Ahsanullah University of Science and Technology), Ka-Ho Chow (University of Hong Kong), Xuyu Wang (Florida International University), Wenbin Zhang (Florida International University), and Yanzhao Wu (Florida International University)</i>	
Federated Hyperparameter Optimization Through Reward-Based Strategies: Challenges and Insights	4236
<i>Krishna Kanth Nakka (Huawei Munich Research Center), Ahmed Frikha (Huawei Munich Research Center), Ricardo Mendis (Huawei Munich Research Center), Xue Jiang (Huawei Munich Research Center), and Xuebing Zhou (Huawei Munich Research Center)</i>	

Image Matching: Local Features and Beyond

DeDoDe v2: Analyzing and Improving the DeDoDe Keypoint Detector	4245
<i>Johan Edstedt (Linköping University, Sweden), Georg Bökman (Chalmers University of Technology, Sweden), and Zhenjun Zhao (The Chinese University of Hong Kong, Hong Kong)</i>	
Affine-based Deformable Attention and Selective Fusion for Semi-dense Matching	4254
<i>Hongkai Chen (Apple Inc.), Zixin Luo (Apple Inc.), Yurun Tian (Apple Inc.), Xuyang Bai (Apple Inc.), Ziyu Wang (Apple Inc.), Lei Zhou (Apple Inc.), Mingmin Zhen (Apple Inc.), Tian Fang (Apple Inc.), David McKinnon (Apple Inc.), Yanghai Tsin (Apple Inc.), and Long Quan (Hong Kong University of Science and Technology)</i>	
EarthMatch: Iterative Coregistration for Fine-grained Localization of Astronaut Photography	4264
<i>Gabriele Berton (Politecnico di Torino), Gabriele Goletto (Politecnico di Torino), Gabriele Trivigno (Politecnico di Torino), Alex Stoken (Jacobs Technology, NASA Johnson Space Center), Barbara Caputo (Politecnico di Torino), and Carlo Masone (Politecnico di Torino)</i>	

XoFTR: Cross-modal Feature Matching Transformer	4275
<i>Önder Tuzcuoğlu (Middle East Technical University), Aybora Köksal (Middle East Technical University), Buğra Sofu (ROKETSAN Inc.), Sinan Kalkan (Middle East Technical University), and A. Aydin Alatan (Middle East Technical University)</i>	
Are Deep Learning Models Pre-trained on RGB Data Good Enough for RGB-Thermal Image Retrieval?	4287
<i>Amulya Pendota (Indian Institute of Technology Hyderabad, India) and Sumohana S. Channappayya (Indian Institute of Technology Hyderabad, India)</i>	

8th Workshop on Media Forensics

Finding AI-Generated Faces in the Wild	4297
<i>Gonzalo J. Aniano Porcile (LinkedIn), Jack Gindi (LinkedIn), Shivansh Mundra (LinkedIn), James R. Verbus (LinkedIn), and Hany Farid (University of California, Berkeley)</i>	
An Investigation into the Impact of AI-Powered Image Enhancement on Forensic Facial Recognition	4306
<i>Justin Norman (UC Berkeley) and Hany Farid (UC Berkeley)</i>	
Lost in Translation: Lip-Sync Deepfake Detection from Audio-Video Mismatch	4315
<i>Matyas Bohacek (Stanford University) and Hany Farid (University of California, Berkeley)</i>	
Can ChatGPT Detect DeepFakes? A Study of Using Multimodal Large Language Models for Media Forensics	4324
<i>Shan Jia (University at Buffalo), Reilin Lyu (Williamsville East High School), Kangran Zhao (The Chinese University of Hong Kong), Yize Chen (The Chinese University of Hong Kong), Zhiyuan Yan (The Chinese University of Hong Kong), Yan Ju (University at Buffalo), Chuanbo Hu (University at Albany), Xin Li (University at Albany), Baoyuan Wu (The Chinese University of Hong Kong), and Siwei Lyu (University at Buffalo)</i>	
E3: Ensemble of Expert Embedders for Adapting Synthetic Image Detectors to New Generators Using Limited Data	4334
<i>Aref Azizpour (Drexel University, USA), Tai D. Nguyen (Drexel University, USA), Manil Shrestha (Drexel University, USA), Kaidi Xu (Drexel University, USA), Edward Kim (Drexel University, USA), and Matthew C. Stamm (Drexel University, USA)</i>	
Fusion Transformer with Object Mask Guidance for Image Forgery Analysis	4345
<i>Dimitrios Karageorgiou (Centre for Research and Technology Hellas, Greece), Giorgos Kordopatis-Zilos (Czech Technical University in Prague, Czech), and Symeon Papadopoulos (Centre for Research and Technology Hellas, Greece)</i>	
Raising the Bar of AI-generated Image Detection with CLIP	4356
<i>Davide Cozzolino (University Federico II of Naples), Giovanni Poggi (University Federico II of Naples), Riccardo Corvi (University Federico II of Naples), Matthias Nießner (Technical University of Munich), and Luisa Verdoliva (University Federico II of Naples)</i>	

StampOne: Addressing Frequency Balance in Printer-proof Steganography	4367
<i>Farhad Shadmand (University of Coimbra, Portugal), Iurii Medvedev (University of Coimbra, Portugal), Luiz Schirmer (University of the Sinos River Valley Rio de Janeiro, Brazil), João Marcos (University of Coimbra, Portugal), and Nuno Gonçalves (University of Coimbra, Portugal)</i>	
Building Secure and Engaging Video Communication by Using Monitor Illumination	4377
<i>Jun Myeong Choi (University of North Carolina at Chapel Hill), Johnathan Leung (University of North Carolina at Chapel Hill), Noah Frahm (University of North Carolina at Chapel Hill), Max Christman (University of North Carolina at Chapel Hill), Gedas Bertasius (University of North Carolina at Chapel Hill), and Roni Sengupta (University of North Carolina at Chapel Hill)</i>	
Audio Provenance Analysis in Heterogeneous Media Sets	4387
<i>Milica Gerhardt (Fraunhofer IDMT, Germany), Luca Cuccovillo (Fraunhofer IDMT, Germany), and Patrick Aichroth (Fraunhofer IDMT, Germany)</i>	
Beyond Deepfake Images: Detecting AI-Generated Videos	4397
<i>Danial Samadi Vahdati (Drexel University, USA), Tai D. Nguyen (Drexel University, USA), Aref Azizpour (Drexel University, USA), and Matthew C. Stamm (Drexel University, USA)</i>	
Audio Transformer for Synthetic Speech Detection via Multi-Formant Analysis	4409
<i>Luca Cuccovillo (Fraunhofer IDMT), Milica Gerhardt (Fraunhofer IDMT), and Patrick Aichroth (Fraunhofer IDMT)</i>	
FairSSD: Understanding Bias in Synthetic Speech Detectors	4418
<i>Amit Kumar Singh Yadav (Purdue University, USA), Kratika Bhagtani (Purdue University, USA), Davide Salvi (Politecnico di Milano, Italy), Paolo Bestagini (Politecnico di Milano, Italy), and Edward J. Delp (Purdue University, USA)</i>	
Beyond the Screen: Evaluating Deepfake Detectors under Moiré Pattern Effects	4429
<i>Razaib Tariq (Sungkyunkwan University, South Korea), Minji Heo (Sungkyunkwan University, South Korea), Simon S. Woo (Sungkyunkwan University, South Korea), and Shahroz Tariq (CSIRO's Data61, Australia)</i>	

7th Workshop on Autonomous Driving (WAD)

Do More With What You Have: Transferring Depth-Scale from Labeled to Unlabeled Domains ...	4440
<i>Alexandra Dana (Samsung), Nadav Carmel (Samsung), Amit Shomer (Samsung), Ofer Manela (Samsung), and Tomer Peleg (Samsung)</i>	
CenterPoint Transformer for BEV Object Detection with Automotive Radar	4451
<i>Loveneet Saini (University of Wuppertal, Aptiv, Wuppertal, Germany), Yu Su (Aptiv, Wuppertal, Germany), Hasan Tercan (University of Wuppertal, Wuppertal, Germany), and Tobias Meisen (University of Wuppertal, Wuppertal, Germany)</i>	

Are NeRFs Ready for Autonomous Driving? Towards Closing the Real-to-simulation Gap	4461
<i>Carl Lindström (Zenseact; Chalmers University of Technology), Georg Hess (Zenseact; Chalmers University of Technology), Adam Lilja (Zenseact; Chalmers University of Technology), Maryam Fatemi (Zenseact), Lars Hammarstrand (Chalmers University of Technology), Christoffer Petersson (Zenseact; Chalmers University of Technology), and Lennart Svensson (Chalmers University of Technology)</i>	
Multi-Stream Cellular Test-Time Adaptation of Real-Time Models Evolving in Dynamic Environments	4472
<i>Benoît Gérin (UCLouvain), Anaïs Halin (ULiège), Anthony Cioppa (ULiège), Maxim Henry (ULiège), Bernard Ghanem (KAUST), Benoît Macq (UCLouvain), Christophe De Vleeschouwer (UCLouvain), and Marc Van Droogenbroeck (ULiège)</i>	
TrajFine: Predicted Trajectory Refinement for Pedestrian Trajectory Forecasting	4483
<i>Kuan-Lin Wang (National Yang Ming Chiao Tung University, Taiwan), Li-Wu Tsao (National Yang Ming Chiao Tung University, Taiwan), Jhih-Ciang Wu (National Taiwan University, Taiwan), Hong-Han Shuai (National Yang Ming Chiao Tung University, Taiwan), and Wen-Huang Cheng (National Taiwan University, Taiwan)</i>	
OccFeat: Self-supervised Occupancy Feature Prediction for Pretraining BEV Segmentation Networks	4493
<i>Sophia Sirko-Galouchenko (Valeo.ai, France), Alexandre Boulch (Valeo.ai, France), Spyros Gidaris (Valeo.ai, France), Andrei Bursuc (Valeo.ai, France), Antonin Vobecky (Czech Technical University, Czech Republic), Patrick Pérez (Kyutai, France), and Renaud Marlet (Ecole des Ponts ParisTech / Valeo.ai, France)</i>	
Potential Risk Localization via Weak Labeling out of Blind Spot	4504
<i>Kota Shimomura (Chubu University), Tsubasa Hirakawa (Chubu University), Takayoshi Yamashita (Chubu University), and Hironobu Fujiyoshi (Chubu University)</i>	
Click, Crop & Detect: One-Click Offline Annotation for Human-in-the-Loop 3D Object Detection on Point Clouds	4514
<i>Nitin Kumar Saravana Kannan (Valeo Schalter und Sensoren GmbH, Germany), Matthias Reuse (Valeo Schalter und Sensoren GmbH, Germany), and Martin Simon (Valeo Schalter und Sensoren GmbH, Germany)</i>	
Lift-Attend-Splat: Bird’s-eye-view Camera-lidar Fusion using Transformers	4526
<i>James Gunn (FiveAI, UK), Zygmunt Lenyk (FiveAI, UK), Anuj Sharma (FiveAI, UK), Andrea Donati (FiveAI, UK), Alexandru Buburuzan (FiveAI, UK), John Redford (FiveAI, UK), and Romain Mueller (FiveAI, UK)</i>	
DuST: Dual Swin Transformer for Multi-modal Video and Time-Series Modeling	4537
<i>Liang Shi (Virginia Polytechnic Institute and State University, USA), Yixin Chen (Virginia Polytechnic Institute and State University, USA), Meimei Liu (Virginia Polytechnic Institute and State University, USA), and Feng Guo (Virginia Polytechnic Institute and State University, USA)</i>	

TFNet: Exploiting Temporal Cues for Fast and Accurate LiDAR Semantic Segmentation	4547
<i>Rong Li (HKUST(GZ), China), Shijie Li (University of Bonn, Germany), Xieyuanli Chen (University of Bonn, Germany), Teli Ma (HKUST(GZ), China), Juergen Gall (University of Bonn, Germany; The Lamarr Institute for Machine Learning and Artificial Intelligence, Germany), and Junwei Liang (HKUST(GZ), China; HKUST, China)</i>	
CaBins: CLIP-based Adaptive Bins for Monocular Depth Estimation	4557
<i>Eunjin Son (Jeonbuk National University) and Sang Jun Lee (Jeonbuk National University)</i>	
Exploring Real World Map Change Generalization of Prior-Informed HD Map Prediction Models	4568
<i>Samuel M. Bateman (Nuro, Inc.), Ning Xu (Nuro, Inc.), H. Charles Zhao (Nuro, Inc.), Yael Ben Shalom (Nuro, Inc.), Vince Gong (Nuro, Inc.), Greg Long (Nuro, Inc.), and Will Maddern (Nuro, Inc.)</i>	
MULi-Ev: Maintaining Unperturbed LiDAR-Event Calibration	4579
<i>Mathieu Cocheteux (Université de technologie de Compiègne, CNRS, Heudiasyc, France), Julien Moreau (Université de technologie de Compiègne, CNRS, Heudiasyc, France), and Franck Davoine (CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, France)</i>	

6th Workshop and Competition on Affective Behavior Analysis in-the-wild

The 6th Affective Behavior Analysis In-the-wild (ABAW) Competition	4587
<i>Dimitrios Kollias (Queen Mary University of London, UK), Panagiotis Tzirakis (Hume AI, USA), Alan Cowen (Hume AI, USA), Stefanos Zafeiriou (Imperial College London, UK), Irene Kotsia (Cogitat, UK), Alice Baird (Hume AI, USA), Chris Gagne (Hume AI, USA), Chunchang Shao (Queen Mary University of London, UK), and Guanyu Hu (Queen Mary University of London, UK; Xi'an Jiaotong University, China)</i>	
Unsupervised Multi-Person 3D Human Pose Estimation From 2D Poses Alone	4599
<i>Peter Hardy (University of Southampton) and Hansung Kim (University of Southampton)</i>	
Multi-Task Multi-Modal Self-Supervised Learning for Facial Expression Recognition	4604
<i>Marah Halawa (Technical University of Berlin, Germany), Florian Blume (Technical University of Berlin, Germany), Pia Bideau (Inria, France), Martin Maier (Humboldt-University of Berlin, Germany), Rasha Abdel Rahman (Humboldt-University of Berlin, Germany), and Olaf Hellwich (Technical University of Berlin, Germany)</i>	
Purposeful Regularization with Reinforcement Learning for Facial Expression Recognition In-the-Wild	4615
<i>SangHwa Hong (Seoul National University of Science and Technology, South Korea)</i>	

Joint Multimodal Transformer for Emotion Recognition in the Wild	4625
<i>Paul Waligora (École de technologie supérieure), Muhammad Haseeb Aslam (École de technologie supérieure), Muhammad Osama Zeeshan (École de technologie supérieure), Soufiane Belharbi (École de technologie supérieure), Alessandro Lameiras Koerich (École de technologie supérieure), Marco Pedersoli (École de technologie supérieure), Simon Bacon (Concordia University), and Eric Granger (École de technologie supérieure)</i>	
CMOSE: Comprehensive Multi-Modality Online Student Engagement Dataset with High-Quality Labels	4636
<i>Chi-Hsuan Wu (Hong Kong University of Science and Technology, Hong Kong), Shih-Yang Liu (Hong Kong University of Science and Technology, Hong Kong), Xijie Huang (Hong Kong University of Science and Technology, Hong Kong), Xingbo Wang (Hong Kong University of Science and Technology, Hong Kong), Rong Zhang (Hong Kong University of Science and Technology, Hong Kong), Luca Minciullo (LifeHikes, Hong Kong), Wong Kai Yiu (LifeHikes, Hong Kong), Kenny Kwan (LifeHikes, Hong Kong), and Kwang-Ting Cheng (Hong Kong University of Science and Technology, Hong Kong)</i>	
3D Human Pose Estimation with Occlusions: Introducing BlendMimic3D Dataset and GCN Refinement	4646
<i>Filipa Lino (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), Carlos Santiago (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), and Manuel Marques (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal)</i>	
Unimodal Multi-Task Fusion for Emotional Mimicry Intensity Prediction	4657
<i>Tobias Hallmen (University of Augsburg, Germany), Fabian Deuser (University of the Bundeswehr Munich, Germany), Norbert Oswald (University of the Bundeswehr Munich, Germany), and Elisabeth André (University of Augsburg)</i>	
Enhancing Emotion Recognition with Pre-trained Masked Autoencoders and Sequential Learning.....	4666
<i>Weiwei Zhou (Chinatelecom Cloud), Jiada Lu (Chinatelecom Cloud), Chengkun Ling (Chinatelecom Cloud), Weifeng Wang (Chinatelecom Cloud), and Shaowei Liu (Chinatelecom Cloud)</i>	
MMA-DFER: MultiModal Adaptation of Unimodal Models for Dynamic Facial Expression Recognition In-the-wild	4673
<i>Kateryna Chumachenko (Tampere University), Alexandros Iosifidis (Aarhus University), and Moncef Gabbouj (Tampere University)</i>	
CAGE: Circumplex Affect Guided Expression Inference	4683
<i>Niklas Wagner (Karlsruhe Institute of Technology (KIT), Germany), Felix Mätzler (Karlsruhe Institute of Technology (KIT), Germany), Samed R. Vossberg (Karlsruhe Institute of Technology (KIT), Germany), Helen Schneider (Karlsruhe Institute of Technology (KIT), Germany), Svetlana Pavlitska (FZI Research Center for Information Technology, Germany), and J. Marius Zöllner (Karlsruhe Institute of Technology (KIT), Germany)</i>	

Video Representation Learning for Conversational Facial Expression Recognition Guided by Multiple View Reconstruction	4693
<i>Valeriya Strizhkova (INRIA, France), Laura M. Ferrari (INRIA, France), Hadi Kachmar (INRIA, France), Antitza Dantcheva (INRIA, France), and Francois Bremond (INRIA, France)</i>	
Leveraging Pre-trained Multi-task Deep Models for Trustworthy Facial Analysis in Affective Behaviour Analysis In-the-Wild	4703
<i>Andrey V. Savchenko (HSE University)</i>	
Drone-HAT: Hybrid Attention Transformer for Complex Action Recognition in Drone Surveillance Videos	4713
<i>Mustaqeem Khan (Mohamed Bin Zayed University of Artificial Intelligence, UAE), Jamil Ahmad (Mohamed Bin Zayed University of Artificial Intelligence, UAE), Abdulmotaleb El Saddik (University of Ottawa, Canada), Wail Gueaieb (University of Ottawa, Canada), Giulia De Masi (Technology Innovation Institute, UAE), and Fakhri Karray (University of Waterloo, Canada)</i>	
TCCT-Net: Two-Stream Network Architecture for Fast and Efficient Engagement Estimation via Behavioral Feature Signals	4723
<i>Alexander Vedernikov (University of Oulu, Finland), Puneet Kumar (University of Oulu, Finland), Haoyu Chen (University of Oulu, Finland), Tapio Seppänen (University of Oulu, Finland), and Xiaobai Li (Zhejiang University, China; University of Oulu, Finland)</i>	
Learning Transferable Compound Expressions from Masked AutoEncoder Pretraining	4733
<i>Feng Qiu (Netease, China), Heming Du (The University of Queensland, Australia), Wei Zhang (Netease, China), Chen Liu (The University of Queensland, Australia), Lincheng Li (Netease, China), Tianchen Guo (The University of Queensland, Australia), and Xin Yu (The University of Queensland, Australia)</i>	
Language-guided Multi-modal Emotional Mimicry Intensity Estimation	4742
<i>Feng Qiu (Netease, China), Wei Zhang (Netease, China), Chen Liu (The University of Queensland, Australia), Lincheng Li (Netease, China), Heming Du (The University of Queensland, Australia), Tianchen Guo (The University of Queensland, Australia), and Xin Yu (The University of Queensland, Australia)</i>	
Zero-Shot Audio-Visual Compound Expression Recognition Method based on Emotion Probability Fusion	4752
<i>Elena Ryumina (St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, Russia), Maxim Markitantov (St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, Russia), Dmitry Ryumin (St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, Russia), Heysem Kaya (Utrecht University, The Netherlands), and Alexey Karpov (St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, Russia)</i>	

An Effective Ensemble Learning Framework for Affective Behaviour Analysis	4761
<i>Wei Zhang (Netease Fuxi AI Lab, China), Feng Qiu (Netease Fuxi AI Lab, China), Chen Liu (The University of Queensland, Australia), Lincheng Li (Netease Fuxi AI Lab, China), Heming Du (The University of Queensland, Australia), Tianchen Guo (The University of Queensland, Australia), and Xin Yu (The University of Queensland, Australia)</i>	
Multi-modal Arousal and Valence Estimation under Noisy Conditions	4773
<i>Denis Dresvyanskiy (Ulm University, Germany; ITMO University, Russia), Maxim Markitantov (St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, Russia), Jiawei Yu (Utrecht University, The Netherlands), Heysem Kaya (Utrecht University, The Netherlands), and Alexey Karpov (St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, Russia)</i>	
Emotic Masked Autoencoder on Dual-views with Attention Fusion for Facial Expression Recognition	4784
<i>Xuan-Bach Nguyen (Ho Chi Minh City University of Technology, Vietnam), Hoang-Thien Nguyen (Posts and Telecommunications Institute of Technology, Vietnam), Thanh-Huy Nguyen (Ho Chi Minh City University of Education, Vietnam), Nhu-Tai Do (University of Economics Ho Chi Minh City-UEH Vietnam), and Quang Vinh Dinh (Vietnamese-German University, Vietnam)</i>	
REFA: Real-time Egocentric Facial Animations for Virtual Reality	4793
<i>Qiang Zhang (Reality Labs at Meta), Tong Xiao (Reality Labs at Meta), Haroun Habeeb (Reality Labs at Meta), Larissa Laich (Reality Labs at Meta), Sofien Bouaziz (Reality Labs at Meta), Patrick Snape (Reality Labs at Meta), Wenjing Zhang (Reality Labs at Meta), Matthew Cioffi (Reality Labs at Meta), Peizhao Zhang (Reality Labs at Meta), Pavel Pidlypenskyi (Reality Labs at Meta), Winnie Lin (Reality Labs at Meta), Luming Ma (Reality Labs at Meta), Mengjiao Wang (Reality Labs at Meta), Kunpeng Li (Reality Labs at Meta), Chengjiang Long (Reality Labs at Meta), Steven Song (Reality Labs at Meta), Martin Prazak (Reality Labs at Meta), Alexander Sjöholm (Reality Labs at Meta), Ajinkya Deogade (Reality Labs at Meta), Jaebong Lee (Reality Labs at Meta), Julio Delgado Mangas (Reality Labs at Meta), and Amaury Aubel (Reality Labs at Meta)</i>	
Recursive Joint Cross-Modal Attention for Multimodal Fusion in Dimensional Emotion Recognition	4803
<i>R. Gnana Praveen (Computer Research Institute of Montreal) and Jahangir Alam (Computer Research Institute of Montreal)</i>	

AUD-TGN: Advancing Action Unit Detection with Temporal Convolution and GPT-2 in Wild Audiovisual Contexts	4814
<i>Jun Yu (University of Science and Technology of China), Zerui Zhang (University of Science and Technology of China), Zhihong Wei (University of Science and Technology of China), Gongpeng Zhao (University of Science and Technology of China), Zhongpeng Cai (University of Science and Technology of China), Yongqi Wang (University of Science and Technology of China), Guochen Xie (University of Science and Technology of China), Jichao Zhu (University of Science and Technology of China), Wangyuan Zhu (University of Science and Technology of China), Qingsong Liu (Unisound AI Technology Co., Ltd.), and Jiaen Liang (Unisound AI Technology Co., Ltd.)</i>	
Uncovering Hidden Emotions with Adaptive Multi-Attention Graph Networks	4822
<i>Ankith Jain Rakesh Kumar (University of California Riverside, USA) and Bir Bhanu (University of California Riverside, USA)</i>	
Evaluating the Effectiveness of Video Anomaly Detection in the Wild: Online Learning and Inference for Real-world Deployment	4832
<i>Shanle Yao (University of North Carolina at Charlotte, USA), Ghazal Alinezhad Noghre (University of North Carolina at Charlotte, USA), Armin Danesh Pazho (University of North Carolina at Charlotte, USA), and Hamed Tabkhi (University of North Carolina at Charlotte, USA)</i>	
Unravelling Robustness of Deep Face Recognition Networks Against Illicit Drug Abuse Images...	4842
<i>Hruturaj Dhake (IISER Bhopal) and Akshay Agarwal (IISER Bhopal)</i>	
EmotiEffNet and Temporal Convolutional Networks in Video-based Facial Expression Recognition and Action Unit Detection	4849
<i>Andrey V. Savchenko (HSE University) and Anna P. Sidorova (HSE University)</i>	
Emotion Recognition Using Transformers with Random Masking	4860
<i>Seongjae Min (Kookmin University, Korea), Junseok Yang (Kookmin University, Korea), and Sejoon Lim (Kookmin University, Korea)</i>	
Efficient Feature Extraction and Late Fusion Strategy for Audiovisual Emotional Mimicry Intensity Estimation	4866
<i>Jun Yu (University of Science and Technology of China, China), Wangyuan Zhu (University of Science and Technology of China, China), Jichao Zhu (University of Science and Technology of China, China), Zhongpeng Cai (University of Science and Technology of China, China), Gongpeng Zhao (University of Science and Technology of China, China), Zerui Zhang (University of Science and Technology of China, China), Guochen Xie (University of Science and Technology of China, China), Zhihong Wei (University of Science and Technology of China, China), Qingsong Liu (Unisound AI Technology Co., Ltd., China), and Jiaen Liang (Unisound AI Technology Co., Ltd., China)</i>	

Multi Model Ensemble for Compound Expression Recognition	4873
<i>Jun Yu (University of Science and Technology of China, China), Jichao Zhu (University of Science and Technology of China, China), Wangyuan Zhu (University of Science and Technology of China, China), Zhongpeng Cai (University of Science and Technology of China, China), Gongpeng Zhao (University of Science and Technology of China, China), Zhihong Wei (University of Science and Technology of China, China), Guochen Xie (University of Science and Technology of China, China), Zerui Zhang (University of Science and Technology of China, China), Qingsong Liu (Unisound AI Technology Co., Ltd, China), and Jiaen Liang (Unisound AI Technology Co., Ltd, China)</i>	
Exploring Facial Expression Recognition through Semi-Supervised Pre-training and Temporal Modeling	4880
<i>Jun Yu (University of Science and Technology of China, China), Zhihong Wei (University of Science and Technology of China, China), Zhongpeng Cai (University of Science and Technology of China, China), Gongpeng Zhao (University of Science and Technology of China, China), Zerui Zhang (University of Science and Technology of China, China), Yongqi Wang (University of Science and Technology of China, China), Guochen Xie (University of Science and Technology of China, China), Jichao Zhu (University of Science and Technology of China, China), Wangyuan Zhu (University of Science and Technology of China, China), Qingsong Liu (Unisound AI Technology Co., Ltd, China), and Jiaen Liang (Unisound AI Technology Co., Ltd, China)</i>	
CUE-Net: Violence Detection Video Analytics with Spatial Cropping, Enhanced UniformerV2 and Modified Efficient Additive Attention	4888
<i>Damith Chamalke Senadeera (Queen Mary University of London, UK), Xiaoyun Yang (Remark AI UK Limited, UK), Dimitrios Kollias (Queen Mary University of London, UK), and Gregory Slabaugh (Queen Mary University of London, UK)</i>	

