

# **2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW 2023)**

**Paris, France  
2-6 October 2023**

**Pages 1-684**



**IEEE Catalog Number: CFP2391A-POD  
ISBN: 979-8-3503-0745-0**

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

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

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

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

IEEE Catalog Number:	CFP2391A-POD
ISBN (Print-On-Demand):	979-8-3503-0745-0
ISBN (Online):	979-8-3503-0744-3
ISSN:	2473-9936

**Additional Copies of This Publication Are Available From:**

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

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# 2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW) ICCVW 2023

## Table of Contents

Organizers .....	lxxii
Keynotes .....	lxxiv

### 1st Workshop on Scene Graphs and Graph Representation Learning (SG2RL)

Polygon Detection for Room Layout Estimation using Heterogeneous Graphs and Wireframes .....	1
<i>David Gillsjö (Lund University, Sweden), Gabrielle Flood (Lund University, Sweden), and Kalle Åström (Lund University, Sweden)</i>	
Fine-Grained is Too Coarse: A Novel Data-Centric Approach for Efficient Scene Graph Generation .....	11
<i>Maëlic Neau (Flinders University, Australia; Ecole Nationale d'Ingénieurs de Brest, France), Paulo E. Santos (Flinders University, Australia), Anne-Gwenn Bosser (Ecole Nationale d'Ingénieurs de Brest, France), and Cédric Buche (Ecole Nationale d'Ingénieurs de Brest, France)</i>	
Knowledge Informed Sequential Scene Graph Verification Using VQA .....	21
<i>Dao Thauwin (ONERA - DTIS, Université Paris Saclay) and Stéphane Herbin (ONERA - DTIS, Université Paris Saclay)</i>	
DeepCut: Unsupervised Segmentation Using Graph Neural Networks Clustering .....	32
<i>Amit Aflalo (Weizmann Institute, Israel), Shai Bagon (Weizmann Institute, Israel), Tamar Kashti (Weizmann Institute, Israel), and Yonina Eldar (Weizmann Institute, Israel)</i>	
nuScenes Knowledge Graph - A Comprehensive Semantic Representation of Traffic Scenes for Trajectory Prediction .....	42
<i>Leon Mlodzian (Heinrich Heine University Düsseldorf), Zhigang Sun (Bosch Center for Artificial Intelligence), Hendrik Berkemeyer (Robert Bosch GmbH; University of Osnabrück), Sebastian Monka (Bosch Center for Artificial Intelligence; Robert Bosch GmbH), Zixu Wang (Bosch Center for Artificial Intelligence; Technical University of Munich), Stefan Dietze (Heinrich Heine University Düsseldorf), Lavdim Halilaj (Bosch Center for Artificial Intelligence; Robert Bosch GmbH), and Juergen Luetttin (Bosch Center for Artificial Intelligence; Robert Bosch GmbH)</i>	

Relational Prior Knowledge Graphs for Detection and Instance Segmentation .....	53
<i>Osman Ülger (University of Amsterdam), Yu Wang (TomTom), Ysbrand Galama (TomTom), Sezer Karaoglu (University of Amsterdam), Theo Gevers (University of Amsterdam), and Martin R. Oswald (University of Amsterdam)</i>	
Haystack: A Panoptic Scene Graph Dataset to Evaluate Rare Predicate Classes .....	62
<i>Julian Lorenz (University of Augsburg, Germany), Florian Barthel (University of Augsburg, Germany), Daniel Kienzle (University of Augsburg, Germany), and Rainer Lienhart (University of Augsburg, Germany)</i>	
Exploring the Road Graph in Trajectory Forecasting for Autonomous Driving .....	71
<i>Rémy Sun (Université Côte d’Azur, France), Diane Lingrand (Université Côte d’Azur, France), and Frédéric Precioso (Université Côte d’Azur, France)</i>	
Dynamic Scene Graph Representation for Surgical Video .....	81
<i>Felix Holm (Technical University Munich, Germany), Ghazal Ghazaei (Carl Zeiss AG, Germany), Tobias Czempiel (Technical University Munich, Germany), Ege Özsoy (Technical University Munich, Germany), Stefan Saur (Carl Zeiss Meditec AG, Germany), and Nassir Navab (Technical University Munich, Germany)</i>	
SceneGenie: Scene Graph Guided Diffusion Models for Image Synthesis .....	88
<i>Azade Farshad (Technical University of Munich; MCML), Yousef Yeganeh (Technical University of Munich; MCML), Yu Chi (Technical University of Munich), Chengzhi Shen (Technical University of Munich), Böjrn Ommer (Ludwig Maximilian University of Munich), and Nassir Navab (Technical University of Munich)</i>	

## **4th Visual Inductive Priors for Data-Efficient Deep Learning Workshop (VIPriors)**

Padding Aware Neurons .....	99
<i>Dario Garcia-Gasulla (Universitat Politècnica de Catalunya - Barcelona TECH (UPC), Barcelona Supercomputing Center (BSC)), Victor Gimenez-Abalos (Barcelona Supercomputing Center (BSC)), and Pablo Martin-Torres (Barcelona Supercomputing Center (BSC))</i>	
Geometric Superpixel Representations for Efficient Image Classification with Graph Neural Networks .....	109
<i>Radu A. Cosma (University of Amsterdam, Netherlands), Lukas Knobel (University of Amsterdam, Netherlands), Putri van der Linden (University of Amsterdam, Netherlands), David M. Knigge (University of Amsterdam, Netherlands), and Erik J. Bekkers (University of Amsterdam, Netherlands)</i>	
Using and Abusing Equivariance .....	119
<i>Tom Edixhoven (Delft University of Technology), Attila Lengyel (Delft University of Technology), and Jan C. van Gemert (Delft University of Technology)</i>	

DFM-X: Augmentation by Leveraging Prior Knowledge of Shortcut Learning .....	129
<i>Shunxin Wang (University of Twente, The Netherlands), Christoph Brune (University of Twente, The Netherlands), Raymond Veldhuis (University of Twente, The Netherlands), and Nicola Strisciuglio (University of Twente, The Netherlands)</i>	
No Data Augmentation? Alternative Regularizations for Effective Training on Small Datasets.....	139
<i>Lorenzo Brigato (University of Bern, Switzerland) and Stavroula Mougiakakou (University of Bern, Switzerland)</i>	
COSE: A Consistency-Sensitivity Metric for Saliency on Image Classification .....	149
<i>Rangel Daroya (University of Massachusetts Amherst), Aaron Sun (University of Massachusetts Amherst), and Subhransu Maji (University of Massachusetts Amherst)</i>	
Video BagNet: Short Temporal Receptive Fields Increase Robustness in Long-Term Action Recognition .....	159
<i>Ombretta Strafforello (Delft University of Technology, TNO), Xin Liu (Delft University of Technology), Klamer Schutte (TNO), and Jan van Gemert (Delft University of Technology)</i>	
PARTICLE: Part Discovery and Contrastive Learning for Fine-Grained Recognition .....	167
<i>Oindrila Saha (UMass Amherst, Amherst) and Subhransu Maji (UMass Amherst, Amherst)</i>	
Self-Supervised Learning of Contextualized Local Visual Embeddings .....	177
<i>Thalles Silva (University of Campinas, Brazil), Helio Pedrini (University of Campinas, Brazil), and Adín Ramírez (University of Oslo, Norway)</i>	
DeepVAT: A Self-Supervised Technique for Cluster Assessment in Image Datasets .....	187
<i>Alokendu Mazumder (Indian Institute of Science Bangalore, India), Tirthajit Baruah (Indian Institute of Science Bangalore, India), Akash Kumar Singh (Indian Institute of Science Education and Research Bhopal, India), Pagadala Krishna Murthy (Indian Institute of Science Bangalore, India), Vishwajeet Pattanaik (Indian Institute of Science Bangalore, India), and Punit Rathore (Indian Institute of Science Bangalore, India)</i>	
RV-VAE: Integrating Random Variable Algebra into Variational Autoencoders .....	196
<i>Vassilis C. Nicodemou (University of Crete, Greece AND Foundation for Research and Technology Hellas, Greece), Iason Oikonomidis (Foundation for Research and Technology Hellas, Greece), and Antonis Argyros (University of Crete, Greece AND Foundation for Research and Technology Hellas, Greece)</i>	
Geometric Contrastive Learning .....	206
<i>Yeskendir Koishchenov (University of Amsterdam, Netherlands), Sharvaree Vadgama (University of Amsterdam, Netherlands), Riccardo Valperga (University of Amsterdam, Netherlands), and Erik J. Bekkers (University of Amsterdam, Netherlands)</i>	
Good Fences Make Good Neighbours .....	216
<i>Imanol G. Estepa (Universitat de Barcelona, Spain), Jesús Rodríguez-de-Vera (Universitat de Barcelona, Spain), Bhalaji Nagarajan (Universitat de Barcelona, Spain), and Petia Radeva (Universitat de Barcelona, Computer Vision Center, Spain)</i>	

Data Efficient Single Image Dehazing via Adversarial Auto-Augmentation and Extended Atmospheric Scattering Model .....	227
<i>Pranjay Shyam (Faurecia IRYStec Inc., Canada) and HyunJin Yoo (Faurecia IRYStec Inc., Canada)</i>	
Distilling Part-Whole Hierarchical Knowledge from a Huge Pretrained Class Agnostic Segmentation Framework .....	238
<i>Ahmed Radwan (University of British Columbia, Canada) and Mohamed S. Shehata (University of British Columbia, Canada)</i>	
Logarithm-Transform Aided Gaussian Sampling for Few-Shot Learning .....	247
<i>Vaibhav Ganatra (BITS Pilani, India)</i>	
InterAug: A Tuning-Free Augmentation Policy for Data-Efficient and Robust Object Detection.....	253
<i>Kowshik Thopalli (Lawrence Livermore National Laboratory, USA, USA), Devi S (SRM Institute of Science and Technology, India), and Jayaraman J. Thiagarajan (Lawrence Livermore National Laboratory, USA)</i>	

## What is Next in Multimodal Foundation Models? (MMFM)

Enhancing CLIP with GPT-4: Harnessing Visual Descriptions as Prompts .....	262
<i>Mayug Maniparambil (Dublin City University, Ireland), Chris Vorster (Dublin City University, Ireland), Derek Molloy (Dublin City University, Ireland), Noel Murphy (Dublin City University, Ireland), Kevin McGuinness (Dublin City University, Ireland), and Noel E. O'Connor (Dublin City University, Ireland)</i>	
Retrieving-to-Answer: Zero-Shot Video Question Answering with Frozen Large Language Models.....	272
<i>Junting Pan (The Chinese University of Hong Kong; Shanghai AI Laboratory), Ziyi Lin (The Chinese University of Hong Kong), Yuying Ge (University of Hong Kong), Xiatian Zhu (University of Surrey), Renrui Zhang (The Chinese University of Hong Kong; Shanghai AI Laboratory), Yi Wang (Shanghai AI Laboratory), Yu Qiao (Shanghai AI Laboratory), and Hongsheng Li (The Chinese University of Hong Kong)</i>	
Interaction-Aware Prompting for Zero-Shot Spatio-Temporal Action Detection .....	284
<i>Wei-Jhe Huang (National Tsing Hua University, Taiwan), Jheng-Hsien Yeh (National Tsing Hua University, Taiwan), Min-Hung Chen (NVIDIA), Gueter Josmy Faure (National Taiwan University), and Shang-Hong Lai (National Tsing Hua University, Taiwan)</i>	
ClipCrop: Conditioned Cropping Driven by Vision-Language Model .....	294
<i>Zhihang Zhong (The University of Tokyo, Japan; National Institute of Informatics, Japan), Mingxi Cheng (Microsoft Research), Zhirong Wu (Microsoft Research), Yuhui Yuan (Microsoft Research), Yinqiang Zheng (The University of Tokyo, Japan), Ji Li (Microsoft Research), Han Hu (Microsoft Research), Stephen Lin (Microsoft Research), Yoichi Sato (The University of Tokyo, Japan), and Imari Sato (The University of Tokyo, Japan; National Institute of Informatics, Japan)</i>	

Painter: Teaching Auto-Regressive Language Models to Draw Sketches .....	305
<i>Reza Pourreza (Qualcomm AI Research, USA), Apratim Bhattacharyya (Qualcomm AI Research, USA), Sunny Panchal (Qualcomm AI Research, Canada), Mingu Lee (Qualcomm AI Research, USA), Pulkit Madan (Qualcomm AI Research, Canada), and Roland Memisevic (Qualcomm AI Research, Canada)</i>	
Video Attribute Prototype Network: A New Perspective for Zero-Shot Video Classification .....	315
<i>Bo Wang (Beijing University of Posts and Telecommunications, China), Kaili Zhao (Beijing University of Posts and Telecommunications, China), Hongyang Zhao (Beijing University of Posts and Telecommunications, China), Shi Pu (Independent Researcher, China), Bo Xiao (Beijing University of Posts and Telecommunications, China), and Jun Guo (Beijing University of Posts and Telecommunications, China)</i>	
Video-and-Language (VidL) models and their cognitive relevance .....	325
<i>Anne Zonneveld (University of Amsterdam), Albert Gatt (Utrecht University), and Iacer Calixto (Amsterdam UMC, University of Amsterdam)</i>	
Towards an Exhaustive Evaluation of Vision-Language Foundation Models .....	339
<i>Emmanuelle Salin (Aix Marseille Univ, Université de Toulon, CNRS, LIS, France), Stéphane Ayache (Aix Marseille Univ, Université de Toulon, CNRS, LIS, France), and Benoit Favre (Aix Marseille Univ, Université de Toulon, CNRS, LIS, France)</i>	
Coarse to Fine Frame Selection for Online Open-Ended Video Question Answering .....	353
<i>Vidyaranya Nuthalapati (n/a) and Anirudh Tunga (Blue Wave AI Labs)</i>	

## **Workshop and Challenge on DeepFake Analysis and Detection (DFAD)**

FIVA: Facial Image and Video Anonymization and Anonymization Defense .....	362
<i>Felix Rosberg (Berge Consulting, Sweden), Eren Erdal Aksoy (Halmstad University, Sweden), Cristofer Englund (Halmstad University, Sweden), and Fernando Alonso-Fernandez (Halmstad University, Sweden)</i>	
A Comprehensive Framework for Evaluating Deepfake Generators: Dataset, Metrics Performance, and Comparative Analysis .....	372
<i>Sahar Hussein (Eurecom Research Center, France; Docaposte Biometrics Lab, France) and Jean-Luc Dugelay (Eurecom Research Center, France)</i>	
Online Detection of AI-Generated Images .....	382
<i>David C. Epstein (Adobe Inc.), Ishan Jain (Adobe Inc.), Oliver Wang (Adobe Inc.), and Richard Zhang (Adobe Inc.)</i>	
WaterLo: Protect Images from Deepfakes Using Localized Semi-Fragile Watermark .....	393
<i>Nicolas Beuve (Univ. Rennes, INSA Rennes, CNRS, France), Wassim Hamidouche (Technology Innovation Institute, UAE), and Olivier Déforges (Univ. Rennes, INSA Rennes, CNRS, France)</i>	
TrainFors: A Large Benchmark Training Dataset for Image Manipulation Detection and Localization .....	403
<i>Soumyaroop Nandi (USC Information Sciences Institute, USA), Prem Natarajan (USC Information Sciences Institute, USA), and Wael Abd-Almageed (USC Information Sciences Institute, USA)</i>	

Undercover Deepfakes: Detecting Fake Segments in Videos .....	415
<i>Sanjay Saha (National University of Singapore, Singapore), Rashindrie Perera (University of Melbourne, Australia), Sachith Seneviratne (University of Melbourne, Australia), Tamasha Malepathirana (University of Melbourne, Australia), Sanka Rasnayaka (National University of Singapore, Singapore), Deshani Geethika (University of Melbourne, Australia), Terence Sim (National University of Singapore, Singapore), and Saman Halgamuge (University of Melbourne, Australia)</i>	
Revisiting Generalizability in Deepfake Detection: Improving Metrics and Stabilizing Transfer .....	426
<i>Sarthak Kamat (University of California, Berkeley), Shruti Agarwal (Adobe Research), Trevor Darrell (University of California, Berkeley), and Anna Rohrbach (University of California, Berkeley)</i>	
Learning Interpretable Forensic Representations via Local Window Modulation .....	436
<i>Sowmen Das (University of Cambridge, UK) and Md. Ruhul Amin (Fordham University, USA)</i>	
Detecting Images Generated by Deep Diffusion Models Using Their Local Intrinsic Dimensionality .....	448
<i>Peter Lorenz (Fraunhofer ITWM), Ricard L. Durall (Fraunhofer ITWM), and Janis Keuper (Fraunhofer ITWM; IMLA - Offenburg University)</i>	
Deepfakes Signatures Detection in the Handcrafted Features Space .....	460
<i>Assia Hamadene (Morsli Abdellah University of Tipaza, Algeria), Abdeldjalil Ouahabi (UMr 1253, iBrain, Inserm, Université de Tours, France), and Abdenour Hadid (Sorbonne Center for Artificial Intelligence, Sorbonne University Abu Dhabi, UAE)</i>	
Interpretable-Through-Prototypes Deepfake Detection for Diffusion Models .....	467
<i>Agil Aghasanli (Lancaster University, United Kingdom), Dmitry Kangin (Lancaster University, United Kingdom), and Plamen Angelov (Lancaster University, United Kingdom)</i>	
Attending Generalizability in Course of Deep Fake Detection by Exploring Multi-Task Learning .....	475
<i>Pranav Balaji (BITS Pilani, India), Abhijit Das (BITS Pilani, India), Srijan Das (University of North Carolina at Charlotte, United States), and Antitza Dantcheva (INRIA, France)</i>	