Domain adaptation, Explainability and Fairness in AI for Medical Image Analysis (DEF-AI-MIA)

One Class Classification-based Quality Assurance of Organs-at-risk Delineation in Radiotherapy	4898
<i>Yihao Zhao (Sun Yat-sen University), Cuiyun Yuan (Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen), Ying Liang (Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen), Yang Li (Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen), Chunxia Li (Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen), Man Zhao (Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen), Jun Hu (Sun Yat-sen University), Ningze Zhong (Sun Yat-sen University), and Chenbin Liu (Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen)</i>	

Domain Adaptation, Explainability & Fairness in AI for Medical Image Analysis: Diagnosis of COVID-19 based on 3-D Chest CT-scans	4907
<i>Dimitrios Kollias (Queen Mary University of London, UK), Anastasios Arsenos (National Technical University of Athens, Greece), and Stefanos Kollias (National Technical University of Athens, Greece; National Infrastructures for Research and Technology, Greece)</i>	
Comparative Analysis of Generalization and Harmonization Methods for 3D Brain fMRI Images: A Case Study on OpenBHB Dataset	4915
<i>Soroosh Safari Loaliyan (University of California Riverside, USA) and Greg Ver Steeg (University of California Riverside, USA)</i>	
A Closer Look at Spatial-Slice Features Learning for COVID-19 Detection	4924
<i>Chih-Chung Hsu (National Cheng Kung University, Taiwan), Chia-Ming Lee (National Cheng Kung University, Taiwan), Yang Fan Chiang (National Cheng Kung University, Taiwan), Yi-Shiuan Chou (National Cheng Kung University, Taiwan), Chih-Yu Jiang (National Cheng Kung University, Taiwan), Shen-Chieh Tai (National Cheng Kung University, Taiwan), and Chi-Han Tsai (National Cheng Kung University, Taiwan)</i>	
Interpreting COVID Lateral Flow Tests' Results with Foundation Models	4935
<i>Stuti Pandey (University of Colorado Boulder, USA), Josh Myers-Dean (University of Colorado Boulder, USA), Jarek Reynolds (University of Colorado Boulder, USA), and Danna Gurari (University of Colorado Boulder; The University of Texas at Austin, USA)</i>	
Fetal ECG Extraction on Time-Frequency Domain using Conditional GAN	4943
<i>Vuong D. Nguyen (University of Houston)</i>	
Focusing on What Matters: Fine-grained Medical Activity Recognition for Trauma Resuscitation via Actor Tracking	4950
<i>Wenjin Zhang (Rutgers University, USA), Keyi Li (Rutgers University, USA), Sen Yang (Waymo, USA), Sifan Yuan (Rutgers University, USA), Ioan Marsic (Rutgers University, USA), Genevieve J. Sippel (Children's National Hospital, USA), Mary S. Kim (Children's National Hospital, USA), and Randall S. Burd (Children's National Medical Center, USA)</i>	
How SAM Perceives Different mp-MRI Brain Tumor Domains?	4959
<i>Cecilia Diana-Albelda (Universidad Autónoma de Madrid, Spain), Roberto Alcover-Couso (Universidad Autónoma de Madrid, Spain), Álvaro García-Martín (Universidad Autónoma de Madrid, Spain), and Jesus Bescos (Universidad Autónoma de Madrid, Spain)</i>	
LaPA: Latent Prompt Assist Model For Medical Visual Question Answering	4971
<i>Tiancheng Gu (The University of Sydney, Australia), Kaicheng Yang (DeepGlint, China), Dongnan Liu (The University of Sydney, Australia), and Weidong Cai (The University of Sydney, Australia)</i>	
SegFormer3D: An Efficient Transformer for 3D Medical Image Segmentation	4981
<i>Shehan Perera (The Ohio State University), Pouyan Navard (The Ohio State University), and Alper Yilmaz (The Ohio State University)</i>	

PP-SAM: Perturbed Prompts for Robust Adaption of Segment Anything Model for Polyp Segmentation	4989
<i>Md Mostafijur Rahman (The University of Texas at Austin, USA), Mustafa Munir (The University of Texas at Austin, USA), Debesh Jha (Northwestern University, USA), Ulas Bagci (Northwestern University, USA), and Radu Marculescu (The University of Texas at Austin, USA)</i>	
Using Counterfactual Information for Breast Classification Diagnosis	4996
<i>Miguel Cardoso (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), Carlos Santiago (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), and Jacinto C. Nascimento (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal)</i>	
FPN-IAIA-BL: A Multi-Scale Interpretable Deep Learning Model for Classification of Mass Margins in Digital Mammography	5003
<i>Julia Yang (Duke University), Alina Jade Barnett (Duke University), Jon Donnelly (Duke University), Satvik Kishore (Duke University), Jerry Fang (Duke University), Fides Regina Schwartz (Brigham and Women's Hospital), Chaofan Chen (University of Maine), Joseph Y. Lo (Duke University), and Cynthia Rudin (Duke University)</i>	
Source-free Domain Adaptation for Video Object Detection Under Adverse Image Conditions	5010
<i>Xingguang Zhang (Purdue University) and Chih-Hsien Chou (Futurewei Technologies, Inc)</i>	
Evaluating Confidence Calibration in Endoscopic Diagnosis Models	5020
<i>Nikoo Dehghani (Eindhoven University of Technology, The Netherlands), Ayla Thijssen (Maastricht University Medical Center+, The Netherlands), Quirine E. W. van der Zander (Maastricht University Medical Center+, The Netherlands), Ramon-Michel Schreuder (Catharina Hospital, The Netherlands), Erik J. Schoon (Catharina Hospital, The Netherlands), Fons van der Sommen (Eindhoven University of Technology, The Netherlands), and Peter H. N. de With (Eindhoven University of Technology, The Netherlands)</i>	
Enhancing Ki-67 Cell Segmentation with Dual U-Net Models: A Step Towards Uncertainty-Informed Active Learning	5026
<i>David Anglada-Rotger (Image Processing Group - Universitat Politècnica de Catalunya), Julia Sala (Image Processing Group - Universitat Politècnica de Catalunya), Ferran Marques (Image Processing Group - Universitat Politècnica de Catalunya), Philippe Salembier (Image Processing Group - Universitat Politècnica de Catalunya), and Montse Pardàs (Image Processing Group - Universitat Politècnica de Catalunya)</i>	
Complex Style Image Transformations for Domain Generalization in Medical Images	5036
<i>Nikolaos Spanos (National Technical University of Athens, Greece), Anastasios Arsenos (National Technical University of Athens, Greece), Paraskevi-Antonia Theofilou (National Technical University of Athens, Greece), Paraskevi Tzouveli (National Technical University of Athens, Greece), Athanasios Voulodimos (National Technical University of Athens, Greece), and Stefanos Kollias (National Technical University of Athens, Greece)</i>	

Medical Image Segmentation with InTEnt: Integrated Entropy Weighting for Single Image Test-Time Adaptation	5046
<i>Haoyu Dong (Duke University, USA), Nicholas Konz (Duke University, USA), Hanxue Gu (Duke University, USA), and Maciej A. Mazurowski (Duke University, USA)</i>	
Prototype-based Interpretable Model for Glaucoma Detection	5056
<i>Mohana Singh (TCS Research, India), B S Vivek (TCS Research, India), Jayavardhana Gubbi (TCS Research, India), and Arpan Pal (TCS Research, India)</i>	
Unsupervised Domain Adaptation for Multi-Stain Cell Detection in Breast Cancer with Transformers	5066
<i>Oscar Pina (Universitat Politècnica de Catalunya - BarcelonaTech (UPC), Spain) and Verónica Vilaplana (Universitat Politècnica de Catalunya - BarcelonaTech (UPC), Spain)</i>	
A Deep Biclustering Framework for Brain Network Analysis	5075
<i>Md Abdur Rahaman (Georgia Institute of Technology), Zening Fu (Center for Translational Research in Neuroimaging and Data Science (TRenDS)), Armin Iraj (Center for Translational Research in Neuroimaging and Data Science (TRenDS)), and Vince Calhoun (Center for Translational Research in Neuroimaging and Data Science (TRenDS))</i>	
Residual-based Language Models are Free Boosters for Biomedical Imaging Tasks	5086
<i>Zhixin Lai (Cornel University), Jing Wu (University of Illinois at Urbana-Champaign), Suiyao Chen (University of South Florida), Yucheng Zhou (University of Macau), and Naira Hovakimyan (University of Illinois at Urbana-Champaign)</i>	
Deep-Adaptation: Ensembling and Test Augmentation for Covid-19 Detection and Covid-19 Domain Adaptation from 3D CT-Scans	5097
<i>Fares Bougourzi (Univerity of Polytechnique Hauts-de-France, France), Feryal Windal Moulai (Univerity of Polytechnique Hauts-de-France, France), Halim Benhabiles (Univ. Lille, Centre for Digital Systems, France), Fadi Dornaika (University of the Basque Country UPV/EHU, San Sebastian, Spain; IKERBASQUE, Basque Foundation for Science, Bilbao, Spain), and Abdelmalik Taleb-Ahmed (Universite Polytechnique Hauts-de-France, France; Universite de Lille, France)</i>	
ConPro: Learning Severity Representation for Medical Images using Contrastive Learning and Preference Optimization	5105
<i>Hong Nguyen (University of Southern California, USA), Hoang Nguyen (Vinuni-Illinoise Smart Health Center, Vietnam), Melinda Chang (University of Southern California, USA), Hieu Pham (Vinuni-Illinoise Smart Health Center, Vietnam), Shrikanth Narayanan (University of Southern California, USA), and Michael Pazzani (University of Southern California, USA)</i>	

Dr-SAM: An End-to-End Framework for Vascular Segmentation, Diameter Estimation, and Anomaly Detection on Angiography Images.	5113
<i>Vazgen Zohranyan (Yerevan State University (YSU), Servicetitan Inc.), Vagner Navasardyan (Johannes Wesling University Hospital, Ruhr University Bochum, Germany), Hayk Navasardyan (Synopsys Armenia CJSC), Jan Borggreffe (Johannes Wesling University Hospital, Ruhr University Bochum, Germany), and Shant Navasardyan (Picsart AI Research (PAIR), Armenia)</i>	
Cluster Triplet Loss for Unsupervised Domain Adaptation on Histology Images	5122
<i>Ruby Wood (University of Oxford, UK), Enric Domingo (University of Oxford, UK), Viktor Hendrik Koelzer (University of Zurich, Switzerland; University of Oxford, UK; University Hospital Basel, Switzerland), Timothy S. Maughan (University of Oxford, UK; University of Liverpool, UK), and Jens Rittscher (University of Oxford, UK)</i>	
Bridging Domains in Melanoma Diagnostics: Predicting BRAF Mutations and Sentinel Lymph Node Positivity with Attention-Based Models in Histological Images	5132
<i>Carlos Hernandez-Perez (Universitat Politecnica de Catalunya, Spain), Lauren Jimenez-Martin (Universitat Politecnica de Catalunya, Spain), and Veronica Vilaplana (Universitat Politecnica de Catalunya, Spain)</i>	
Domain Adaptation Using Pseudo Labels for COVID-19 Detection	5141
<i>Runtian Yuan (Fudan University, China), Qingqiu Li (Fudan University, China), Junlin Hou (The Hong Kong University of Science and Technology, China), Jilan Xu (Fudan University, China), Yuejie Zhang (Fudan University, China), Rui Feng (Fudan University, China), and Hao Chen (The Hong Kong University of Science and Technology, China)</i>	
Advancing COVID-19 Detection in 3D CT Scans	5149
<i>Qingqiu Li (Fudan University, China), Runtian Yuan (Fudan University, China), Junlin Hou (The Hong Kong University of Science and Technology, China), Jilan Xu (Fudan University, China), Yuejie Zhang (Fudan University, China), Rui Feng (Fudan University, China), and Hao Chen (The Hong Kong University of Science and Technology, China)</i>	
Achieving Reliable and Fair Skin Lesion Diagnosis via Unsupervised Domain Adaptation	5157
<i>Janet Wang (Tulane University), Yunbei Zhang (Tulane University), Zhengming Ding (Tulane University), and Jihun Hamm (Tulane University)</i>	
Classification of 2D Ultrasound Breast Cancer Images with Deep Learning	5167
<i>Jack Ellis (University of York, UK), Kofi Appiah (University of York, UK), Emmanuel Amankwaa-Frempong (Sweden-Ghana Medical Center, Ghana), and Sze Chai Kwok (Duke Kunshan University, China)</i>	
DCE-diff: Diffusion Model for Synthesis of Early and Late Dynamic Contrast-Enhanced MR Images from Non-Contrast Multimodal Inputs	5174
<i>Kishore Kumar M (Indian Institute of Technology Madras (IITM), India), Sriprabha Ramanarayanan (Indian Institute of Technology Madras (IITM), India), Sadhana S (Indian Institute of Technology Madras (IITM), India), Arunima Sarkar (Indian Institute of Technology Madras (IITM), India), Matcha Naga Gayathri (Indian Institute of Technology Madras (IITM), India), Keerthi Ram (Healthcare Technology Innovation Center(HTIC), India), and Mohanasankar Sivaprakasam (Indian Institute of Technology Madras (IITM), India; Healthcare Technology Innovation Center (HTIC), India)</i>	

Test-Time Adaptation with SaLIP: A Cascade of SAM and CLIP for Zero-shot Medical Image Segmentation	5184
<i>Sidra Aleem (ML-Labs, Dublin City University), Fangyijie Wang (ML-Labs, University College Dublin), Mayug Maniparambil (ML-Labs, Dublin City University), Eric Arazo (Centre for Applied AI (CeADAR), University College Dublin), Julia Dietlmeier (Insight SFI Centre for Data Analytics, Dublin City University), Kathleen Curran (ML-Labs, University College Dublin), Noel E. O' Connor (ML-Labs, Dublin City University), and Suzanne Little (Insight SFI Centre for Data Analytics, Dublin City University)</i>	
Improving Consistency in Cardiovascular Disease Risk Assessment: Cross-Camera Adaptation for Retinal Images	5194
<i>Weiye Zhang (The Hong Kong Polytechnic University, China), Danli Shi (The Hong Kong Polytechnic University, China), and Mingguang He (The Hong Kong Polytechnic University, China)</i>	
EfficientNet-SAM: A Novel EffecientNet with Spatial Attention Mechanism for COVID-19 Detection in Pulmonary CT Scans	5200
<i>Ramy Farag (University of Missouri, USA), Parth Upadhyay (University of Missouri, USA), Jacket Dembys (University of Missouri, USA), Yixiang Gao (University of Missouri, USA), Katherin Garces Montoya (University of Missouri, USA), Seyed Mohamad Ali Tousi (University of Missouri, USA), Gbenga Omotara (University of Missouri, USA), and Guilherme DeSouza (University of Missouri, USA)</i>	
A Multimodal Approach Integrating Convolutional and Recurrent Neural Networks for Alzheimer's Disease Temporal Progression Prediction	5207
<i>Durga Supriya HL (National Institute of Technology Karnataka, India), Swetha Mary Thomas (National Institute of Technology Karnataka, India), and Sowmya Kamath S (National Institute of Technology Karnataka, India)</i>	
Separating Lungs in CT Scans for Improved COVID19 Detection	5216
<i>Robert Turnbull (University of Melbourne, Australia) and Simon Mutch (University of Melbourne, Australia)</i>	
Blurry-Consistency Segmentation Framework with Selective Stacking on Differential Interference Contrast 3D Breast Cancer Spheroid	5223
<i>Thanh-Huy Nguyen (National Cheng Kung University), Thi Kim Ngan Ngo (National Cheng Kung University), Mai Anh Vu (National Cheng Kung University), and Ting-Yuan Tu (National Cheng Kung University)</i>	
Key Patches Are All You Need: A Multiple Instance Learning Framework For Robust Medical Diagnosis	5231
<i>D.J. Araújo (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), M.R. Verdelho (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), A. Bissoto (University of Campinas, Brazil), J.C. Nascimento (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), C. Santiago (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal), and C. Barata (Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Portugal)</i>	

IMIL: Interactive Medical Image Learning Framework	5241
<i>Adrit Rao (Stanford University), Andrea Fisher (Stanford University), Ken Chang (Stanford University), John Christopher Panagides (Stanford University), Katherine McNamara (Stanford University), Joon-Young Lee (Adobe Research), and Oliver Aalami (Stanford University)</i>	

The Sixth Workshop on Deep Learning for Geometric Computing (DLGC 2024)

RDPN6D: Residual-based Dense Point-wise Network for 6Dof Object Pose Estimation Based on RGB-D Images	5251
<i>Zong-Wei Hong (National Taiwan University), Yen-Yang Hung (National Taiwan University), and Chu-Song Chen (National Taiwan University)</i>	
LGAfford-Net: A Local Geometry Aware Affordance Detection Network for 3D Point Clouds	5261
<i>Ramesh Ashok Tabib (KLE Technological University, India), Dikshit Hegde (KLE Technological University, India), and Uma Mudenagudi (KLE Technological University, India)</i>	
SDFConnect: Neural Implicit Surface Reconstruction of a Sparse Point Cloud with Topological Constraints	5271
<i>Anushrut Jignasu (Iowa State University, USA), Aditya Balu (Iowa State University, USA), Soumik Sarkar (Iowa State University, USA), Chinmay Hegde (New York University, USA), Baskar Ganapathysubramanian (Iowa State University, USA), and Adarsh Krishnamurthy (Iowa State University, USA)</i>	

The First Workshop on the Evaluation of Generative Foundation Models

Diagnostic Benchmark and Iterative Inpainting for Layout-Guided Image Generation	5280
<i>Jaemin Cho (UNC Chapel Hill), Linjie Li (Microsoft Research), Zhengyuan Yang (Microsoft Research), Zhe Gan (Microsoft Research), Lijuan Wang (Microsoft Research), and Mohit Bansal (UNC Chapel Hill)</i>	
Evaluating and Improving Compositional Text-to-Visual Generation	5290
<i>Baiqi Li (Carnegie Mellon University), Zhiqiu Lin (Carnegie Mellon University), Deepak Pathak (Carnegie Mellon University), Jiayao Li (Carnegie Mellon University), Yixin Fei (Carnegie Mellon University), Kewen Wu (Carnegie Mellon University), Xide Xia (Meta), Pengchuan Zhang (Meta), Graham Neubig (Carnegie Mellon University), and Deva Ramanan (Carnegie Mellon University)</i>	
TITScore: Towards Long-Tail Effects in Text-to-Visual Evaluation with Generative Foundation Models	5302
<i>Pengliang Ji (Carnegie Mellon University, USA) and Junchen Liu (UC Berkeley, USA)</i>	

Evaluating Multimodal Large Language Models Across Distribution Shifts and Augmentations ..	5314
<i>Aayush Atul Verma (Arizona State University), Amir Saeidi (Arizona State University), Shamanthak Hegde (Arizona State University), Ajay Therala (Arizona State University), Fenil Denish Bardoliya (Arizona State University), Nagaraju Machavarapu (Arizona State University), Shri Ajay Kumar Ravindhiran (Arizona State University), Srija Malyala (Arizona State University), Agneet Chatterjee (Arizona State University), Yezhou Yang (Arizona State University), and Chitta Baral (Arizona State University)</i>	
T2VBench: Benchmarking Temporal Dynamics for Text-to-Video Generation	5325
<i>Pengliang Ji (Carnegie Mellon University, USA), Chuyang Xiao (ShanghaiTech University, China), Huilin Tai (McGill University, Canada), and Mingxiao Huo (Carnegie Mellon University, USA)</i>	

Agriculture-Vision: Challenges & Opportunities for Computer Vision in Agriculture

Improved Crop and Weed Detection with Diverse Data Ensemble Learning	5336
<i>Muhammad Hamza Asad (University of Regina, Canada), Saeed Anwar (King Fahd University of Petroleum and Minerals & SDAIA-KFUPM Joint Research Center for Artificial Intelligence, Saudi Arabia), and Abdul Bais (University of Regina, Canada)</i>	
The New Agronomists: Language Models are Experts in Crop Management	5346
<i>Jing Wu (University of Illinois at Urbana-Champaign), Zhixin Lai (Cornel University), Suiyao Chen (University of South Florida), Ran Tao (University of Illinois at Urbana-Champaign), Pan Zhao (University of Alabama), and Naira Hovakimyan (University of Illinois at Urbana-Champaign)</i>	
Energy-Efficient Uncertainty-Aware Biomass Composition Prediction at the Edge	5357
<i>Muhammad Zawish (Walton Institute, South East Technological University, Ireland), Paul Albert (Center for Augmented Reasoning, Australian Institute of Machine Learning, Australia), Flavio Esposito (Saint Louis University, USA), Steven Davy (Centre for Sustainable Digital Technologies at Technological University Dublin, Ireland), and Lizy Abraham (Walton Institute, South East Technological University, Ireland)</i>	
HarvestNet: A Dataset for Detecting Smallholder Farming Activity Using Harvest Piles and Remote Sensing	5366
<i>Jonathan Xu (University of Waterloo), Amna Elmustafa (Stanford University), Liya Weldegebriel (Stanford University), Emnet Negash (Ghent University), Richard Lee (Stanford University), Chenlin Meng (Stanford University), Stefano Ermon (Stanford University), and David Lobell (Stanford University)</i>	
Domain Targeted Synthetic Plant Style Transfer using Stable Diffusion, LoRA and ControlNet.....	5375
<i>Zane K.J. Hartley (University of Nottingham), Rob J. Lind (Syngenta), Michael P. Pound (University of Nottingham), and Andrew P. French (University of Nottingham)</i>	

Lacunarity Pooling Layers for Plant Image Classification using Texture Analysis	5384
<i>Akshatha Mohan (Texas A&M University) and Joshua Peebles (Texas A&M University)</i>	
Label Efficient Lifelong Multi-View Broiler Detection	5393
<i>Thorsten Cardoen (Ghent University - Imec, Belgium), Sam Leroux (Ghent University - Imec, Belgium), and Pieter Simoens (Ghent University - Imec, Belgium)</i>	
End-to-End Deep Learning Models for Gap Identification in Maize Fields	5403
<i>Rana Waqar (BioSense Institute, Serbia), Željana Grbović (BioSense Institute, Serbia), Maryam Khan (Farmevo Technologies, New York, USA), Nina Pajević (BioSense Institute, Serbia), Dimitrije Stefanović (BioSense Institute), Vladan Filipović (BioSense Institute, Serbia), Marko Panić (BioSense Institute, Serbia), and Nemanja Djuric (BioSense Institute, Serbia)</i>	
Tracking and Counting Apples in Orchards Under Intermittent Occlusions and Low Frame Rates	5413
<i>Gonçalo P. Matos (SISCOG - Sistemas Cognitivos, SA, Lisbon, Portugal; University of Lisbon, Portugal), Carlos Santiago (University of Lisbon, Portugal), João P. Costeira (University of Lisbon, Portugal), Ricardo L. Saldanha (SISCOG - Sistemas Cognitivos, SA, Lisbon, Portugal), and Ernesto M. Morgado (SISCOG - Sistemas Cognitivos SA, Lisbon, Portugal)</i>	
Generating Diverse Agricultural Data for Vision-Based Farming Applications	5422
<i>Mikolaj Cieslak (GreenMatterAI), Umabharathi Govindarajan (Blue River Technology), Alejandro Garcia (GreenMatterAI), Anuradha Chandrashekar (Blue River Technology), Torsten Hadrich (GreenMatterAI), Aleksander Mendoza-Drosik (GreenMatterAI), Dominik L. Michels (GreenMatterAI; KAUST / TU Darmstadt), Soren Pirk (GreenMatterAI; CAU), Chia-Chun Fu (Blue River Technology), and Wojciech Palubicki (GreenMatterAI; AMU)</i>	
Knowledge Distillation for Efficient Instance Semantic Segmentation with Transformers	5432
<i>Maohui Li (University of Bonn, Germany), Michael Halstead (University of Bonn, Germany), and Chris McCool (University of Bonn; Lamarr Institute for Machine Learning and Artificial Intelligence, Germany)</i>	
Label-free Anomaly Detection in Aerial Agricultural Images with Masked Image Modeling	5440
<i>Sambal Shikhar (Plaksha University) and Anupam Sobti (Plaksha University)</i>	
Domain Generalization for Crop Segmentation with Standardized Ensemble Knowledge Distillation	5450
<i>Simone Angarano (Politecnico di Torino), Mauro Martini (Politecnico di Torino), Alessandro Navone (Politecnico di Torino), and Marcello Chiaberge (Politecnico di Torino)</i>	
IrrNet: Advancing Irrigation Mapping with Incremental Patch Size Training on Remote Sensing Imagery	5460
<i>Oishee Binte Hoque (University of Virginia, USA), Samarth Swarup (University of Virginia, USA), Abhijin Adiga (University of Virginia, USA), Sayjro Kossi Nouwakpo (US Department of Agriculture, Agricultural Research Service, USA), and Madhav Marathe (University of Virginia, USA)</i>	

VisTA-SR: Improving the Accuracy and Resolution of Low-Cost Thermal Imaging Cameras for Agriculture	5470
<i>Heesup Yun (University of California, Davis), Sassoum Lo (University of California, Davis), Christine H. Diepenbrock (University of California, Davis), Brian N. Bailey (University of California, Davis), and J. Mason Earles (University of California, Davis)</i>	
Photorealistic Arm Robot Simulation for 3D Plant Reconstruction and Automatic Annotation using Unreal Engine 5	5480
<i>Xingjian Li (North Carolina State University, USA), Jeremy Park (North Carolina State University, USA), Chris Reberg-Horton (North Carolina State University, USA), Steven Mirsky (USDA Agricultural Research Service, USA), Edgar Lobaton (North Carolina State University, USA), and Lirong Xiang (North Carolina State University, USA)</i>	
Gasformer: A Transformer-based Architecture for Segmenting Methane Emissions from Livestock in Optical Gas Imaging	5489
<i>Toqi Tahamid Sarker (Southern Illinois University, USA), Mohamed G Embaby (Southern Illinois University, USA), Khaled R Ahmed (Southern Illinois University, USA), and Amer AbuGhazaleh (Southern Illinois University, USA)</i>	
End-to-end Solution for Tenebrio Molitor Rearing Monitoring with Uncertainty Estimation and Domain Shift Detection	5498
<i>Paweł Majewski (Wrocław University of Science and Technology, Poland), Piotr Lampa (Wrocław University of Science and Technology, Poland), Robert Burduk (Wrocław University of Science and Technology, Poland), and Jacek Reiner (Wrocław University of Science and Technology, Poland)</i>	

5th Workshop on Robot Visual Perception in Human Crowded Environments

InViG: Benchmarking Open-Ended Interactive Visual Grounding with 500K Dialogues	5508
<i>Hanbo Zhang (ByteDance Research), Jie Xu (Xi'an Jiaotong University, ByteDance Research), Yuchen Mo (ByteDance Research), and Tao Kong (ByteDance Research)</i>	
Must Unsupervised Continual Learning Relies on Previous Information?	5519
<i>Haoyang Cheng (University of Electronic Science and Technology of China, Chengdu, China), Haitao Wen (University of Electronic Science and Technology of China, Chengdu, China), Heqian Qiu (University of Electronic Science and Technology of China, Chengdu, China), Lanxiao Wang (University of Electronic Science and Technology of China, Chengdu, China), Minjian Zhang (University of Electronic Science and Technology of China, Chengdu, China), and Hongliang Li (University of Electronic Science and Technology of China, Chengdu, China)</i>	

HumanFormer: Human-centric Prompting Multi-modal Perception Transformer for Referring Crowd Detection	5530
<i>Heqian Qiu (University of Electronic Science and Technology of China, China), Lanxiao Wang (University of Electronic Science and Technology of China, China), Taijin Zhao (University of Electronic Science and Technology of China, China), Fanman Meng (University of Electronic Science and Technology of China, China), and Hongliang Li (University of Electronic Science and Technology of China, China)</i>	
GM-DETR: Generalized Multispectral DETection TRansformer with Efficient Fusion Encoder for Visible-Infrared Detection	5541
<i>Yiming Xiao (University of Electronic Science and Technology of China), Fanman Meng (University of Electronic Science and Technology of China), Qingbo Wu (University of Electronic Science and Technology of China), Linfeng Xu (University of Electronic Science and Technology of China), Mingzhou He (University of Electronic Science and Technology of China), and Hongliang Li (University of Electronic Science and Technology of China)</i>	
Pre-trained Bidirectional Dynamic Memory Network For Long Video Question Answering	5550
<i>Jinmeng Wu (Wuhan Institute of Technology, China), Pengcheng Shu (Wuhan Institute of Technology, China), Hanyu Hong (Wuhan Institute of Technology, China), Lei Ma (Wuhan Institute of Technology, China), Ying Zhu (Wuhan Institute of Technology, China), and Lei Wang (Wuhan Institute of Technology, China)</i>	
DSTCFuse: A Method based on Dual-cycled Cross-awareness of Structure Tensor for Semantic Segmentation via Infrared and Visible Image Fusion	5558
<i>Xuan Li (Wuhan Institute of Technology, China), Rongfu Chen (Wuhan Institute of Technology, China), Jie Wang (Wuhan Institute of Technology, China), Lei Ma (Wuhan Institute of Technology, China), Li Cheng (Wuhan Institute of Technology, China), and Haiwen Yuan (Wuhan Institute of Technology, China)</i>	
Is Our Continual Learner Reliable? Investigating Its Decision Attribution Stability through SHAP Value Consistency	5568
<i>Yusong Cai (University of Electronic Science and Technology of China), Shimou Ling (University of Electronic Science and Technology of China), Liang Zhang (University of Electronic Science and Technology of China), Lili Pan (University of Electronic Science and Technology of China), and Hongliang Li (University of Electronic Science and Technology of China)</i>	

Computer Vision for Mixed Reality

GRIB: Combining Global Reception and Inductive Bias For Human Segmentation and Matting	5576
<i>Yezhi Shen (Purdue University), Weichen Xu (Purdue University), Qian Lin (HP Inc.), Jan P. Allebach (Purdue University), and Fengqing Zhu (Purdue University)</i>	
3D Human Scan With A Moving Event Camera	5586
<i>Kai Kohyama (Keio University, Japan), Shintaro Shiba (Keio University, Japan), and Yoshimitsu Aoki (Keio University, Japan)</i>	

fMPI: Fast Novel View Synthesis in the Wild with Layered Scene Representations	N/A
<i>Jonas Kohler (Meta, Zurich), Nicolas Griffiths Sanchez (Meta, Switzerland), Luca Cavalli (Meta, Switzerland), Catherine Herold (Meta, Switzerland), Alberto Garcia Garcia (Google, Switzerland), Albert Pumerola (Meta, Switzerland), and Ali Thabet (Meta, Switzerland)</i>	
BOP Challenge 2023 on Detection, Segmentation and Pose Estimation of Seen and Unseen Rigid Objects	5610
<i>Tomas Hodan (Meta, Switzerland), Martin Sundermeyer (Google, Germany), Yann Labbe (Meta, Switzerland), Van Nguyen Nguyen (ENPC, France), Gu Wang (Tsinghua University, China), Eric Brachmann (Niantic, Germany), Bertram Drost (MVTec, Germany), Vincent Lepetit (ENPC, France), Carsten Rother (Heidelberg University, Germany), and Jiri Matas (CTU in Prague, Czech Republic)</i>	
Modeling Detailed Human Geometry with Adaptive Local Refinement	5620
<i>Bang Du (University of California San Diego), Kunyao Chen (University of California San Diego), Haochen Zhang (University of California San Diego), Fei Yin (University of California San Diego), Baichuan Wu (University of California San Diego), and Truong Nguyen (University of California San Diego)</i>	

The Sixth Workshop on Precognition: Seeing through the Future

H ³ Net: Irregular Posture Detection by Understanding Human Character and Core Structures ...	5631
<i>Seungha Noh (Kyonggi University), Kangmin Bae (ETRI), Yuseok Bae (ETRI), and Byong-Dai Lee (Kyonggi University)</i>	
CONDA: Continual Unsupervised Domain Adaptation Learning in Visual Perception for Self-Driving Cars	5642
<i>Thanh-Dat Truong (University of Arkansas), Pierce Helton (University of Arkansas), Ahmed Moustafa (University of Arkansas), Jackson David Cothren (University of Arkansas), and Khoa Luu (University of Arkansas)</i>	
VT-Former: An Exploratory Study on Vehicle Trajectory Prediction for Highway Surveillance through Graph Isomorphism and Transformer	5651
<i>Armin Danesh Pazho (University of North Carolina Charlotte), Ghazal Alinezhad Noghre (University of North Carolina Charlotte), Vinit Katariya (University of North Carolina Charlotte), and Hamed Tabkhi (University of North Carolina Charlotte)</i>	
VMRNN: Integrating Vision Mamba and LSTM for Efficient and Accurate Spatiotemporal Forecasting	5663
<i>Yujin Tang (The Hong Kong University of Science and Technology (Guangzhou), China), Peijie Dong (The Hong Kong University of Science and Technology (Guangzhou), China), Zhenheng Tang (Hong Kong Baptist University, China), Xiaowen Chu (The Hong Kong University of Science and Technology (Guangzhou), China), and Junwei Liang (The Hong Kong University of Science and Technology (Guangzhou), China)</i>	

Exploration of Data Augmentation Techniques for Bush Detection in Blueberry Orchards	5674
<i>Boris Čuljak (BioSense Institute, Serbia), Nina Pajević (BioSense Institute, Serbia), Vladan Filipović (BioSense Institute, Serbia), Dimitrije Stefanović (BioSense Institute, Serbia), Zeljana Grbović (BioSense Institute, Serbia), Nemanja Djuric (BioSense Institute, Serbia), and Marko Panić (BioSense Institute, Serbia)</i>	

AIS: Vision, Graphics and AI for Streaming

Retina : Low-Power Eye Tracking with Event Camera and Spiking Hardware	5684
<i>Pietro Bonazzi (ETH Zurich), Sizhen Bian (ETH Zurich), Giovanni Lippolis (Inivation AG), Yawei Li (ETH Zurich), Sadique Sheik (Synsense AG), and Michele Magno (ETH Zurich)</i>	
Joint Motion Detection in Neural Videos Training	5693
<i>Niloufar Pourian (Intel) and Alexey Supikov (Intel)</i>	
A Hybrid ANN-SNN Architecture for Low-Power and Low-Latency Visual Perception	5701
<i>Asude Aydin (University of Zurich, Switzerland), Mathias Gehrig (University of Zurich, Switzerland), Daniel Gehrig (University of Zurich, Switzerland), and Davide Scaramuzza (University of Zurich, Switzerland)</i>	
A Perspective on Deep Vision Performance with Standard Image and Video Codecs	5712
<i>Christoph Reich (TU Munich, Germany; TU Darmstadt, Germany), Oliver Hahn (TU Darmstadt, Germany), Daniel Cremers (TU Munich, Germany), Stefan Roth (TU Darmstadt, Germany), and Biplob Debnath (NEC Laboratories America Inc., USA)</i>	
One-Click Upgrade from 2D to 3D: Sandwiched RGB-D Video Compression for Stereoscopic Teleconferencing	5722
<i>Yueyu Hu (New York University), Onur G. Guleryuz (Google LLC), Philip A. Chou (Google LLC), Danhang Tang (Google LLC), Jonathan Taylor (Google LLC), Rus Maxham (Google LLC), and Yao Wang (New York University)</i>	
Deep Video Codec Control for Vision Models	5732
<i>Christoph Reich (TU Munich), Biplob Debnath (NEC Laboratories America, Inc.), Deep Patel (NEC Laboratories America, Inc.), Tim Prangemeier (TU Darmstadt), Daniel Cremers (TU Munich), and Srimat Chakradhar (NEC Laboratories America, Inc.)</i>	
Adaptive Render-Video Streaming for Virtual Environments	5742
<i>Jia-Jie Lim (Sony Interactive Entertainment), Matthias S. Treder (Sony Interactive Entertainment), Aaron Chadha (Sony Interactive Entertainment), and Yiannis Andreopoulos (Sony Interactive Entertainment)</i>	
Low Latency Point Cloud Rendering with Learned Splatting	5752
<i>Yueyu Hu (New York University), Ran Gong (Tsinghua University), Qi Sun (New York University), and Yao Wang (New York University)</i>	

MambaPupil: Bidirectional Selective Recurrent Model for Event-based Eye Tracking	5762
<i>Zhong Wang (University of Science and Technology of China), Zengyu Wan (University of Science and Technology of China), Han Han (University of Science and Technology of China), Bohao Liao (University of Science and Technology of China), Yuliang Wu (University of Science and Technology of China), Wei Zhai (University of Science and Technology of China), Yang Cao (University of Science and Technology of China), and Zheng-jun Zha (University of Science and Technology of China)</i>	
Co-designing a Sub-millisecond Latency Event-based Eye Tracking System with Submanifold Sparse CNN	5771
<i>Baoheng Zhang (The University of Hong Kong), Yizhao Gao (The University of Hong Kong), Jingyuan Li (The University of Hong Kong), and Hayden Kwok-Hay So (The University of Hong Kong)</i>	
A Lightweight Spatiotemporal Network for Online Eye Tracking with Event Camera	5780
<i>Yan Ru Pei (Brainchip Inc., USA), Sasskia Brüers (Brainchip Inc., France), Sébastien Crouzet (Brainchip Inc., France), Douglas McLelland (Brainchip Inc., France), and Olivier Coenen (Brainchip Inc., USA)</i>	
FAPNet: An Effective Frequency Adaptive Point-based Eye Tracker	5789
<i>Xiaopeng Lin (The Hong Kong University of Science and Technology(Guangzhou), China), Hongwei Ren (The Hong Kong University of Science and Technology(Guangzhou), China), and Bojun Cheng (The Hong Kong University of Science and Technology(Guangzhou), China)</i>	
COVER: A Comprehensive Video Quality Evaluator	5799
<i>Chenlong He (Fudan University, China), Qi Zheng (Fudan University, China), Ruoxi Zhu (Fudan University, China), Xiaoyang Zeng (Fudan University, China), Yibo Fan (Fudan University, China), and Zhengzhong Tu (University of Texas at Austin, USA)</i>	

Event-Based Eye Tracking. AIS 2024 Challenge Survey	5810
<p><i>Zuowen Wang (University of Zurich, Switzerland), Chang Gao (Delft University of Technology, Netherlands), Zongwei Wu (University of Würzburg, Germany), Marcos V. Conde (University of Würzburg, Germany), Radu Timofte (University of Würzburg, Germany), Shih-Chii Liu (University of Zurich, Switzerland), Qinyu Chen (University of Leiden, Netherlands), Zheng-jun Zha (University of Science and Technology of China), Wei Zhai (University of Science and Technology of China), Han Han (University of Science and Technology of China), Bohao Liao (University of Science and Technology of China), Yuliang Wu (University of Science and Technology of China), Zengyu Wan (University of Science and Technology of China), Zhong Wang (University of Science and Technology of China), Yang Cao (University of Science and Technology of China), Ganchao Tan (University of Science and Technology of China), Jinze Chen (University of Science and Technology of China), Yan Ru Pei (Brainchip Inc.), Sasskia Bruers (Brainchip Inc.), Sebastien Crouzet (Brainchip Inc.), Douglas McLelland (Brainchip Inc.), Oliver Coenen (Brainchip Inc.), Baoheng Zhang (The University of Hong Kong), Yizhao Gao (The University of Hong Kong), Jingyuan Li (The University of Hong Kong), Hayden Kwok-Hay So (The University of Hong Kong), Philippe Bich (The University of Hong Kong), Chiara Boretti (The University of Hong Kong), Luciano Prono (The University of Hong Kong), Mircea Lica (Delft University of Technology), David Dinucu-Jianu (Delft University of Technology), Catalin Griu (Delft University of Technology), Xiaopeng Lin (The Hong Kong University of Science and Technology (Guangzhou)), Hongwei Ren (The Hong Kong University of Science and Technology (Guangzhou)), Bojun Cheng (The Hong Kong University of Science and Technology (Guangzhou)), Xinan Zhang (Georgia Institute of Technology), Valentin Vial (Georgia Institute of Technology), Anthony Yezzi (Georgia Institute of Technology), and James Tsai (Georgia Institute of Technology)</i></p>	
AIS 2024 Challenge on Video Quality Assessment of User-Generated Content: Methods and Results	5826
<p><i>Marcos V. Conde (University of Würzburg; Sony Interactive Entertainment), Saman Zadtootaghaj (Sony Interactive Entertainment), Nabajeet Barman (Sony Interactive Entertainment), Radu Timofte (University of Würzburg), Chenlong He, Qi Zheng, Ruoxi Zhu, Zhengzhong Tu, Haiqiang Wang, Xiangguang Chen, Wenhui Meng, Xiang Pan, Huiying Shi, Han Zhu, Xiaozhong Xu, Lei Sun, Zhenzhong Chen, Shan Liu, Zicheng Zhang, Haoning Wu, Yingjie Zhou, Chunyi Li, Xiaohong Liu, Weisi Lin, Guangtao Zhai, Wei Sun, Yuqin Cao, Yanwei Jiang, Jun Jia, Zhichao Zhang, Zijian Chen, Weixia Zhang, Xiongkuo Min, Steve Goring, Zihao Qi, and Chen Feng</i></p>	

Real-Time 4K Super-Resolution of Compressed AVIF Images. AIS 2024 Challenge Survey	5838
<i>Marcos V. Conde (University of Würzburg), Zhijun Lei (Meta), Wen Li (Meta), Ioannis Katsavounidis (Meta), Radu Timofte (University of Würzburg), Min Yan, Xin Liu, Qian Wang, Xiaoqian Ye, Zhan Du, Tiansen Zhang, Zhiyuan Li, Hao Wei, Chenyang Ge, Jiangtao Lv, Long Sun, Jinshian Pan, Jiangxin Dong, Jinhui Tang, Menghan Zhou, Yiqiang Yan, Kihwan Yoon, Ganzorig Gankhuyag, Jae-Hyeon Lee, Ui-Jin Choi, Hyeon-Cheol Moon, Tae-hyun Jeong, Yoonmo Yang, Jae-Gon Kim, Jinwoo Jeong, Sunjei Kim, Xintao Qiu, Yuanbo Zhou, Kongxian Wu, Xinwei Dai, Hui Tang, Wei Deng, Qingquan Gao, Tong Tong, Long Peng, Jiaming Guo, Xin Di, Bohao Liao, Zhibo Du, Peize Xia, Renjing Pei, Yang Wang, Yang Cao, Zhengjun Zha, Bingnan Han, Hongyuan Yu, Zhuoyuan Wu, Cheng Wan, Yuqing Liu, Haodong Yu, Jizhe Li, Zhijuan Huang, Yuan Huang, Yajun Zou, Xianyu Guan, Qi Jia, Heng Zhang, Xuanwu Yin, Kunlong Zuo, Dongyang Zhang, Tianle Liu, Huaian Chen, and Yi Jin</i>	

4th Mobile AI Workshop and Challenges

Scaling Graph Convolutions for Mobile Vision	5857
<i>William Avery (The University of Texas at Austin), Mustafa Munir (The University of Texas at Austin), and Radu Marculescu (The University of Texas at Austin)</i>	
End-to-End Neural Network Compression via ℓ_1/ℓ_2 Regularized Latency Surrogates	5866
<i>Anshul Nasery (University of Washington), Hardik Shah (ETH, Zurich), Arun Sai Suggala (Google Research, India), and Prateek Jain (Google Research, India)</i>	
CoDISP: Exploring Compressed Domain Camera ISP with RGB-guided Encoder	5878
<i>Molin Zhang (Massachusetts Institute of Technology, USA), Soumendu Majee (Samsung Research America, USA), Chengyu Wang (Samsung Research America, USA), Seok-Jun Lee (Samsung Research America, USA), and Hamid Sheikh (Samsung Research America, USA)</i>	
Efficient Skeleton-Based Action Recognition for Real-Time Embedded Systems	5889
<i>Nadhira Noor (Inha University), Fabianaugie Jametoni (Inha University), Jinbeom Kim (Finedigital Inc.), Hyunsu Hong (Finedigital Inc.), and In Kyu Park (Inha University)</i>	

9th New Trends in Image Restoration and Enhancement Workshop and Challenges

S3R-Net: A Single-Stage Approach to Self-Supervised Shadow Removal	5898
<i>Nikolina Kubiak (University of Surrey, UK), Armin Mustafa (University of Surrey, UK), Graeme Phillipson (BBC R&D, UK), Stephen Jolly (BBC R&D, UK), and Simon Hadfield (University of Surrey, UK)</i>	
DCDR-UNet: Deformable Convolution Based Detail Restoration via U-shape Network for Single Image HDR Reconstruction	5909
<i>Joonsoo Kim (Samsung Research America, USA), Zhe Zhu (Samsung Research America, USA), Tien Bau (Samsung Research America, USA), and Chenguang Liu (Samsung Research America, USA)</i>	

Image Restoration Refinement with Uformer GAN	5919
<i>Xu Ouyang (Illinois Institute of Technology), Ying Chen (Illinois Institute of Technology), Kaiyue Zhu (Illinois Institute of Technology), and Gady Agam (Illinois Institute of Technology)</i>	
Towards Real-world Video Face Restoration: A New Benchmark	5929
<i>Ziyan Chen (Shanghai AI Laboratory, China), Jingwen He (The Chinese University of Hong Kong, China), Xinqi Lin (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China), Yu Qiao (Shanghai AI Laboratory, China), and Chao Dong (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences)</i>	
Burst Image Super-Resolution with Base Frame Selection	5940
<i>Sanghyun Kim (POSTECH), Minjung Lee (POSTECH), Woohyeok Kim (POSTECH), Deunsol Jung (POSTECH), Jaesung Rim (POSTECH), Sunghyun Cho (POSTECH), and Minsu Cho (POSTECH)</i>	
Multi-scale Attention Network for Single Image Super-Resolution	5950
<i>Yan Wang (Nankai University, China), Yusen Li (Nankai University, China), Gang Wang (Nankai University, China), and Xiaoguang Liu (Nankai University, China)</i>	
Semantic Pre-supplement for Exposure Correction	5961
<i>Zhen Zou (University of Science and Technology of China, China), Wei Yu (University of Science and Technology of China, China), Jie Huang (University of Science and Technology of China, China), and Feng Zhao (University of Science and Technology of China, Chain)</i>	
Unsupervised Image Prior via Prompt Learning and CLIP Semantic Guidance for Low-Light Image Enhancement	5971
<i>Igor Morawski (National Taiwan University, Taiwan), Kai He (Qualcomm Inc., USA), Shusil Dangi (Qualcomm Inc., USA), and Winston H. Hsu (National Taiwan University; Mobile Drive Technology, Taiwan)</i>	
Diffusion-Based Adaptation for Classification of Unknown Degraded Images	5982
<i>Dinesh Daultani (Tokyo Institute of Technology), Masayuki Tanaka (Tokyo Institute of Technology), Masatoshi Okutomi (Tokyo Institute of Technology), and Kazuki Endo (Teikyo Heisei University)</i>	
Reciprocal Attention Mixing Transformer for Lightweight Image Restoration	5992
<i>Haram Choi (RippleAI), Cheolwoong Na (Sogang University), Jihyeon Oh (Sogang University), Seungjae Lee (Sogang University), Jinseop Kim (Sogang University), Subeen Choe (Sogang University), Jeongmin Lee (LG Innotek), Taehoon Kim (LG AI Research), and Jihoon Yang (Sogang University)</i>	
Audio-Visual Speech Representation Expert for Enhanced Talking Face Video Generation and Evaluation	6003
<i>Dogucan Yaman (Karlsruhe Institute of Technology), Fevziye Irem Eyiokur (Karlsruhe Institute of Technology), Leonard Bärman (Karlsruhe Institute of Technology), Seymanur Akti (Karlsruhe Institute of Technology), Hazım Kemal Ekenel (Istanbul Technical University), and Alexander Waibel (Karlsruhe Institute of Technology)</i>	

Fourier Prior-Based Two-Stage Architecture for Image Restoration	6014
<i>Hemkant Nehete (Indian Institute of Technology Roorkee, India), Amit Monga (Indian Institute of Technology Roorkee, India), Partha Kaushik (Indian Institute of Technology Roorkee, India), and Brajesh Kumar Kaushik (Indian Institute of Technology Roorkee, India)</i>	
Sketch-guided Image Inpainting with Partial Discrete Diffusion Process	6024
<i>Nakul Sharma (IIT Jodhpur), Aditay Tripathi (IISc, Bengaluru), Anirban Chakraborty (IISc, Bengaluru), and Anand Mishra (IIT Jodhpur)</i>	
Towards Online Real-Time Memory-based Video Inpainting Transformers	6035
<i>Guillaume Thiry (ETH Zurich), Hao Tang (ETH Zurich), Radu Timofte (University of Wurzburg; ETH Zurich), and Luc Van Gool (ETH Zurich)</i>	
Shadow Removal based on Diffusion, Segmentation and Super-resolution Models	6045
<i>Chenghua Li (Institute of Automation, Chinese Academy of Sciences (CASIA), China), Bo Yang (Nanjing University of Information Science and Technology, China), Zhiqi Wu (Xi'an Jiaotong-Liverpool University, China), Gao Chen (University of Wollongong, Australia), Yihan Yu (High School Affiliated to Nanjing Normal University Jiangning Campus (NSFJ-CN), China), and Shengxiao Zhou (Nanjing University of Information Science and Technology, China)</i>	
Hybrid Cross-View Attention Network for Lightweight Stereo Image Super-Resolution	6055
<i>Yuqiang Yang (Xiaomi Inc., China), Zhiming Zhang (Xiaomi Inc., China), Yao Du (Xiaomi Inc., China), Jingjing Yang (Xiaomi Inc., China), Long Bao (Xiaomi Inc., China), and Heng Sun (Xiaomi Inc., China)</i>	
IDENet: Implicit Degradation Estimation Network for Efficient Blind Super Resolution	6065
<i>Asif Hussain Khan (University of Udine Italy), Christian Micheloni (University of Udine Italy), and Niki Martinel (University of Udine Italy)</i>	
ISSR-DIL: Image Specific Super-Resolution Using Deep Identity Learning	6076
<i>Sree Rama Vamsidhar S (Indian Institute of Technology (IIT) Tirupati), Jayadeep D (Indian Institute of Technology (IIT) Tirupati), and Rama Krishna Gorthi (Indian Institute of Technology (IIT) Tirupati)</i>	
CRNet: A Detail-Preserving Network for Unified Image Restoration and Enhancement Task	6086
<i>Kangzhen Yang (Northwestern Polytechnical University), Tao Hu (Northwestern Polytechnical University), Kexin Dai (Northwestern Polytechnical University), Genggeng Chen (Xi'an University of Architecture and Technology), Yu Cao (Xi'an Institute of Optics and Precision Mechanics of CAS), Wei Dong (Xi'an University of Architecture and Technology), Peng Wu (Northwestern Polytechnical University), Yanning Zhang (Northwestern Polytechnical University), and Qingsen Yan (Northwestern Polytechnical University)</i>	

Bracketing Image Restoration and Enhancement with High-Low Frequency Decomposition	6097
<i>Genggeng Chen (Xi'an University of Architecture and Technology), Kexin Dai (Northwestern Polytechnical University), Kangzhen Yang (Northwestern Polytechnical University), Tao Hu (Northwestern Polytechnical University), Xiangyu Chen (University of Macau), Yongqing Yang (Xi'an Institute of Optics and Precision Mechanics of CAS), Wei Dong (Xi'an University of Architecture and Technology), Peng Wu (Northwestern Polytechnical University), Yanning Zhang (Northwestern Polytechnical University), and Qingsen Yan (Northwestern Polytechnical University)</i>	

NTIRE 2024 Challenge on Image Super-Resolution (×4): Methods and Results	6108
--	------