## **8th Workshop on Computer Vision in Plant Phenotyping and Agriculture (CVPPA)**

Improving Deep Learning on Hyperspectral Images of Grain by Incorporating Domain Knowledge from Chemometrics .....	485
<i>Ole-Christian Galbo Engström (University of Copenhagen, Denmark; FOSS Analytical A/S, Denmark), Erik Schou Dreier (FOSS Analytical A/S, Denmark), Birthe Møller Jespersen (University of Copenhagen, Denmark), and Kim Steenstrup Pedersen (University of Copenhagen, Denmark)</i>	
An Interpretable Framework to Characterize Compound Treatments on Filamentous Fungi Using Cell Painting and Deep Metric Learning .....	495
<i>Laurent Lejeune (Bayer Crop Science, France), Morgane Roussin (Bayer Crop Science, France), Bruno Leggio (Bayer Crop Science, France), and Aurelia Vernay (Bayer Crop Science, France)</i>	



Weed Mapping with Convolutional Neural Networks on High Resolution Whole-Field Images ....	505
<i>Yuemin Wang (University of Saskatchewan, Canada), Thuan Ha (University of Saskatchewan, Canada), Kathryn Aldridge (University of Saskatchewan, Canada), Hema Duddu (University of Saskatchewan, Canada), Steve Shirtliffe (University of Saskatchewan, Canada), and Ian Stavness (University of Saskatchewan, Canada)</i>	
Non-Destructive Infield Quality Estimation of Strawberries Using Deep Architectures .....	515
<i>Cees Jol (Delft University of Technology, The Netherlands), Junhan Wen (Delft University of Technology, The Netherlands), and Jan van Gemert (Delft University of Technology, The Netherlands)</i>	
Estimation of Crop Production by Fusing Images and Crop Features .....	525
<i>Ángela Casado-García (Universidad de La Rioja, Spain), Jónathan Heras (Universidad de La Rioja, Spain), Jon Miranda-Apodaca (University of the Basque Country, Spain), Xabier Simon Martínez-Goñi (University of the Basque Country, Spain), and Usue Pérez-López (University of the Basque Country, Spain)</i>	
Plant Root Occlusion Inpainting with Generative Adversarial Network .....	531
<i>Hao Song (University of Saskatchewan, Canada), Karim Panjvani (Global Institute For Food Security, Canada), Zhigang Liu (Global Institute For Food Security, Canada), Huzaifa Amar (Global Institute For Food Security, Canada), Leon Kochian (Global Institute For Food Security, Canada), Shengjian Ye (Aquatic and Crop Resource Development, National Research Council Canada), Xuan Yang (Aquatic and Crop Resource Development, National Research Council Canada, Canada), J. Allan Feurtado (Aquatic and Crop Resource Development, National Research Council Canada, Canada), Krunal Chavda (University of Saskatchewan, Canada), Karina Angela Chimbo Huatatoca (University of Saskatchewan, Canada), and Mark Eramian (University of Saskatchewan, Canada)</i>	
A new Large Dataset and a Transfer Learning Methodology for Plant Phenotyping in Vertical Farms .....	540
<i>Nico Sama (DeepPlants), Etienne David (Agricola Moderna), Simone Rossetti (University of Rome Sapienza; DeepPlants), Alessandro Antona (Agricola Moderna), Benjamin Franchetti (Agricola Moderna), and Fiora Pirri (DeepPlants; University of Rome Sapienza)</i>	
Deep Learning for Apple Fruit Quality Inspection Using X-Ray Imaging .....	552
<i>Astrid Tempelaere (KU Leuven, Belgium), Leen Van Doorselaer (KU Leuven, Belgium), Jiaqi He (KU Leuven, Belgium), Pieter Verboven (KU Leuven, Belgium), Tinne Tuytelaars (KU Leuven, Belgium), and Bart Nicolai (KU Leuven, Belgium)</i>	
Deep Learning Based 3d Reconstruction for Phenotyping of Wheat Seeds: a Dataset, Challenge, and Baseline Method .....	561
<i>Vsevolod Cherepashkin (Forschungszentrum Jülich IBG-2, Germany), Erenus Yildiz (Forschungszentrum Jülich IAS-8, Germany), Andreas Fischbach (Forschungszentrum Jülich IBG-2, Germany), Leif Kobbelt (RWTH Aachen University, Germany), and Hanno Scharr (Forschungszentrum Jülich IAS-8, Germany)</i>	

Analyzing the Behavior of Cauliflower Harvest-Readiness Models by Investigating Feature Relevance	572
<i>Niklas Penzel (Friedrich Schiller University Jena, Germany), Jana Kierdorf (Institute of Geodesy and Geoinformation, University of Bonn, Germany), Ribana Roscher (Institute of Geodesy and Geoinformation, University of Bonn; Forschungszentrum Jülich GmbH, Germany), and Joachim Denzler (Friedrich Schiller University Jena, Germany)</i>	
Semantic Segmentation of Crops and Weeds with Probabilistic Modeling and Uncertainty Quantification	582
<i>Ekin Celikkan (GFZ German Research Centre for Geosciences; Humboldt-Universität zu Berlin, Germany), Mohammadmehdi Saberioon (GFZ German Research Centre for Geosciences, Germany), Martin Herold (GFZ German Research Centre for Geosciences, Germany), and Nadja Klein (Technische Universität Dortmund, Germany)</i>	
Class-Incremental Learning of Plant and Disease Detection: Growing Branches with Knowledge Distillation	593
<i>Mathieu Pagé-Fortin (Laval University, Canada)</i>	
Adapting Vision Foundation Models for Plant Phenotyping	604
<i>Feng Chen (The University of Edinburgh, U.K.), Mario Valerio Giuffrida (The University of Nottingham, U.K.), and Sotirios A. Tsaftaris (The University of Edinburgh, U.K.)</i>	
Group-Conditional Conformal Prediction via Quantile Regression Calibration for Crop and Weed Classification	614
<i>Paul Melki (IMS, CNRS, University of Bordeaux; EXXACT Robotics, France), Lionel Bombrun (IMS, CNRS, University of Bordeaux; Bordeaux Sciences Agro, France), Boubacar Diallo (EXXACT Robotics, France), Jérôme Dias (EXXACT Robotics, France), and Jean-Pierre Da Costa (IMS, CNRS, University of Bordeaux; Bordeaux Sciences Agro, France)</i>	
Vision-Based Monitoring of the Short-Term Dynamic Behaviour of Plants for Automated Phenotyping	624
<i>Nikolaus Wagner (University of Lincoln, United Kingdom) and Grzegorz Cielniak (University of Lincoln, United Kingdom)</i>	
Rapid Tomato DUS Trait Analysis Using an Optimized Mobile-Based Coarse-to-Fine Instance Segmentation Algorithm	634
<i>Dan Jeric Arcega Rustia (Wageningen University &amp; Research, The Netherlands), Guido Alexander Jansen (Wageningen University &amp; Research, The Netherlands), Selwin Hageraats (Wageningen University &amp; Research, The Netherlands), Joseph Peller (Wageningen University &amp; Research, The Netherlands), Rick van de Zedde (Wageningen University &amp; Research, The Netherlands), Cécile Marchennay (Naktuinbouw, The Netherlands), Wim Sangster (Naktuinbouw, The Netherlands), and Gosia Blokker (Naktuinbouw, The Netherlands)</i>	
Pollinators as Data Collectors: Estimating Floral Diversity with Bees and Computer Vision	643
<i>Frederic Tausch (Karlsruhe Institute of Technology, Germany; apic.ai, Germany), Jan Wagner (Karlsruhe Institute of Technology, Germany; apic.ai, Germany), and Simon Klaus (Karlsruhe Institute of Technology, Germany; apic.ai, Germany)</i>	

Inductive Conformal Prediction for Harvest-Readiness Classification of Cauliflower Plants: A Comparative Study of Uncertainty Quantification Methods .....	651
<i>Mohamed Farag (University of Bonn, Germany), Jana Kierdorf (University of Bonn, Germany), and Ribana Roscher (Forschungszentrum Jülich GmbH, Germany; University of Bonn, Germany)</i>	
Detection of Fusarium Damaged Kernels in Wheat Using Deep Semi-Supervised Learning on a Novel WheatSeedBelt Dataset .....	660
<i>Keyhan Najafian (University of Saskatchewan, Canada), Lingling Jin (University of Saskatchewan, Canada), H. Randy Kutcher (University of Saskatchewan, Canada), Mackenzie Hladun (University of Saskatchewan, Canada), Samuel Horovatin (University of Saskatchewan, Canada), Maria Alejandra Oviedo-Ludena (University of Saskatchewan, Canada), Sheila Maria Pereira de Andrade (University of Saskatchewan, Canada), Lipu Wang (University of Saskatchewan, Canada), and Ian Stavness (University of Saskatchewan, Canada)</i>	
Embedded Plant Recognition: A Benchmark for low Footprint Deep Neural Networks .....	670
<i>Sehaba Mohammed El Amine (Univ Lyon, Lyon 2, LIRIS, France), Crispim-Junior Carlos (Univ Lyon, Lyon 2, LIRIS, France), and Tougne Rodet Laure (Univ Lyon, Lyon 2, LIRIS, France)</i>	
Unlocking Comparative Plant Scoring with Siamese Neural Networks and Pairwise Pseudo Labelling .....	678
<i>Zane K. J. Hartley (University of Nottingham, United Kingdom), Rob J. Lind (Syngenta, United Kingdom), Nicholas Smith (Syngenta, United Kingdom), Bob Collison (Syngenta, United Kingdom), and Andrew P. French (University of Nottingham, United Kingdom)</i>	
Unified Automatic Plant Cover and Phenology Prediction .....	685
<i>Matthias Körschens (Friedrich Schiller University; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Germany), Solveig Franziska Bucher (Friedrich Schiller University; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig; Michael Stifel Center Jena, Germany), Christine Römermann (Friedrich Schiller University; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig; Michael Stifel Center Jena, Germany), and Joachim Denzler (Friedrich Schiller University; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig; Michael Stifel Center Jena, Germany)</i>	
Reinforcement Learning with Space Carving for Plant Scanning .....	694
<i>Antonio Pico Villalpando (Adaptive Systems Group, Humboldt-Universität zu Berlin, Germany), Matthias Kubisch (Adaptive Systems Group, Humboldt-Universität zu Berlin, Germany), David Colliaux (Sony CSL Paris, France), Peter Hanappe (Sony CSL Paris, France), and Verena V. Hafner (Adaptive Systems Group, Humboldt-Universität zu Berlin, Germany)</i>	
Towards Automated Regulation of <i>Jacobaea Vulgaris</i> in Grassland Using Deep Neural Networks .	702
<i>Moritz Schauer (University of Applied Sciences Mittelhessen), Renke Hohl (University of Applied Sciences Mittelhessen), Dennis Vaupel (University of Applied Sciences Mittelhessen), Diethelm Bienhaus (University of Applied Sciences Mittelhessen), and Seyed Eghbal Ghobadi (University of Applied Sciences Mittelhessen)</i>	

Efficient Grapevine Structure Estimation in Vineyards Conditions .....	712
<i>Théophile Gentilhomme (Idiap Research Institute), Michael Villamizar (Idiap Research Institute), Jerome Corre (3D2cut SA), and Jean-Marc Odobez (Idiap Research Institute)</i>	

## ICCV Workshop on New Ideas in Vision Transformers (NIVT)

A Hybrid Visual Transformer for Efficient Deep Human Activity Recognition .....	721
<i>Youcef Djenouri (University of South-Eastern Norway; NORCE Norwegian Research Centre, Norway) and Ahmed Nabil Belbachir (Norwegian Research Center, Norway)</i>	
SCSC: Spatial Cross-Scale Convolution Module to Strengthen Both CNNs and Transformers .....	731
<i>Xijun Wang (The University of Maryland, USA), Xiaojie Chu (MEGVII Technology, China), Chunrui Han (MEGVII Technology, China), and Xiangyu Zhang (MEGVII Technology, China)</i>	
TSOSVNet: Teacher-Student Collaborative Knowledge Distillation for Online Signature Verification .....	742
<i>Chandra Sekhar V (IIIT-SriCity, India), Avinash Gautam (BITS-Pilani, India), Viswanath P (IIIT-SriCity, India), Sreeja SR (IIIT-SriCity, India), and Rama Krishna Sai G (IIT-Tirupati, India)</i>	
SeMask: Semantically Masked Transformers for Semantic Segmentation .....	752
<i>Jitesh Jain (SHI Labs, Georgia Tech &amp; UIUC &amp; Oregon), Anukriti Singh (SHI Labs, Georgia Tech &amp; UIUC &amp; Oregon), Nikita Orlov (Picsart AI Research (PAIR)), Zilong Huang (SHI Labs, Georgia Tech &amp; UIUC &amp; Oregon), Jiachen Li (SHI Labs, Georgia Tech &amp; UIUC &amp; Oregon), Steven Walton (SHI Labs, Georgia Tech &amp; UIUC &amp; Oregon), and Humphrey Shi (SHI Labs, Georgia Tech &amp; UIUC &amp; Oregon; Picsart AI Research (PAIR))</i>	
Interactive Image Segmentation with Cross-Modality Vision Transformers .....	762
<i>Kun Li (University of Twente, The Netherlands), George Vosselman (University of Twente, The Netherlands), and Michael Ying Yang (University of Twente, The Netherlands)</i>	
Which Tokens to Use? Investigating Token Reduction in Vision Transformers .....	773
<i>Joakim Bruslund Haurum (Aalborg University &amp; Pioneer Centre for AI, Denmark), Sergio Escalera (Aalborg University &amp; Pioneer Centre for AI, Denmark; Universitat de Barcelona &amp; Computer Vision Center, Spain), Graham W. Taylor (University of Guelph &amp; Vector Institute for AI, Canada), and Thomas B. Moeslund (Aalborg University &amp; Pioneer Centre for AI, Denmark)</i>	
Actor-Agnostic Multi-Label Action Recognition with Multi-Modal Query .....	784
<i>Anindya Mondal (University of Surrey; CVSSP; Surrey Institute for People-Centred AI), Sauradip Nag (University of Surrey; CVSSP; iFlyTek-Surrey Joint Research Center on AI), Joaquín M Prada (University of Surrey; School of Veterinary Medicine), Xiatian Zhu (University of Surrey; CVSSP; Surrey Institute for People-Centred AI), and Anjan Dutta (University of Surrey; CVSSP; Surrey Institute for People-Centred AI; School of Veterinary Medicine)</i>	

Hierarchical Spatiotemporal Transformers for Video Object Segmentation .....	795
<i>Jun-Sang Yoo (Korea University, Republic of Korea), Hongjae Lee (Korea University, Republic of Korea), and Seung-Won Jung (Korea University, Republic of Korea)</i>	
Explaining Through Transformer Input Sampling .....	806
<i>Alexandre Engleburt (UCLouvain, Belgium), Sédric Stassin (UMons, Belgium), Géraldin Nanfack (University of Concordia, Canada), Sidi Ahmed Mahmoudi (UMons, Belgium), Xavier Siebert (UMons, Belgium), Olivier Cornu (UCLouvain, Belgium), and Christophe De Vleeschouwer (UCLouvain, Belgium)</i>	
IDTransformer: Transformer for Intrinsic Image Decomposition .....	816
<i>Partha Das (University of Amsterdam; 3DUniversum, The Netherlands), Maxime Gevers (Concordia University, Canada; 3DUniversum, The Netherlands), Sezer Karaoglu (University of Amsterdam; 3DUniversum, The Netherlands), and Theo Gevers (University of Amsterdam; 3DUniversum, The Netherlands)</i>	
All-pairs Consistency Learning for Weakly Supervised Semantic Segmentation .....	826
<i>Weixuan Sun (Australian National University; OpenNLPLab; Shanghai AI Lab), Yanhao Zhang (OPPO Research Institute), Zhen Qin (OpenNLPLab), Zheyuan Liu (Australian National University), Lin Cheng (Xiamen University), Fanyi Wang (OPPO Research Institute), Yiran Zhong (OpenNLPLab; Shanghai AI Lab), and Nick Barnes (Australian National University)</i>	
MSViT: Dynamic Mixed-Scale Tokenization for Vision Transformers .....	838
<i>Jakob Drachmann Havtorn (Technical University of Denmark &amp; Corti.ai, Denmark), Amélie Royer (Qualcomm AI Research, The Netherlands), Tijmen Blankevoort (Qualcomm AI Research, The Netherlands), and Babak Ehteshami Bejnordi (Qualcomm AI Research, The Netherlands)</i>	
TransInpaint: Transformer-Based Image Inpainting with Context Adaptation .....	849
<i>Pourya Shamsolmoali (East China Normal University, China), Masoumeh Zareapoor (Shanghai JiaoTong University, China), and Eric Granger (ETS Montreal, Canada)</i>	
Spatio-Temporal Convolution-Attention Video Network .....	859
<i>Ali Diba (KU Leuven), Vivek Sharma (Sony AI; Massachusetts Institute of Technology), Mohammad.M Arzani (n/a), and Luc Van Gool (KU Leuven; ETH Zurich)</i>	
Dual-Contrastive Dual-Consistency Dual-Transformer: A Semi-Supervised Approach to Medical Image Segmentation .....	870
<i>Ziyang Wang (University of Oxford, UK) and Congying Ma (University of Bath, UK)</i>	
On Moving Object Segmentation from Monocular Video with Transformers .....	880
<i>Christian Homeyer (Robert Bosch GmbH; Heidelberg University, Germany) and Christoph Schnörr (Heidelberg University)</i>	

MOSAIC: Multi-Object Segmented Arbitrary Stylization Using CLIP .....	892
<i>Prajwal Ganugula (OPPO Mobiles R &amp; D Center; IIT Hyderabad, India), Y S S Santosh Kumar (OPPO Mobiles R &amp; D Center, India), N K Sagar Reddy (OPPO Mobiles R &amp; D Center, India), Prabhath Chellingi (IIT Hyderabad, India), Avinash Thakur (OPPO Mobiles R &amp; D Center, India), Neeraj Kasera (OPPO Mobiles R &amp; D Center, India), and C Shyam Anand (OPPO Mobiles R &amp; D Center, India)</i>	