Zheng Chen (Shanghai Jiao Tong University, China), Zongwei Wu (University of Wurzburg, Germany), Eduard Zamfir (University of Wurzburg, Germany), Kai Zhang (ETH, Zurich), Yulun Zhang (Shanghai Jiao Tong University, China), Radu Timofte (University of Wurzburg, Germany), Xiaokang Yang (Shanghai Jiao Tong University, China), Hongyuan Yu (Xiaomi Inc), Cheng Wan (Georgia Institute of Technology), Yuxin Hong (Lanzhou University), Zhijuan Huang (Xiaomi Inc), Yajun Zou (Xiaomi Inc), Yuan Huang (Xiaomi Inc), Jiamin Lin (Xiaomi Inc), Bingnan Han (Xiaomi Inc), Xianyu Guan (Xiaomi Inc), Yongsheng Yu (University of Rochester), Daoan Zhang (University of Rochester), Xuanwu Yin (Xiaomi Inc), Kunlong Zuo (Xiaomi Inc), Jinhua Hao (Kuaishou Technology), Kai Zhao (Kuaishou Technology), Kun Yuan (Kuaishou Technology), Ming Sun (Kuaishou Technology), Chao Zhou (Kuaishou Technology), Hongyu An (University of Chinese Academy of Sciences, China), Xinfeng Zhang (University of Chinese Academy of Sciences, China), Zhiyuan Song (Sun Yat-sen University, China), Ziyue Dong (Xi'an Jiaotong University, China), Qing Zhao (Sun Yat-sen University, China), Xiaogang Xu (Zhejiang University, China), Pengxu Wei (Sun Yat-sen University, China), Zhi-chao Dou (Shandong University of Science and Technology Qingdao, China), Gui-ling Wang (Shandong University of Science and Technology Qingdao, China), Chih-Chung Hsu (National Cheng Kung University, Taiwan), Chia-Ming Lee (National Cheng Kung University, Taiwan), Yi-Shiuan Chou (National Cheng Kung University, Taiwan), Cansu Korkmaz (Koc University), A. Murat Tekalp (Koc University), Yubin Wei (Xiamen University), Xiaole Yan (Xiamen University), Binren Li (Xiamen University), Haonan Chen (Xiamen University), Siqi Zhang (Xiamen University), Sihan Chen (Xiamen University), Amogh Joshi (KLE Technological University), Nikhil Akalwadi (KLE Technological University), Sampada Malagi (KLE Technological University), Palani Yashaswini (KLE Technological University), Chaitra Desai (KLE Technological University), Ramesh Ashok Tabib (KLE Technological University), Ujwala Patil (KLE Technological University), Uma Mudenagudi (KLE Technological University), Anjali Sarvaiya (Sardar Vallabhbhai National Institute of Technology, India), Pooja Choksy (Sardar Vallabhbhai National Institute of Technology, India), Jagrit Joshi (Sardar Vallabhbhai National Institute of Technology, India), Shubh Kawa (Sardar Vallabhbhai National Institute of Technology, India), Kishor Upla (Sardar Vallabhbhai National Institute of Technology, India), Sushrut Patwardhan (Norwegian University of Science and Technology, Norway), Raghavendra Ramachandra (Norwegian University of Science and Technology, Norway), Sadat Hossain (DeltaX, Seoul, South Korea), Geongi Park (DeltaX, Seoul, South Korea), S.M. Nadim Uddin (DeltaX, Seoul, South Korea), Hao Xu (McMaster University Hamilton, ON, Canada), Yanhui Guo (McMaster University Hamilton, ON, Canada), Aman Urumbekov (Kyrgyz State Technical University(KSTU), Kyrgyzstan), Xingzhuo Yan (Bosch Investment Ltd.), Wei Hao (Bosch Investment Ltd.), Minghan Fu (Bosch Investment Ltd.), Isaac Orais (Santa Clara University, Santa Clara, California, USA), Samuel Smith (Santa Clara University, Santa Clara, California, USA), Ying Liu (Santa Clara University, Santa Clara, California, USA), Wangwang Jia (National

University of Defense Technology, China), Qisheng Xu (National University of Defense Technology, China), Kele Xu (National University of Defense Technology, China), Weijun Yuan (Jinan University, China), Zhan Li (Jinan University, China), Wenqin Kuang (Jinan University, China), Ruijin Guan (Jinan University, China), Rufing Deng (Jinan University, China), Zhao Zhang (Hefei University of Technology, China), Bo Wang (Hefei University of Technology, China), Suiyi Zhao (Hefei University of Technology, China), Yan Luo (Hefei University of Technology, China), Yanyan Wei (Hefei University of Technology, China), Asif Hussain Khan (University of Udine, Italy), Christian Micheloni (University of Udine, Italy), and Niki Martinel (University of Udine, Italy)

DRCT: Saving Image Super-Resolution Away from Information Bottleneck 6133
 Chih-Chung Hsu (National Cheng Kung University, Taiwan), Chia-Ming Lee (National Cheng Kung University, Taiwan), and Yi-Shiuan Chou (National Cheng Kung University, Taiwan)

DiffLight: Integrating Content and Detail for Low-light Image Enhancement 6143
 Yixu Feng (Northwestern Polytechnical University), Shuo Hou (Northwestern Polytechnical University), Haotian Lin (Northwestern Polytechnical University), Yu Zhu (Northwestern Polytechnical University), Peng Wu (Northwestern Polytechnical University), Wei Dong (Xi'an University of Architecture and Technology), Jinqiu Sun (Northwestern Polytechnical University), Qingsen Yan (Northwestern Polytechnical University), and Yanning Zhang (Northwestern Polytechnical University)

NTIRE 2024 Challenge on Bracketing Image Restoration and Enhancement: Datasets, Methods and Results	6153
<p>Zhilu Zhang (Harbin Institute of Technology, China), Shuohao Zhang (Harbin Institute of Technology, China), Renlong Wu (Harbin Institute of Technology, China), Wangmeng Zuo (Harbin Institute of Technology, China), Radu Timofte (University of Wurzburg & ETH Zurich), Xiaoxia Xing (Samsung Research China - Beijing), Hyunhee Park (Samsung Electronics), Sejun Song (Samsung Electronics), Changho Kim (Samsung Electronics), Xiangyu Kong (Samsung Research China - Beijing), Jinlong Wu (Samsung Research China - Beijing), Jianxing Zhang (Samsung Research China - Beijing), Jingfan Tan (Samsung Research China - Beijing), Zikun Liu (Samsung Research China - Beijing), Wenhan Luo (Sun Yat-sen University), Wenjie Lin (Megvii Technology), Chengzhi Jiang (Megvii Technology), Mingyan Han (Megvii Technology), Zhen Liu (Megvii Technology), Ting Jiang (Megvii Technology), Jinting Luo (Megvii Technology), Shen Cheng (Megvii Technology), Linze Li (Megvii Technology), Xinhao Niu (Megvii Technology), Shuaicheng Liu (Megvii Technology), Kexin Dai (Northwestern Polytechnical University, Xi'an, China), Kangzhen Yang (Northwestern Polytechnical University, Xi'an, China), Tao Hu (Northwestern Polytechnical University, Xi'an, China), Xiangyu Chen (Shenzhen Institute of Advanced Technology), Yu Cao (Xi'an University of Architecture and Technology), Qingsen Yan (Northwestern Polytechnical University, Xi'an, China), Yanning Zhang (Northwestern Polytechnical University, Xi'an, China), Genggen Chen (Northwestern Polytechnical University, Xi'an, China), Yongqing Yang (Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences), Wei Dong (Xi'an University of Architecture and Technology), Xinwei Dai (Fuzhou University), Yuanbo Zhou (Fuzhou University), Xintao Qiu (Fuzhou University), Hui Tang (Fuzhou University), Wei Deng (Imperial Vision Technology), Qingquan Gao (Fuzhou University), Tong Tong (Fuzhou University), Peng Zhang (Mi Lab), Yifei Chen (Mi Lab), Wenbo Xiong (Mi Lab), Zhijun Song (Mi Lab), Pu Cheng (Mi Lab), Taolue Feng (Mi Lab), Yunqing He (Mi Lab), Daiguo Zhou (Mi Lab), Ying Huang (Mi Lab), Xiaowen Ma (Northwestern Polytechnical University, Xi'an, China), and Peng Wu (Northwestern Polytechnical University, Xi'an, China)</p>	
PCQA: A Strong Baseline for AIGC Quality Assessment Based on Prompt Condition	6167
<p>Xi Fang (DP Technology), Weigang Wang (Cisco), Xiaoxin Lv (Shopee), and Jun Yan (Tonji University)</p>	
Virtually Enriched NYU Depth V2 Dataset for Monocular Depth Estimation: Do We Need Artificial Augmentation?	6177
<p>Dmitry Ignatov (University of Würzburg, Germany), Andrey Ignatov (ETH Zurich, Switzerland), and Radu Timofte (University of Würzburg, Germany)</p>	
BigEPIT: Scaling EPIT for Light Field Image Super-Resolution	6187
<p>Wentao Chao (Beijing Normal University, China), Yiming Kan (Beijing Normal University, China), Xuechun Wang (Beijing Normal University, China), Fuqing Duan (Beijing Normal University, China), and Guanghui Wang (Toronto Metropolitan University, Canada)</p>	

NTIRE 2024 Challenge on Stereo Image Super-Resolution: Methods and Results	6198
<p>Longguang Wang (Aviation University of Air Force), Yulan Guo (The Shenzhen Campus of Sun Yat-sen University), Juncheng Li (Shanghai University), Hongda Liu (Sun Yat-sen University), Yang Zhao (Shanghai University), Yingqian Wang (National University of Defense Technology), Zhi Jin (Sun Yat-sen University), Shuhang Gu (University of Electronic Science and Technology), and Radu Timofte (University of Wurzburg, Germany)</p>	
ShadowRefiner: Towards Mask-free Shadow Removal via Fast Fourier Transformer	6208
<p>Wei Dong (McMaster University), Han Zhou (McMaster University), Yuqiong Tian (McMaster University), Jingke Sun (McMaster University), Xiaohong Liu (Shanghai Jiao Tong University), Guangtao Zhai (Shanghai Jiao Tong University), and Jun Chen (McMaster University)</p>	
NTIRE 2024 Challenge on Light Field Image Super-Resolution: Methods and Results	6218
<p>Yingqian Wang (National University of Defense Technology), Zhengyu Liang (National University of Defense Technology), Qianyu Chen (National University of Defense Technology), Longguang Wang (Aviation University of Air Force), Jungang Yang (National University of Defense Technology), Radu Timofte (University of Wurzburg & ETH Zurich), Yulan Guo (National University of Defense Technology), Wentao Chao (Beijing Normal University), Yiming Kan (Beijing Normal University), Xuechun Wang (Beijing Normal University), Fuqing Duan (Beijing Normal University), Guanghui Wang (Toronto Metropolitan University), Wang Xia (Beijing Institute of Technology; Shenzhen MSU-BIT University), Ziqi Wang (Beijing Institute of Technology; Shenzhen MSU-BIT University), Yue Yan (Beijing Institute of Technology; Shenzhen MSU-BIT University), Peiqi Xia (Beijing Institute of Technology; Shenzhen MSU-BIT University), Shunzhou Wang (Peking University Shenzhen Graduate School), Yao Lu (Beijing Institute of Technology; Shenzhen MSU-BIT University), Angulia Yang (Bigo Technology Pte. Ltd.), Kai Jin (Bigo Technology Pte. Ltd.), Zeqiang Wei (Beijing University of Posts and Telecommunications), Sha Guo (Institute of Digital Media, Peking University), Mingzhi Gao (Bigo Technology Pte. Ltd.), Xiuzhuang Zhou (Beijing University of Posts and Telecommunications), Zhongxin Yu (Fujian Normal University), Shaofei Luo (Fujian Normal University), Cheng Zhong (Fujian Normal University), Shaorui Chen (Fujian Normal University), Long Peng (University of Science and Technology of China), Yuhong He (Northeastern University), Gaosheng Liu (Tianjin University), Huanjing Yue (Tianjin University), Jingyu Yang (Tianjin University), Zhengjian Yao (Peking University), Jiakui Hu (Peking University), Lujia Jin (Peking University), Zhi-Song Liu (Lappeenranta–Lahti University of Technology LUT), Chenhang He (The Hong Kong Polytechnic University), Jun Xiao (The Hong Kong Polytechnic University), Xiuyuan Wang (The Hong Kong Polytechnic University), Zonglin Tian (Beijing Normal University), Yifan Mao (Anqing Normal University), Deyang Liu (Anqing Normal University), Shizheng Li (Anqing Normal University), and Ping An (Shanghai University)</p>	

Learnable Global Spatio-Temporal Adaptive Aggregation for Bracketing Image Restoration and Enhancement	6235
<i>Xinwei Dai (Fuzhou University, China), Yuanbo Zhou (Fuzhou University, China), Xintao Qiu (Fuzhou University, China), Hui Tang (Fuzhou University, China), Wei Deng (Imperial Vision Technology, China), Qinquan Gao (Fuzhou University, China; Imperial Vision Technology, China), and Tong Tong (Fuzhou University, China; Imperial Vision Technology, China)</i>	
Swift Parameter-free Attention Network for Efficient Super-Resolution	6246
<i>Cheng Wan (Georgia Institute of Technology), Hongyuan Yu (Xiaomi Inc), Zhiqi Li (Georgia Institute of Technology), Yihang Chen (Georgia Institute of Technology), Yajun Zou (Xiaomi Inc), Yuqing Liu (Xiaomi Inc), Xuanwu Yin (Xiaomi Inc), and Kunlong Zuo (Xiaomi Inc)</i>	
HMANet: Hybrid Multi-Axis Aggregation Network for Image Super-Resolution	6257
<i>Shu-Chuan Chu (Shandong University of Science and Technology), Zhi-Chao Dou (Shandong University of Science and Technology), Jeng-Shyang Pan (Nanjing University of Information Science and Technology), Shaowei Weng (Guangdong University of Technology), and Junbao Li (Harbin Institute of Technology)</i>	
High Quality Reference Feature for Two Stage Bracketing Image Restoration and Enhancement ..	6267
<i>Xiaoxia Xing (Samsung R&D Institute China-Beijing), HyunHee Park (Samsung Electronics), Fan Wang (Samsung R&D Institute China-Beijing), Ying Zhang (Samsung R&D Institute China-Beijing), Sejun Song (Samsung Electronics), Changho Kim (Samsung Electronics), and Xiangyu Kong (Samsung R&D Institute China-Beijing)</i>	
Efficient Light Field Image Super-Resolution via Progressive Disentangling	6277
<i>Gaosheng Liu (Tianjin University, China), Huanjing Yue (Tianjin University, China), and Jingyu Yang (Tianjin University, China)</i>	
Attention Guidance Distillation Network for Efficient Image Super-Resolution	6287
<i>Hongyuan Wang (Xinjiang University, China), Ziyan Wei (Xinjiang University, China), Qingting Tang (Xinjiang University, China), Shuli Cheng (Xinjiang University, China), Liejun Wang (Xinjiang University, China), and Yongming Li (Xinjiang University, China)</i>	
Short-form UGC Video Quality Assessment Based on Multi-Level Video Fusion with Rank-Aware	6297
<i>Haoran Xu (Zhejiang University), Mengduo Yang (Zhejiang University), Jie Zhou (Zhejiang University), and Jiaze Li (Zhejiang University)</i>	
Shadow Removal via Global Residual Free Unet and Shadow Generation	6307
<i>Dong Li (University of Science and Technology of China), Xin Lu (University of Science and Technology of China), Yurui Zhu (University of Science and Technology of China), Xi Wang (University of Science and Technology of China), Jie Xiao (University of Science and Technology of China), Yunpeng Zhang (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>	

Large Kernel Frequency-enhanced Network for Efficient Single Image Super-Resolution	6317
<i>Jiadi Chen (Zhejiang Normal University, China), Chunjiang Duanmu (Zhejiang Normal University, China), and Huanhuan Long (Zhejiang Normal University, China)</i>	
AIGIQA-20K: A Large Database for AI-Generated Image Quality Assessment	6327
<i>Chunyi Li (Shanghai Jiao Tong University), Tengchuan Kou (Shanghai Jiao Tong University), Yixuan Gao (Shanghai Jiao Tong University), Yuqin Cao (Shanghai Jiao Tong University), Wei Sun (Shanghai Jiao Tong University), Zicheng Zhang (Shanghai Jiao Tong University), Yingjie Zhou (Shanghai Jiao Tong University), Zhichao Zhang (Shanghai Jiao Tong University), Weixia Zhang (Shanghai Jiao Tong University), Haoning Wu (Nanyang Technological University), Xiaohong Liu (Shanghai Jiao Tong University), Xiongkuo Min (Shanghai Jiao Tong University), and Guangtao Zhai (Shanghai Jiao Tong University)</i>	

NTIRE 2024 Quality Assessment of AI-Generated Content Challenge	6337
---	------

Xiaohong Liu (Shanghai Jiao Tong University, China), Xiongkuo Min (Shanghai Jiao Tong University, China), Guangtao Zhai (Shanghai Jiao Tong University, China), Chunyi Li (Shanghai Jiao Tong University, China), Tengchuan Kou (Shanghai Jiao Tong University, China), Wei Sun (Shanghai Jiao Tong University, China), Haoning Wu (Nanyang Technological University, Singapore), Yixuan Gao (Shanghai Jiao Tong University, China), Yuqin Cao (Shanghai Jiao Tong University, China), Zicheng Zhang (Shanghai Jiao Tong University, China), Xiele Wu (Shanghai Jiao Tong University, China), Radu Timofte (University of Wurzburg, Germany; ETH Zurich, Switzerland), Fei Peng (Beijing University of Posts and Telecommunications, China), Huiyuan Fu (Beijing University of Posts and Telecommunications, China), Anlong Ming (Beijing University of Posts and Telecommunications, China), Chuanming Wang (Beijing University of Posts and Telecommunications, China), Huadong Ma (Beijing University of Posts and Telecommunications, China), Shuai He (Beijing University of Posts and Telecommunications, China), Zifei Dou (Beijing Xiaomi Mobile Software Co., Ltd.), Shu Chen (Beijing Xiaomi Mobile Software Co., Ltd.), Huacong Zhang (Sun Yat-sen University), Haiyi Xie (Sun Yat-sen University), Chengwei Wang (Sun Yat-sen University), Baoying Chen (Alibaba Group), Jishen Zeng (Alibaba Group), Jianquan Yang (Sun Yat-sen University), Weigang Wang (Cisco Systems, Inc.), Xi Fang (DP Technology, Ltd.), Xiaoxin Lv (Shopee Pte. Ltd.), Jun Yan (Tongji University), Tianwu Zhi (Bytedance Inc.), Yabin Zhang (Bytedance Inc.), Yaohui Li (Bytedance Inc.), Yang Li (Bytedance Inc.), Jingwen Xu (Bytedance Inc.), Jianzhao Liu (Bytedance Inc.), Yiting Liao (Bytedance Inc.), Junlin Li (Bytedance Inc.), Zihao Yu (University of Science and Technology of China), Fengbin Guan (University of Science and Technology of China), Yiting Lu (University of Science and Technology of China), Xin Li (University of Science and Technology of China), Hossein Motamednia (Institute for Research in Fundamental Sciences, Tehran, Iran), S. Farhad Hosseini-Benvidi (Kharazmi University, Tehran, Iran), Ahmad Mahmoudi-Aznavah (Shahid Beheshti University, Tehran, Iran), Azadeh Mansouri (Kharazmi University, Tehran, Iran), Ganzorig Gankhuyag (Korea Electronics Technology Institute), Kihwan Yoon (Korea Electronics Technology Institute), Yifang Xu (ByteDance), Haotian Fan (ByteDance), Fangyuan Kong (ByteDance), Shiling Zhao (Hangzhou Dianzi University), Weifeng Dong (Hangzhou Dianzi University), Haibing Yin (Hangzhou Dianzi University), Li Zhu (Sankuai), Zhiling Wang (Sankuai), Bingchen Huang (Sankuai), Avinab Saha (University of Texas at Austin), Sandeep Mishra (University of Texas at Austin), Shashank Gupta (University of Texas at Austin), Rajesh Suresh (University of Texas at Austin), Oindrila Saha (University of Massachusetts Amherst), Luigi Celona (University of Milano - Bicocca), Simone Bianco (University of Milano - Bicocca), Paolo Napoletano (University of Milano - Bicocca), Raimondo Schettini (University of Milano - Bicocca), Junfeng Yang (Xiangjiang Laboratory and Hunan University of Technology and Business), Jing Fu (Xiangjiang Laboratory and Hunan University of Technology and Business), Wei Zhang (Xiangjiang Laboratory and Hunan University of Technology and Business), Wenzhi Cao (Xiangjiang Laboratory and Hunan University of

Technology and Business), Limei Liu (Xiangjiang Laboratory and Hunan University of Technology and Business), Han Peng (Xiangjiang Laboratory and Hunan University of Technology and Business), Weijun Yuan (Jinan University), Zhan Li (Jinan University), Yihang Cheng (Jinan University), Yifan Deng (Jinan University), Haohui Li (Peking University), Bowen Qu (Peking University), Yao Li (Peking University), Shuqing Luo (Peking University), Shunzhou Wang (Peking University), Wei Gao (Peking University), Zihao Lu (University of Wurzburg), Marcos V. Conde (University of Wurzburg), Radu Timofte (University of Wurzburg), Xinrui Wang (University of Science and Technology of China), Zhibo Chen (University of Science and Technology of China), Ruling Liao (University of Science and Technology of China), Yan Ye (University of Science and Technology of China), Qiulin Wang (Kuaishou Technology), Bing Li (University of Science and Technology of China), Zhaokun Zhou (Peking University), Miao Geng (Kuaishou Technology), Rui Chen (Kuaishou Technology), Xin Tao (Kuaishou Technology), Xiaoyu Liang (Peking University), Shangkun Sun (Peking University), Xingyuan Ma (Beijing University of Posts and Telecommunications), Jiaze Li (Zhejiang University), Mengduo Yang (Zhejiang University), Haoran Xu (Zhejiang University), Jie Zhou (Zhejiang University), Shiding Zhu (Zhejiang University), Bohan Yu (Zhejiang University), Pengfei Chen (Xidian University), Xinrui Xu (Xidian University), Jiabin Shen (Xidian University), Zhichao Duan (Xidian University), Erfan Asadi (Kharazmi University), Jiahe Liu (University of British Columbia), Qi Yan (University of British Columbia), Youran Qu (Peking University), Xiaohui Zeng (University of Toronto), Lele Wang (University of British Columbia), and Renjie Liao (University of British Columbia)

Dformer: Learning Efficient Image Restoration with Perceptual Guidance	6363
Nodirkhujja Khudjaev (Opt-AI Inc., LG Sciencepark, Seoul, South Korea), Roman Tsoy (Opt-AI Inc., LG Sciencepark, Seoul, South Korea), S M A Sharif (Opt-AI Inc., LG Sciencepark, Seoul, South Korea), Azamat Myrzabekov (Opt-AI Inc., LG Sciencepark, Seoul, South Korea), Seongwan Kim (Opt-AI Inc., LG Sciencepark, Seoul, South Korea), and Jaeho Lee (Opt-AI Inc., LG Sciencepark, Seoul, South Korea)	
Learning Optimized Low-Light Image Enhancement for Edge Vision Tasks	6373
S M A Sharif (Opt-AI Inc., LG Sciencepark, Seoul, South Korea), Azamat Myrzabekov (Opt-AI Inc., LG Sciencepark, Seoul, South Korea), Nodirkhujja Khudjaev (Opt-AI Inc., LG Sciencepark, Seoul, South Korea), Roman Tsoy (Opt-AI Inc., LG Sciencepark, Seoul, South Korea), Seongwan Kim (Opt-AI Inc., LG Sciencepark, Seoul, South Korea), and Jaeho Lee (Opt-AI Inc., LG Sciencepark, Seoul, South Korea)	
AIGC-VQA: A Holistic Perception Metric for AIGC Video Quality Assessment	6384
Yiting Lu (University of Science and Technology of China), Xin Li (University of Science and Technology of China), Bingchen Li (University of Science and Technology of China), Zihao Yu (University of Science and Technology of China), Fengbin Guan (University of Science and Technology of China), Xinrui Wang (University of Science and Technology of China), Ruling Liao (Alibaba Group), Yan Ye (Alibaba Group), and Zhibo Chen (University of Science and Technology of China)	

MoE-AGIQA: Mixture-of-Experts Boosted Visual Perception-Driven and Semantic-Aware Quality Assessment for AI-Generated Images	6395
<i>Junfeng Yang (Hunan University of Technology and Business, China), Jing Fu (Hunan University of Technology and Business, China), Wei Zhang (ByteDance, China), Wenzhi Cao (Hunan University of Technology and Business, China), Limei Liu (Hunan University of Technology and Business, China), and Han Peng (Hunan University of Technology and Business, China)</i>	
DehazeDCT: Towards Effective Non-Homogeneous Dehazing via Deformable Convolutional Transformer	6405
<i>Wei Dong (McMaster University), Han Zhou (McMaster University), Ruiyi Wang (Shanghai Jiao Tong University), Xiaohong Liu (Shanghai Jiao Tong University), Guangtao Zhai (Shanghai Jiao Tong University), and Jun Chen (McMaster University)</i>	

Xin Li (University of Science and Technology of China, China), Kun Yuan (Kuaishou Technology, China), Yajing Pei (University of Science and Technology of China, China), Yiting Lu (University of Science and Technology of China, China), Ming Sun (Kuaishou Technology, China), Chao Zhou (Kuaishou Technology, China), Zhibo Chen (University of Science and Technology of China, China), Radu Timofte (University of Wurzburg & ETH Zurich), Wei Sun (Shanghai Jiao Tong University), Haoning Wu (Nanyang Technological University), Zicheng Zhang (Shanghai Jiao Tong University), Jun Jia (Shanghai Jiao Tong University), Zhichao Zhang (Shanghai Jiao Tong University), Linhan Cao (Shanghai Jiao Tong University), Qiubo Chen (Nanyang Technological University), Xionguo Min (Shanghai Jiao Tong University), Weisi Lin (Nanyang Technological University), Guangtao Zhai (Shanghai Jiao Tong University), Jianhui Sun (WeChat), Tianyi Wang (WeChat), Lei Li (WeChat), Han Kong (WeChat), Wenxuan Wang (WeChat), Bing Li (WeChat), Cheng Luo (WeChat), Haiqiang Wang (Tencent), Xiangguang Chen (Tencent), Wenhui Meng (Tencent), Xiang Pan (Tencent), Huiying Shi (Wuhan University), Han Zhu (Wuhan University), Xiaozhong Xu (Tencent), Lei Sun (Tencent), Zhenzhong Chen (Wuhan University), Shan Liu (Tencent), Fangyuan Kong (ByteDance, Shenzhen), Haotian Fan (ByteDance, Shenzhen), Yifang Xu (ByteDance, Shenzhen), Haoran Xu (Zhejiang University), Mengduo Yang (Zhejiang University), Jie Zhou (Zhejiang University), Jiaze Li (Zhejiang University), Shijie Wen (Beihang University, Beijing, China), Mai Xu (Beihang University, Beijing, China), Da Li (Southern University of Science and Technology), Shunyu Yao (Harbin Institute of Technology), Jiazhi Du (Harbin Institute of Technology), Wangmeng Zuo (Harbin Institute of Technology), Zhibo Li (Beijing University of Posts and Telecommunications, China), Shuai He (Beijing University of Posts and Telecommunications, China), Anlong Ming (Beijing University of Posts and Telecommunications, China), Huiyuan Fu (Beijing University of Posts and Telecommunications, China), Huadong Ma (Beijing University of Posts and Telecommunications, China), Yong Wu (China Merchants Bank), Fie Xue (China Merchants Bank), Guozhi Zhao (China Merchants Bank), Lina Du (Shandong Jianzhu University), Jie Guo (Shandong Jianzhu University), Yu Zhang (Shandong Jianzhu University), Huimin Zheng (Shandong Jianzhu University), Junhao Chen (Shandong Jianzhu University), Yue Liu (Shandong Jianzhu University), Dulan Zhou (Key Laboratory for Parallel and Distributed Processing, Changsha, China), Kele Xu (Key Laboratory for Parallel and Distributed Processing, Changsha, China), Qisheng Xu (Key Laboratory for Parallel and Distributed Processing, Changsha, China), Tao Sun (Key Laboratory for Parallel and Distributed Processing, Changsha, China), Zhixiang Ding (Institute of Automation, CAS), and Yuhang Hu (University of Chinese Academy of Sciences)

AIGC Image Quality Assessment via Image-Prompt Correspondence	6432
<i>Fei Peng (Beijing University of Posts and Telecommunications, China), Huiyuan Fu (Beijing University of Posts and Telecommunications, China), Anlong Ming (Beijing University of Posts and Telecommunications, China), Chuanming Wang (Beijing University of Posts and Telecommunications, China), Huadong Ma (Beijing University of Posts and Telecommunications, China), Shuai He (Beijing University of Posts and Telecommunications, China), Zifei Dou (Beijing Xiaomi Mobile Software Co., Ltd.), and Shu Chen (Beijing Xiaomi Mobile Software Co., Ltd.)</i>	
PromptCIR: Blind Compressed Image Restoration with Prompt Learning	6442
<i>Bingchen Li (University of Science and Technology of China), Xin Li (University of Science and Technology of China), Yiting Lu (University of Science and Technology of China), Ruoyu Feng (University of Science and Technology of China), Mengxi Guo (Bytedance Inc.), Shijie Zhao (Bytedance Inc.), Li Zhang (Bytedance Inc.), and Zhibo Chen (University of Science and Technology of China)</i>	

NTIRE 2024 Dense and Non-Homogeneous Dehazing Challenge Report	6453
--	------

Codruta O. Ancuti (University Politehnica Timisoara), Cosmin Ancuti (UCL), Florin-Alexandru Vasluianu (Computer Vision Lab, University of Wurzburg), Radu Timofte (University of Wurzburg; ETH Zurich), Yidi Liu (University of Science and Technology of China, Hefei, China), Xingbo Wang (University of Science and Technology of China, Hefei, China), Yurui Zhu (University of Science and Technology of China, Hefei, China), Gege Shi (University of Science and Technology of China, Hefei, China), Xin Lu (University of Science and Technology of China, Hefei, China), Xueyang Fu (University of Science and Technology of China, Hefei, China), Zheng-jun Zha (University of Science and Technology of China, Hefei, China), Wei Dong (McMaster University, Canada), Han Zhou (McMaster University, Canada), Ruiyi Wang (Shanghai Jiao Tong University, China), Xiaohong Liu (Shanghai Jiao Tong University, China), Guangtao Zhai (Shanghai Jiao Tong University, China), Jun Chen (McMaster University, Canada), Wei Song (Guangdong University of Technology, China), Yichang Gao (Guangdong University of Technology, China), Jiahao Xiong (Guangdong University of Technology, China), Hualiang Lin (Guangdong University of Technology, China), Xiang Li (Guangdong University of Technology, China), Dong Li (Guangdong University of Technology, China), Mohab Kishawy (McMaster University), Ruibin Li (McMaster University), Seyed Amirreza Mousavi (McMaster University), Rana Rauf (McMaster University), Yangyi Liu (McMaster University), Huan Liu (McMaster University), MingSheng Tu (Chongqing University of Posts and Telecommunications, China), Kele Xu (National University of Defense Technology, China), JiaWen Chen (Chongqing University of Posts and Telecommunications, China), Qisheng Xu (National University of Defense Technology, China), Tao Sun (National University of Defense Technology, China), Jin Guo (Zhejiang Dahua Technology Co.,Ltd.), Ben Shao (Zhejiang Dahua Technology Co.,Ltd.), Tianli Liu (Zhejiang Dahua Technology Co.,Ltd.), Mohao Wu (Zhejiang Dahua Technology Co.,Ltd.), Xingzhuo Yan (Bosch Investment Ltd.), Minghan Fu (University of Saskatchewan), Lehan Yang (University of Sydney), Xin Lin (Sichuan University), Lu Qi (University of California Merced), Jincen Song (University of Sydney), Xiaoqian Hu (University of New South Wales), Linwai Tao (University of Sydney), Hongming Chen (Shenyang Aerospace University), Xiang Chen (Nanjing University of Science and Technology), Chuanlong Xie (Shenyang Aerospace University), Zhao Zhang (Hefei University of Technology), Junhu Wang (Hefei University of Technology), Yanyan Wei (Hefei University of Technology), Suiyi Zhao (Hefei University of Technology), Shengeng Tang (Hefei University of Technology), Sampada Malagi (KLE Technological University, Hubballi, Karnataka, India), Amogh Joshi (KLE Technological University, Hubballi, Karnataka, India), Nikhil Akalwadi (KLE Technological University, Hubballi, Karnataka, India), Chaitra Desai (KLE Technological University, Hubballi, Karnataka, India), Ramesh Ashok Tabib (KLE Technological University, Hubballi, Karnataka, India), Uma Mudenagudi (KLE Technological University, Hubballi, Karnataka, India), Wenjing Jiang (Hebei University, Baoding, China), Jagadeesh Kalyanshetti (KLE Technological University, Hubballi, Karnataka, India), Vijayalaxmi Ashok Aralikatti (KLE Technological University, Hubballi, Karnataka,

India), Yashaswini P (KLE Technological University, Hubballi, Karnataka, India), Nitish Upasi (KLE Technological University, Hubballi, Karnataka, India), Dikshit Hegde (KLE Technological University, Hubballi, Karnataka, India), Ujwala Patil (KLE Technological University, Hubballi, Karnataka, India), and Sujata C (KLE Technological University, Hubballi, Karnataka, India)

Cross-view Aggregation Network For Stereo Image Super-Resolution	6469
<i>Zhitao Chen (Wuhan Institute of Technology), Tao Lu (Wuhan Institute of Technology), Kanghui Zhao (Wuhan Institute of Technology), Bolin Zhu (Wuhan Institute of Technology), Zhen Li (Wuhan Institute of Technology), Jiaming Wang (Wuhan Institute of Technology), and Yanduo Zhang (Wuhan Institute of Technology)</i>	
RBSFormer: Enhanced Transformer Network for Raw Image Super-Resolution	6479
<i>Siyuan Jiang (University of Science and Technology of China), Senyan Xu (University of Science and Technology of China), and Xingfu Wang (University of Science and Technology of China)</i>	
Multi-Level Feature Fusion Network for Lightweight Stereo Image Super-Resolution	6489
<i>Yunxiang Li (Fuzhou University), Wenbin Zou (South China University of Technology), Qiaomu Wei (Chengdu University of Information Technology), Feng Huang (Fuzhou University), and Jing Wu (Fuzhou University)</i>	

NTIRE 2024 Challenge on HR Depth from Images of Specular and Transparent Surfaces 6499

Pierluigi Zama Ramirez (University of Bologna), Fabio Tosi (University of Bologna), Luigi Di Stefano (University of Bologna), Radu Timofte (Computer Vision Lab, University of Wurzburg, Germany), Alex Costanzino (University of Bologna), Matteo Poggi (University of Bologna), Samuele Salti (University of Bologna), Stefano Mattoccia (University of Bologna), Yangyang Zhang (Xiaomi Inc., China), Cailin Wu (Xiaomi Inc., China), Zhuangda He (Xiaomi Inc., China), Shuangshuang Yin (Xiaomi Inc., China), Jiaxu Dong (Xiaomi Inc., China), Yangchenxu Liu (Xiaomi Inc., China), Hao Jiang (Xiaomi Inc., China), Jun Shi (Samsung R&D Institute China-Beijing), Yong A (Samsung R&D Institute China-Beijing), Yixiang Jin (Samsung R&D Institute China-Beijing), Dingzhe Li (Samsung R&D Institute China-Beijing), Bingxin Ke (Photogrammetry and Remote Sensing, ETH Zürich), Anton Obukhov (Photogrammetry and Remote Sensing, ETH Zürich), Tinafu Wang (Photogrammetry and Remote Sensing, ETH Zürich), Nando Metzger (Photogrammetry and Remote Sensing, ETH Zürich), Shengyu Huang (Photogrammetry and Remote Sensing, ETH Zürich), Konrad Schindler (Photogrammetry and Remote Sensing, ETH Zürich), Yachuan Huang (Huazhong University of Science and Technology, China), Jiaqi Li (Huazhong University of Science and Technology, China), Junrui Zhang (Huazhong University of Science and Technology, China), Yiran Wang (Huazhong University of Science and Technology, China), Zihao Huang (Huazhong University of Science and Technology, China), Tianqi Liu (Huazhong University of Science and Technology, China), Zhiguo Cao (Huazhong University of Science and Technology, China), Pengzhi Li (Tsinghua University), Jui-Lin Wang (Tsinghua University), Wenjie Zhu (National University of Defense Technology, Changsha, China), Hui Geng (National University of Defense Technology, Changsha, China), Yuxin Zhang (National University of Defense Technology, Changsha, China), Long Lan (National University of Defense Technology, Changsha, China), Kele Xu (National University of Defense Technology, Changsha, China), Tao Sun (National University of Defense Technology, Changsha, China), Qisheng Xu (National University of Defense Technology, Changsha, China), Sourav Saini (Indian Institute of Technology Jammu), Aashray Gupta (Indian Institute of Technology Jammu), Sahaj K. Mistry (Indian Institute of Technology Jammu), Aryan Shukla (Indian Institute of Technology Jammu), Vinit Jakheti (Indian Institute of Technology Jammu), Sunil Jaiswal (K \ Lens GmbH), Yuejin Sun (Xiaomi Inc., China), Zhuofan Zheng (Xiaomi Inc., China), Yi Ning (Xiaomi Inc., China), Jen-Hao Cheng (University of Washington), Hou-I Liu (University of Washington, National Yang Ming Chiao Tung University, Carnegie Mellon University, University of Illinois Urbana-Champaign), Hsiang-Wei Huang (University of Washington, National Yang Ming Chiao Tung University, Carnegie Mellon University, University of Illinois Urbana-Champaign), Cheng-Yen Yang (University of Washington, National Yang Ming Chiao Tung University, Carnegie Mellon University, University of Illinois Urbana-Champaign), Zhongyu Jiang (University of Washington, National Yang Ming Chiao Tung University, Carnegie Mellon University, University of Illinois Urbana-Champaign), Yi-Hao Peng (University of Washington, National Yang Ming Chiao Tung University, Carnegie Mellon University, University of Illinois Urbana-Champaign),

Aishi Huang (University of Washington, National Yang Ming Chiao Tung University, Carnegie Mellon University, University of Illinois Urbana-Champaign), and Jenq-Neng Hwang (University of Washington, National Yang Ming Chiao Tung University, Carnegie Mellon University, University of Illinois Urbana-Champaign)

HirFormer: Dynamic High Resolution Transformer for Large-Scale Image Shadow Removal 6513

Xin Lu (University of Science and Technology of China), Yurui Zhu (University of Science and Technology of China), Xi Wang (University of Science and Technology of China), Dong Li (University of Science and Technology of China), Jie Xiao (University of Science and Technology of China), Yunpeng Zhang (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)

NTIRE 2024 Challenge on Blind Enhancement of Compressed Image: Methods and Results	6524
--	------

Ren Yang (ETH Zurich, Switzerland), Radu Timofte (Julius Maximilian University of Würzburg, Germany), Bingchen Li, Xin Li (University of Science and Technology of China, ByteDance Inc.), Mengxi Guo (University of Science and Technology of China, ByteDance Inc.), Shijie Zhao (University of Science and Technology of China, ByteDance Inc.), Li Zhang (University of Science and Technology of China, ByteDance Inc.), Zhibo Chen (University of Science and Technology of China, ByteDance Inc.), Dongyang Zhang (MGTV, Changsha, China), Yash Arora (Amrita Vishwa Vidyapeetham, Kerala, India; York University, Toronto, Canada), Aditya Arora (Amrita Vishwa Vidyapeetham, Kerala, India; York University, Toronto, Canada), Yuanbin Chen (Fuzhou University, Fuzhou, China), Hui Tang (Fuzhou University, Fuzhou, China), Tao Wang (Fuzhou University, Fuzhou, China), Longxuan Zhao (Fuzhou University, Fuzhou, China), Bin Chen (Fuzhou University, Fuzhou, China), Tong Tong (Fuzhou University, Fuzhou, China), Qiao Mo (Kuaishou Technology, UESTC), Jingwei Bao (Kuaishou Technology, UESTC), Jinhua Hao (Kuaishou Technology, UESTC), Yukang Ding (Kuaishou Technology, UESTC), Hantang Li (Kuaishou Technology, UESTC), Ming Sun (Kuaishou Technology, UESTC), Chao Zhou (Kuaishou Technology, UESTC), Shuyuan Zhu (Kuaishou Technology, UESTC), Zhi Jin (Sun Yat-sen University, Shenzhen, China), Wei Wang (Sun Yat-sen University, Shenzhen, China), Dandan Zhan (Sun Yat-sen University, Shenzhen, China), Jiawei Wu (Sun Yat-sen University, Shenzhen, China), Jiahao Wu (Sun Yat-sen University, Shenzhen, China), Luwei Tu (Sun Yat-sen University, Shenzhen, China), Hongyu An (University of Chinese Academy of Sciences, China), Xinfeng Zhang (University of Chinese Academy of Sciences, China), Woon-Ha Yeo (Sahmyook University, Republic of Korea), Wang-Taek Oh (Sahmyook University, Republic of Korea), Young-Il Kim (Sahmyook University, Republic of Korea), Han-Cheol Ryu (Sahmyook University, Republic of Korea), Long Sun (Nanjing University of Science and Technology, Nanjing, China), Mingjun Zhen (Nanjing University of Science and Technology, Nanjing, China), Jinshan Pan (Nanjing University of Science and Technology, Nanjing, China), Jiangxin Dong (Nanjing University of Science and Technology, Nanjing, China), Jinhui Tang (Nanjing University of Science and Technology, Nanjing, China), Yapeng Du (University of Electronic Science and Technology of China, China), Ao Li (University of Electronic Science and Technology of China, China), Ziyang He (University of Electronic Science and Technology of China, China), Lei Luo (University of Electronic Science and Technology of China, China), Ce Zhu (University of Electronic Science and Technology of China, China), Xin Yao (Politecnico di Torino, Italy), Sunder Ali Khowaja (Technological University Dublin (TU Dublin), Ireland), IK Hyun Lee (Tech University of Korea, Siheung-Si, Republic of Korea), Jaeho Lee (Opt-AI), Seongwan Kim (Opt-AI), Sharif S M A (Opt-AI), Nodirkhujja Khujaev (Opt-AI), and Roman Tsou (Opt-AI)

DVMSR: Distillated Vision Mamba for Efficient Super-Resolution	6536
--	------

Xiaoyan Lei (Zhengzhou University of Light Industry, China), Wenlong Zhang (The HongKong Polytechnic University, China), and Weifeng Cao (Zhengzhou University of Light Industry, China)

Florin-Alexandru Vasluianu (Computer Vision Lab, University of Wurzburg), Tim Seizinger (University of Wurzburg), Zhuyun Zhou (Univ. Bourgogne Franche-Comte, France), Zongwei Wu (University of Wurzburg), Cailian Chen (Shanghai Jiao Tong University), Radu Timofte (University of Wurzburg & ETH Zurich), Wei Dong (McMaster University, Canada), Han Zhou (McMaster University, Canada), Yuqiong Tian (McMaster University, Canada), Jun Chen (McMaster University, Canada), Xueyang Fu (University of Science and Technology of China, Hefei, China), Xin Lu (University of Science and Technology of China, Hefei, China), Yurui Zhu (University of Science and Technology of China, Hefei, China), Xi Wang (University of Science and Technology of China, Hefei, China), Dong Li (University of Science and Technology of China, Hefei, China), Jie Xiao (University of Science and Technology of China, Hefei, China), Yunpeng Zhang (University of Science and Technology of China, Hefei, China), Zheng-Jun Zha (University of Science and Technology of China, Hefei, China), Zhao Zhang (Hefei University of Technology, China), Suiyi Zhao (Hefei University of Technology, China), Bo Wang (Hefei University of Technology, China), Yan Luo (Hefei University of Technology, China), Yanyan Wei (Hefei University of Technology, China), Zhihao Zhao (Nanjing University of Science and Technology, Nanjing, Jiangsu Province, China), Long Sun (Nanjing University of Science and Technology, Nanjing, Jiangsu Province, China), Tingting Yang (Nanjing University of Science and Technology, Nanjing, Jiangsu Province, China), Jinshan Pan (Nanjing University of Science and Technology, Nanjing, Jiangsu Province, China), Jiangxin Dong (Nanjing University of Science and Technology, Nanjing, Jiangsu Province, China), Jinhui Tang (Nanjing University of Science and Technology, Nanjing, Jiangsu Province, China), Bilel Benjdira (Prince Sultan University, Riyadh, Saudi Arabia), Mohammed Nassif (Prince Sultan University, Riyadh, Saudi Arabia), Anis Koubaa (Prince Sultan University, Riyadh, Saudi Arabia), Ahmed Elhayek (Prince Muqrin University, Medinah, Saudi Arabia), Anas M. Ali (Prince Sultan University, Riyadh, Saudi Arabia), Kyotaro Tokoro (Toyota Technological Institute, Japan), Kento Kawai (Toyota Technological Institute, Japan), Kaname Yokoyama (Toyota Technological Institute, Japan), Takuya Seno (Toyota Technological Institute, Japan), Yuki Kondo (Toyota Technological Institute, Japan), Norimichi Ukita (Toyota Technological Institute, Japan), Chenghua Li (Nanjing Artificial Intelligence Research of IA), Bo Yang (Nanjing Artificial Intelligence Research of IA), Zhiqi Wu (Nanjing Artificial Intelligence Research of IA), Gao Chen (Nanjing Artificial Intelligence Research of IA), Yihan Yu (High School Affiliated to Nanjing Normal University Jiangning Campus), Sixiang Chen (Hong Kong University of Science and Technology (Guangzhou)), Kai Zhang (Hong Kong University of Science and Technology (Guangzhou)), Tian Ye (Hong Kong University of Science and Technology (Guangzhou)), Wenbin Zou (South China University of Technology), Yunlong Lin (Xiamen University), Zhaohu Xing (Hong Kong University of Science and Technology (Guangzhou)), Jinbin Bai (National University of Singapore), Wenhao Chai (University of Washington), Lei Zhu (Hong Kong University of Science and Technology (Guangzhou)), Ritik Maheshwari (GEC Ajmer), Rakshank Verma (GEC

Ajmer), Rahul Tekchandani (GEC Ajmer), Praful Hambarde (CVPR Lab IIT Ropar), Satya Narayan Tazi (GEC Ajmer), Santosh Kumar Vipparthi (CVPR Lab IIT Ropar), Subrahmanyam Murala (SCSS Trinity College Dublin), Jaeho Lee (Opt-AI), Seongwan Kim (Opt-AI), Sharif S M A (Opt-AI), Nodirkhujja Khujaev (Opt-AI), Roman Tsoy (Opt-AI), Fan Gao (Nanjing University of Posts and Telecommunications, China), Weidan Yan (Nanjing University of Posts and Telecommunications, China), Wenzhe Shao (Nanjing University of Posts and Telecommunications, China), Dengyin Zhang (Nanjing University of Posts and Telecommunications, China), Bin Chen (Fuzhou University, China), Siqi Zhang (Xiamen University, China), Yanxin Qian (Xiamen University, China), Yuanbin Chen (Fuzhou University, China), Yuanbo Zhou (Fuzhou University, China), Tong Tong (Fuzhou University, China), Rongfeng Wei (University of Hong Kong), Ruiqi Sun (University of Hong Kong), Yue Liu (Sun Yat-sen University), Nikhil Akalwadi (KLE Technological University, Hubballi, Karnataka, India), Amogh Joshi (KLE Technological University, Hubballi, Karnataka, India), Sampada Malagi (KLE Technological University, Hubballi, Karnataka, India), Chaitra Desai (KLE Technological University, Hubballi, Karnataka, India), Ramesh Ashok Tabib (KLE Technological University, Hubballi, Karnataka, India), Uma Mudenagudi (KLE Technological University, Hubballi, Karnataka, India), Ali Murtaza (University Teknologi Malaysia, Kuala Lumpur, , Malaysia), Uswah Khairuddin (University Teknologi Malaysia, Kuala Lumpur, , Malaysia), Ahmad Athif Mohd Faudzi (Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia), Adinath Dukre (Shri Guru Gobind Singhji Institute Of Engineering and Technology, Nanded, India), Vivek Deshmukh (Shri Guru Gobind Singhji Institute Of Engineering and Technology, Nanded, India), Shruti S. Phutke (Indian Institute of Technology Ropar, India), Ashutosh Kulkarni (Indian Institute of Technology Ropar, India), Santosh Kumar Vipparthi (Indian Institute of Technology Ropar, India), Anil Gonde (Shri Guru Gobind Singhji Institute Of Engineering and Technology, Nanded, India), Subrahmanyam Murala (Trinity College Dublin, Ireland), Arun Karthik K (Shiv Nadar University, Chennai, India), Manasa N (Shiv Nadar University, Chennai, India), Shri Hari Priya (Shiv Nadar University, Chennai, India), Wei Hao (Fortinet, Inc.), Xingzhuo Yan (Bosch Investment Ltd.), and Minghan Fu (University of Saskatchewan)

NTIRE 2024 Challenge on Low Light Image Enhancement: Methods and Results	6571
--	------

Xiaoning Liu (University of Electronic Science and Technology of China, China), Zongwei Wu (University of Wurzburg, Germany), Ao Li (University of Electronic Science and Technology of China, China), Florin-Alexandru Vasluianu (University of Wurzburg, Germany), Yulun Zhang (Shanghai Jiao Tong University, China), Shuhang Gu (University of Electronic Science and Technology of China, China), Le Zhang (University of Electronic Science and Technology of China, China), Ce Zhu (University of Electronic Science and Technology of China, China), Radu Timofte (University of Wurzburg, Germany), Zhi Jin (Shenzhen Campus of Sun Yat-sen University, China), Hongjun Wu (Shenzhen Campus of Sun Yat-sen University, China), Chenxi Wang (Shenzhen Campus of Sun Yat-sen University, China), Haitao Ling (Shenzhen Campus of Sun Yat-sen University, China), Yuanhao Cai (Johns Hopkins University), Hao Bian (Tsinghua University), Yuxin Zheng (Tsinghua University), Jing Lin (Tsinghua University), Alan Yuille (Johns Hopkins University), Ben Shao (Zhejiang Dahua Technology Co.,Ltd.), Jin Guo (Zhejiang Dahua Technology Co.,Ltd.), Tian Liu (Zhejiang Dahua Technology Co.,Ltd.), Mohao Wu (Zhejiang Dahua Technology Co.,Ltd.), Yixu Feng (Northwestern Polytechnical University, China), Shuo Hou (Northwestern Polytechnical University, China), Haotian Lin (Northwestern Polytechnical University, China), Yu Zhu (Northwestern Polytechnical University, China), Peng Wu (Northwestern Polytechnical University, China), Wei Dong (Xi'an University of Architecture and Technology, China), Jinqiu Sun (Northwestern Polytechnical University, China), Yanning Zhang (Northwestern Polytechnical University, China), Qingsen Yan (Northwestern Polytechnical University, China), Wenbin Zou (South China University of Technology, China), Weipeng Yang (South China University of Technology, China), Yunxiang Li (Fuzhou University, China), Qiaomu Wei (Chengdu University of Information Technology, China), Tian Ye (Hong Kong University of Science and Technology (Guangzhou), China), Sixiang Chen (Chengdu University of Information Technology, China; Hong Kong University of Science and Technology (Guangzhou), China), Zhao Zhang (Hefei University of Technology, China), Suiyi Zhao (Hefei University of Technology, China), Bo Wang (Hefei University of Technology, China), Yan Luo (Hefei University of Technology, China), Zhichao Zuo (Hefei University of Technology, China), Mingshen Wang (Hefei University of Technology, China), Junhu Wang (Hefei University of Technology, China), Yanyan Wei (Hefei University of Technology, China), Xiaopeng Sun (Individual Researcher), Yu Gao (Individual Researcher), Jiancheng Huang (Individual Researcher), Hongming Chen (Shenyang Aerospace University, China), Xiang Chen (Nanjing University of Science and Technology, China), Hui Tang (Fuzhou University, Fuzhou, China), Yuanbin Chen (Fuzhou University, Fuzhou, China), Yuanbo Zhou (Fuzhou University, Fuzhou, China), Xinwei Dai (Fuzhou University, Fuzhou, China), Xintao Qiu (Fuzhou University, Fuzhou, China), Wei Deng (Imperial Vision Technology, Fuzhou, China), Qinquan Gao (Fuzhou University, Fuzhou, China; Imperial Vision Technology, Fuzhou, China), Tong Tong (Fuzhou University, Fuzhou, China; Imperial Vision Technology, Fuzhou, China), Mingjia Li (Tianjin University, China), Jin Hu (Tianjin University, China), Xinyu He (Tianjin University, China), Xiaojie Guo (Tianjin