Template-Guided Illumination Correction for Document Images with Imperfect Geometric Reconstruction .....	904
<i>Felix Hertlein (FZI Research Center for Information Technology, Germany) and Alexander Naumann (FZI Research Center for Information Technology, Germany)</i>	

## **Representation learning with very limited images: the potential of self-, synthetic- and formula-supervision (LIMIT)**

Adaptive Self-Training for Object Detection .....	914
<i>Renaud Vandeghen (University of Liège), Gilles Louppe (University of Liège), and Marc Van Droogenbroeck (University of Liège)</i>	

Tensor Factorization for Leveraging Cross-Modal Knowledge in Data-Constrained Infrared Object Detection .....	924
<i>Manish Sharma (Rochester Institute of Technology, USA), Moitreya Chatterjee (Mitsubishi Electric Research Laboratories, USA), Kuan-Chuan Peng (Mitsubishi Electric Research Laboratories, USA), Suhas Lohit (Mitsubishi Electric Research Laboratories, USA), and Michael Jones (Mitsubishi Electric Research Laboratories, USA)</i>	

Learning Universal Semantic Correspondences with No Supervision and Automatic Data Curation .....	933
<i>Aleksandar Shtedritski (University of Oxford), Andrea Vedaldi (University of Oxford), and Christian Rupprecht (University of Oxford)</i>	

Semantic RGB-D Image Synthesis .....	944
<i>Shijie Li (University of Bonn, Germany), Rong Li (HKUST (Guangzhou), China), and Juergen Gall (University of Bonn, Germany)</i>	

JEDI: Joint Expert Distillation in a Semi-Supervised Multi-Dataset Student-Teacher Scenario for Video Action Recognition .....	953
<i>Lucian Bicsi (University of Bucharest), Bogdan Alexe (University of Bucharest), Radu Tudor Ionescu (University of Bucharest), and Marius Leordeanu (Politehnica University of Bucharest, Institute of Mathematics of the Romanian Academy)</i>	

Frequency-Aware Self-Supervised Long-Tailed Learning .....	963
<i>Ci-Siang Lin (National Taiwan University; NVIDIA, Taiwan), Min-Hung Chen (NVIDIA, Taiwan), and Yu-Chiang Frank Wang (National Taiwan University; NVIDIA, Taiwan)</i>	

SelectNAdapt: Support Set Selection for Few-Shot Domain Adaptation .....	973
<i>Youssef Dawoud (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Gustavo Carneiro (University of Surrey, United Kingdom), and Vasileios Belagiannis (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)</i>	

Self-Supervised Hypergraphs for Learning Multiple World Interpretations .....	983
<i>Alina Marcu (Politehnica University of Bucharest; Institute of Mathematics of the Romanian Academy), Mihai Pirvu (Politehnica University of Bucharest; Institute of Mathematics of the Romanian Academy), Dragos Costea (Politehnica University of Bucharest; Institute of Mathematics of the Romanian Academy), Emanuela Haller (Bitdefender), Emil Slusanschi (Politehnica University of Bucharest), Ahmed Nabil Belbachir (NORCE), Rahul Sukthankar (Google Research), and Marius Leordeanu (Politehnica University of Bucharest; Institute of Mathematics of the Romanian Academy; NORCE)</i>	
MIAD: A Maintenance Inspection Dataset for Unsupervised Anomaly Detection .....	993
<i>Tianpeng Bao (SenseTime Research, China), Jiadong Chen (SenseTime Research, China), Wei Li (SenseTime Research, China), Xiang Wang (SenseTime Research, China), Jingjing Fei (SenseTime Research, China), Liwei Wu (SenseTime Research, China), Rui Zhao (SenseTime Research; Shanghai Jiao Tong University, China), and Ye Zheng (JD.com, Inc., China)</i>	
Self-Training and Multi-Task Learning for Limited Data: Evaluation Study on Object Detection .....	1003
<i>Hoàng-Ân Lê (Université Bretagne Sud, France) and Minh-Tan Pham (Université Bretagne Sud, France)</i>	
Augmenting Features via Contrastive Learning-Based Generative Model for Long-Tailed Classification .....	1010
<i>Minho Park (Electronics and Telecommunications Research Institute (ETRI), South Korea), Hyung-Il Kim (Electronics and Telecommunications Research Institute (ETRI), South Korea), Hwa Jeon Song (Electronics and Telecommunications Research Institute (ETRI), South Korea), and Dong-oh Kang (Electronics and Telecommunications Research Institute (ETRI), South Korea)</i>	
Boosting Semi-Supervised Learning by Bridging high and low-Confidence Predictions .....	1020
<i>Khanh-Binh Nguyen (Sungkyunkwan University, South Korea) and Joon-Sung Yang (Yonsei University, South Korea)</i>	
FedLID: Self-Supervised Federated Learning for Leveraging Limited Image Data .....	1031
<i>Athanasios Psaltis (Centre for Research and Technology Hellas; University of West Attica, Greece), Anestis Kastellos (Centre for Research and Technology Hellas, Greece), Charalampos Z. Patrikakis (University of West Attica, Greece), and Petros Daras (Centre for Research and Technology Hellas, Greece)</i>	
A Horse with no Labels: Self-Supervised Horse Pose Estimation from Unlabelled Images and Synthetic Prior .....	1041
<i>Jose Sosa (University of Leeds) and David Hogg (University of Leeds)</i>	
Enhancing Classification Accuracy on Limited Data via Unconditional GAN .....	1049
<i>Chunsan Hong (KAIST, South Korea), Byunghee Cha (Seoul National University, South Korea), Bohyung Kim (CNAI, South Korea), and Tae-Hyun Oh (POSTECH; Yonsei Univ., South Korea)</i>	
Deep Generative Networks for Heterogeneous Augmentation of Cranial Defects .....	1058
<i>Kamil Kwarciak (AGH University of Krakow, Poland) and Marek Wodziński (AGH University of Krakow, Poland; University of Applied Sciences Western Switzerland (HES-SO Valais), Switzerland)</i>	

360° from a Single Camera: A Few-Shot Approach for LiDAR Segmentation .....	1067
<i>Laurenz Reichardt (Mannheim University for Applied Science, Germany), Nikolas Ebert (Mannheim University for Applied Science, Germany; RPTU Kaiserslautern-Landau, Germany), and Oliver Wasenmüller (Mannheim University for Applied Science, Germany)</i>	
Guiding Video Prediction with Explicit Procedural Knowledge .....	1076
<i>Patrick Takenaka (Hochschule der Medien Stuttgart; University of Stuttgart, Germany), Johannes Maucher (Hochschule der Medien Stuttgart, Germany), and Marco F. Huber (University of Stuttgart; Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Germany)</i>	
G2L: A High-Dimensional Geometric Approach for Automatic Generation of Highly Accurate Pseudo-Labels .....	1085
<i>John R. Kender (Columbia University, USA), Parijat Dube (IBM Research, USA), Zhengyang Han (New York University, USA), and Bishwaranjan Bhattacharjee (IBM Research, USA)</i>	
Image Guided Inpainting with Parameter Efficient Learning .....	1095
<i>Sangbeom Lim (Korea University; NCSOFT) and Seungryong Kim (Korea University)</i>	

## **To NeRF or not to NeRF: A View Synthesis Challenge for Human Heads (RHWC)**

ILSH: The Imperial Light-Stage Head Dataset for Human Head View Synthesis .....	1104
<i>Jiali Zheng (Imperial College London, UK), Youngkyoon Jang (Imperial College London, UK), Athanasios Papaioannou (Imperial College London, UK), Christos Kampouris (Imperial College London, UK), Rolandos Alexandros Potamias (Imperial College London, UK), Foivos Paraperas Papantoniou (Imperial College London, UK), Efstathios Galanakis (Imperial College London, UK), Aleš Leonardis (University of Birmingham, UK), and Stefanos Zafeiriou (Imperial College London, UK)</i>	



VSCHH 2023: A Benchmark for the View Synthesis Challenge of Human Heads .....	1113
---	------

*Youngkyoon Jang (Imperial College London, United Kingdom), Jiali Zheng (Imperial College London, United Kingdom), Jifei Song (Huawei Noah's Ark Lab, United Kingdom), Helisa Dhamo (Huawei Noah's Ark Lab, United Kingdom), Eduardo Pérez-Pellitero (Huawei Noah's Ark Lab, United Kingdom), Thomas Tanay (Huawei Noah's Ark Lab, United Kingdom), Matteo Maggioni (Huawei Noah's Ark Lab, United Kingdom), Richard Shaw (Huawei Noah's Ark Lab, United Kingdom), Sibi Catley-Chandar (Huawei Noah's Ark Lab, United Kingdom; Queen Mary University of London, United Kingdom), Yiren Zhou (Huawei Noah's Ark Lab, United Kingdom), Jiankang Deng (Imperial College London, United Kingdom), Ruijie Zhu (University of Science and Technology of China, China), Jiahao Chang (University of Science and Technology of China, China), Ziyang Song (University of Science and Technology of China, China), Jiahuan Yu (University of Science and Technology of China, China), Tianzhu Zhang (University of Science and Technology of China, China), Khanh-Binh Nguyen (Sungkyunkwan University, South Korea), Joon-Sung Yang (Yonsei University, South Korea), Andreea Dogaru (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Bernhard Egger (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Heng Yu (Carnegie Mellon University, United States), Aarush Gupta (Carnegie Mellon University, United States), Joel Julin (Carnegie Mellon University, United States), László A. Jeni (Carnegie Mellon University, United States), Hyeseong Kim (Yonsei University, South Korea), Jungbin Cho (Yonsei University, South Korea), Dosik Hwang (Yonsei University, South Korea), Deukhee Lee (Yonsei University, South Korea), Doyeon Kim (LG Electronics, South Korea), Dongseong Seo (Seoul National University, South Korea), SeungJin Jeon (Dongguk University, South Korea), YoungDon Choi (Korea Water Resources Corporation, South Korea), Jun Seok Kang (University of Science and Technology, South Korea), Ahmet Cagatay Seker (Institute of Science and Technology, South Korea), Sang Chul Ahn (Korea Institute of Science and Technology, South Korea), Ales Leonardis (University of Birmingham, United Kingdom), and Stefanos Zafeiriou (Imperial College London, United Kingdom)*

## **Workshop on Resource Efficient Deep Learning for Computer Vision (RCV)**

A Simple and Generic Framework for Feature Distillation via Channel-Wise Transformation .....	1121
---	------

*Ziwei Liu (Wangxuan Institute of Computer Technology, Peking University), Yongtao Wang (Wangxuan Institute of Computer Technology, Peking University), Xiaojie Chu (Wangxuan Institute of Computer Technology, Peking University), Nan Dong (Chong Qing Changan Automobile Co., Ltd), Shengxiang Qi (Chong Qing Changan Automobile Co., Ltd), and Haibin Ling (Stony Brook University)*

Deterministic Neural Illumination Mapping for Efficient Auto-White Balance Correction .....	1131
---	------

*Furkan Kınlı (Özyeğin University, Türkiye), Doğa Yılmaz (Özyeğin University, Türkiye), Barış Özcan (Özyeğin University, Türkiye), and Furkan Kırac (Özyeğin University, Türkiye)*

A Comprehensive Study of Transfer Learning Under Constraints .....	1140
<i>Tom Pégeot (Université Paris-Saclay, CEA, LIST, France), Inna Kucher (Université Paris-Saclay, CEA, LIST, France), Adrian Popescu (Université Paris-Saclay, CEA, LIST, France), and Bertrand Delezoide (Amanda, France)</i>	
Ray-Patch: An Efficient Querying for Light Field Transformers .....	1150
<i>Tomás Berriel Martins (Universidad de Zaragoza, Spain) and Javier Civera (Universidad de Zaragoza, Spain)</i>	
Cross-Model Temporal Cooperation via Saliency Maps for Efficient Frame Classification .....	1156
<i>Tomaso Trinci (Verizon Connect; University of Florence), Tommaso Bianconcini (Verizon Connect), Leonardo Sarti (Verizon Connect), Leonardo Taccari (Verizon Connect), and Francesco Sambo (Verizon Connect)</i>	
YOLOBench: Benchmarking Efficient Object Detectors on Embedded Systems .....	1161
<i>Ivan Lazarevich (Deeplite), Matteo Grimaldi (Deeplite), Ravish Kumar (Deeplite), Saptarshi Mitra (Deeplite), Shahrukh Khan (Deeplite), and Sudhakar Sah (Deeplite)</i>	
Accelerating Deep Neural Networks via Semi-Structured Activation Sparsity .....	1171
<i>Matteo Grimaldi (Deeplite), Darshan C. Ganji (Deeplite), Ivan Lazarevich (Deeplite), and Sudhakar Sah (Deeplite)</i>	
Revisiting Kernel Temporal Segmentation as an Adaptive Tokenizer for Long-form Video Understanding .....	1181
<i>Mohamed Afham (Meta AI), Satya Narayan Shukla (Meta AI), Omid Poursaeed (Meta AI), Pengchuan Zhang (Meta AI), Ashish Shah (Meta AI), and Sernam Lim (Meta AI)</i>	
NCQS: Nonlinear Convex Quadrature Surrogate Hyperparameter Optimization .....	1187
<i>Sophia Abraham (University of Notre Dame, USA), Kehelwala Dewage Gayan Maduranga (Tennessee Technological University), Jeffery Kinnison (University of Notre Dame, USA), Jonathan Hauenstein (University of Notre Dame, USA), and Walter Scheirer (University of Notre Dame, USA)</i>	
When Layers Play the Lottery, all Tickets Win at Initialization .....	1196
<i>Artur Jordao (Escola Politécnica, Universidade de São Paulo), George de Araújo (University of Campinas), Helena de Almeida Maia (University of Campinas), and Helio Pedrini (University of Campinas)</i>	
Video Action Recognition with Adaptive Zooming Using Motion Residuals .....	1206
<i>Mostafa Shahabinejad (Noah's Ark Laboratories, Huawei Technologies, Canada), Irina Kezele (Noah's Ark Laboratories, Huawei Technologies, Canada), Seyed Shahabeddin Nabavi (Noah's Ark Laboratories, Huawei Technologies, Canada), Wentao Liu (Noah's Ark Laboratories, Huawei Technologies, Canada), Seel Patel (Noah's Ark Laboratories, Huawei Technologies, Canada), Yuanhao Yu (Noah's Ark Laboratories, Huawei Technologies, Canada), Yang Wang (Concordia University, Canada), and Jin Tang (Noah's Ark Laboratories, Huawei Technologies, Canada)</i>	
Shapley Deep Learning: A Consensus for General-Purpose Vision Systems .....	1216
<i>Youcef Djenouri (University of South-Eastern Norway; NORCE Norwegian Research Centre, Norway), Ahmed Nabil Belbachir (Norwegian Research Center, Norway), Tomasz Michalak (IDEAS NCBR, Poland), and Anis Yazidi (OsloMet University, Norway)</i>	

HyperSparse Neural Networks: Shifting Exploration to Exploitation Through Adaptive Regularization .....	1226
<i>Patrick Glandorf (Leibniz Universität Hannover, Germany), Timo Kaiser (Leibniz Universität Hannover, Germany), and Bodo Rosenhahn (Leibniz Universität Hannover, Germany)</i>	
Reconstructing Pruned Filters Using Cheap Spatial Transformations .....	1236
<i>Roy Miles (Imperial College London) and Krystian Mikolajczyk (Imperial College London)</i>	
Enhancing Differentiable Architecture Search: A Study on Small Number of Cell Blocks in the Search Stage, and Important Branches-Based Cells Selection .....	1245
<i>Bedionita Soro (KAIST AI, South Korea) and Chong Song (KAIST AI, South Korea)</i>	
CoroNetGAN: Controlled Pruning of GANs via Hypernetworks .....	1254
<i>Aman Kumar (OPPO Mobiles R &amp; D Center, India), Khushboo Anand (OPPO Mobiles R &amp; D Center, India), Shubham Mandloi (OPPO Mobiles R &amp; D Center, India), Ashutosh Mishra (OPPO Mobiles R &amp; D Center, India), Avinash Thakur (OPPO Mobiles R &amp; D Center, India), Neeraj Kasera (OPPO Mobiles R &amp; D Center, India), and Prathosh A P (Indian Institute of Science, India)</i>	
Developing Robust and Lightweight Adversarial Defenders by Enforcing Orthogonality on Attack-Agnostic Denoising Autoencoders .....	1264
<i>Aristeidis Bifis (University of Patras, GR), Emmanouil Z. Psarakis (University of Patras, GR), and Dimitrios Kosmopoulos (University of Patras, GR)</i>	
QBitOpt: Fast and Accurate Bitwidth Reallocation During Training .....	1274
<i>Jorn Peters (Qualcomm AI Research, The Netherlands), Marios Fournarakis (Qualcomm AI Research, The Netherlands), Markus Nagel (Qualcomm AI Research, The Netherlands), Mart van Baalen (Qualcomm AI Research, The Netherlands), and Tijmen Blankevoort (Qualcomm AI Research, The Netherlands)</i>	
MGiaD: Multigrid in all Dimensions. Efficiency and Robustness by Weight Sharing and Coarsening in Resolution and Channel Dimensions .....	1284
<i>Antonia van Betteray (University of Wuppertal), Matthias Rottmann (University of Wuppertal), and Karsten Kahl (University of Wuppertal)</i>	
Accumulation Knowledge Distillation for Conditional GAN Compression .....	1294
<i>Tingwei Gao (Alibaba Group) and Rujiao Long (Alibaba Group)</i>	
Characterizing Face Recognition for Resource Efficient Deployment on Edge .....	1304
<i>Ayan Biswas (IIIT Hyderabad, India), Sai Amrit Patnaik (IIIT Hyderabad, India), A. H. Abdul Hafez (Hasan Kalyoncu University, Turkey), and Anoop M. Namboodiri (IIIT Hyderabad, India)</i>	
MOFA: A Model Simplification Roadmap for Image Restoration on Mobile Devices .....	1314
<i>Xiangyu Chen (University of Kansas; SenseBrain Technology, USA), Ruiwen Zhen (SenseBrain Technology, USA), Shuai Li (SenseBrain Technology, USA), Xiaotian Li (SenseBrain Technology, USA), and Guanghui Wang (Toronto Metropolitan University, Canada)</i>	