University, China), Sabarinathan Sabarinathan (Couger Inc, Japan), K Uma (Sasi Institute of Technology & Engineering, India), A Sasithradevi (Vellore Institute of Technology, India), B Sathya Rama (Thiagarajar college of engineering, India), S. Mohamed Mansoor Roomi (Thiagarajar college of engineering, India), V. Srivatsav (Coventry University, United Kingdom), Jinjuan Wang (Independent Researchers), Long Sun (Independent Researchers), Qiuying Chen (Independent Researchers), Jiahong Shao (Independent Researchers), Yizhi Zhang (Independent Researchers), Marcos V. Conde (Cidaut AI; CVLab, University of Wuerzburg), Daniel Feijoo (Cidaut AI), Juan C. Benito (Cidaut AI), Alvaro Garcia (Cidaut AI), Jaeho Lee (Opt-AI), Seongwan Kim (Opt-AI), Sharif S M A (Opt-AI), Nodirkhujja Khujaev (Opt-AI), Roman Tsoy (Opt-AI), Ali Murtaza (University Teknologi Malaysia, Kuala Lumpur, , Malaysia; Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia), Uswah Khairuddin (University Teknologi Malaysia, Kuala Lumpur, , Malaysia; Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia), Amad Athif Mohd Faudzi (Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia), Sampada Malagi (KLE Technological University, Hubballi, Karnataka, India), Amogh Joshi (KLE Technological University, Hubballi, Karnataka, India), Nikhil Akalwadi (KLE Technological University, Hubballi, Karnataka, India), Chaitra Desai (KLE Technological University, Hubballi, Karnataka, India), Ramesh Ashok Tabib (KLE Technological University, Hubballi, Karnataka, India), Uma Mudenagudi (KLE Technological University, Hubballi, Karnataka, India), Palani Yashaswini (KLE Technological University, Hubballi, Karnataka, India), Nitish Upasi (KLE Technological University, Hubballi, Karnataka, India), Dikshit Hegde (KLE Technological University, Hubballi, Karnataka, India), Ujwala Patil (KLE Technological University, Hubballi, Karnataka, India), Sujata C (KLE Technological University, Hubballi, Karnataka, India), Xingzhuo Yan (Bosch Investment Ltd.), Wei Hao (Fortinet, Inc.), Minghan Fu (University of Saskatchewan), Pooja Choksy (Sardar Vallabhbhai National Institute of Technology, India), Anjali Sarvaiya (Sardar Vallabhbhai National Institute of Technology, India), Kishor Upla (Sardar Vallabhbhai National Institute of Technology, India), Kiran Raja (Norwegian University of Science and Technology, Norway), Hailong Yan (University of Electronic Science and Technology of China, China), Yunkai Zhang (Hefei University of Technology, China), Baiang Li (Hefei University of Technology, China), Jingyi Zhang (Hefei University of Technology, China), and Huan Zheng (University of Macau, China)

The Ninth NTIRE 2024 Efficient Super-Resolution Challenge Report	6595
--	------

Bin Ren (University of Pisa, Italy University of Trento, Italy), Yawei Li (ETH Zurich, Switzerland), Nancy Mehta (University of Würzburg, Germany), Radu Timofte (University of Würzburg, Germany), Hongyuan Yu (Xiaomi Inc.), Cheng Wan (Georgia Institute of Technology), Yuxin Hong (Lanzhou University), Bingnan Han (Xiaomi Inc.), Zhuoyuan Wu (Xiaomi Inc.), Yajun Zou (Xiaomi Inc.), Yuqing Liu (Xiaomi Inc.), Jizhe Li (Xiaomi Inc.), Keji He (Institute of Automation, Chinese Academy of Sciences), Chao Fan (Beijing University of Technology), Heng Zhang (Xiaomi Inc.), Xiaolin Zhang (Xiaomi Inc.), Xuanwu Yin (Xiaomi Inc.), Kunlong Zuo (Xiaomi Inc.), Bohao Liao (University of Science and Technology of China), Peizhe Xia (University of Science and Technology of China), Long Peng (University of Science and Technology of China), Zhibo Du (University of Science and Technology of China), Xin Di (University of Science and Technology of China), Wangkai Li (University of Science and Technology of China), Yang Wang (University of Science and Technology of China), Wei Zhai (University of Science and Technology of China), Renjing Pei (Huawei Noah's Ark Lab), Jiaming Guo (Huawei Noah's Ark Lab), Songcen Xu (Huawei Noah's Ark Lab), Yang Cao (University of Science and Technology of China), Zhengjun Zha (University of Science and Technology of China), Yan Wang (Nankai University), Yi Liu (ByteDance Inc.), Qing Wang (ByteDance Inc.), Gang Zhang (ByteDance Inc.), Liou Zhang (ByteDance Inc.), Shijie Zhao (ByteDance Inc.), Long Sun (Nanjing University of Science and Technology), Jinshan Pan (Nanjing University of Science and Technology), Jiangxin Dong (Nanjing University of Science and Technology), Jinhui Tang (Nanjing University of Science and Technology), Xin Liu (China Mobile Research Institute), Min Yan (China Mobile Research Institute), Qian Wang (China Mobile Research Institute), Menghan Zhou (Lenovo Research), Yiqiang Yan (Lenovo Research), Yixuan Liu (Advanced Micro Devices, Inc., Beijing, China), Wensong Chan (Advanced Micro Devices, Inc., Beijing, China), Dehua Tang (Advanced Micro Devices, Inc., Beijing, China), Dong Zhou (Advanced Micro Devices, Inc., Beijing, China), Li Wang (Advanced Micro Devices, Inc., Beijing, China), Lu Tian (Advanced Micro Devices, Inc., Beijing, China), Barsoum Emad (Advanced Micro Devices, Inc., Beijing, China), Bohan Jia (East China Normal University), Junbo Qiao (East China Normal University; Huawei Noah's Ark Lab), Yunshuai Zhou (East China Normal University), Yun Zhang (Huawei Noah's Ark Lab; The Hong Kong University of Science and Technology), Wei Li (Huawei Noah's Ark Lab), Shaohui Lin (East China Normal University), Shenglong Zhou (University of Science and Technology of China), Binbin Chen (Huazhong University of Science and Technology), Jincheng Liao (East China Normal University), Suiyi Zhao (Hefei University of Technology), Zhao Zhang (Hefei University of Technology), Bo Wang (Hefei University of Technology), Yan Luo (Hefei University of Technology), Yanyan Wei (Hefei University of Technology), Feng Li (Hefei University of Technology), Mingshen Wang (Hefei University of Technology), Yawei Li (Hefei University of Technology), Jinhan Guan (Hefei University of Technology), Dehua Hu (Hefei University of Technology), Jiawei Yu (National University of Defense Technology), Qisheng Xu (National University of Defense Technology), Tao Sun (National University of

Defense Technology), Long Lan (National University of Defense Technology), Kele Xu (National University of Defense Technology), Xin Lin (Sichuan University), Jingtong Yue (Sichuan University), Lehan Yang (The University of Sydney), Shiyi Du (Carnegie Mellon University), Lu Qi (The University of California, Merced), Chao Ren (Sichuan University), Zeyu Han (Sichuan University), Yuhan Wang (Sichuan University), Chaolin Chen (Sichuan University), Haobo Li (Independent Researcher), Mingjun Zheng (Nanjing University of Science and Technology), Zhongbao Yang (Nanjing University of Science and Technology), Lianhong Song (Nanjing University of Science and Technology), Xingzhuo Yan (Bosch Investment Ltd.), Minghan Fu (University of Saskatchewan), Jingyi Zhang (Hefei University of Technology), Baiang Li (Hefei University of Technology), Qi Zhu (University of Science and Technology of China), Xiaogang Xu (The Chinese University of Hong Kong; Zhejiang University), Dan Guo (Hefei University of Technology), Chunle Guo (Nankai University), Jiadi Chen (Zhejiang Normal University, Jinhua, China), Huanhuan Long (Zhejiang Normal University, Jinhua, China), Chunjiang Duanmu (Zhejiang Normal University, Jinhua, China), Xiaoyan Lei (Zhengzhou University of Light Industry), Jie Liu (Zhengzhou University of Light Industry), Weilin Jia (Zhengzhou University of Light Industry), Weifeng Cao (Zhengzhou University of Light Industry), Wenlong Zhang (The Hong Kong Polytechnic University), Yanyu Mao (Xian University of Posts and Telecommunications, Xi'an, China), Ruilong Guo (Xian University of Posts and Telecommunications, Xi'an, China), Nihao Zhang (Xian University of Posts and Telecommunications, Xi'an, China), Qian Wang (Xian University of Posts and Telecommunications, Xi'an, China; National Engineering Laboratory for Cyber Event Warning and Control Technologies), Manoj Pandey (DeltaX), Maksym Chernozhukov (DeltaX), Giang Le (DeltaX), Shuli Cheng (Xinjiang University, Urumqi, China), Hongyuan Wang (Xinjiang University, Urumqi, China), Ziyang Wei (Xinjiang University, Urumqi, China), Qingting Tang (Xinjiang University, Urumqi, China), Liejun Wang (Xinjiang University, Urumqi, China), Yongming Li (Xinjiang University, Urumqi, China), Yanhui Guo (McMaster University), Hao Xu (McMaster University), Akram Khatami-Rizi (Cyberspace Research Institute of Shahid Beheshti University of Iran), Ahamad Mahmoudi-Aznavah (Cyberspace Research Institute of Shahid Beheshti University of Iran), Chih-Chung Hsu (National Cheng Kung University), Chia-Ming Lee (National Cheng Kung University), Yi-Shiuan Chou (National Cheng Kung University), Amogh Joshi (KLE Technological University, Hubballi, Karnataka, India), Nikhil Akalwadi (KLE Technological University, Hubballi, Karnataka, India), Sampada Malagi (KLE Technological University, Hubballi, Karnataka, India), Palani Yashaswini (KLE Technological University, Hubballi, Karnataka, India), Chaitra Desai (KLE Technological University, Hubballi, Karnataka, India), Ramesh Ashok Tabib (KLE Technological University, Hubballi, Karnataka, India), Ujwala Patil (KLE Technological University, Hubballi, Karnataka, India), and Uma Mudanagudi (KLE Technological University, Hubballi, Karnataka, India)

NTIRE 2024 Restore Any Image Model (RAIM) in the Wild Challenge	6632
<i>Jie Liang (OPPO Research Institute), Radu Timofte (Computer Vision Lab, University of Wurzburg, Germany), Qiaosi Yi (OPPO Research Institute and The Hong Kong Polytechnic University), Shuaizheng Liu (OPPO Research Institute and The Hong Kong Polytechnic University), Lingchen Sun (OPPO Research Institute and The Hong Kong Polytechnic University), Rongyuan Wu (OPPO Research Institute and The Hong Kong Polytechnic University), Xindong Zhang (OPPO Research Institute), Hui Zeng (OPPO Research Institute), Lei Zhang (OPPO Research Institute and The Hong Kong Polytechnic University), Yibin Huang (Xiaomi Inc., China), Shai Liu (Xiaomi Inc., China), Yongqiang Li (Xiaomi Inc., China), Chaoyu Feng (Xiaomi Inc., China), Xiaotao Wang (Xiaomi Inc., China), Lei Lei (Xiaomi Inc., China), Yuxiang Chen (Xiaohongshu), Xiangyu Chen (University of Macau; Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Qiubo Chen (Xiaohongshu), Fengyu Sun (China University of Petroleum (East China)), Mengying Cui (China University of Petroleum (East China)), Jiaxu Chen (China University of Petroleum (East China)), Zhenyu Hu (Wuhan University), Jingyun Liu (Wuhan University), Wenzhuo Ma (Wuhan University), Ce Wang (Wuhan University), Hanyou Zheng (Wuhan University), Wanjie Sun (Wuhan University), Zhenzhong Chen (Wuhan University), Ziwei Luo (Uppsala University), Fredrik K. Gustafsson (Uppsala University), Zheng Zhao (Uppsala University), Jens Sjolund (Uppsala University), Thomas B. Schon (Uppsala University), Xiong Dun (Tongji University), Pengzhou Ji (Tongji University), Yujie Xing (Tongji University), Xuquan Wang (Tongji University), Zhanshan Wang (Tongji University), Xinbin Cheng (Tongji University), Jun Xiao (The Hong Kong Polytechnic University), Chenhang He (The Hong Kong Polytechnic University), Xiuyuan Wang (The Hong Kong Polytechnic University), Zhi-Song Liu (Lappeenranta-Lahti University of Technology), Zimeng Miao (Harbin Institute of Technology), Zhicun Yin (Harbin Institute of Technology), Ming Liu (Harbin Institute of Technology), Wangmeng Zuo (Harbin Institute of Technology), and Shuai Li (The Hong Kong Polytechnic University)</i>	
Photo-Realistic Image Restoration in the Wild with Controlled Vision-Language Models	6641
<i>Ziwei Luo (Uppsala University, Sweden), Fredrik K. Gustafsson (Karolinska Institutet, Sweden), Zheng Zhao (Uppsala University, Sweden), Jens Sjölund (Uppsala University, Sweden), and Thomas B. Schön (Uppsala University, Sweden)</i>	
Exploring AIGC Video Quality: A Focus on Visual Harmony, Video-Text Consistency and Domain Distribution Gap	6652
<i>Bowen Qu (Peking University, China), Xiaoyu Liang (Peking University, China), Shangkun Sun (Peking University, China; Peng Cheng Laboratory, China), and Wei Gao (Peking University, China; Peng Cheng Laboratory, China)</i>	
Training Transformer Models by Wavelet Losses Improves Quantitative and Visual Performance in Single Image Super-Resolution	6661
<i>Cansu Korkmaz (Koc University, Turkey) and A. Murat Tekalp (Koc University, Turkey)</i>	

Equipping Diffusion Models with Differentiable Spatial Entropy for Low-Light Image Enhancement	6671
<i>Wenyi Lian (Uppsala University, Sweden), Wenjing Lian (Northeastern University, China), and Ziwei Luo (Uppsala University, Sweden)</i>	
SDCNet:Spatially-Adaptive Deformable Convolution Networks for HR NonHomogeneous Dehazing .	6682
<i>Yidi Liu (University of Science and Technology of China), Xingbo Wang (University of Science and Technology of China), Yurui Zhu (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), and Zheng-Jun Zha (University of Science and Technology of China)</i>	
SF-IQA: Quality and Similarity Integration for AI Generated Image Quality Assessment	6692
<i>Zihao Yu (University of Science and Technology of China, China), Fengbin Guan (University of Science and Technology of China, China), Yiting Lu (University of Science and Technology of China, China), Xin Li (University of Science and Technology of China, China), and Zhibo Chen (University of Science and Technology of China, China)</i>	
NTIRE 2024 Challenge on Night Photography Rendering	6702
<i>Egor Ershov (Institute for Information Transmission Problems (Kharkevich institute), Russia), Artyom Panshin (Institute for Information Transmission Problems (Kharkevich institute), Russia), Oleg Karasev (Institute for Information Transmission Problems (Kharkevich institute), Russia), Sergey Korchagin (Institute for Information Transmission Problems (Kharkevich institute), Russia), Shepelev Lev (Institute for Information Transmission Problems (Kharkevich institute), Russia), Alexandr Startsev (Institute for Information Transmission Problems (Kharkevich institute), Russia), Daniil Vladimirov (Institute for Information Transmission Problems (Kharkevich institute), Russia), Ekaterina Zaychenkova (Institute for Information Transmission Problems (Kharkevich institute), Russia), Nikola Banić (Gideon Brothers), Dmitrii R Iarchuk (Institute for Information Transmission Problems (Kharkevich institute), Russia), Maria Efimove (Institute for Information Transmission Problems (Kharkevich institute), Russia), Radu Timofte (University of Würzburg), and Arseniy Terekhin (Institute for Information Transmission Problems (Kharkevich institute), Russia)</i>	
LGFN: Lightweight Light Field Image Super-Resolution using Local Convolution Modulation and Global Attention Feature Extraction	6712
<i>Zhongxin Yu (Fujian Normal University), Liang Chen (Fujian Normal University), Zhiyun Zeng (Fujian Normal University), Kunping Yang (Fujian Normal University), Shaofei Luo (Fujian Normal University), Shaorui Chen (Fujian Normal University), and Cheng Zhong (Fujian Normal University)</i>	
Two Stage Dehazing Framework for Dense and Non-Homogeneous Dehazing	6722
<i>Wei Song (Guangdong University of Technology), Yichang Gao (Guangdong University of Technology), Jiahao Xiong (Guangdong University of Technology), Hualiang Lin (Guangdong University of Technology), Dong Li (Guangdong University of Technology), and Yun Zhang (Guangdong University of Technology)</i>	

Deep Portrait Quality Assessment. A NTIRE 2024 Challenge Survey	6732
<i>Nicolas Chahine (DXOMARK), Marcos V. Conde (University of Würzburg), Daniela Carfora (DXOMARK), Gabriel Pacianotto (DXOMARK), Benoit Pochon (DXOMARK), Sira Ferradans (DXOMARK), Radu Timofte (University of Würzburg), Zhichao Duan, Xinrui Xu, Yipo Huang, Quan Yuan, Xiangfei Sheng, Zhichao Yang, Leida Li, Haotian Fan, Fangyuan Kong, Yifang Xu, Wei Sun, Weixia Zhang, Yanwei Jiang, Haoning Wu, Zicheng Zhang, Jun Jia, Yingjie Zhou, Zhongpeng Ji, Xiongkuo Min, Weisi Lin, Guangtao Zhai, Xiaoqi Wang, Junqi Liu, Zixi Guo, Yun Zhang, Zewen Chen, Wen Wang, Juan Wang, and Bing Li</i>	
Deep RAW Image Super-Resolution. A NTIRE 2024 Challenge Survey	6745
<i>Marcos V. Conde (University of Würzburg), Florin-Alexandru Vasluianu (University of Würzburg), Radu Timofte (University of Würzburg), Jianxing Zhang, Jia Li, Fan Wang, Xiaopeng Li, Zikun Liu, Hyunhee Park, Sejun Song, Changho Kim, Zhijuan Huang, Hongyuan Yu, Cheng Wan, Wending Xiang, Jiamin Lin, Hang Zhong, Qiaosong Zhang, Yue Sun, Xuanwu Yin, Kunlong Zuo, Senyan Xu, Siyuan Jiang, Zhijing Sun, Jiaying Zhu, Liangyan Li, Ke Chen, Yunzhe Li, Yimo Ning, Guanhua Zhao, Jun Chen, Jinyang Yu, Kele Xu, Qisheng Xu, and Yong Dou</i>	

AI4Space 2024

Robust Perspective-n-Crater for Crater-based Camera Pose Estimation	6760
<i>Sofia McLeod (The University of Adelaide, Australia), Chee Kheng Chng (The University of Adelaide, Australia), Tatsuharu Ono (Hokkaido University, Japan), Yuta Shimizu (University of Tokyo, Japan), Ryodo Hemmi (University of Tokyo, Japan), Lachlan Holden (The University of Adelaide, Australia), Matthew Rodda (The University of Adelaide, Australia), Feras Dayoub (The University of Adelaide, Australia), Hirdy Miyamoto (University of Tokyo, Japan), Yukihiro Takahashi (Hokkaido University, Japan), Yasuko Kasai (Tokyo Institute of Technology, Japan), and Tat-Jun Chin (The University of Adelaide, Australia)</i>	
Exploring AI-Based Satellite Pose Estimation: from Novel Synthetic Dataset to In-Depth Performance Evaluation	6770
<i>Fabien Gallet (IRT Saint Exupery), Christophe Marabotto (IRT Saint Exupery), and Thomas Chambon (IRT Saint Exupery)</i>	
Optimized Martian Dust Displacement Detection Using Explainable Machine Learning	6779
<i>Ana Lomashvili (German Aerospace Center (DLR), Germany), Kristin Rammelkamp (German Aerospace Center (DLR), Germany), Olivier Gasnault (Institut de Recherche en Astrophysique et Plan écologique (IRAP), France), Protim Bhattacharjee (German Aerospace Center (DLR), Germany), Elise Clavé (German Aerospace Center (DLR), Germany), Christoph H. Egerland (German Aerospace Center (DLR), Germany), Susanne Schröder (German Aerospace Center (DLR), Germany), Begüm Demir (Technical University Berlin, Germany), and Nina L. Lanza (Los Alamos National Laboratory, USA)</i>	

Mitigating Challenges of the Space Environment for Onboard Artificial Intelligence: Design Overview of the Imaging Payload on SPIRIT	6789
<i>Miguel Ortiz del Castillo (The University of Melbourne, Australia), Jonathan Morgan (The University of Melbourne, Australia), Jack McRobbie (The University of Melbourne, Australia), Clint Therakam (The University of Melbourne, Australia), Zaher Joukhadar (The University of Melbourne, Australia), Robert Mearns (The University of Melbourne, Australia), Simon Barraclough (The University of Melbourne, Australia), Richard Sinnott (The University of Melbourne, Australia), Andrew Woods (The University of Melbourne, Australia), Chris Bayliss (The University of Melbourne, Australia), Kris Ehinger (The University of Melbourne, Australia), Ben Rubinstein (The University of Melbourne, Australia), James Bailey (The University of Melbourne, Australia), Airlie Chapman (The University of Melbourne, Australia), and Michele Trenti (The University of Melbourne, Australia)</i>	
A Dual-Mode Approach for Vision-Based Navigation in a Lunar Landing Scenario	6799
<i>Luca Ostrogovich (University of Naples Federico II), Roberto Del Prete (University of Naples Federico II), Giuseppe Tomasichio (Telespazio SRL), Nicolas Longepe (Φ-lab, ESA ESRIN), and Alfredo Renga (University of Naples Federico II)</i>	
Tackling the Satellite Downlink Bottleneck with Federated Onboard Learning of Image Compression	6809
<i>Pablo Gómez (Advanced Concepts Team, ESA/ESTEC; AI Sweden, Göteborg, Sweden) and Gabriele Meoni (Φ-lab, ESA/ESRIN; Aerospace Engineering, TU Delft; AI Sweden, Göteborg, Sweden)</i>	
Transformers for Orbit Determination Anomaly Detection and Classification	6819
<i>Nathan Parrish Ré (Advanced Space, LLC), Matthew Popplewell (Advanced Space, LLC), Michael Caudill (Advanced Space, LLC), Timothy Sullivan (Advanced Space, LLC), Tyler Hanf (Advanced Space, LLC), Benjamin Tatman (Advanced Space, LLC), Kanak Parmar (Advanced Space, LLC), Tyler Presser (Advanced Space, LLC), Sai Chikine (Advanced Space, LLC), Michael Grant (Advanced Space, LLC), and Richard Poulson (Advanced Space, LLC)</i>	
Deploying Machine Learning Anomaly Detection Models to Flight Ready AI Boards	6828
<i>James Murphy (Realtra Space Systems Engineering, Ireland), Maria Buckley (Ubotica Technologies, Ireland), Leonie Buckley (Ubotica Technologies, Ireland), Adam Taylor (Aduvo Engineering & Training, Ireland), Jake O'Brien (Realtra Space Systems Engineering, Ireland), and Brian Mac Namee (University College Dublin, Ireland)</i>	
Cross-Temporal Spectrogram Autoencoder (CTSAE): Unsupervised Dimensionality Reduction for Clustering Gravitational Wave Glitches	6837
<i>Yi Li (Northwestern University, Evanston, IL), Yunan Wu (Northwestern University, Evanston, IL), and Aggelos K. Katsaggelos (Northwestern University, Evanston, IL)</i>	

Monocular 6-DoF Pose Estimation of Spacecrafts Utilizing Self-iterative Optimization and Motion Consistency	6847
<i>Yunfeng Zhang (Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China), Linjing You (MAIS, Institute of Automation, Chinese Academy of Sciences, China), Luyu Yang (MAIS, Institute of Automation, Chinese Academy of Sciences, China), Zhiwei Zhang (MAIS, Institute of Automation, Chinese Academy of Sciences, China), Xiangli Nie (MAIS, Institute of Automation, Chinese Academy of Sciences, China), and Bo Zhang (Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China)</i>	
CroSpace6D: Leveraging Geometric and Motion Cues for High-Precision Cross-Domain 6DoF Pose Estimation for Non-Cooperative Spacecrafts	6857
<i>Jianhong Zuo (Nanjing University of Aeronautics and Astronautics, China; Technology and Engineering Center for Space Utilization, China), Shengyang Zhang (University of Chinese Academy of Sciences, China), Qianyu Zhang (University of Chinese Academy of Sciences, China), Yutao Zhao (University of Chinese Academy of Sciences, China), Baichuan Liu (University of Chinese Academy of Sciences, China), Aodi Wu (University of Chinese Academy of Sciences, China), Xue Wan (Technology and Engineering Center for Space Utilization, China), Leizheng Shu (Technology and Engineering Center for Space Utilization, China), and Guohua Kang (Nanjing University of Aeronautics and Astronautics, China)</i>	
Revisiting the Domain Gap Issue in Non-cooperative Spacecraft Pose Tracking	6864
<i>Kun Liu (Nanjing University of Science and Technology, China) and Yongjun Yu (Nanjing University of Science and Technology, China)</i>	

9th Workshop on Computer Vision for Microscopy Image Analysis

Unsupervised Microscopy Video Denoising	6874
<i>Mary Aiyetigbo (Clemson University, USA), Alexander Korte (Clemson University, USA), Ethan Anderson (Clemson University, USA), Reda Chalhoub (The Medical University of South Carolina, USA), Peter Kalivas (The Medical University of South Carolina, USA), Feng Luo (Clemson University, USA), and Nianyi Li (Clemson University, USA)</i>	
Discovering Interpretable Models of Scientific Image Data with Deep Learning	6884
<i>Christopher J. Soelistyo (The Alan Turing Institute, UK) and Alan R. Lowe (The Alan Turing Institute, UK)</i>	
Vim4Path: Self-Supervised Vision Mamba for Histopathology Images	6894
<i>Ali Nasiri-Sarvi (Concordia University), Vincent Quoc-Huy Trinh (University of Montreal), Hassan Rivaz (Concordia University), and Mahdi S. Hosseini (Concordia University)</i>	
Refining Biologically Inconsistent Segmentation Masks with Masked Autoencoders	6904
<i>Alexander Sauer (University of Oxford, UK), Yuan Tian (Yale School of Medicine, USA), Joerg Bewersdorf (Yale School of Medicine, USA), and Jens Rittscher (University of Oxford, UK)</i>	
Histopathological Image Classification with Cell Morphology Aware Deep Neural Networks	6913
<i>Andrey Ignatov (ETH Zurich), Josephine Yates (ETH Zurich), and Valentina Boeva (ETH Zurich)</i>	