DetOFA: Efficient Training of Once-for-All Networks for Object Detection Using Path Filter .....	1325
<i>Yuiko Sakuma (Sony Group Corporation, Japan), Masato Ishii (Sony Group Corporation, Japan), and Takuya Narihira (Sony Group Corporation, Japan)</i>	
Post Training Mixed Precision Quantization of Neural Networks Using First-Order Information .....	1335
<i>Arun Chauhan (Samsung Research Institute, India), Utsav Tiwari (Samsung Research Institute, India), and Vikram N R (Samsung Research Institute, India)</i>	
ZiCo-BC: A Bias Corrected Zero-Shot NAS for Vision Tasks .....	1345
<i>Kartikaya Bhardwaj (Qualcomm AI Research, USA), Hsin-Pai Cheng (Qualcomm AI Research, USA), Sweta Priyadarshi (Qualcomm AI Research, USA), and Zhuojin Li (Qualcomm AI Research, USA)</i>	
Bi-Encoder Cascades for Efficient Image Search .....	1350
<i>Robert Hönig (Data Analytics Lab, ETH Zurich), Jan Ackermann (Data Analytics Lab, ETH Zurich), and Mingyuan Chi (Data Analytics Lab, ETH Zurich)</i>	
Tiny and Efficient Model for the Edge Detection Generalization .....	1356
<i>Xavier Soria (National University of Chimborazo, Ecuador), Yachuan Li (China University of Petroleum (East China), China), Mohammad Rouhani (INRIA Paris, France), and Angel D. Sappa (ESPOL Polytechnic University, Ecuador; Computer Vision Center, Spain)</i>	
Factorized Dynamic Fully-Connected Layers for Neural Networks .....	1366
<i>Francesca Babiloni (Huawei, Noah's Ark Lab; Imperial College London), Thomas Tanay (Huawei, Noah's Ark Lab), Jiankang Deng (Huawei, Noah's Ark Lab; Imperial College London), Matteo Maggioni (Huawei, Noah's Ark Lab), and Stefanos Zafeiriou (Imperial College London)</i>	
DONNAv2 - Lightweight Neural Architecture Search for Vision Tasks .....	1376
<i>Sweta Priyadarshi (Qualcomm AI Research, USA), Tianyu Jiang (Qualcomm AI Research, USA), Hsin-Pai Cheng (Qualcomm AI Research, USA), Sendil Krishna (Qualcomm AI Research, USA), Viswanath Ganapathy (Qualcomm AI Research, USA), and Chirag Patel (Qualcomm AI Research, USA)</i>	
RCD-SGD: Resource-Constrained Distributed SGD in Heterogeneous Environment Via Submodular Partitioning .....	1385
<i>Haoze He (New York University, USA) and Parijat Dube (IBM Research, USA)</i>	
Can Unstructured Pruning Reduce the Depth in Deep Neural Networks? .....	1394
<i>Zhu Liao (LTCI, Télécom Paris, Institut Polytechnique de Paris, France), Victor Quéru (LTCI, Télécom Paris, Institut Polytechnique de Paris, France), Van-Tam Nguyen (LTCI, Télécom Paris, Institut Polytechnique de Paris, France), and Enzo Tartaglione (LTCI, Télécom Paris, Institut Polytechnique de Paris, France)</i>	
Surround the Nonlinearity: Inserting Foldable Convolutional Autoencoders to Reduce Activation Footprint .....	1399
<i>Baptiste Rossignaux (Université de Rennes, CEA, France), Inna Kucher (CEA, France), Vincent Lorrain (CEA, France), and Emmanuel Casseau (Université de Rennes, CNRS, INRIA, France)</i>	

Cross-Domain Transfer Learning with CoRTE: Consistent and Reliable Transfer from Black-Box to Lightweight Segmentation Model .....	1404
<i>Claudia Cuttano (Politecnico di Torino, ITALIA), Antonio Tavera (Politecnico di Torino, ITALIA), Fabio Cermelli (Politecnico di Torino, ITALIA), Giuseppe Averta (Politecnico di Torino, ITALIA), and Barbara Caputo (Politecnico di Torino, ITALIA)</i>	
Efficient Neural PDE-Solvers Using Quantization Aware Training .....	1415
<i>Winfried van den Dool (University of Amsterdam, Netherlands), Tijmen Blankevoort (Qualcomm AI Research, Netherlands), Max Welling (University of Amsterdam, Netherlands), and Yuki Asano (University of Amsterdam, Netherlands)</i>	
Single-Shot Pruning for Pre-Trained Models: Rethinking the Importance of Magnitude Pruning...	1425
<i>Hirokazu Kohama (Chubu University), Hiroaki Minoura (Chubu University), Tsubasa Hirakawa (Chubu University), Takayoshi Yamashita (Chubu University), and Hironobu Fujiyoshi (Chubu University)</i>	
SCoTTi: Save Computation at Training Time with an Adaptive Framework .....	1435
<i>Ziyu Li (Institut Polytechnique de Paris, France), Enzo Tartaglione (Institut Polytechnique de Paris, France), and Van-Tam Nguyen (Institut Polytechnique de Paris, France)</i>	
Softmax Bias Correction for Quantized Generative Models .....	1445
<i>Nilesh Prasad Pandey (Qualcomm AI Research), Marios Fournarakis (Qualcomm AI Research), Chirag Patel (Qualcomm AI Research), and Markus Nagel (Qualcomm AI Research)</i>	
Entropic Score Metric: Decoupling Topology and Size in Training-Free NAS .....	1451
<i>Niccolò Cavagnero (Politecnico di Torino, ITALIA), Luca Robbiano (Politecnico di Torino, ITALIA), Francesca Pistilli (Politecnico di Torino, ITALIA), Barbara Caputo (Politecnico di Torino, ITALIA), and Giuseppe Averta (Politecnico di Torino, ITALIA)</i>	
Fast Object Detection in High-Resolution Videos .....	1461
<i>Ryan Tran (Percipient.ai, USA), Atul Kanaujia (Percipient.ai, USA), and Vasu Parameswaran (Percipient.ai, USA)</i>	
Extending TrOCR for Text Localization-Free OCR of Full-Page Scanned Receipt Images .....	1471
<i>Hongkuan Zhang (Nagoya University), Edward Whittaker (K.K. Best Path Research), and Ikuo Kitagishi (Money Forward, Inc.)</i>	
Multi-Exit Resource-Efficient Neural Architecture for Image Classification with Optimized Fusion Block .....	1478
<i>Youva Addad (Normandy University, France), Alexis Lechervy (Normandy University, France), and Frédéric Jurie (Normandy University, France)</i>	
Lightweight Vision Transformer with Spatial and Channel Enhanced Self-Attention .....	1484
<i>Jiahao Zheng (University of Electronic Science and Technology of China, China), Longqi Yang (Defense Innovation Institute, Academy of Military Sciences, China), Yiyi Li (Defense Innovation Institute, Academy of Military Sciences, China), Ke Yang (Defense Innovation Institute, Academy of Military Sciences, China), Zhiyuan Wang (Defense Innovation Institute, Academy of Military Sciences, China), and Jun Zhou (University of Electronic Science and Technology of China, China)</i>	

Dynamic Neural Network is All You Need: Understanding the Robustness of Dynamic Mechanisms in Neural Networks .....	1489
<i>Mirazul Haque (University of Texas at Dallas, USA) and Wei Yang (University of Texas at Dallas, USA)</i>	
AntiNODE: Evaluating Efficiency Robustness of Neural ODEs .....	1499
<i>Mirazul Haque (University of Texas at Dallas, USA), Simin Chen (University of Texas at Dallas, USA), Wasif Haque (Argo Data), Cong Liu (University of California, USA), and Wei Yang (University of Texas at Dallas, USA)</i>	
Shannon Strikes Again! Entropy-Based Pruning in Deep Neural Networks for Transfer Learning Under Extreme Memory and Computation Budgets .....	1510
<i>Gabriele Spadaro (University of Turin, Italy), Riccardo Renzulli (University of Turin, Italy), Andrea Bragagnolo (Independent Researcher), Jhony H. Giraldo (LTCI, Télécom Paris - Institut Polytechnique de Paris, France), Attilio Fiandrotti (University of Turin, Italy), Marco Grangetto (University of Turin, Italy), and Enzo Tartaglione (LTCI, Télécom Paris - Institut Polytechnique de Paris, France)</i>	
InstaTune: Instantaneous Neural Architecture Search During Fine-Tuning .....	1515
<i>Sharath Nittur Sridhar (Intel Labs, USA), Souvik Kundu (Intel Labs, USA), Sairam Sundaresan (Intel Labs, USA), Maciej Szankin (Intel Labs, USA), and Anthony Sarah (Intel Labs, USA)</i>	
Quantized Generative Models for Solving Inverse Problems .....	1520
<i>Nareddy Kartheek Kumar Reddy (Indian Institute of Science, India), Vinayak Killedar (Indian Institute of Science, India), and Chandra Sekhar Seelamantula (Indian Institute of Science, India)</i>	
RCV2023 Challenges: Benchmarking Model Training and Inference for Resource-Constrained Deep Learning .....	1526
<i>Rishabh Tiwari (Transmute AI Lab (Texmin Hub); Google Research), Arnav Chavan (Transmute AI Lab (Texmin Hub); NyunAI), Deepak Gupta (Transmute AI Lab (Texmin Hub); UiT Tromso), Gowreesh Mago (Transmute AI Lab (Texmin Hub)), Animesh Gupta (Transmute AI Lab (Texmin Hub)), Akash Gupta (NyunAI), Suraj Sharan (Transmute AI Lab (Texmin Hub)), Yukun Yang (Ant Group), Shanwei Zhao (Ant Group), Shihao Wang (Ant Group), Youngjun Kwak (KakaoBank Corp.; KAIST), Seonghun Jeong (KAIST), Yunseung Lee (KAIST), Changick Kim (KakaoBank Corp.), Subin Kim (KakaoBank Corp.), Ganzorig Gankhuyag (KETI), Ho Jung (Hanwha Systems), Junwhan Ryu (Hanwha Systems), HaeMoon Kim (Hanwha Systems), Byeong H. Kim (Korea Institute of Industrial Technology), Tu Vo (KC Machine Learning Lab), Sheir Zaheer (KC Machine Learning Lab), Alexander Holston (KC Machine Learning Lab), Chan Park (KC Machine Learning Lab), Dheemant Dixit (IIT ISM Dhanbad), Nahush Lele (IIT ISM Dhanbad), Kushagra Bhushan (IIT ISM Dhanbad), Debjani Bhowmick (Transmute AI Lab (Texmin Hub)), Devanshu Arya (Serket BV), Sadaf Gulshad (University of Amsterdam), Amirhossein Habibian (Qualcomm AI Research), Amir Ghodrati (Qualcomm AI Research), Babak Bejnordi (Qualcomm AI Research), Jai Gupta (Google Research), Zhuang Liu (Meta AI Research), Jiahui Yu (Google Brain), Dilip Prasad (UiT Tromso), and Zhiqiang Shen (MBUZAI)</i>	

## 1st Workshop on Computer Vision Aided Architectural Design (CVAAD)

Scalable MAV Indoor Reconstruction with Neural Implicit Surfaces .....	1536
<i>Haoda Li (University of California, Berkeley), Puyuan Yi (University of California, Berkeley), Yunhao Liu (University of California, Berkeley), and Avideh Zakhor (University of California, Berkeley)</i>	
PanoStyle: Semantic, Geometry-Aware and Shading Independent Photorealistic Style Transfer for Indoor Panoramic Scenes .....	1545
<i>M. Tukur (ICT, CSE, HBKU, Qatar), A. Ur Rehman (ICT, CSE, HBKU, Qatar), G. Pintore (CRS4, Italy), E. Gobbetti (, Italy), J. Schneider (ICT, CSE, HBKU, Qatar), and M. Agus (ICT, CSE, HBKU, Qatar)</i>	
MARL: Multi-scale Archetype Representation Learning for Urban Building Energy Modeling .....	1557
<i>Xinwei Zhuang (University of California, Berkeley), Zixun Huang (University of California, Berkeley), Wentao Zeng (University of California, Berkeley), and Luisa Caldas (University of California, Berkeley)</i>	
SSIG: A Visually-Guided Graph Edit Distance for Floor Plan Similarity .....	1565
<i>Casper C. J. van Engelenburg (Delft University of Technology, The Netherlands), Seyran Khademi (Delft University of Technology, The Netherlands), and Jan C. van Gemert (Delft University of Technology, The Netherlands)</i>	
Floor Plan Reconstruction from Sparse Views: Combining Graph Neural Network with Constrained Diffusion .....	1575
<i>Arnaud Gueze (LIX, École Polytechnique, Institut Polytechnique de Paris, France; Homiwoo, France), Matthieu Ospici (Homiwoo, France), Damien Rohmer (LIX, École Polytechnique, Institut Polytechnique de Paris, France), and Marie-Paule Cani (LIX, École Polytechnique, Institut Polytechnique de Paris, France)</i>	

## 4th ICCV Workshop on Electronic Cultural Heritage (e-Heritage)

3D Surface Approximation of the Entire Bayeux Tapestry for Improved Pedagogical Access .....	1585
<i>Marjorie Redon (Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, France), Matthieu Pizenberg (Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, France), Yvain Quéau (Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, France), and Abderrahim Elmoataz (Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, France)</i>	
DeFi: Detection and Filling of Holes in Point Clouds Towards Restoration of Digitized Cultural Heritage Models .....	1595
<i>Ramesh Ashok Tabib (KLE Technological University, India), Dikshit Hegde (KLE Technological University, India), Tejas Anvekar (KLE Technological University, India), and Uma Mudenagudi (KLE Technological University, India)</i>	

Facsimiles-Based Deep Learning for Matching Relief-Printed Decorations on Medieval Ceramic Sherds .....	1605
<i>Khawla Brahim (Université d'Orléans, Laboratoire PRISME), Sylvie Treuillet (Université d'Orléans, Laboratoire PRISME), Matthieu Exbrayat (Université d'Orléans, Laboratoire LIFO), and Sebastien Jesset (Service Archéologique Municipal d'Orléans)</i>	
Learning to Rank Approach for Refining Image Retrieval in Visual Arts .....	1615
<i>Tetiana Yemelianenko (Univ Lyon, Univ Lyon 2, CNRS, INSA Lyon, UCBL, LIRIS, France), Iuliia Tkachenko (Univ Lyon, Univ Lyon 2, CNRS, INSA Lyon, UCBL, LIRIS, France), Tess Masclef (Univ Lyon, Univ Lyon 2, CNRS, INSA Lyon, UCBL, LIRIS, France), Mihaela Scuturici (Univ Lyon, Univ Lyon 2, CNRS, INSA Lyon, UCBL, LIRIS, France), and Serge Miguet (Univ Lyon, Univ Lyon 2, CNRS, INSA Lyon, UCBL, LIRIS, France)</i>	
MatchMakerNet: Enabling Fragment Matching for Cultural Heritage Analysis .....	1624
<i>Ariana M. Villegas-Suarez (Universidad de Ingenieria y Tecnologia – UTEC, Peru), Cristian Lopez (Universidad de Ingenieria y Tecnologia – UTEC, Peru), and Ivan Sipiran (University of Chile, Chile)</i>	
Evaluation of 3D Reconstruction for Cultural Heritage Applications .....	1634
<i>Cristián Llull (University of Chile), Nelson Baloian (University of Chile), Benjamin Bustos (University of Chile), Kornelius Kupczik (Universidad de Chile), Ivan Sipiran (University of Chile), and Andrés Baloian (University of Chile)</i>	
ASUR3D: Arbitrary Scale Upsampling and Refinement of 3D Point Clouds Using Local Occupancy Fields .....	1644
<i>Akash Kumbar (KLE Technological University, India), Tejas Anvekar (KLE Technological University, India), Ramesh Ashok Tabib (KLE Technological University, India), and Uma Mudenagudi (KLE Technological University, India)</i>	
Hyperspectral Imaging of In-Site Stained Glasses: Illumination Variation Compensation Using Two Perpendicular Scans .....	1654
<i>Suzan Joseph Kessy (Nara Institute of Science and Technology, Japan), Takuya Funatomi (Nara Institute of Science and Technology, Japan), Kazuya Kitano (Nara Institute of Science and Technology, Japan), Yuki Fujimura (Nara Institute of Science and Technology, Japan), Guillaume Caron (Université de Picardie Jules Verne, France, CNRS-AIST Joint Robotics Laboratory (JRL), Japan), El Mustapha Mouaddib (Université de Picardie Jules Verne, France; CNRS-AIST JRL (Joint Robotics Laboratory), IRL), and Yasuhiro Mukaigawa (Nara Institute of Science and Technology, Japan)</i>	
Pigment Mapping for Tomb Murals Using Neural Representation and Physics-Based Model .....	1663
<i>Mayuka Tsuji (Nara Institute of Science and Technology, Japan), Yuki Fujimura (Nara Institute of Science and Technology, Japan), Takuya Funatomi (Nara Institute of Science and Technology, Japan), Yasuhiro Mukaigawa (Nara Institute of Science and Technology, Japan), Tetsuro Morimoto (The University of Tokyo, Japan), Takeshi Oishi (The University of Tokyo, Japan), Jun Takamatsu (Microsoft), and Katsushi Ikeuchi (Microsoft)</i>	



CNN Based Cuneiform Sign Detection Learned from Annotated 3D Renderings and Mapped Photographs with Illumination Augmentation .....	1672
<i>Ernst Stötzner (Martin-Luther-Universität Halle-Wittenberg, Germany), Timo Homburg (University of Applied Sciences Mainz, Germany), and Hubert Mara (Martin-Luther-Universität Halle-Wittenberg, Germany)</i>	
Semantic Segmentation Using Foundation Models for Cultural Heritage: an Experimental Study on Notre-Dame de Paris .....	1681
<i>Kévin Réby (UMR CNRS, France), Anaïs Guilhelm (UMR CNRS, France), and Livio De Luca (UMR CNRS, France)</i>	
An Interactive Method for Adaptive Acquisition in Reflectance Transformation Imaging for Cultural Heritage .....	1690
<i>Muhammad Arsalan Khawaja (Université de Bourgogne &amp; Norwegian University of Science and Technology), Sony George (Norwegian University of Science and Technology), Franck Marzani (Université de Bourgogne), Jon Yngve Hardeberg (Norwegian University of Science and Technology), and Alamin Mansouri (Université de Bourgogne)</i>	
Diffusion Based Augmentation for Captioning and Retrieval in Cultural Heritage .....	1699
<i>Dario Cioni (University of Florence), Lorenzo Berlincioni (University of Florence), Federico Becattini (University of Siena), and Alberto Del Bimbo (University of Florence)</i>	
Semantic Motif Segmentation of Archaeological Fresco Fragments .....	1709
<i>Aref Enayati (Ca' Foscari University of Venice, Italy), Luca Palmieri (Ca' Foscari University of Venice, Italy), Sebastiano Vascon (Ca' Foscari University of Venice, Italy), Marcello Pelillo (Ca' Foscari University of Venice, Italy), and Sinem Aslan (Ca' Foscari University of Venice, Italy)</i>	
Volumetric Fast Fourier Convolution for Detecting Ink on the Carbonized Herculaneum Papyri... 1718	
<i>Fabio Quattrini (University of Modena and Reggio Emilia, Italy), Vittorio Pippi (University of Modena and Reggio Emilia, Italy), Silvia Cascianelli (University of Modena and Reggio Emilia, Italy), and Rita Cucchiara (University of Modena and Reggio Emilia, Italy)</i>	

## **Solving CAD History and pArAmeters Recovery from Point clouds and 3D scans (SHARP)**

Building CAD Model Reconstruction from Point Clouds via Instance Segmentation, Signed Distance Function, and Graph Cut .....	1727
<i>Takayuki Shinohara (PASCO Corporation, Japan), Li Yonghe (PASCO, Japan), Mitsuteru Sakamoto (PASCO, Japan), and Toshiaki Satoh (PASCO, Japan)</i>	
2D Cross-View Object Segmentation and Perceptual Grouping in Computer-Aided Design Drawings .....	1737
<i>Mohamed Dhia Elhak Besbes (Capgemini), Zahra Vahidi Ferdousi (Capgemini), Hedi Tabia (Univ Paris Saclay), and Mouna Fradi (Capgemini)</i>	

APNet: Urban-Level Scene Segmentation of Aerial Images and Point Clouds .....	1747
<i>Weijie Wei (University of Amsterdam, The Netherlands), Martin R. Oswald (University of Amsterdam, The Netherlands), Fatemeh Karimi Nejadasl (University of Amsterdam, The Netherlands), and Theo Gevers (University of Amsterdam, The Netherlands)</i>	
Rotation-Invariant Hierarchical Segmentation on Poincaré Ball for 3D Point Cloud .....	1757
<i>Pierre Onghena (TheCrossProduct), Leonardo Gigli (TheCrossProduct), and Santiago Velasco-Forero (PSL Research University / Mines Paris)</i>	
Fine-Tuned but Zero-Shot 3D Shape Sketch View Similarity and Retrieval .....	1767
<i>Gianluca Berardi (University of Bologna, Italy; University of Surrey, UK) and Yulia Gryaditskaya (University of Surrey, UK)</i>	
SHARP Challenge 2023: Solving CAD History and pParameters Recovery from Point Clouds and 3D Scans. Overview, Datasets, Metrics, and Baselines. ....	1778
<i>Dimitrios Mallis (University of Luxembourg), Ali Sk Aziz (University of Luxembourg), Elona Dupont (University of Luxembourg), Kseniya Cherenkova (University of Luxembourg, Artec3D), Ahmet Serdar Karadeniz (University of Luxembourg), Mohammad Sadil Khan (University of Luxembourg), Anis Kacem (University of Luxembourg), Gleb Gusev (Artec 3D), and Djamila Aouada (University of Luxembourg)</i>	

# Visual Object Tracking and Segmentation Challenge 2023 (VOTS)

The First Visual Object Tracking Segmentation VOTS2023 Challenge Results ..... 1788

*Matej Kristan (University of Ljubljana, Slovenia), Jiří Matas (Czech Technical University, Czech Republic), Martin Danelljan (ETH Zürich, Switzerland), Michael Felsberg (Linköping University, Sweden), Hyung Jin Chang (University of Birmingham, United Kingdom), Luka Čehovin Zajc (University of Ljubljana, Slovenia), Alan Lukežič (University of Ljubljana, Slovenia), Ondřej Drbohlav (Czech Technical University, Czech Republic), Zhongqun Zhang (University of Birmingham, United Kingdom), Khanh-Tung Tran (Umeå University, Sweden), Xuan-Son Vu (Umeå University, Sweden), Johanna Björklund (Umeå University, Sweden), Christoph Mayer (ETH Zürich, Switzerland), Yushan Zhang (Linköping University, Sweden), Lei Ke (ETH Zürich, Switzerland), Jie Zhao (ETH Zürich, Switzerland), Gustavo Fernández (Austrian Institute of Technology, Austria), Noor Al-Shakarji (University of Missouri, USA), Dong An (Institute of Automation, China), Michael Arens (Fraunhofer IOSB, Germany), Stefan Becker (Fraunhofer IOSB, Germany), Goutam Bhat (ETH Zürich, Switzerland), Sebastian Bullinger (Fraunhofer IOSB, Germany), Antoni B. Chan (City University of Hong Kong, Hong Kong, China), Shijie Chang (Dalian University of Technology, China), Hanyuan Chen (DAMO Academy, China), Xin Chen (Dalian University of Technology, China), Yan Chen (HuaQiao University, China), Zhenyu Chen (Dalian University of Technology, China), Yangming Cheng (Zhejiang University, China), Yutao Cui (Nanjing University, China), Chunyuan Deng (Georgia Institute of Technology, USA), Jiahua Dong (Shenyang Institute of Automation, Chinese Academy of Science, China), Matteo Dunnhofer (University of Udine, Italy), Wei Feng (Tianjin University, China), Jianlong Fu (Microsoft Research Asia, China), Jie Gao (HuaQiao University, China), Ruize Han (Tianjin University, China), Zeqi Hao (Dalian University of Technology, China), Jun-Yan He (DAMO Academy, China), Keji He (Institute of Automation, China), Zhenyu He (Harbin Institute of Technology (Shenzhen), China), Xiantao Hu (Guangxi Normal University, China), Kaer Huang (Lenovo Research, China), Yuqing Huang (Harbin Institute of Technology (Shenzhen), China), Yi Jiang (ByteDance, China), Ben Kang (Dalian University of Technology, China), Jin-Peng Lan (DAMO Academy, China), Hyungjun Lee (Nota AI, South Korea), Chenyang Li (DAMO Academy, China), Jiahao Li (Zhejiang University, China), Ning Li (Guangxi Normal University, China), Wangkai Li (University of Science and Technology of China, China), Xiaodi Li (Zhejiang University, China), Xin Li (Peng Cheng Laboratory, China), Pengyu Liu (Dalian University of Technology, China), Yue Liu (Jiangnan University, China), Huchuan Lu (Dalian University of Technology, China), Bin Luo (DAMO Academy, China), Ping Luo (The University of Hong Kong, Hong Kong, China), Yinchao Ma (University of Science and Technology of China, China), Deshui Miao (Harbin Institute of Technology (Shenzhen), China), Christian Micheloni (University of Udine, Italy), Kannappan Palaniappan (University of Missouri, USA), Hancheol Park (Nota AI, South Korea), Matthieu Paul (ETH Zürich, Switzerland), HouWen Peng (Microsoft Research, China), Zekun Qian (Tianjin University, China), Gani Rahmon (University of Missouri, USA), Norbert Scherer-Negenborn (Fraunhofer IOSB, Germany), Pengcheng Shao (Jiangnan University, China), Wooksu Shin (Nota AI, South Korea), Elham Soltani Kazemi (University of Missouri, USA), Tianhui Song (Nanjing University, China), Rainer Stiefelhagen (Karlsruhe Institute of Technology (KIT), Germany), Rui Sun (University of Science and Technology of China, China), Chuanming Tang (University of Chinese Academy of Sciences, China), Zhangyong Tang (Jiangnan University, China), Imad Eddine Toubal (University of Missouri, USA), Jack Valmadre (University of Adelaide, Australia), Joost van de Weijer (Computer Vision Center, Spain), Luc Van Gool (ETH Zürich, Switzerland), Jash Vira (University of Adelaide, Australia), Stéphane Vujasinović (Fraunhofer IOSB, Germany), Cheng Wan (Georgia Institute of Technology, USA), Jia Wan (Boston College, USA), Dong Wang (Dalian University of Technology, China), Fei Wang (University of Science and*

Technology of China, China), Feifan Wang (Tianjin University, China), He Wang (Jiangnan University, China), Limin Wang (Nanjing University, China), Song Wang (University of South Carolina, USA), Yaowei Wang (Peng Cheng Laboratory, China), Zhepeng Wang (Lenovo Research, China), Gangshan Wu (Nanjing University, China), Jiannan Wu (The University of Hong Kong, Hong Kong, China), Qiangqiang Wu (City University of Hong Kong, Hong Kong, China), Xiaojun Wu (Jiangnan University, China), Anqi Xiao (Institute of Automation, China), Jinxia Xie (Guangxi Normal University, China), Chenlong Xu (Guangxi Normal University, China), Min Xu (Carnegie Mellon University, USA), Tianyang Xu (Jiangnan University, China), Yuanyou Xu (Zhejiang University, China), Bin Yan (Dalian University of Technology, China), Dawei Yang (University of Science and Technology of China, China), Ming-Hsuan Yang (University of California at Merced, USA), Tianyu Yang (International Digital Economy Academy, China), Yi Yang (Zhejiang University, China), Zongxin Yang (Zhejiang University, China), Xuanwu Yin (Multimedia Department Xiaomi Inc., China), Fisher Yu (ETH Zürich, Switzerland), Hongyuan Yu (Multimedia Department Xiaomi Inc., China), Qianjin Yu (University of Science and Technology of China, China), Weichen Yu (Carnegie Mellon University, USA), Yongsheng Yuan (Dalian University of Technology, China), Zehuan Yuan (ByteDance, China), Jianlin Zhang (University of Chinese Academy of Sciences, China), Lu Zhang (Dalian University of Technology, China), Tianzhu Zhang (University of Science and Technology of China, China), Guodongfang Zhao (Institute of Computing Technology, Chinese Academy of Sciences, China), Shaochuan Zhao (Jiangnan University, China), Yaozong Zheng (Guangxi Normal University; HuaQiao University, China), Bineng Zhong (Guangxi Normal University, China), Jiawen Zhu (Dalian University of Technology, China), Xuefeng Zhu (Jiangnan University, China), Yueting Zhuang (Zhejiang University, China), ChengAo Zong (Dalian University of Technology, China), and Kunlong Zuo (Multimedia Department Xiaomi Inc., China)

## 11th Workshop on Assistive Computer Vision and Robotics (ACVR)

Vision-Based Treatment Localization with Limited Data: Automated Documentation of Military Emergency Medical Procedures .....	1811
<i>Trevor Powers (MIT Lincoln Laboratories, USA; Northeastern University, USA), Elaheh Hatamimajoumerd (Northeastern University, USA), William Chu (MIT Lincoln Laboratories, USA), Vishakk Rajendran (Massachusetts Institute of Technology, USA), Rishi Shah (Massachusetts Institute of Technology, USA; MIT Lincoln Laboratories, USA), Frank Diabour (Texas Tech University, USA), Marc Vaillant (MIT Lincoln Laboratories, USA), Richard Fletcher (Massachusetts Institute of Technology, USA; MIT Lincoln Laboratories, USA), and Sarah Ostadabbas (Northeastern University, USA)</i>	
Autonomous Mobile Robot for Automatic out of Stock Detection in a Supermarket .....	1821
<i>Giuseppe De Simone (University of Salerno), Pasquale Foggia (University of Salerno), Alessia Saggese (University of Salerno), and Mario Vento (University of Salerno)</i>	
Real-Time Optimisation-Based Path Planning for Visually Impaired People in Dynamic Environments .....	1831
<i>Hadeel R. Surougi (Imperial College London, UK) and Julie A. McCann (Imperial College London, UK)</i>	

Enhancing Human-Robot Collaborative Object Search Through Human Behavior Observation and Dialog .....	1841
<i>Takahiro Ishii (Toyohashi University of Technology), Jun Miura (Toyohashi University of Technology), and Kotaro Hayashi (Toyohashi University of Technology)</i>	
Open Scene Understanding: Grounded Situation Recognition Meets Segment Anything for Helping People with Visual Impairments .....	1849
<i>Ruiping Liu (Karlsruhe Institute of Technology), Jiaming Zhang (Karlsruhe Institute of Technology, Germany), Kunyu Peng (Karlsruhe Institute of Technology, Germany), Junwei Zheng (Karlsruhe Institute of Technology, Germany), Ke Cao (Karlsruhe Institute of Technology, Germany), Yufan Chen (Karlsruhe Institute of Technology, Germany), Kailun Yang (Hunan University, China), and Rainer Stiefelhagen (Karlsruhe Institute of Technology, Germany)</i>	
Personalized Monitoring in Home Healthcare: An Assistive System for Post Hip Replacement Rehabilitation .....	1860
<i>Alaa Kryeem (University of Haifa), Shmuel Raz (University of Haifa), Dana Eluz (Galilee Medical Center), Dorit Itah (Galilee Medical Center), Hagit Hel-Or (University of Haifa), and Ilan Shimshoni (University of Haifa)</i>	
Repetition-Aware Image Sequence Sampling for Recognizing Repetitive Human Actions .....	1870
<i>Konstantinos Bacharidis (University of Crete; Institute of Computer Science, FORTH) and Antonis Argyros (University of Crete; Institute of Computer Science, FORTH)</i>	
IFPNet: Integrated Feature Pyramid Network with Fusion Factor for Lane Detection .....	1880
<i>Zinan Lv (Harbin Engineering University, China), Dong Han (Harbin Engineering University, China), Wenzhe Wang (Zhejiang University; Westlake University, China), and Cheng Chen (Harbin Engineering University, China)</i>	
Affordance Segmentation of Hand-Occluded Containers from Exocentric Images .....	1890
<i>Tommaso Apicella (University of Genoa, Italy; Queen Mary University of London, U.K.), Alessio Xompero (Queen Mary University of London, U.K.), Edoardo Ragusa (University of Genoa, Italy), Riccardo Berta (University of Genoa, Italy), Andrea Cavallaro (Queen Mary University of London, U.K.; Idiap Research Institute, Switzerland; École Polytechnique Fédérale de Lausanne, Switzerland), and Paolo Gastaldo (University of Genoa, Italy)</i>	
Multi-Camera 3D Position Estimation Using Conditional Random Field .....	1900
<i>Shusuke Matsuda (Osaka University), Nattaon Techasartikul (Osaka University), and Hideyuki Shimonishi (Osaka University)</i>	
VLMAH: Visual-Linguistic Modeling of Action History for Effective Action Anticipation .....	1909
<i>Victoria Manousaki (University of Crete; Institute of Computer Science, FORTH), Konstantinos Bacharidis (University of Crete; Institute of Computer Science, FORTH), Konstantinos Papoutsakis (Hellenic Mediterranean University), and Antonis Argyros (University of Crete; Institute of Computer Science, FORTH)</i>	