NOISe: Nuclei-Aware Osteoclast Instance Segmentation for Mouse-to-Human Domain Transfer .	6926
<i>Sai Kumar Reddy Manne (Northeastern University, USA), Brendan Martin (Northeastern University, USA), Tyler Roy (MaineHealth Institute for Research, USA), Ryan Neilson (MaineHealth Institute for Research, USA), Rebecca Peters (MaineHealth Institute for Research, USA; University of Maine, USA), Meghana Chillara (Northeastern University, USA), Christine W. Lary (Northeastern University, USA), Katherine J. Motyl (MaineHealth Institute for Research, USA; University of Maine, USA; Tufts University School of Medicine, USA), and Michael Wan (Northeastern University, USA)</i>	
Low-Resolution-Only Microscopy Super-Resolution Models Generalizing to Non-Periodicities at Atomic Scale	6936
<i>Björn Möller (Technische Universität Braunschweig), Zhengyang Li (Technische Universität Braunschweig), Markus Etzkorn (Technische Universität Braunschweig), and Tim Fingscheidt (Technische Universität Braunschweig)</i>	
Uncertainty Estimation for Tumor Prediction with Unlabeled Data	6946
<i>Juyoung Yun (Stony Brook University, USA), Shahira Abousamra (Stony Brook University, USA), Chen Li (Stony Brook University, USA), Rajarsi Gupta (Stony Brook University, USA), Tahsin Kurc (Stony Brook University, USA), Dimitris Samaras (Stony Brook University, USA), Alison Van Dyke (National Cancer Institute, USA), Joel Saltz (Stony Brook University, USA), and Chao Chen (Stony Brook University, USA)</i>	
Triage of 3D Pathology Data via 2.5D Multiple-instance Learning to Guide Pathologist Assessments	6955
<i>Gan Gao (University of Washington, USA), Andrew H. Song (Mass General Brigham, USA; Harvard University, USA), Fiona Wang (University of Washington, USA), David Brenes (University of Washington, USA), Rui Wang (University of Washington, USA), Sarah S.L. Chow (University of Washington, USA), Kevin W. Bishop (University of Washington, USA), Lawrence D. True (University of Washington, USA), Faisal Mahmood (Mass General Brigham, USA; Harvard University, USA), and Jonathan T.C. Liu (University of Washington, USA)</i>	
Super-resolution of Biomedical Volumes with 2D Supervision	6966
<i>Cheng Jiang (University of Michigan, USA), Alexander Gedeon (University of Michigan, USA), Yiwei Lyu (University of Michigan, USA), Eric Landgraf (University of Michigan, USA), Yufeng Zhang (University of Michigan, USA), Xinhai Hou (University of Michigan, USA), Akhil Kondepudi (University of Michigan, USA), Asadur Chowdury (University of Michigan, USA), Honglak Lee (University of Michigan, USA), and Todd Hollon (University of Michigan, USA)</i>	
Weakly Supervised Set-Consistency Learning Improves Morphological Profiling of Single-Cell Images	6978
<i>Heming Yao (Genentech), Phil Hanslovsky (Genentech), Jan-Christian Huetter (Genentech), Burkhard Hoeckendorf (Genentech), and David Richmond (Genentech)</i>	
Grad-CAMO: Learning Interpretable Single-Cell Morphological Profiles from 3D Cell Painting Images	6988
<i>Vivek Gopalakrishnan (MIT), Jingzhe Ma (Xellar Biosystems), and Zhiyong Xie (Xellar Biosystems)</i>	

2nd Workshop on Multimodal Content Moderation

- An End-to-End Vision Transformer Approach for Image Copy Detection 6997
Jiahe Steven Lee (National University of Singapore, Singapore), Wynne Hsu (National University of Singapore, Singapore), and Mong Li Lee (National University of Singapore, Singapore)

8th AI City Challenge

- A Robust Online Multi-Camera People Tracking System With Geometric Consistency and State-aware Re-ID Correction 7007
Zhenyu Xie (Shanghai Jiao Tong University, China), Zelin Ni (Shanghai Jiao Tong University, China), Wenjie Yang (Shanghai Jiao Tong University, China), Yuang Zhang (Shanghai Jiao Tong University, China), Yihang Chen (Shanghai Jiao Tong University, China; Monash University, Australia), Yang Zhang (AI Lab, Lenovo Research, China), and Xiao Ma (AI Lab, Lenovo Research, China)
- Robust Data Augmentation and Ensemble Method for Object Detection in Fisheye Camera Images..... 7017
Viet Hung Duong (VNPT AI, VNPT Group, Hanoi, Vietnam), Duc Quyen Nguyen (VNPT AI, VNPT Group, Hanoi, Vietnam), Thien Van Luong (Phenikaa University, Hanoi, Vietnam), Huan Vu (University of Transport and Communications, Hanoi, Vietnam), and Tien Cuong Nguyen (VNPT AI, VNPT Group, Hanoi, Vietnam)
- Motorcyclist Helmet Violation Detection Framework by Leveraging Robust Ensemble and Augmentation Methods 7027
Thien Van Luong (Phenikaa University, Hanoi, Vietnam), Huu Si Phuc Nguyen (VNPT AI, VNPT Group, Hanoi, Vietnam), Duy Khanh Dinh (VNPT AI, VNPT Group, Hanoi, Vietnam), Viet Hung Duong (VNPT AI, VNPT Group, Hanoi, Vietnam), Duy Hong Sam Vo (VNPT AI, VNPT Group, Hanoi, Vietnam), Huan Vu (University of Transport and Communications, Hanoi, Vietnam), Minh Tuan Hoang (VNPT AI, VNPT Group, Hanoi, Vietnam), and Tien Cuong Nguyen (VNPT AI, VNPT Group, Hanoi, Vietnam)
- An Online Approach and Evaluation Method for Tracking People Across Cameras in Extremely Long Video Sequence 7037
Cheng-Yen Yang (University of Washington), Hsiang-Wei Huang (University of Washington), Pyong-Kun Kim (Electronics and Telecommunications Research Institute), Zhongyu Jiang (University of Washington), Kwang-Ju Kim (Electronics and Telecommunications Research Institute), Chung-I Huang (National Center for High-performance Computing), Haiqing Du (Beijing University of Posts and Telecommunications), and Jenq-Neng Hwang (University of Washington)

Divide and Conquer Boosting for Enhanced Traffic Safety Description and Analysis with Large Vision Language Model	7046
<i>Khai Trinh Xuan (Ho Chi Minh City University of Technology, Vietnam), Khoi Nguyen Nguyen (Ho Chi Minh City University of Technology, Vietnam), Bach Hoang Ngo (University of Science, Vietnam), Vu Dinh Xuan (University of Information Technology, Vietnam), Minh-Hung An (FPT Telecom, Vietnam), and Quang-Vinh Dinh (Vietnamese German University, Vietnam)</i>	
Low-Light Image Enhancement Framework for Improved Object Detection in Fisheye Lens Datasets	7056
<i>Dai Quoc Tran (Sungkyunkwan University, South Korea), Armstrong Aboah (North Dakota State University, USA), Yuntae Jeon (Sungkyunkwan University, South Korea), Maged Shoman (University of Central Florida, USA), Minsoo Park (Sungkyunkwan University, South Korea), and Seunghee Park (Sungkyunkwan University, South Korea)</i>	
A Coarse-to-fine Two-stage Helmet Detection Method for Motorcyclists	7066
<i>Hongpu Zhang (Beijing University of Posts and Telecommunications, China), Zhe Cui (Beijing University of Posts and Telecommunications, China), and Fei Su (Beijing University of Posts and Telecommunications, China)</i>	
Multi-perspective Traffic Video Description Model with Fine-grained Refinement Approach	7075
<i>Tuan-An To (University of Science - VNUHCM), Minh-Nam Tran (University of Science - VNUHCM), Trong-Bao Ho (University of Science - VNUHCM), Thien-Loc Ha (University of Science - VNUHCM), Quang-Tan Nguyen (University of Science - VNUHCM), Hoang-Chau Luong (University of Science - VNUHCM), Thanh-Duy Cao (University of Science - VNUHCM), and Minh-Triet Tran (University of Science - VNUHCM)</i>	
An Effective Method for Detecting Violation of Helmet Rule for Motorcyclists	7085
<i>Yunliang Chen (China Mobile Shanghai ICT Co.,Ltd), Wei Zhou (China Mobile Shanghai ICT Co.,Ltd), Zicen Zhou (China Mobile Shanghai ICT Co.,Ltd), Bing Ma (China Mobile Shanghai ICT Co.,Ltd), Chen Wang (China Mobile Shanghai ICT Co.,Ltd), Yingda Shang (China Mobile Shanghai ICT Co.,Ltd), An Guo (China Mobile Shanghai ICT Co.,Ltd), and Tianshu Chu (China Mobile Shanghai ICT Co.,Ltd)</i>	
FE-Det: An Effective Traffic Object Detection Framework for Fish-Eye Cameras	7091
<i>Xingshuang Luo (Beijing University of Posts and Telecommunications, China), Zhe Cui (Beijing University of Posts and Telecommunications, China), and Fei Su (Beijing University of Posts and Telecommunications, China)</i>	
Improving Object Detection to Fisheye Cameras with Open-Vocabulary Pseudo-Label Approach	7100
<i>Long Hoang Pham (Sungkyunkwan University, South Korea), Quoc Pham-Nam Ho (Sungkyunkwan University, South Korea), Duong Nguyen-Ngoc Tran (Sungkyunkwan University, South Korea), Tai Huu-Phuong Tran (Sungkyunkwan University, South Korea), Huy-Hung Nguyen (Sungkyunkwan University, South Korea), Duong Khac Vu (Sungkyunkwan University, South Korea), Chi Dai Tran (Sungkyunkwan University, South Korea), Ngoc Doan-Minh Huynh (Sungkyunkwan University, South Korea), Hyung-Min Jeon (Sungkyunkwan University, South Korea), Hyung-Joon Jeon (Sungkyunkwan University, South Korea), and Jae Wook Jeon (Sungkyunkwan University, South Korea)</i>	

Augmented Self-Mask Attention Transformer for Naturalistic Driving Action Recognition	7108
<i>Tiantian Zhang (China Telecom Artificial Intelligence Technology (Beijing) Co., Ltd.), Qingtian Wang (China Telecom Artificial Intelligence Technology (Beijing) Co., Ltd.), Xiaodong Dong (China Telecom Artificial Intelligence Technology (Beijing) Co., Ltd.), Wenqing Yu (China Telecom Artificial Intelligence Technology (Beijing) Co., Ltd.), Hao Sun (China Telecom Artificial Intelligence Technology (Beijing) Co., Ltd.), Xuyang Zhou (China Telecom Artificial Intelligence Technology (Beijing) Co., Ltd.), Aigong Zhen (China Telecom Artificial Intelligence Technology (Beijing) Co., Ltd.), Shun Cui (China Telecom Artificial Intelligence Technology (Beijing) Co., Ltd.), Dong Wu (China Telecom Artificial Intelligence Technology (Beijing) Co., Ltd.), and Zhongjiang He (China Telecom Artificial Intelligence Technology (Beijing) Co., Ltd.)</i>	
KI-GAN: Knowledge-Informed Generative Adversarial Networks for Enhanced Multi-Vehicle Trajectory Forecasting at Signalized Intersections	7115
<i>Chuheng Wei (University of California Riverside), Guoyuan Wu (University of California Riverside), Matthew J. Barth (University of California Riverside), Amr Abdelraouf (InfoTech Labs, Toyota Motor North America), Rohit Gupta (InfoTech Labs, Toyota Motor North America), and Kyungtae Han (InfoTech Labs, Toyota Motor North America)</i>	
Enhancing Traffic Safety with Parallel Dense Video Captioning for End-to-End Event Analysis	7125
<i>Maged Shoman (University of Central Florida, USA), Dongdong Wang (University of Central Florida, USA), Armstrong Aboah (North Dakota State University, USA), and Mohamed Abdel-Aty (University of Central Florida, USA)</i>	
TrafficVLM: A Controllable Visual Language Model for Traffic Video Captioning	7134
<i>Quang Minh Dinh (Simon Fraser University, Canada), Minh Khoi Ho (Hanoi University of Science and Technology, Vietnam), Anh Quan Dang (Hanoi University of Science and Technology, Vietnam), and Hung Phong Tran (Hanoi University of Science and Technology, Vietnam)</i>	
Multi-View Spatial-Temporal Learning for Understanding Unusual Behaviors in Untrimmed Naturalistic Driving Videos	7144
<i>Huy-Hung Nguyen (Sungkyunkwan University, South Korea), Chi Dai Tran (Sungkyunkwan University, South Korea), Long Hoang Pham (Sungkyunkwan University, South Korea), Duong Nguyen-Ngoc Tran (Sungkyunkwan University, South Korea), Tai Huu-Phuong Tran (Sungkyunkwan University, South Korea), Duong Khac Vu (Sungkyunkwan University, South Korea), Quoc Pham-Nam Ho (Sungkyunkwan University, South Korea), Ngoc Doan-Minh Huynh (Sungkyunkwan University, South Korea), Hyung-Min Jeon (Sungkyunkwan University, South Korea), Hyung-Joon Jeon (Sungkyunkwan University, South Korea), and Jae Wook Jeon (Sungkyunkwan University, South Korea)</i>	
Overlap Suppression Clustering for Offline Multi-Camera People Tracking	7153
<i>Ryuto Yoshida (Yachiyo Engineering Co., Ltd., Japan), Junichi Okubo (Yachiyo Engineering Co., Ltd., Japan), Junichiro Fujii (Yachiyo Engineering Co., Ltd., Japan), Masazumi Amakata (Yachiyo Engineering Co., Ltd., Japan), and Takayoshi Yamashita (Chubu University, Japan)</i>	

Robust Motorcycle Helmet Detection in Real-World Scenarios: Using Co-DETR and Minority Class Enhancement	7163
<i>Hao Vo (University of Information Technology, VNU-HCM, Vietnam), Sieu Tran (University of Information Technology, VNU-HCM, Vietnam), Duc Minh Nguyen (University of Information Technology, VNU-HCM, Vietnam), Thua Nguyen (University of Information Technology, VNU-HCM, Vietnam), Tien Do (University of Information Technology, VNU-HCM, Vietnam), Duy-Dinh Le (University of Information Technology, VNU-HCM, Vietnam), and Thanh Duc Ngo (University of Information Technology, VNU-HCM, Vietnam)</i>	
Multi-View Action Recognition for Distracted Driver Behavior Localization	7172
<i>Yuehuan Xu (Beijing University of Posts and Telecommunications, China), Shuai Jiang (Beijing University of Posts and Telecommunications, China), Zhe Cui (Beijing University of Posts and Telecommunications, China), and Fei Su (Beijing University of Posts and Telecommunications, China)</i>	
CityLLaVA: Efficient Fine-Tuning for VLMs in City Scenario	7180
<i>Zhizhao Duan (Alibaba Group), Hao Cheng (Alibaba Group), Duo Xu (Alibaba Group), Xi Wu (Alibaba Group), Xiangxie Zhang (Alibaba Group), Xi Ye (Alibaba Group), and Zhen Xie (Alibaba Group)</i>	
Cluster Self-Refinement for Enhanced Online Multi-Camera People Tracking	7190
<i>Jeongho Kim (Nota Inc., Republic of Korea), Wooksu Shin (Nota Inc., Republic of Korea), Hancheol Park (Nota Inc., Republic of Korea), and Donghyuk Choi (Nota Inc., Republic of Korea)</i>	
Online Multi-camera People Tracking with Spatial-temporal Mechanism and Anchor-feature Hierarchical Clustering	7198
<i>Riu Cherdchusakulchai (AI and Robotics Ventures, Thailand), Sasin Phimsiri (AI and Robotics Ventures, Thailand), Visarut Trairattanapa (AI and Robotics Ventures, Thailand), Suchat Tungjitnob (AI and Robotics Ventures, Thailand), Wasu Kudisthalert (AI and Robotics Ventures, Thailand), Pornprom Kiawjak (AI and Robotics Ventures, Thailand), Ek Thamwiwatthana (AI and Robotics Ventures, Thailand), Phawat Borisuitsawat (AI and Robotics Ventures, Thailand), Teepakorn Tosawadi (AI and Robotics Ventures, Thailand), Pakcheera Choppradi (AI and Robotics Ventures, Thailand), Kasisdis Mahakijdechachai (AI and Robotics Ventures, Thailand), Supawit Vatathanavaro (AI and Robotics Ventures, Thailand), Worawit Saetan (AI and Robotics Ventures, Thailand), and Vasin Suttichaya (AI and Robotics Ventures, Thailand)</i>	
Simple In-place Data Augmentation for Surveillance Object Detection	7208
<i>Munkh-Erdene Otgonbold (United Arab Emirates University, United Arab Emirates), Ganzorig Batnasan (United Arab Emirates University, United Arab Emirates), and Munkhjargal Gochoo (United Arab Emirates University, United Arab Emirates)</i>	
Efficient Online Multi-Camera Tracking with Memory-Efficient Accumulated Appearance Features and Trajectory Validation	7217
<i>Lap Quoc Tran (Asilla Inc) and Huan Duc Vi (Asilla Inc)</i>	

Enhancing Road Object Detection in Fisheye Cameras: An Effective Framework Integrating SAHI and Hybrid Inference	7227
<i>Bao Tran Gia (University of Information Technology, VNU-HCM), Tuong Bui Cong Khanh (University of Information Technology, VNU-HCM), Hien Ho Trong (University of Information Technology, VNU-HCM), Thuyen Tran Doan (University of Information Technology, VNU-HCM), Tien Do (University of Information Technology, VNU-HCM), Duy-Dinh Le (University of Information Technology, VNU-HCM), and Thanh Duc Ngo (University of Information Technology, VNU-HCM)</i>	
OCMCTrack: Online Multi-Target Multi-Camera Tracking with Corrective Matching Cascade	7236
<i>Andreas Specker (Fraunhofer IOSB, Germany)</i>	
Road Object Detection Robust to Distorted Objects at the Edge Regions of Images	7245
<i>Wooksu Shin (Nota Inc., Republic of Korea), Donghyuk Choi (Nota Inc., Republic of Korea), Hancheol Park (Nota Inc., Republic of Korea), and Jeongho Kim (Nota Inc., Republic of Korea)</i>	
DeepLocalization: Using Change Point Detection for Temporal Action Localization	7252
<i>Mohammed Shaiqur Rahman (Iowa State University), Ibne Farabi Shihab (Iowa State University), Lynna Chu (Iowa State University), and Anuj Sharma (Iowa State University)</i>	
The 8th AI City Challenge	7261
<i>Shuo Wang (NVIDIA Corporation), David C. Anastasiu (Santa Clara University), Zheng Tang (NVIDIA Corporation), Ming-Ching Chang (University at Albany, SUNY), Yue Yao (Australian National University), Liang Zheng (Australian National University), Mohammed Shaiqur Rahman (Iowa State University), Meenakshi S. Arya (Iowa State University), Anuj Sharma (Iowa State University), Pranamesh Chakraborty (Indian Institute of Technology Kanpur), Sanjita Prajapati (Indian Institute of Technology Kanpur), Quan Kong (Woven by Toyota), Norimasa Kobori (Woven by Toyota), Munkhjargal Gochoo (United Arab Emirates University), Munkh-Erdene Otgonbold (United Arab Emirates University), Fady Alhajjar (United Arab Emirates University), Ganzorig Batnasan (United Arab Emirates University), Ping-Yang Chen (National Yang-Ming Chiao-Tung University), Jun-Wei Hsieh (National Yang-Ming Chiao-Tung University), Xunlei Wu (NVIDIA Corporation), Sameer Satish Pusegaonkar (NVIDIA Corporation), Yizhou Wang (NVIDIA Corporation), Sujit Biswas (NVIDIA Corporation), and Rama Chellappa (Johns Hopkins University)</i>	

3rd Workshop on Vision Datasets Understanding and DataCV Challenge

Improving Noisy Fine-Grained Datasets using Active Label Cleaning Framework	N/A
<i>Avik Pal (University of Amsterdam, Netherlands)</i>	

DTLLM-VLT: Diverse Text Generation for Visual Language Tracking Based on LLM	7283
<i>Xuchen Li (Institute of Automation, Chinese Academy of Sciences, China), Xiaokun Feng (Institute of Automation, Chinese Academy of Sciences, China), Shiyu Hu (University of Chinese Academy of Sciences, China), Meiqi Wu (University of Chinese Academy of Sciences, China), Dailing Zhang (Institute of Automation, Chinese Academy of Sciences, China), Jing Zhang (Institute of Automation, Chinese Academy of Sciences, China), and Kaiqi Huang (Chinese Academy of Sciences, China)</i>	
ALINA: Advanced Line Identification and Notation Algorithm	7293
<i>Mohammed Abdul Hafeez Khan (Florida Institute of Technology, USA), Parth Ganeriwala (Florida Institute of Technology, USA), Siddhartha Bhattacharyya (Florida Institute of Technology, USA), Natasha Neogi (NASA Langley Research Center, USA), and Raja Muthalagu (Birla Institute of Technology and Science Pilani, Dubai Campus, UAE)</i>	
Grounding Stylistic Domain Generalization with Quantitative Domain Shift Measures and Synthetic Scene Images	7303
<i>Yiran Luo (Arizona State University), Joshua Feinglass (Arizona State University), Tejas Gokhale (University of Maryland, Baltimore County), Kuan-Cheng Lee (Arizona State University), Chitta Baral (Arizona State University), and Yezhou Yang (Arizona State University)</i>	
A Survey of Video Datasets for Grounded Event Understanding	7314
<i>Kate Sanders (Johns Hopkins University, USA) and Benjamin Van Durme (Johns Hopkins University, USA)</i>	
DDOS: The Drone Depth and Obstacle Segmentation Dataset	7328
<i>Benedikt Kolbeinsson (Imperial College London) and Krystian Mikolajczyk (Imperial College London)</i>	
Classifier Guided Cluster Density Reduction for Dataset Selection	7338
<i>Cheng Chang (Layer 6 AI), Keyu Long (Layer 6 AI), Zijian Li (Layer 6 AI), and Himanshu Rai (Layer 6 AI)</i>	
Optimizing Object Detection via Metric-driven Training Data Selection	7348
<i>Changyuan Zhou (Onewo Space-Tech Service Co., Ltd.), Yumin Guo (Onewo Space-Tech Service Co., Ltd.), Qinxue Lv (Onewo Space-Tech Service Co., Ltd.), and Ji Yuan (Onewo Space-Tech Service Co., Ltd.)</i>	

New frontiers for zero-shot Image Captioning Evaluation (NICE)

NICE: CVPR 2023 Challenge on Zero-shot Image Captioning	7356
<i>Taehoon Kim (LG AI Research), Pyunghwan Ahn (LG AI Research), Sangyun Kim (LG AI Research), Sihaeng Lee (LG AI Research), Mark Marsden (Shutterstock), Alessandra Sala (Shutterstock), Seung Hwan Kim (LG AI Research), Bohyung Han (Seoul National University), Kyoung Mu Lee (Seoul National University), Honglak Lee (LG AI Research), Kyounghoon Bae (LG AI Research), Xiangyu Wu (Nanjing University of Science and Technology), Yi Gao (Nanjing University of Science and Technology), Hailiang Zhang (Nanjing University of Science and Technology), Yang Yang (Nanjing University of Science and Technology), Weili Guo (Nanjing University of Science and Technology), Jianfeng Lu (Nanjing University of Science and Technology), Youngtaek Oh (Korea Advanced Institute of Science and Technology), Jae Won Cho (Korea Advanced Institute of Science and Technology), Dong-jin Kim (Hanyang University), In So Kweon (Korea Advanced Institute of Science and Technology), Junmo Kim (Korea Advanced Institute of Science and Technology), Wooyoung Kang (Kakao Brain), Won Young Jhoo (Kakao Brain), Byungseok Roh (Kakao Brain), Jonghwan Mun (Kakao Brain), Solgil Oh (Wooribank), Kenan Emir Ak (Amazon Inc.), Gwang-Gook Lee (Amazon Inc.), Yan Xu (Amazon Inc.), Mingwei Shen (Amazon Inc.), Kyomin Hwang (Seoul National University), Wonsik Shin (Seoul National University), Kamin Lee (Seoul National University), Wonhark Park (Seoul National University), Dongkwan Lee (Seoul National University), Nojun Kwak (Seoul National University), Yujin Wang (Tsinghua University), Yimu Wang (University of Waterloo), Tiancheng Gu (University of Sydney), Xingchang Lv (Illinois Institute of Technology), and Mingmao Sun (University of California, Berkeley)</i>	
Technical Report of NICE Challenge at CVPR 2024: Caption Re-ranking Evaluation Using Ensembled CLIP and Consensus Scores	7366
<i>Kiyoong Jeong (Korea University), Woojun Lee (Korea University), Woongchan Nam (Korea University), Minjeong Ma (Korea University), and Pilsung Kang (Korea University)</i>	
Large-Scale Bidirectional Training for Zero-Shot Image Captioning	7373
<i>Taehoon Kim (LG AI Research), Mark Marsden (Shutterstock), Pyunghwan Ahn (LG AI Research), Sangyun Kim (LG AI Research), Sihaeng Lee (LG AI Research), Alessandra Sala (Shutterstock), and Seung Hwan Kim (LG AI Research)</i>	

2nd Workshop on Generative Models for Computer Vision

An Empty Room is All We Want: Automatic Defurnishing of Indoor Panoramas	7384
<i>Mira Slavcheva (Matterport), Dave Gausebeck (Matterport), Kevin Chen (Matterport), David Buchhofer (Matterport), Azwad Sabik (Matterport), Chen Ma (Matterport), Sachal Dhillon (Matterport), Olaf Brandt (Matterport), and Alan Dolhasz (Matterport)</i>	

ART-V: Auto-Regressive Text-to-Video Generation with Diffusion Models	7395
<i>Wenming Weng (University of Science and Technology of China), Ruoyu Feng (University of Science and Technology of China), Yanhui Wang (University of Science and Technology of China), Qi Dai (Microsoft Research Asia), Chunyu Wang (Microsoft Research Asia), Dacheng Yin (University of Science and Technology of China), Zhiyuan Zhao (Microsoft Research Asia), Kai Qiu (Microsoft Research Asia), Jianmin Bao (Microsoft Research Asia), Yuhui Yuan (Microsoft Research Asia), Chong Luo (Microsoft Research Asia), Yueyi Zhang (University of Science and Technology of China), and Zhiwei Xiong (University of Science and Technology of China)</i>	
Investigating the Effectiveness of Cross-Attention to Unlock Zero-Shot Editing of Text-to-Video Diffusion Models	7406
<i>Saman Motamed (INSAT Sofia University, Bulgaria), Wouter Van Gansbeke (INSAT Sofia University, Bulgaria), and Luc Van Gool (INSAT Sofia University, Bulgaria; ETH Zurich, Switzerland; KU Leuven, Belgium)</i>	
StereoDiffusion: Training-Free Stereo Image Generation Using Latent Diffusion Models	7416
<i>Lezhong Wang (Technical University of Denmark), Jeppe Revall Frisvad (Technical University of Denmark), Mark Bo Jensen (Technical University of Denmark), and Siavash Arjomand Bigdeli (Technical University of Denmark)</i>	
iEdit: Localised Text-guided Image Editing with Weak Supervision	7426
<i>Rumeysa Bodur (Imperial College London, UK), Erhan Gundogdu (Amazon), Binod Bhattarai (University of Aberdeen, UK), Tae-Kyun Kim (Imperial College London, UK; KAIST, South Korea), Michael Donoser (Amazon), and Loris Bazzani (Amazon)</i>	
OmniControlNet: Dual-stage Integration for Conditional Image Generation	7436
<i>Yilin Wang (Tsinghua University, China), Haiyang Xu (University of California San Diego, USA), Xiang Zhang (University of California San Diego, USA), Zeyuan Chen (University of California San Diego, USA), Zhizhou Sha (Tsinghua University, China), Zirui Wang (Princeton University, USA), and Zhuowen Tu (University of California San Diego, USA)</i>	
Robust Disaster Assessment from Aerial Imagery Using Text-to-Image Synthetic Data	7449
<i>Tarun Kalluri (UC San Diego), Jihyeon Lee (Google Research), Kihyuk Sohn (Google Research), Sahil Singla (Google Research), Manmohan Chandraker (UC San Diego), Joseph Xu (Google Research), and Jeremiah Liu (Google Research)</i>	
MixSyn: Compositional Image Synthesis with Fuzzy Masks and Style Fusion	7460
<i>Ilke Demir (Intel Labs) and Umur Aybars Ciftci (Binghamton University)</i>	
AI Art Neural Constellation: Revealing the Collective and Contrastive State of AI-Generated and Human Art	7470
<i>Faizan Farooq Khan (KAUST), Diana Kim (KAUST), Divyansh Jha (KAUST), Youssef Mohamed (KAUST), Hanna H Chang (KAUST), Ahmed Elgammal (Rutgers University), Luba Elliott (ELLUBA), and Mohamed Elhoseiny (KAUST)</i>	