Towards Estimation of Human Intent in Assistive Robotic Teleoperation Using Kinaesthetic and Visual Feedback .....	1920
<i>Muneeb Ahmed (Indian Institute of Technology Delhi, India), Brejesh Lall (Indian Institute of Technology Delhi, India), Rajesh Kumar (Addverb Technologies, India), and Arzad A. Kherani (Indian Institute of Technology Bhilai, India)</i>	
SHOWMe: Benchmarking Object-Agnostic Hand-Object 3D Reconstruction .....	1927
<i>Anilkumar Swamy (NAVER LABS Europe; Inria centre at the University Grenoble Alpes, France), Vincent Leroy (NAVER LABS Europe, France), Philippe Weinzaepfel (NAVER LABS Europe, France), Fabien Baradel (NAVER LABS Europe, France), Salma Galaoui (NAVER LABS Europe, France), Romain Brégier (NAVER LABS Europe, France), Matthieu Armando (NAVER LABS Europe, France), Jean-Sebastien Franco (Inria centre at the University Grenoble Alpes, France), and Grégory Rogez (NAVER LABS Europe, France)</i>	
Learnt Contrastive Concept Embeddings for Sign Recognition .....	1937
<i>Ryan Wong (University of Surrey, United Kingdom), Necati Cihan Camgoz (Meta Reality Labs, Switzerland), and Richard Bowden (University of Surrey, United Kingdom)</i>	
Is Context all you Need? Scaling Neural Sign Language Translation to Large Domains of Discourse .....	1947
<i>Ozge Mercanoglu Sincan (University of Surrey, United Kingdom), Necati Cihan Camgoz (Meta Reality Labs, Switzerland), and Richard Bowden (University of Surrey, United Kingdom)</i>	
A New Dataset for End-to-End Sign Language Translation: The Greek Elementary School Dataset .....	1958
<i>Andreas Voskou (Cyprus University of Technology, Cyprus), Konstantinos P. Panousis (Cyprus University of Technology, Cyprus), Harris Partaourides (AI Cyprus Ethical Novelties Ltd, Cyprus), Kyriakos Toliias (Cyprus University of Technology, Cyprus), and Sotirios Chatziz (Cyprus University of Technology, Cyprus)</i>	
Multimodal Error Correction with Natural Language and Pointing Gestures .....	1968
<i>Stefan Constantin (Karlsruhe Institute of Technology, Germany), Feoziye Irem Eyiokur (Karlsruhe Institute of Technology, Germany), Dogucan Yaman (Karlsruhe Institute of Technology, Germany), Leonard Bärmann (Karlsruhe Institute of Technology, Germany), and Alex Waibel (Karlsruhe Institute of Technology, Germany)</i>	
Modeling Visual Impairments with Artificial Neural Networks: a Review .....	1979
<i>Lucia Schiatti (Massachusetts Institute of Technology, USA; Istituto Italiano di Tecnologia, Italy), Monica Gori (Istituto Italiano di Tecnologia, Italy), Martin Schrimpf (NeuroX Institute, EPFL, Switzerland), Giulia Cappagli (Istituto Italiano di Tecnologia, Italy), Federica Morelli (IRCCS Mondino Foundation, Italy; University of Pavia, Italy), Sabrina Signorini (IRCCS Mondino Foundation, Italy), Boris Katz (Massachusetts Institute of Technology, USA), and Andrei Barbu (Massachusetts Institute of Technology, USA)</i>	
Continuous Hand Gesture Recognition for Human-Robot Collaborative Assembly .....	1992
<i>Bogdan Kwolek (AGH University of Krakow, Poland)</i>	

From Scarcity to Understanding: Transfer Learning for the Extremely Low Resource Irish Sign Language .....	2000
<i>Ruth Holmes (SFI Lero – Trinity College Dublin), Ellen Rushe (SFI Lero – Trinity College Dublin), Mathieu De Coster (IDLab-AIRO – Ghent University – imec), Maxim Bonnaerens (IDLab-AIRO – Ghent University – imec), Shin'ichi Satoh (National Institute of Informatics), Akihiro Sugimoto (National Institute of Informatics), and Anthony Ventresque (SFI Lero, Trinity College Dublin)</i>	
FewFaceNet: A Lightweight Few-Shot Learning-Based Incremental Face Authentication for Edge Cameras .....	2010
<i>Abu Sufian (University of Gour Banga, India), Anirudha Ghosh (Visva-Bharati, India), Debaditya Barman (Visva-Bharati, India), Marco Leo (CNR-ISASI, Italy), Cosimo Distante (CNR-ISASI, Italy), and Baihua Li (Loughborough University, UK)</i>	

## 1st Workshop on Open-Vocabulary 3D Scene Understanding (OpenSUN3D)

CLIP Goes 3D: Leveraging Prompt Tuning for Language Grounded 3D Recognition .....	2020
<i>Deepti Hegde (Johns Hopkins University), Jeya Maria Jose Valanarasu (Johns Hopkins University), and Vishal Patel (Johns Hopkins University)</i>	
Learning to Prompt CLIP for Monocular Depth Estimation: Exploring the Limits of Human Language .....	2031
<i>Dylan Auty (Imperial College London, United Kingdom) and Krystian Mikolajczyk (Imperial College London, United Kingdom)</i>	
CLIP-FO3D: Learning Free Open-World 3D Scene Representations from 2D Dense CLIP .....	2040
<i>Junbo Zhang (Tsinghua University, China), Runpei Dong (Xi'an Jiaotong University, China), and Kaisheng Ma (Tsinghua University, China)</i>	
The Change You Want to See (Now in 3D) .....	2052
<i>Ragav Sachdeva (University of Oxford) and Andrew Zisserman (University of Oxford)</i>	
Dynamic Texts From UAV Perspective Natural Images .....	2062
<i>Hidetomo Sakaino (Visual Recognition Group, Weather Transportation Lab, Weathernews Inc., Japan)</i>	

## Recovering 6D Object Pose (R6D)

SpyroPose: SE(3) Pyramids for Object Pose Distribution Estimation .....	2074
<i>Rasmus Laurvig Haugaard (University of Southern Denmark, Denmark), Frederik Hagelskjær (University of Southern Denmark, Denmark), and Thorbjørn Mosekjær Iversen (University of Southern Denmark, Denmark)</i>	
Diff3DHPE: A Diffusion Model for 3D Human Pose Estimation .....	2084
<i>Jieming Zhou (Australian National University, Australia), Tong Zhang (EPFL, Switzerland), Zeeshan Hayder (CSIRO, Australia), Lars Petersson (CSIRO, Australia), and Mehrtash Harandi (Monash University, Australia)</i>	

Revisiting Fully Convolutional Geometric Features for Object 6D Pose Estimation .....	2095
<i>Jaime Corsetti (Fondazione Bruno Kessler, Italy), Davide Boscaini (Fondazione Bruno Kessler, Italy), and Fabio Poiesi (Fondazione Bruno Kessler, Italy)</i>	
Accidental Turntables: Learning 3D Pose by Watching Objects Turn .....	2105
<i>Zezhou Cheng (UMass Amherst), Matheus Gadelha (Adobe Research), and Subhransu Maji (UMass Amherst)</i>	
NeRF-Pose: A First-Reconstruct-Then-Regress Approach for Weakly-Supervised 6D Object Pose Estimation .....	2115
<i>Fu Li (National University of Defense Technology; Technical University of Munich), Shishir Reddy Vutukur (Technical University of Munich; Siemens AG), Hao Yu (Technical University of Munich), Ivan Shugurov (Technical University of Munich; Siemens AG), Benjamin Busam (Technical University of Munich), Shaowu Yang (National University of Defense Technology), and Slobodan Ilic (Technical University of Munich; Siemens AG)</i>	
CNOS: A Strong Baseline for CAD-Based Novel Object Segmentation .....	2126
<i>Van Nguyen Nguyen (Ecole des Ponts ParisTech, France), Thibault Groueix (Adobe, USA), Georgy Ponimatkin (Ecole des Ponts ParisTech, France), Vincent Lepetit (Ecole des Ponts ParisTech, France), and Tomas Hodan (Meta Reality Labs, USA)</i>	
Reconstruction of 3D Interaction Models from Images Using Shape Prior .....	2133
<i>Mehrshad Mirmohammadi (Sharif University of Technology), Parham Saremi (Sharif University of Technology), Yen-Ling Kuo (University of Virginia), and Xi Wang (ETH Zurich)</i>	
PoseMatcher: One-Shot 6D Object Pose Estimation by Deep Feature Matching .....	2140
<i>Pedro Castro (Imperial College London, United Kingdom) and Tae-Kyun Kim (Imperial College London, United Kingdom)</i>	

## **Visual Perception for Navigation in Human Environments: The JackRabbit Human Motion Forecasting Dataset and Benchmark (JRDB)**

SegDA: Maximum Separable Segment Mask with Pseudo Labels for Domain Adaptive Semantic Segmentation .....	2150
<i>Anant Khandelwal (Applied Scientist, Amazon)</i>	
Cross-Dimensional Refined Learning for Real-Time 3D Visual Perception from Monocular Video. ....	2161
<i>Ziyang Hong (Hong Kong University of Science and Technology) and C. Patrick Yue (Hong Kong University of Science and Technology)</i>	
A Lightweight Skeleton-Based 3D-CNN for Real-Time Fall Detection and Action Recognition .....	2171
<i>Nadhira Noor (Inha University, Korea) and In Kyu Park (Inha University)</i>	
A Dual Perspective of Human Motion Analysis - 3D Pose Estimation and 2D Trajectory Prediction .....	2181
<i>Mayssa Zaier (IMT Nord Europe, University of Lille, France), Hazem Wannous (IMT Nord Europe, University of Lille, France), Hassen Drira (University of Strasbourg, France), and Jacques Boonaert (IMT Nord Europe, University of Lille, France)</i>	



THÖR-Magni: Comparative Analysis of Deep Learning Models for Role-Conditioned Human Motion Prediction .....	2192
<i>Tiago Rodrigues de Almeida (Örebro University, Sweden), Andrey Rudenko (Robert Bosch GmbH, Germany), Tim Schreiter (Örebro University, Sweden), Yufei Zhu (Örebro University, Sweden), Eduardo Gutierrez Maestro (Örebro University, Sweden), Lucas Morillo-Mendez (Örebro University, Sweden), Tomasz P. Kucner (Aalto University; Finnish Center for Artificial Intelligence, Finland), Oscar Martinez Mozos (Örebro University, Sweden), Martin Magnusson (Örebro University, Sweden), Luigi Palmieri (Robert Bosch GmbH, Germany), Kai O. Arras (Robert Bosch GmbH, Germany), and Achim J. Lilienthal (Örebro University, Sweden)</i>	

## Women in Computer Vision (WiCV)

SynDrone - Multi-Modal UAV Dataset for Urban Scenarios .....	2202
<i>Giulia Rizzoli (University of Padova, Italy), Francesco Barbato (University of Padova, Italy), Matteo Caligiuri (University of Padova, Italy), and Pietro Zanuttigh (University of Padova, Italy)</i>	
Do Planar Constraints Improve Camera Pose Estimation in Monocular SLAM? .....	2213
<i>Charlotte Arndt (Robert Bosch GmbH &amp; University of Zaragoza), Reza Sabzevari (Center for Artificial Intelligence Bosch Research, Germany), and Javier Civera (University of Zaragoza, Spain)</i>	
LightNet: Generative Model for Enhancement of Low-Light Images .....	2223
<i>Chaitra Desai (KLE Technological University, India), Nikhil Akalwadi (KLE Technological University, India), Amogh Joshi (KLE Technological University, India), Sampada Malagi (KLE Technological University, India), Chinmayee Mandi (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>	
Domain Adversarial Learning Towards Underwater Image Enhancement .....	2233
<i>Meghna Kapoor (Indian Institute of Technology Jammu, India), Rohan Baghel (Indian Institute of Technology Jammu, India), Badri Narayan Subudhi (Indian Institute of Technology Jammu, India), Vinit Jakhetiya (Indian Institute of Technology Jammu, India), and Ankur Bansal (Indian Institute of Technology Jammu, India)</i>	
Embedded Deformation-Based Compression for Human 3D Dynamic Meshes with Changing Topology .....	2244
<i>Huong Hoang (University of California, USA), Kunyao Chen (University of California; Qualcomm Institute, USA), Truong Nguyen (University of California, USA), and Pamela Cosman (University of California, USA)</i>	
Window-Based Model Averaging Improves Generalization in Heterogeneous Federated Learning .....	2255
<i>Debora Caldarola (Politecnico di Torino), Barbara Caputo (Politecnico di Torino), and Marco Ciccone (Politecnico di Torino)</i>	

TP-NoDe: Topology-Aware Progressive Noising and Denoising of Point Clouds Towards Upsampling .....	2264
<i>Akash Kumbar (KLE Technological University, India), Tejas Anvekar (KLE Technological University, India), Tulasi Amitha Vikrama (KLE Technological University, India), Ramesh Ashok Tabib (KLE Technological University, India), and Uma Mudenagudi (KLE Technological University, India)</i>	

## 2nd Workshop on Computer Vision for Automated Medical Diagnosis (CVAMD)

Mirror U-Net: Marrying Multimodal Fission with Multi-Task Learning for Semantic Segmentation in Medical Imaging .....	2275
<i>Zdravko Marinov (Karlsruhe Institute of Technology; HIDSS4Health - Helmholtz Information and Data Science School for Health, Germany), Simon Reiß (Karlsruhe Institute of Technology, Germany), David Kersting (German Cancer Consortium (DKTK); University Hospital Essen; West German Cancer Center, Germany), Jens Kleesiek (University Hospital Essen, Germany), and Rainer Stiefelhagen (Karlsruhe Institute of Technology, Germany)</i>	
SEPAL: Spatial Gene Expression Prediction from Local Graphs .....	2286
<i>Gabriel Mejia (Universidad de Los Andes, Colombia), Paula Cárdenas (Universidad de Los Andes, Colombia), Daniela Ruiz (Universidad de Los Andes, Colombia), Angela Castillo (Universidad de Los Andes, Colombia), and Pablo Arbeláez (Universidad de Los Andes, Colombia)</i>	
Transformers Pay Attention to Convolutions: Leveraging Emerging Properties of ViTs by Dual Attention-Image Network .....	2296
<i>Yousef Yeganeh (Technical University of Munich, Germany; MCML), Azade Farshad (Technical University of Munich, Germany; MCML), Peter Weinberger (Technical University of Munich, Germany), Seyed-Ahmad Ahmadi (NVIDIA, Germany), Ehsan Adeli (Stanford University, USA), and Nassir Navab (Technical University of Munich, Germany; Johns Hopkins University)</i>	
RRc-UNet 3D for Lung Tumor Segmentation from CT Scans of Non-Small Cell Lung Cancer Patients .....	2308
<i>Van-Linh Le (MONC team - INRIA Bordeaux Sud-Ouest; University of Bordeaux; Institute Bergonié, France) and Olivier Saut (MONC team - INRIA Bordeaux Sud-Ouest; University of Bordeaux, France)</i>	
Topo-CXR: Chest X-ray TB and Pneumonia Screening with Topological Machine Learning .....	2318
<i>Faisal Ahmed (UT Dallas), Brighton Nuwagira (UT Dallas), Furkan Torlak (UT Southwestern), and Baris Coskunuzer (UT Dallas)</i>	
Contrastive Image Synthesis and Self-Supervised Feature Adaptation for Cross-Modality Biomedical Image Segmentation .....	2329
<i>Xinrong Hu (University of Notre Dame), Corey Wang (Northwestern University), and Yiyu Shi (University of Notre Dame)</i>	

Cross-Grained Contrastive Representation for Unsupervised Lesion Segmentation in Medical Images .....	2339
<i>Ziqi Yu (Fudan University, China), Botao Zhao (Zhejiang Lab, China), Yipin Zhang (Fudan University, China), Shengjie Zhang (Fudan University, China), Xiang Chen (Fudan University, China), Haibo Yang (Fudan University, China), Tingying Peng (Helmholtz zentrum Muenchen, Germany), and Xiao-Yong Zhang (Fudan University, China)</i>	
Semi-Supervised Quality Evaluation of Colonoscopy Procedures .....	2347
<i>Idan Kligovasser (Verily, Israel), George Leifman (Verily, Israel), Roman Goldenberg (Verily, Israel), Ehud Riolin (Verily, Israel), and Michael Elad (Verily, Israel)</i>	
Ensuring a Connected Structure for Retinal Vessels Deep-Learning Segmentation .....	2356
<i>Idris Dulau (Bordeaux University, France), Catherine Helmer (Bordeaux University, France), Cecile Delcourt (Bordeaux University, France), and Marie Beurton-Aimar (Bordeaux University, France)</i>	
CLIPath: Fine-Tune CLIP with Visual Feature Fusion for Pathology Image Analysis Towards Minimizing Data Collection Efforts .....	2366
<i>Zhengfeng Lai (University of California Davis), Zhuoheng Li (University of California Davis), Luca Cerny Oliveira (University of California Davis), Joohi Chauhan (University of California Davis), Brittany N. Dugger (University of California Davis), and Chen-Nee Chuah (University of California Davis)</i>	
Implicit Neural Representation in Medical Imaging: A Comparative Survey .....	2373
<i>Amirali Molaei (Iran University of Science and Technology), Amirhossein Aminimehr (Iran University of Science and Technology), Armin Tavakoli (Iran University of Science and Technology), Amirhossein Kazerouni (Iran University of Science and Technology), Bobby Azad (South Dakota State University, USA), Reza Azad (RWTH Aachen University, Germany), and Dorit Merhof (University of Regensburg; Fraunhofer Institute for Digital Medicine MEVIS, Germany)</i>	
HyperCoil-Recon: A Hypernetwork-Based Adaptive Coil Configuration Task Switching Network for MRI Reconstruction .....	2384
<i>Sriprabha Ramanarayanan (Indian Institute of Technology Madras (IITM); Healthcare Technology Innovation Centre, IITM, India), Mohammad Al Fahim (Indian Institute of Technology Madras (IITM), India), Rahul G S (Indian Institute of Technology Madras (IITM); Healthcare Technology Innovation Centre, IITM, India), Amrit Kumar Jethi (Indian Institute of Technology Madras (IITM), India), Keerthi Ram (Healthcare Technology Innovation Centre, IITM, India), and Mohanasankar Sivaprakasam (Indian Institute of Technology Madras (IITM); Healthcare Technology Innovation Centre, IITM, India)</i>	
Self-Supervised Anomaly Detection from Anomalous Training Data via Iterative Latent Token Masking .....	2394
<i>Ashay Patel (Kings College London), Petru-Daniel Tudosiu (Kings College London), Walter H.L. Pinaya (Kings College London), Mark S. Graham (Kings College London), Olusola Adeleke (Kings College London), Gary Cook (Kings College London), Vicky Goh (Kings College London), Sebastien Ourselin (Kings College London), and M. Jorge Cardoso (Kings College London)</i>	

Robust AMD Stage Grading with Exclusively OCTA Modality Leveraging 3D Volume .....	2403
<i>Haochen Zhang (UC San Diego), Anna Heinke (UC San Diego), Carlo Miguel B. Galang (UC San Diego), Daniel N. Deussen (UC San Diego), Bo Wen (UC San Diego), Dirk-Uwe G. Bartsch (UC San Diego), William R. Freeman (UC San Diego), Truong Q. Nguyen (UC San Diego), and Cheolhong An (UC San Diego)</i>	
Segmentation-Based Assessment of Tumor-Vessel Involvement for Surgical Resectability	
Prediction of Pancreatic Ductal Adenocarcinoma .....	2413
<i>Christiaan Viviers (Eindhoven University of Technology, The Netherlands), Mark Ramaekers (Catharina Ziekenhuis, The Netherlands), Amaan Valiuddin (Eindhoven University of Technology, The Netherlands), Terese Hellström (Eindhoven University of Technology, The Netherlands), Nick Tasios (Philips, The Netherlands), John van der Ven (Philips, The Netherlands), Igor Jacobs (Philips, The Netherlands), Lotte Ewals (Catharina Ziekenhuis), Joost Nederend (Catharina Ziekenhuis, The Netherlands), Peter de With (Eindhoven University of Technology, The Netherlands), Misha Luyer (Catharina Ziekenhuis, The Netherlands), and Fons van der Sommen (Eindhoven University of Technology, The Netherlands)</i>	
Sharing is Caring: Concurrent Interactive Segmentation and Model Training Using a Joint Model .....	2424
<i>Ioan Mikhailov (Université Clermont Auvergne; SURGAR, France), Benoit Chauveau (SURGAR; CHU de Clermont-Ferrand, France), Nicolas Bourdel (Université Clermont Auvergne; SURGAR; CHU de Clermont-Ferrand, France), and Adrien Bartoli (Université Clermont Auvergne; SURGAR; CHU de Clermont-Ferrand, France)</i>	
Robust MSFM Learning Network for Classification and Weakly Supervised Localization .....	2434
<i>Komal Kumar (Jio Institute, India), Balakrishna Pailla (Reliance Jio - Artificial Intelligence Centre of Excellence (AICoE), India), Kalyan Tadepalli (Reliance Jio - Artificial Intelligence Centre of Excellence (AICoE); Sir HN Reliance Foundation Hospital, India), and Sudipta Roy (Jio Institute, India)</i>	
DISGAN: Wavelet-Informed Discriminator Guides GAN to MRI Super-Resolution with Noise Cleaning .....	2444
<i>Qi Wang (Max Planck Institute for Biological Cybernetics, Germany), Lucas Mahler (Max Planck Institute for Biological Cybernetics, Germany), Julius Steiglechner (University Hospital Tübingen; Max Planck Institute for Biological Cybernetics, Germany), Florian Birk (University Hospital Tübingen; Max Planck Institute for Biological Cybernetics, Germany), Klaus Scheffler (University Hospital Tübingen; Max Planck Institute for Biological Cybernetics, Germany), and Gabriele Lohmann (University Hospital Tübingen; Max Planck Institute for Biological Cybernetics, Germany)</i>	
Studying the Impact of Augmentations on Medical Confidence Calibration .....	2454
<i>Adrit Rao (Stanford University; Palo Alto High School), Joon-Young Lee (Adobe Research), and Oliver Aalami (Stanford University)</i>	
Multimodal Contrastive Learning and Tabular Attention for Automated Alzheimer’s Disease Prediction .....	2465
<i>Weichen Huang (St Andrew’s College, Ireland)</i>	

Weakly Semi-Supervised Detector-Based Video Classification with Temporal Context for Lung Ultrasound .....	2475
<i>Gary Y. Li (Philips Research North America), Li Chen (Philips Research North America), Mohsen Zahiri (Philips Research North America), Naveen Balaraju (Philips Research North America), Shubham Patil (Philips Research North America), Courosh Mehanian (Global Health Laboratories), Cynthia Gregory (Oregon Health &amp; Science University), Kenton Gregory (Oregon Health &amp; Science University), Balasundar Raju (Philips Research North America), Jochen Kruecker (Philips Research North America), and Alvin Chen (Philips Research North America)</i>	
Order-ViT: Order Learning Vision Transformer for Cancer Classification in Pathology Images.....	2485
<i>Ju Cheon Lee (Korea University, Republic of Korea) and Jin Tae Kwak (Korea University, Republic of Korea)</i>	
Mind the Clot: Automated LVO Detection on CTA Using Deep Learning .....	2495
<i>Shubham Kumar (Qure.ai, India), Arjun Agarwal (Qure.ai, India), Satish Golla (Qure.ai, India), Swetha Tanamala (Qure.ai, India), Ujjwal Upadhyay (Qure.ai, India), Subhankar Chattoraj (Qure.ai, India), Preetham Putha (Qure.ai, India), and Sasank Chilamkurthy (Qure.ai, India)</i>	
A Comparative Study of Vision Transformer Encoders and Few-Shot Learning for Medical Image Classification .....	2505
<i>Maxat Nurgazin (Nazarbayev University, Kazakhstan) and Nguyen Anh Tu (Nazarbayev University, Kazakhstan)</i>	
RheumaViT: Transformer-Based Model for Automated Scoring of Hand Joints in Rheumatoid Arthritis .....	2514
<i>Alexander Stolpoovskiy (Artificial Intelligence Research Institute (AIRI)), Elizaveta Dakhova (Artificial Intelligence Research Institute (AIRI)), Polina Druzhinina (Artificial Intelligence Research Institute (AIRI); Skolkovo Institute of Science and Technology), Polina Postnikova (V.A. Nasonova Research Institute of Rheumatology), Daniil Kudinsky (V.A. Nasonova Research Institute of Rheumatology), Alexander Smirnov (V.A. Nasonova Research Institute of Rheumatology), Anastasia Sukhinina (V.A. Nasonova Research Institute of Rheumatology), Alexander Lila (V.A. Nasonova Research Institute of Rheumatology), and Anvar Kurmukov (Artificial Intelligence Research Institute (AIRI); Institute for Information Transmission Problems)</i>	
AW-Net: A Novel Fully Connected Attention-Based Medical Image Segmentation Model .....	2524
<i>Debojyoti Pal (Jio Institute, India), Tanushree Meena (Jio Institute, India), Dwarikanath Mahapatra (Inception Institute of AI (IIAI), UAE), and Sudipta Roy (Jio Institute, India)</i>	
Geodesic Regression Characterizes 3D Shape Changes in the Female Brain During Menstruation..	2534
<i>Adele Myers (UC Santa Barbara), Caitlin Taylor (UC Santa Barbara), Emily Jacobs (UC Santa Barbara), and Nina Miolane (UC Santa Barbara)</i>	

Computational Evaluation of the Combination of Semi-Supervised and Active Learning for Histopathology Image Segmentation with Missing Annotations .....	2544
<i>Laura Gálvez Jimenez (Université Libre de Bruxelles, Belgium), Lucile Dierckx (Université Catholique de Louvain, Belgium), Maxime Amodei (Université de Liège, Belgium), Hamed Razavi Khosroshahi (Université Libre de Bruxelles, Belgium), Natarajan Chidambaran (Université de Mons, Belgium), Anh-Thu Phan Ho (Multitel, Belgium), and Alberto Franzin (Université Libre de Bruxelles, Belgium)</i>	
Towards Robust Natural-Looking Mammography Lesion Synthesis on Ipsilateral Dual-Views Breast Cancer Analysis .....	2556
<i>Thanh-Huy Nguyen (National Cheng Kung University, Taiwan; Saigon Precision Medicine Research Center (SAIGONMEC), Vietnam), Quang Hien Kha (Taipei Medical University, Taiwan; Saigon Precision Medicine Research Center (SAIGONMEC), Vietnam), Thai Ngoc Toan Truong (Ho Chi Minh City International University, Vietnam), Ba Thinh Lam (Saigon Precision Medicine Research Center (SAIGONMEC), Vietnam), Ba Hung Ngo (Chonnam National University, Korea), Quang Vinh Dinh (Vietnamese-German University, Vietnam), and Nguyen Quoc Khanh Le (Taipei Medical University; Professional Master Program in Artificial Intelligence in Medicine, Taiwan)</i>	
End-to-End Deep Learning for Reconstructing Segmented 3D CT Image from Multi-Energy X-ray Projections .....	2566
<i>Siqi Wang (The University of Tokyo, Japan), Tatsuya Yatagawa (Hitotsubashi University, Japan), Yutaka Ohtake (The University of Tokyo, Japan), Toru Aoki (Shizuoka University, Japan), and Jun Hotta (Zodiac Co., Ltd., Japan)</i>	
Combating Coronary Calcium Scoring Bias for Non-Gated CT by Semantic Learning on Gated CT .....	2575
<i>Jiajian Li (CVTE Research, China), Anwei Li (CVTE Research, China), Jiansheng Fang (CVTE Research, China), Yonghe Hou (Yibicom Health Management Center, CVTE, China), Chao Song (Yibicom Health Management Center, CVTE, China), Huifang Yang (Yibicom Health Management Center, CVTE, China), Jingwen Wang (CVTE Research, China), Hongbo Liu (CVTE Research, China), and Jiang Liu (Southern University of Science and Technology, China)</i>	
Comprehensive Multimodal Segmentation in Medical Imaging: Combining YOLOv8 with SAM and HQ-SAM Models .....	2584
<i>Sumit Pandey (University of Copenhagen), Kuan-Fu Chen (Chang Gung Memorial Hospital), and Erik B. Dam (University of Copenhagen)</i>	
Semantic Parsing of Colonoscopy Videos with Multi-Label Temporal Networks .....	2591
<i>Ori Kelner (Verily Life Sciences, Israel), Or Weinstein (Verily Life Sciences, Israel), Ehud Rivlin (Verily Life Sciences, Israel), and Roman Goldenberg (Verily Life Sciences, Israel)</i>	

Towards Fixing Clever-Hans Predictors with Counterfactual Knowledge Distillation .....	2599
<i>Sidney Bender (Technische Universität Berlin, Germany), Christopher J. Anders (Technische Universität Berlin; BIFOLD – Berlin Institute for the Foundations of Learning and Data, Germany), Pattarawat Chormai (Technische Universität Berlin; Max Planck School of Cognition, Germany), Heike Antje Marxfeld (BASF SE, Germany), Jan Herrmann (BASF SE, Germany), and Grégoire Montavon (Freie Universität Berlin; Technische Universität Berlin, Germany)</i>	
Causality-Driven One-Shot Learning for Prostate Cancer Grading from MRI .....	2608
<i>Gianluca Carloni (Institute of Information Science and Technologies (ISTI-CNR); University of Pisa, Italy), Eva Pachetti (Institute of Information Science and Technologies (ISTI-CNR); University of Pisa, Italy), and Sara Colantonio (Institute of Information Science and Technologies (ISTI-CNR), Italy)</i>	
ShaRPy: Shape Reconstruction and Hand Pose Estimation from RGB-D with Uncertainty .....	2617
<i>Vanessa Wirth (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Germany), Anna-Maria Liphardt (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg; University Hospital Erlangen, Germany), Birte Coppers (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg; University Hospital Erlangen, Germany), Johanna Bräunig (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Germany), Simon Heinrich (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Germany), Sigrid Leyendecker (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Germany), Arnd Kleyer (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg; University Hospital Erlangen, Germany), Georg Schett (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg; University Hospital Erlangen, Germany), Martin Vossiek (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Germany), Bernhard Egger (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Germany), and Marc Stamminger (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Germany)</i>	
An Empirical Analysis for Zero-Shot Multi-Label Classification on COVID-19 CT Scans and Uncurated Reports .....	2626
<i>Ethan Dack (ARTORG Center for Biomedical Engineering Research; University of Bern), Lorenzo Brigato (ARTORG Center for Biomedical Engineering Research; University of Bern), Matthew McMurray (Bern University Hospital), Matthias Fontanellaz (ARTORG Center for Biomedical Engineering Research; University of Bern), Thomas Frauenfelder (Bern University Hospital), Hanno Hoppe (University of Bern; Lindenhofspital; Campus Stiftung Lindenhof Bern (SLB)), Aristomenis Exadaktylos (Bern University Hospital), Thomas Geiser (Bern University Hospital), Manuela Funke-Chambour (Bern University Hospital), Andreas Christe (Bern University Hospital), Lukas Ebner (Bern University Hospital), and Stavroula Mouggiakakou (ARTORG Center for Biomedical Engineering Research; University of Bern)</i>	

Fusion Approaches to Predict Post-Stroke Aphasia Severity from Multimodal Neuroimaging Data .....	2636
<i>Saurav Chennuri (Boston University), Sha Lai (Boston University), Anne Billot (Boston University), Maria Varkanitsa (Boston University), Emily J. Braun (Boston University), Swathi Kiran (Boston University), Archana Venkataraman (Boston University), Janusz Konrad (Boston University), Prakash Ishwar (Boston University), and Margrit Betke (Boston University)</i>	
Self-Supervised Semantic Segmentation: Consistency over Transformation .....	2646
<i>Sanaz Karimijafarbigloo (University of Regensburg, Germany), Reza Azad (RWTH Aachen University, Germany), Amirhossein Kazerouni (Iran University of Science and Technology, Iran), Yury Velichko (Northwestern University, USA), Ulas Bagci (Northwestern University, USA), and Dorit Merhof (University of Regensburg; Fraunhofer Institute for Digital Medicine MEVIS, Germany)</i>	
ALFA – Leveraging All Levels of Feature Abstraction for Enhancing the Generalization of Histopathology Image Classification Across Unseen Hospitals .....	2656
<i>Milad Sikaroudi (Kimia Lab, University of Waterloo, Canada), Maryam Hosseini (Kimia Lab, University of Waterloo, Canada), Shahryar Rahnamayan (Brock University, Canada), and H. R. Tizhoosh (Rhazes Lab, Mayo Clinic, USA)</i>	
Pathology-Based Ischemic Stroke Etiology Classification via Clot Composition Guided Multiple Instance Learning .....	2666
<i>Mara Pleasure (UCLA), Ekaterina Redekop (UCLA), Jennifer S Polson (UCLA), Haoyue Zhang (UCLA), Naoki Kaneko (UCLA), William Speier (UCLA), and Corey W Arnold (UCLA)</i>	
Enhancing Medical Image Segmentation: Optimizing Cross-Entropy Weights and Post-Processing with Autoencoders .....	2676
<i>Pranav Singh (New York University), Luoyao Chen (New York University), Mei Chen (New York University), Jinqian Pan (New York University), Raviteja Chukkapalli (New York University), Shravan Chaudhari (New York University), and Jacopo Cirrone (New York University)</i>	
Using Large Text To Image Models with Structured Prompts for Skin Disease Identification: A Case Study .....	2686
<i>Sajith Rajapaksa (Yuko AI Inc., Canada; SickKids Hospital, Canada), Jean Marie Uwabeza Vianney (Yuko AI Inc., Canada), Renell Castro (Yuko AI Inc., Canada), Farzad Khalvati (University of Toronto, Canada; SickKids Hospital, Canada; Vector Institute, Canada), and Shubhra Aich (Yuko AI Inc., Canada; Carnegie Mellon University, USA)</i>	
CheXFusion: Effective Fusion of Multi-View Features Using Transformers for Long-Tailed Chest X-Ray Classification .....	2694
<i>Dongkyun Kim (Carnegie Mellon University, USA)</i>	
Robust Asymmetric Loss for Multi-Label Long-Tailed Learning .....	2703
<i>Wongi Park (Ajou University), Inhyuk Park (Ajou University), Sungeun Kim (Ajou University), and Jongbin Ryu (Ajou University)</i>	
Effect of Stage Training for Long-Tailed Multi-Label Image Classification .....	2713
<i>Yosuke Yamagishi (The University of Tokyo, Japan) and Shohei Hanaoka (The University of Tokyo, Japan)</i>	



Advanced Augmentation and Ensemble Approaches for Classifying Long-Tailed Multi-Label Chest X-Rays .....	2721
<i>Trong-Hieu Nguyen-Mau (University of Science, VNU-HCM, Vietnam), Tuan-Luc Huynh (University of Science, VNU-HCM, Vietnam), Thanh-Danh Le (University of Science, VNU-HCM, Vietnam), Hai-Dang Nguyen (University of Science, VNU-HCM, Vietnam), and Minh-Triet Tran (University of Science, VNU-HCM, Vietnam)</i>	
An Optimized Ensemble Framework for Multi-Label Classification on Long-Tailed Chest X-ray Data .....	2731
<i>Jaehyup Jeong (KT Research &amp; Development Center, KT Corporation, Korea), Bosoung Jeoun (KT Research &amp; Development Center, KT Corporation, Korea), Yeonju Park (KT Research &amp; Development Center, KT Corporation, Korea), and Bohyung Han (ECE &amp; IPAI, Seoul National University, Korea)</i>	
Enhancing Multi-Label Long-Tailed Classification on Chest X-Rays Through ML-GCN Augmentation .....	2739
<i>HyeRyeong Seo (Seoul National University, South Korea), MinHyuk Lee (Seoul National University, South Korea), WooJin Cheong (Seoul National University, South Korea), HyeKyung Yoon (Seoul National University, South Korea), SoHyung Kim (Seoul National University, South Korea), and MyungJoo Kang (Seoul National University, South Korea)</i>	
Chest X-Ray Feature Pyramid Sum Model with Diseased Area Data Augmentation Method .....	2749
<i>Changhyun Kim (SK Telecom, South Korea), Giyeol Kim (Gacheon University, South Korea), Sooyoung Yang (ChungAng University, South Korea), Hyunsu Kim (SungKyunKwan University, South Korea), Sangyool Lee (SK Telecom, South Korea), and Hansu Cho (SK Telecom, South Korea)</i>	