GeoGen: Geometry-Aware Generative Modeling via Signed Distance Functions	7479
<i>Salvatore Esposito (University of Edinburgh), Qingshan Xu (Huazhong University of Science and Technology), Kacper Kania (Warsaw University of Technology), Charlie Hewitt (Microsoft), Octave Mariotti (University of Edinburgh), Lohit Petikam (Microsoft), Julien Valentin (Microsoft), Arno Onken (University of Edinburgh), and Oisin Mac Aodha (University of Edinburgh)</i>	
Salient Object-Aware Background Generation using Text-Guided Diffusion Models	7489
<i>Amir Erfan Eshratifar (Yahoo Research), Joao V.B. Soares (Yahoo Research), Kapil Thadani (Yahoo Research), Shaunak Mishra (Amazon), Mikhail Kuznetsov (Amazon), Yueh-Ning Ku (ByteDance), and Paloma de Juan (Yahoo Research)</i>	
Style Transfer for 2D Talking Head Generation	7500
<i>Trong Thang Pham (University of Arkansas, USA), Tuong Do (AIOZ, Singapore; University of Liverpool, UK), Nhat Le (AIOZ, Singapore), Ngan Le (University of Arkansas, USA), Hung Nguyen (AIOZ, Singapore), Erman Tjiputra (AIOZ, Singapore), Quang Tran (AIOZ, Singapore), and Anh Nguyen (University of Liverpool, UK)</i>	
LATENTMAN: Generating Consistent Animated Characters using Image Diffusion Models	7510
<i>Abdelrahman Eldesokey (KAUST, Saudi Arabia) and Peter Wonka (KAUST, Saudi Arabia)</i>	
Segmentation-Free Guidance for Text-to-Image Diffusion Models	7520
<i>Kambiz Azarian (Qualcomm AI Research, USA), Debasmit Das (Qualcomm AI Research, USA), Qiqi Hou (Qualcomm AI Research, USA), and Fatih Porikli (Qualcomm AI Research, USA)</i>	
Can Synthetic Plant Images From Generative Models Facilitate Rare Species Identification and Classification?	7530
<i>Debajyoti Dasgupta (Indian Institute of Technology Kharagpur, India), Arijit Mondal (Indian Institute of Technology Patna, India), and Partha P. Chakrabarti (Indian Institute of Technology Kharagpur, India)</i>	
Contrastive Clothing and Pose Generation for Cloth-Changing Person Re-Identification	7541
<i>Vuong D. Nguyen (University of Houston), Pranav Mantini (University of Houston), and Shishir K. Shah (University of Houston)</i>	
PQ-VAE: Learning Hierarchical Discrete Representations with Progressive Quantization	7550
<i>Lun Huang (Duke University, USA), Qiang Qiu (Purdue University, USA), and Guillermo Sapiro (Duke University, USA)</i>	
GenVideo: One-shot Target-image and Shape Aware Video Editing using T2I Diffusion Models ..	7559
<i>Sai Sree Harsha (Adobe Inc, USA), Ambareesh Revanur (Adobe Inc, USA), Dhwanit Agarwal (Adobe Inc, USA), and Shradha Agrawal (Adobe Inc, USA)</i>	
Efficient Exploration of Image Classifier Failures with Bayesian Optimization and Text-to-Image Models	7569
<i>Adrien Le Coz (IRT SystemX, France), Housseem Ouertatani (IRT SystemX, France), Stéphane Herbin (ONERA, France), and Faouzi Adjed (IRT SystemX, France)</i>	

MVDiff: Scalable and Flexible Multi-view Diffusion for 3D Object Reconstruction from Single-View	7579
<i>Emmanuelle Bourigault (University of Oxford) and Pauline Bourigault (Imperial College London)</i>	

OpenSUN3D: 2nd Workshop on Open-Vocabulary 3D Scene Understanding

AffordanceLLM: Grounding Affordance from Vision Language Models	7587
<i>Shengyi Qian (AWS AI), Weifeng Chen (AWS AI), Min Bai (AWS AI), Xiong Zhou (AWS AI), Zhuowen Tu (AWS AI), and Li Erran Li (AWS AI)</i>	
Zero-Shot Dual-Path Integration Framework for Open-Vocabulary 3D Instance Segmentation	7598
<i>Tri Ton (Korea Advanced Institute of Science and Technology, South Korea), Ji Woo Hong (Korea Advanced Institute of Science and Technology, South Korea), SooHwan Eom (Korea Advanced Institute of Science and Technology, South Korea), Jun Yeop Shim (Korea Advanced Institute of Science and Technology, South Korea), Junyeong Kim (Chung-Ang University, South Korea), and Chang D. Yoo (Korea Advanced Institute of Science and Technology, South Korea)</i>	

1st Workshop on Urban Scene Modeling: Where Vision Meets Photogrammetry and Graphics

AsymFormer: Asymmetrical Cross-Modal Representation Learning for Mobile Platform Real-Time RGB-D Semantic Segmentation	7608
<i>Siqi Du (Shenzhen University), Weixi Wang (Shenzhen University), Renzhong Guo (Shenzhen University), Ruisheng Wang (Shenzhen University; University of Calgary), and Shengjun Tang (Shenzhen University)</i>	
SimpliCity: Reconstructing Buildings with Simple Regularized 3D Models	7616
<i>Jean-Philippe Bauchet (LuxCarta Technology, France), Raphael Sulzer (LuxCarta Technology, France), Florent Lafarge (Inria Sophia Antipolis, France), and Yuliya Tarabalka (LuxCarta Technology, France)</i>	
ECLAIR: A High-Fidelity Aerial LiDAR Dataset for Semantic Segmentation	7627
<i>Iaroslav Melekhov (Sharper Shape, Finland), Anand Umashankar (Sharper Shape, Finland), Hyeon-Jin Kim (Sharper Shape, Finland), Vladislav Serkov (Sharper Shape, Finland), and Dusty Argyle (Sharper Shape, USA)</i>	
uTRAND: Unsupervised Anomaly Detection in Traffic Trajectories	7638
<i>Giacomo D'Amicantonio (Eindhoven University of Technology), Egor Bondarau (Eindhoven University of Technology), and Peter H.N. de With (Eindhoven University of Technology)</i>	
OpenTrench3D: A Photogrammetric 3D Point Cloud Dataset for Semantic Segmentation of Underground Utilities	7646
<i>Lasse H. Hansen (Aalborg University, Denmark), Simon B. Jensen (Aalborg University, Denmark), Mark P. Philipsen (Aalborg University, Denmark), Andreas Møgelmoose (Aalborg University, Denmark), Lars Bodum (Aalborg University, Denmark), and Thomas B. Moeslund (Aalborg University, Denmark)</i>	

1st Workshop on Dataset Distillation for Computer Vision

Exploring the Impact of Dataset Bias on Dataset Distillation	7656
<i>Yao Lu (Zhejiang University of Technology), Jianyang Gu (Zhejiang University), Xuguang Chen (Zhejiang University of Technology), Saeed Vahidian (Duke University), and Qi Xuan (Zhejiang University of Technology)</i>	
Generative Dataset Distillation: Balancing Global Structure and Local Details	7664
<i>Longzhen Li (Hokkaido University), Guang Li (Hokkaido University), Ren Togo (Hokkaido University), Keisuke Maeda (Hokkaido University), Takahiro Ogasawa (Hokkaido University), and Miki Haseyama (Hokkaido University)</i>	
AugData Distillation for Monocular 3D Human Pose Estimation	7672
<i>Jiman Kim (Samsung Research)</i>	
Coreset Selection for Object Detection	7682
<i>Hojun Lee (Seoul National University, Korea), Suyoung Kim (Seoul National University, Korea), Junhoo Lee (Seoul National University, Korea), Jaeyoung Yoo (NAVER WEBTOON AI, Korea), and Nojun Kwak (Seoul National University, Korea)</i>	
ATOM: Attention Mixer for Efficient Dataset Distillation	7692
<i>Samir Khaki (University of Toronto, Canada), Ahmad Sajedi (University of Toronto, Canada), Kai Wang (National University of Singapore, Singapore), Lucy Z. Liu (Royal Bank of Canada (RBC), Canada), Yuri A. Lawryshyn (University of Toronto, Canada), and Konstantinos N. Plataniotis (University of Toronto, Canada)</i>	
Dataset Condensation with Latent Quantile Matching	7703
<i>Wei Wei (University of Antwerp, Belgium), Tom De Schepper (University of Antwerp, Belgium), and Kevin Mets (University of Antwerp, Belgium)</i>	
Large-scale Dataset Pruning with Dynamic Uncertainty	7713
<i>Muyang He (Beijing Academy of Artificial Intelligence), Shuo Yang (University of Technology Sydney), Tiejun Huang (Peking University), and Bo Zhao (Beijing Academy of Artificial Intelligence)</i>	
DEEPDISTAL: Deepfake Dataset Distillation using Active Learning	7723
<i>Md Shohel Rana (Florida Gulf Coast University), Mohammad Nur Nobi (The University of Texas at San Antonio), and Andrew Sung (The University of Southern Mississippi)</i>	

Representation Learning with Very Limited Images: Zero-shot, Unsupervised, and Synthetic Learning in the Era of Big Models

Data-free Model Fusion with Generator Assistants	7731
<i>Luyao Shi (IBM Research), Prashanth Vijayaraghavan (IBM Research), and Ehsan Degan (IBM Research)</i>	
i-MAE: Are Latent Representations in Masked Autoencoders Linearly Separable?	7740
<i>Kevin Zhang (Peking University; KNQ.AI) and Zhiqiang Shen (Mohamed bin Zayed University of AI)</i>	

Enhancing 2D Representation Learning with a 3D Prior	7750
<i>Mehmet Aygun (University of Edinburgh), Prithviraj Dhar (Meta), Zhicheng Yan (Meta), Oisin Mac Aodha (University of Edinburgh), and Rakesh Ranjan (Meta)</i>	
Prompt Learning with One-Shot Setting based Feature Space Analysis in Vision-and-Language Models	7761
<i>Yuki Hirohashi (Omron Corp., Japan), Tsubasa Hirakawa (Chubu University, Japan), Takayoshi Yamashita (Chubu University, Japan), and Hironobu Fujiyoshi (Chubu University, Japan)</i>	
POPE: 6-DoF Promptable Pose Estimation of Any Object, in Any Scene, with One Reference	7771
<i>Zhiwen Fan (University of Texas at Austin), Panwang Pan (ByteDance Inc), Peihao Wang (University of Texas at Austin), Yifan Jiang (University of Texas at Austin), Dejie Xu (University of Texas at Austin), and Zhangyang Wang (University of Texas at Austin)</i>	
Federated Learning with a Single Shared Image	7782
<i>Sunny Soni (Universiteit van Amsterdam), Aaqib Saeed (TU Eindhoven), and Yuki M. Asano (Universiteit van Amsterdam)</i>	
'Eyes of a Hawk and Ears of a Fox': Part Prototype Network for Generalized Zero-Shot Learning	7791
<i>Joshua Feinglass (Arizona State University), Jayaraman J. Thiagarajan (Lawrence Livermore National Lab), Rushil Anirudh (Lawrence Livermore National Lab), T.S. Jayram (Lawrence Livermore National Lab), and Yezhou Yang (Arizona State University)</i>	
Vision-Language Pseudo-Labels for Single-Positive Multi-Label Learning	7799
<i>Xin Xing (University of Nebraska Omaha, USA), Zhexiao Xiong (Washington University in St. Louis, USA), Abby Stylianou (Saint Louis University, USA), Srikumar Sastry (Washington University in St. Louis, USA), Liyu Gong (Oracle Inc, USA), and Nathan Jacobs (Washington University in St. Louis, USA)</i>	
A Method of Moments Embedding Constraint and its Application to Semi-Supervised Learning .	7809
<i>Michael Majurski (National Institute of Standards and Technology, USA), Sumeet Menon (University of Maryland, Baltimore County, USA), Parniyan Favardin (University of Miami, USA), and David Chapman (University of Miami, USA)</i>	
PromptSync: Bridging Domain Gaps in Vision-Language Models through Class-Aware Prototype Alignment and Discrimination	7819
<i>Anant Khandelwal (Glance AI)</i>	

The Seventh International Workshop on Computer Vision for Physiological Measurement (CVPM) - Part 2

Toward Motion Robustness: A Masked Attention Regularization Framework in Remote Photoplethysmography	7829
<i>Pengfei Zhao (Shanghai Artificial Intelligence Laboratory, China), Qigong Sun (SenseTime Research, China), Xiaolin Tian (Xidian University, China), Yige Yang (Xidian University, China), Shuo Tao (Xidian University, China), Jie Cheng (Xidian University, China), and Jiantong Chen (Shanghai Artificial Intelligence Laboratory, China)</i>	

EarthVision: Large Scale Computer Vision for Remote Sensing Imagery - Part 2

Good at Captioning, Bad at Counting: Benchmarking GPT-4V on Earth Observation Data 7839
Chenhui Zhang (Massachusetts Institute of Technology) and Sherrie Wang (Massachusetts Institute of Technology)

Neural Rendering Intelligence - Part 2

DiCo-NeRF: Difference of Cosine Similarity for Neural Rendering of Fisheye Driving Scenes 7850
Jiho Choi (Jeonbuk National University, Republic of Korea), Gyutae Hwang (Jeonbuk National University, Republic of Korea), and Sang Jun Lee (Jeonbuk National University, Republic of Korea)

Efficient Large Vision Models - Part 2

EfficientViT-SAM: Accelerated Segment Anything Model Without Performance Loss 7859
Zhuoyang Zhang (Tsinghua University; NVIDIA), Han Cai (MIT; NVIDIA), and Song Han (MIT; NVIDIA)

Parameter Efficient Fine-tuning of Self-supervised ViTs without Catastrophic Forgetting 7864
Reza Akbarian Bafghi (University of Colorado, Boulder, USA), Nidhin Harilal (University of Colorado, Boulder, USA), Claire Monteleoni (University of Colorado, Boulder, USA), and Maziar Raissi (University of California, Riverside, USA)

6th Workshop and Competition on Affective Behavior Analysis in-the-wild - Part 2

Advanced Facial Analysis in Multi-Modal Data with Cascaded Cross-Attention based Transformer 7870
Jun-Hwa Kim (Konyang University, South Korea), Namho Kim (Korean Broadcasting System(KBS), South Korea), Minsoo Hong (Korean Broadcasting System(KBS), South Korea), and Chee Sun Won (Dongguk University, South Korea)

Improving Valence-Arousal Estimation with Spatiotemporal Relationship Learning and Multimodal Fusion	7878
--	------

Jun Yu (University of Science and Technology of China), Gongpeng Zhao (University of Science and Technology of China), Yongqi Wang (University of Science and Technology of China), Zhihong Wei (University of Science and Technology of China), Zerui Zhang (University of Science and Technology of China), Zhongpeng Cai (University of Science and Technology of China), Guochen Xie (University of Science and Technology of China), Jichao Zhu (University of Science and Technology of China), Wangyuan Zhu (University of Science and Technology of China), Shuoping Yang (University of Science and Technology of China), Yang Zheng (University of Science and Technology of China), Qingsong Liu (Unisound AI Technology Co., Ltd.), and Jiaen Liang (Unisound AI Technology Co., Ltd.)

Domain adaptation, Explainability and Fairness in AI for Medical Image Analysis (DEF-AI-MIA) - Part 2

Motion-aware Needle Segmentation in Ultrasound Images	7886
---	------

Raghavv Goel (Qualcomm AI Research, USA), Cecilia Morales (Carnegie Mellon University, USA), Manpreet Singh (AIM Intelligent Machines, USA), Artur Dubrawski (Carnegie Mellon University, CMU), John Galeotti (Carnegie Mellon University, USA), and Howie Choset (Carnegie Mellon University, USA)

The First Workshop on the Evaluation of Generative Foundation Models - Part 2

Towards Quantitative Evaluation Metrics for Image Editing Approaches	7892
--	------

Dana Cohen Hochberg (Amazon, Israel), Oron Anshel (Amazon, Israel), Alon Shoshan (Amazon, Israel), Igor Kviatkovsky (Amazon, Israel), Manoj Aggarwal (Amazon, Israel), and Gerard Medioni (Amazon, Israel)

ReMOVE: A Reference-free Metric for Object Erasure	7901
--	------

Aditya Chandrasekar (Indian Institute of Science, India), Goirik Chakrabarty (TCS Research, India), Jai Bardhan (TCS Research, India), Ramya Hebbalaguppe (TCS Research, India), and Prathosh AP (Indian Institute of Science, India)

AIS: Vision, Graphics and AI for Streaming - Part 2

CASR: Efficient Cascade Network Structure with Channel Aligned method for 4K Real-Time Single Image Super-Resolution	7911
--	------

Kihwan Yoon (The University of Seoul), Ganzorig Gankhuyag (Korea Electronics Technology Institute), Jinman Park (Korea Electronics Technology Institute), Haengseon Son (Korea Electronics Technology Institute), and Kyoungwon Min (Korea Electronics Technology Institute)

9th Workshop on Computer Vision for Microscopy Image Analysis - Part 2

Gene-Level Representation Learning via Interventional Style Transfer in Optical Pooled Screening	7921
<i>Mahtab Bigverdi (University of Washington, USA), Burkhard Hoeckendorf (Genentech, USA), Heming Yao (Genentech, USA), Phil Hanslovsky (Genentech, USA), Romain Lopez (Genentech, USA; Stanford University, USA), and David Richmond (Genentech, USA)</i>	

8th AI City Challenge - Part 2

PV-Cap: 3D Dynamic Scene Understanding Through Open Physics-based Vocabulary	7932
<i>Hidetomo Sakaino (AI-Image Group, Data Solution Dept., (FCJ.ABC) FPT Consulting Japan, FPT Software), Thao Nguyen Phuong (Waseda University), and Vinh Nguyen Duy (AI-Image Group, Data Solution Dept., (FCJ.ABC) FPT Consulting Japan, FPT Software)</i>	

3rd Workshop on Vision Datasets Understanding and DataCV Challenge - Part 2

Collaborative Blind Image Deblurring	7943
<i>Thomas Eboli (ENS Paris-Saclay, France), Jean-Michel Morel (City University of Hong Kong, Hong Kong), and Gabriele Facciolo (ENS Paris-Saclay, France)</i>	
OpenStory: A Large-Scale Open-Domain Dataset for Subject-Driven Visual Storytelling	7953
<i>Zilyu Ye (South China University of Technology), Jinxiu Liu (South China University of Technology), JinJin Cao (Westlake University), Zhiyang Chen (Foundation Model Research Center, CASIA; Westlake University), Ziwei Xuan (OPPO US Research Center), Mingyuan Zhou (OPPO US Research Center), Qi Liu (South China University of Technology), and Guo-Jun Qi (Westlake University)</i>	

2nd Workshop on Generative Models for Computer Vision - Part 2

Gaussian Splatting Decoder for 3D-aware Generative Adversarial Networks	7963
<i>Florian Barthel (Humboldt University of Berlin; Fraunhofer Heinrich Hertz Institute, Germany), Arian Beckmann (Fraunhofer Heinrich Hertz Institute, Germany), Wieland Morgenstern (Fraunhofer Heinrich Hertz Institute, Germany), Anna Hilsmann (Fraunhofer Heinrich Hertz Institute, Germany), and Peter Eisert (Humboldt University of Berlin; Fraunhofer Heinrich Hertz Institute, Germany)</i>	

Agriculture-Vision: Challenges & Opportunities for Computer Vision in Agriculture - Part 2

AnimalFormer: Multimodal Vision Framework for Behavior-based Precision Livestock Farming .	7973
<i>Ahmed Qazi (Tibbling Technologies, USA), Taha Razzaq (Tibbling Technologies, USA), and Asim Iqbal (Tibbling Technologies, USA)</i>	

IrrNet: Spatio-Temporal Segmentation Guided Classification for Irrigation Mapping	7983
<i>Oishee Bintey Hoque (University of Virginia)</i>	

The 7th Workshop on Efficient Deep Learning for Computer Vision (ECV24) - Part 2

Masked Autoencoders are Secretly Efficient Learners	7986
<i>Zihao Wei (University of Michigan, Ann Arbor), Chen Wei (Johns Hopkins University), Jieru Mei (Johns Hopkins University), Yutong Bai (Johns Hopkins University), Zeyu Wang (UC Santa Cruz), Xianhang Li (UC Santa Cruz), Hongru Zhu (Johns Hopkins University), Huiyu Wang (Meta), Alan Yuille (Johns Hopkins University), Yuyin Zhou (UC Santa Cruz), and Cihang Xie (UC Santa Cruz)</i>	
MA-AVT: Modality Alignment for Parameter-Efficient Audio-Visual Transformers	7996
<i>Tanvir Mahmud (The University of Texas at Austin), Shentong Mo (Carnegie Mellon University), Yapeng Tian (The University of Texas at Dallas), and Diana Marculescu (The University of Texas at Austin)</i>	
ELSA: Exploiting Layer-wise N:M Sparsity for Vision Transformer Acceleration	8006
<i>Ning-Chi Huang (National Yang Ming Chiao Tung University, Taiwan), Chi-Chih Chang (National Yang Ming Chiao Tung University, Taiwan), Wei-Cheng Lin (National Yang Ming Chiao Tung University, Taiwan), Endri Taka (University of Texas at Austin, USA), Diana Marculescu (University of Texas at Austin, USA), and Kai-Chiang Wu (National Yang Ming Chiao Tung University, Taiwan)</i>	
FlowIBR: Leveraging Pre-Training for Efficient Neural Image-Based Rendering of Dynamic Scenes	8016
<i>Marcel B��sching (KTH Royal Institute of Technology, Sweden), Josef Bengtson (Chalmers University of Technology, Sweden), David Nilsson (Chalmers University of Technology, Sweden), and M��rten Bj��rkman (KTH Royal Institute of Technology, Sweden)</i>	
Improving the Efficiency-Accuracy Trade-off of DETR-Style Models in Practice	8027
<i>Yumin Suh (NEC Laboratories America), Dongwan Kim (Seoul National University), Abhishek Aich (NEC Laboratories America), Samuel Schuster (NEC Laboratories America), Jong-Chyi Su (NEC Laboratories America), Bohyung Han (Seoul National University), and Manmohan Chandraker (NEC Laboratories America)</i>	
Multi-Objective Hardware Aware Neural Architecture Search using Hardware Cost Diversity	8032
<i>Nilotpal Sinha (University of Luxembourg), Peyman Rostami (University of Luxembourg), Abd El Rahman Shabayek (University of Luxembourg), Anis Kacem (University of Luxembourg), and Djamila Aouada (University of Luxembourg)</i>	
Cache and Reuse: Rethinking the Efficiency of On-device Transfer Learning	8040
<i>Yuedong Yang (The University of Texas at Austin), Hung-Yueh Chiang (The University of Texas at Austin), Guihong Li (The University of Texas at Austin), Diana Marculescu (The University of Texas at Austin), and Radu Marculescu (The University of Texas at Austin)</i>	

SuperLoRA: Parameter-Efficient Unified Adaptation for Large Vision Models	8050
<i>Xiangyu Chen (University of Kansas), Jing Liu (Mitsubishi Electric Research Laboratories), Ye Wang (Mitsubishi Electric Research Laboratories), Pu Wang (Mitsubishi Electric Research Laboratories), Matthew Brand (Mitsubishi Electric Research Laboratories), Guanghui Wang (Mitsubishi Electric Research Laboratories), and Toshiaki Koike-Akino (Mitsubishi Electric Research Laboratories)</i>	
Data-Efficient and Robust Task Selection for Meta-Learning	8056
<i>Donglin Zhan (Columbia University) and James Anderson (Columbia University)</i>	
The Revenge of BiSeNet: Efficient Multi-Task Image Segmentation	8066
<i>Gabriele Rosi (Politecnico di Torino; Focoos AI), Claudia Cuttano (Politecnico di Torino), Niccolò Cavagnero (Politecnico di Torino), Giuseppe Averta (Politecnico di Torino; Focoos AI), and Fabio Cermelli (Focoos AI)</i>	
ShiftAddAug: Augment Multiplication-Free Tiny Neural Network with Hybrid Computation	8075
<i>Yipin Guo (Zhejiang University, China), Zihao Li (Zhejiang University, China), Yilin Lang (Zhejiang University, China), and Qinyuan Ren (Zhejiang University, China)</i>	
LVS: A Learned Video Storage for Fast and Efficient Video Understanding	8085
<i>Yunghee Lee (Korea Advanced Institute of Science and Technology (KAIST)) and Jongse Park (Korea Advanced Institute of Science and Technology (KAIST))</i>	
Block Selective Reprogramming for On-device Training of Vision Transformers	8094
<i>Sreetama Sarkar (University of Southern California, Los Angeles, USA), Souvik Kundu (Intel Labs, San Diego, USA), Kai Zheng (University of Southern California, Los Angeles, USA), and Peter A. Bearel (University of Southern California, Los Angeles, USA)</i>	
Selectively Dilated Convolution for Accuracy-Preserving Sparse Pillar-based Embedded 3D Object Detection	8104
<i>Seongmin Park (Hanyang University), Minjae Lee (Hanyang University), Junwon Choi (Seoul National University), and Jungwook Choi (Hanyang University)</i>	

Author Index