## 5th Workshop on Closing the Loop Between Vision and Language (CLVL)

Sparse Linear Concept Discovery Models .....	2759
<i>Konstantinos Panagiotis Panousis (Inria; University of Montpellier; LIRMM; UMR-Tetis), Dino Ienco (INRAE; University of Montpellier; UMR-Tetis), and Diego Marcos (Inria; University of Montpellier; LIRMM; UMR-Tetis)</i>	
ProVLA: Compositional Image Search with Progressive Vision-Language Alignment and Multimodal Fusion .....	2764
<i>Zhizhang Hu (University of California), Xinliang Zhu (Amazon Visual Search &amp; AR), Son Tran (Amazon M5), René Vidal (Amazon Visual Search &amp; AR), and Arnab Dhua (Amazon Visual Search &amp; AR)</i>	
Vision-Language Models Performing Zero-Shot Tasks Exhibit Disparities Between Gender Groups .....	2770
<i>Melissa Hall (Meta AI), Laura Gustafson (Meta AI), Aaron Adcock (Meta AI), Ishan Misra (Meta AI), and Candace Ross (Meta AI)</i>	
BiLMa: Bidirectional Local-Matching for Text-based Person Re-identification .....	2778
<i>Takuro Fujii (Yokohama National University, Japan) and Shuhei Tarashima (NTT Communications Corporation, Japan)</i>	

Alignment and Generation Adapter for Efficient Video-Text Understanding .....	2783
<i>Han Fang (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), Zhifei Yang (China Telecom Corporation Ltd. Data&amp;AI Technology Company; Beijing University of Posts and Telecommunications, China), Yuhan Wei (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China; Rice University, USA), Xianghao Zang (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), Chao Ban (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), Zerun Feng (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), Zhongjiang He (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), Yongxiang Li (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), and Hao Sun (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China)</i>	
LLaViLo: Boosting Video Moment Retrieval via Adapter-Based Multimodal Modeling .....	2790
<i>Kaijing Ma (China Telecom Corporation Ltd. Data&amp;AI Technology Company; Xi'an Jiaotong University, China), Xianghao Zang (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), Zerun Feng (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), Han Fang (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), Chao Ban (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), Yuhan Wei (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China; Rice University, USA), Zhongjiang He (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), Yongxiang Li (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China), and Hao Sun (China Telecom Corporation Ltd. Data&amp;AI Technology Company, China)</i>	
Zero-Shot and Few-Shot Video Question Answering with Multi-Modal Prompts .....	2796
<i>Deniz Engin (Inria, Univ Rennes, CNRS, IRISA) and Yannis Avrithis (Institute of Advanced Research in Artificial Intelligence (IARAI))</i>	
ECO: Ensembling Context Optimization for Vision-Language Models .....	2803
<i>Lorenzo Agnolucci (University of Florence, Italy), Alberto Baldrati (University of Florence, Italy; University of Pisa, Italy), Francesco Todino (University of Florence, Italy), Federico Becattini (University of Siena, Italy), Marco Bertini (University of Florence, Italy), and Alberto Del Bimbo (University of Florence, Italy)</i>	
A Cross-Dataset Study on the Brazilian Sign Language Translation .....	2808
<i>Amanda Hellen de Avellar Sarmiento (Universidade de São Paulo, Brazil) and Moacir Antonelli Ponti (Universidade de São Paulo, Brazil)</i>	
Context-VQA: Towards Context-Aware and Purposeful Visual Question Answering .....	2813
<i>Nandita Naik (Stanford University), Christopher Potts (Stanford University), and Elisa Kreiss (Stanford University)</i>	
Explaining Vision and Language Through Graphs of Events in Space and Time .....	2818
<i>Mihai Masala (Institute of Mathematics of the Romanian Academy, Romania; University Politehnica of Bucharest, Romania), Nicolae Cudlenco (Institute of Mathematics of the Romanian Academy, Romania), Traian Rebedea (University Politehnica of Bucharest, Romania), and Marius Leordeanu (Institute of Mathematics of the Romanian Academy, Romania; University Politehnica of Bucharest, Romania)</i>	

Mapping Memes to Words for Multimodal Hateful Meme Classification .....	2824
<i>Giovanni Burbi (University of Florence, Italy), Alberto Baldrati (University of Florence, Italy; University of Pisa, Italy), Lorenzo Agnolucci (University of Florence, Italy), Marco Bertini (University of Florence, Italy), and Alberto Del Bimbo (University of Florence, Italy)</i>	
Cross-Modal Dense Passage Retrieval for Outside Knowledge Visual Question Answering .....	2829
<i>Benjamin Reichman (Georgia Institute of Technology) and Larry Heck (Georgia Institute of Technology)</i>	
PatFig: Generating Short and Long Captions for Patent Figures .....	2835
<i>Dana Aubakirova (Université Paris-Saclay, France; Qatent, France), Kim Gerdes (Université Paris-Saclay, France; Qatent, France), and Lufei Liu (Qatent, France)</i>	
An Empirical Study of the Effect of Video Encoders on Temporal Video Grounding .....	2842
<i>Ignacio M. De la Jara (University of Chile, CENIA &amp; IMFD), Cristian Rodriguez-Opazo (Australian Institute for Machine Learning, University of Adelaide), Edison Marrese-Taylor (National Institute of Advanced Industrial Science and Technology), and Felipe Bravo-Marquez (University of Chile, CENIA &amp; IMFD)</i>	
Towards Vision-Language Mechanistic Interpretability: A Causal Tracing Tool for BLIP .....	2848
<i>Vedant Palit (IIT Kharagpur), Rohan Pandey (Reworkd.ai), Aryaman Arora (Georgetown University), and Paul Pu Liang (Carnegie Mellon University)</i>	
Multimodal Neurons in Pretrained Text-Only Transformers .....	2854
<i>Sarah Schwettmann (MIT CSAIL), Neil Chowdhury (MIT CSAIL), Samuel Klein (MIT KFG), David Bau (Northeastern University), and Antonio Torralba (MIT CSAIL)</i>	

## AI for 3D Content Creation (AI3DCC)

Looking at Words and Points with Attention: a Benchmark for text-to-Shape Coherence .....	2860
<i>Andrea Amaduzzi (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>	
BluNF: Blueprint Neural Field .....	2870
<i>Robin Courant (LIX, Ecole Polytechnique), Xi Wang (LIX, Ecole Polytechnique), Marc Christie (Inria, IRISA, CNRS, Univ. Rennes), and Vicky Kalogeiton (LIX, Ecole Polytechnique)</i>	
NeRF-GAN Distillation for Efficient 3D-Aware Generation with Convolutions .....	2880
<i>Mohamad Shahbazi (ETH Zurich, Switzerland), Evangelos Ntavelis (ETH Zurich, Switzerland; ML &amp; Robotics, CSEM, Switzerland), Alessio Tonioni (Google, Switzerland), Edo Collins (Google, Switzerland), Danda Pani Paudel (ETH Zurich, Switzerland), Martin Danelljan (ETH Zurich, Switzerland), and Luc Van Gool (ETH Zurich, Switzerland)</i>	

LatentSwap3D: Semantic Edits on 3D Image GANs .....	2891
<i>Enis Simsar (ETH Zurich, Switzerland; TU Munich, Germany), Alessio Tonioni (Google, Switzerland), Evin Pinar Ornek (Technical University of Munich, Germany), and Federico Tombari (Google, Switzerland; Technical University of Munich, Germany)</i>	
BuilDiff: 3D Building Shape Generation Using Single-Image Conditional Point Cloud Diffusion Models .....	2902
<i>Yao Wei (University of Twente, The Netherlands), George Vosselman (University of Twente, The Netherlands), and Michael Ying Yang (University of Twente, The Netherlands)</i>	
Set-the-Scene: Global-Local Training for Generating Controllable NeRF Scenes .....	2912
<i>Dana Cohen-Bar (Tel Aviv University, Israel), Elad Richardson (Tel Aviv University, Israel), Gal Metzger (Tel Aviv University, Israel), Raja Giryes (Tel Aviv University, Israel), and Daniel Cohen-Or (Tel Aviv University, Israel)</i>	
SPARF: Large-Scale Learning of 3D Sparse Radiance Fields from Few Input Images .....	2922
<i>Abdullah Hamdi (Technical University of Munich (TUM); King Abdullah University of Science and Technology (KAUST); University of Oxford), Bernard Ghanem (King Abdullah University of Science and Technology (KAUST)), and Matthias Nießner (Technical University of Munich (TUM))</i>	
Blended-NeRF: Zero-Shot Object Generation and Blending in Existing Neural Radiance Fields ....	2933
<i>Ori Gordon (The Hebrew University), Omri Avrahami (The Hebrew University), and Dani Lischinski (The Hebrew University)</i>	
S2RF: Semantically Stylized Radiance Fields .....	2944
<i>Moneish Kumar (Carnegie Mellon University), Neeraj Panse (Carnegie Mellon University), and Dishani Lahiri (Carnegie Mellon University)</i>	

## AI for Creative Video Editing and Understanding (CVEU)

Is There Progress in Activity Progress Prediction? .....	2950
<i>Frans de Boer (Delft University of Technology), Jan C. van Gemert (Delft University of Technology), Jouke Dijkstra (Leiden University Medical Center), and Silvia L. Pintea (Delft University of Technology; Leiden University Medical Center)</i>	
Are Current Long-Term Video Understanding Datasets Long-Term? .....	2959
<i>Ombretta Strafforello (TU Delft, TNO), Klammer Schutte (TNO), and Jan van Gemert (TU Delft)</i>	
VAST: Vivify Your Talking Avatar via Zero-Shot Expressive Facial Style Transfer .....	2969
<i>Liyang Chen (Tsinghua University; Microsoft), Zhiyong Wu (Tsinghua University), Runnan Li (Microsoft), Weihong Bao (Tsinghua University), Jun Ling (Shanghai Jiao Tong University), Xu Tan (Microsoft), and Sheng Zhao (Microsoft)</i>	
PAT: Position-Aware Transformer for Dense Multi-Label Action Detection .....	2980
<i>Faegheh Sardari (University of Surrey, UK), Armin Mustafa (University of Surrey, UK), Philip J. B. Jackson (University of Surrey, UK), and Adrian Hilton (University of Surrey, UK)</i>	

Expressive Talking Head Video Encoding in StyleGAN2 Latent Space .....	2990
<i>Trevine Oorloff (University of Maryland, USA) and Yaser Yacoob (University of Maryland, USA)</i>	
Benchmarking Data Efficiency and Computational Efficiency of Temporal Action Localization Models .....	3000
<i>Jan Warchocki (Delft University of Technology, the Netherlands), Teodor Oprescu (Delft University of Technology, the Netherlands), Yunhan Wang (Delft University of Technology, the Netherlands), Alexandru Dămăcuș (Delft University of Technology, the Netherlands), Paul Misterka (Delft University of Technology, the Netherlands), Robert-Jan Bruintjes (Delft University of Technology, the Netherlands), Attila Lengyel (Delft University of Technology, the Netherlands), Ombretta Strafforello (Delft University of Technology, the Netherlands), and Jan van Gemert (Delft University of Technology, the Netherlands)</i>	
InFusion: Inject and Attention Fusion for Multi Concept Zero-Shot Text-Based Video Editing.....	3009
<i>Anant Khandelwal (Glance AI)</i>	
LEMMS: Label Estimation of Multi-Feature Movie Segments .....	3019
<i>Bartolomeo Vacchetti (Politecnico di Torino, Italy), Dawit Mureja (KAIST, South Korea), and Tania Cerquitelli (Politecnico di Torino, Italy)</i>	

## Artificial Social Intelligence Workshop and Challenge (ASI)

Pointing Gesture Recognition via Self-Supervised Regularization for ASD Screening .....	3028
<i>Cheol-Hwan Yoo (ETRI, Republic of Korea), Jang-Hee Yoo (ETRI, Republic of Korea), Ho-Won Kim (ETRI, Republic of Korea), and ByungOk Han (ETRI, Republic of Korea)</i>	
External Commonsense Knowledge as a Modality for Social Intelligence Question-Answering ....	3036
<i>Sanika Natu (Carnegie Mellon University, USA), Shounak Sural (Carnegie Mellon University, USA), and Sulagna Sarkar (Carnegie Mellon University, USA)</i>	
D-ViSA: A Dataset for Detecting Visual Sentiment from Art Images .....	3043
<i>Seoyun Kim (Sungkyunkwan University, Korea), ChaeHee An (Sungkyunkwan University, Korea), Junyeop Cha (Sungkyunkwan University, Korea), Dongjae Kim (Sungkyunkwan University, Korea), and Eunil Park (Sungkyunkwan University, Korea)</i>	
Few Labels are Enough! Semi-Supervised Graph Learning for Social Interaction .....	3052
<i>Nicola Corbellini (University of Genoa, Italy), Jhony H. Giraldo (Télécom Paris - Institut Polytechnique de Paris, France), Giovanna Varni (University of Trento, Italy), and Gualtiero Volpe (University of Genoa, Italy)</i>	
Interaction Acceptance Modelling and Estimation for a Proactive Engagement in the Context of Human-Robot Interactions .....	3061
<i>Timothée Dhaussy (LIA-CERI, Avignon University, France), Bassam Jabaian (LIA-CERI, Avignon University, France), and Fabrice Lefèvre (LIA-CERI, Avignon University, France)</i>	

Multi-Modal Correlated Network with Emotional Reasoning Knowledge for Social Intelligence Question-Answering .....	3067
<i>Baijun Xie (The George Washington University, USA) and Chung Hyuk Park (The George Washington University, USA)</i>	
Just Ask Plus: Using Transcripts for VideoQA .....	3074
<i>Mohammad Javad Pirhadi (Iran University of Science and Technology, Iran), Motahhare Mirzaei (Iran University of Science and Technology, Iran), and Sauleh Eetemadi (Iran University of Science and Technology, Iran)</i>	

## International Workshop on Analysis and Modeling of Faces and Gestures (AMFG)

GANDiffFace: Controllable Generation of Synthetic Datasets for Face Recognition with Realistic Variations .....	3078
<i>Pietro Melzi (Universidad Autonoma de Madrid, Spain), Christian Rathgeb (Hochschule Darmstadt, Germany), Ruben Tolosana (Universidad Autonoma de Madrid, Spain), Ruben Vera-Rodriguez (Universidad Autonoma de Madrid, Spain), Dominik Lawatsch (secunet Security Networks AG, Essen, Germany), Florian Domin (secunet Security Networks AG, Essen, Germany), and Maxim Schaubert (secunet Security Networks AG, Essen, Germany)</i>	
PoseBias: On Dataset Bias and Task Difficulty - Is There an Optimal Camera Position for Facial Image Analysis? .....	3088
<i>Mohit Choithwani (Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)), Sneha Almeida (Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)), and Bernhard Egger (Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU))</i>	
Kinship Representation Learning with Face Componential Relation .....	3097
<i>Wengtai Su (Noyatek), Min-Hung Chen (NVIDIA), Chien-Yi Wang (NVIDIA), Shang-Hong Lai (National Tsing Hua University), and Trista Pei-chun Chen (Microsoft Corporation)</i>	
Disjoint Pose and Shape for 3D Face Reconstruction .....	3107
<i>Raja Kumar (University of California Santa Cruz), Jiahao Luo (University of California Santa Cruz), Alex Pang (University of California Santa Cruz), and James Davis (University of California Santa Cruz)</i>	
M2C: Concise Music Representation for 3D Dance Generation .....	3118
<i>Matthew Marchellus (Inha University, South Korea) and In Kyu Park (Inha University, South Korea)</i>	
Denosing Diffusion for 3D Hand Pose Estimation from Images .....	3128
<i>Maksym Ivashechkin (University of Surrey, United Kingdom), Oscar Mendez (University of Surrey, United Kingdom), and Richard Bowden (University of Surrey, United Kingdom)</i>	
POSTER: A Pyramid Cross-Fusion Transformer Network for Facial Expression Recognition .....	3138
<i>Ce Zheng (University of Central Florida, USA), Matias Mendieta (University of Central Florida, USA), and Chen Chen (University of Central Florida, USA)</i>	

Dynamic Multiview Refinement of 3D Hand Datasets Using Differentiable Ray Tracing .....	3148
<i>Giorgos Karvounas (Institute of Computer Science, FORTH; University of Crete, Greece), Nikolaos Kyriazis (Institute of Computer Science, FORTH, Greece), Iason Oikonomidis (Institute of Computer Science, FORTH, Greece), and Antonis Argyros (Institute of Computer Science, FORTH; University of Crete, Greece)</i>	
Controllable Inversion of Black-Box Face Recognition Models via Diffusion .....	3159
<i>Manuel Kansy (ETH Zurich, Switzerland; DisneyResearch, Studios, Switzerland), Anton Raël (ETH Zurich, Switzerland), Graziana Mignone (DisneyResearch, Studios, Switzerland), Jacek Naruniec (DisneyResearch, Studios, Switzerland), Christopher Schroers (DisneyResearch, Studios, Switzerland), Markus Gross (ETH Zurich, Switzerland; DisneyResearch, Studios, Switzerland), and Romann M. Weber (DisneyResearch, Studios, Switzerland)</i>	
A Unified Approach for Occlusion Tolerant 3D Facial Pose Capture and Gaze Estimation Using MocapNETs .....	3170
<i>Ammar Qammar (University of Crete; Institute of Computer Science, FORTH, Greece) and Antonis A. Argyros (University of Crete; Institute of Computer Science, FORTH, Greece)</i>	
A Gated Attention Transformer for Multi-Person Pose Tracking .....	3181
<i>Andreas Doering (University of Bonn, Germany; Lamarr Institute for Machine Learning and Artificial Intelligence) and Juergen Gall (University of Bonn, Germany; Lamarr Institute for Machine Learning and Artificial Intelligence)</i>	
Occluded Gait Recognition via Silhouette Registration Guided by Automated Occlusion Degree Estimation .....	3191
<i>Chi Xu (Osaka University, Japan), Shogo Tsuji (Osaka University, Japan), Yasushi Makihara (Osaka University, Japan), Xiang Li (Osaka University, Japan), and Yasushi Yagi (Osaka University, Japan)</i>	
Unraveling a Decade: A Comprehensive Survey on Isolated Sign Language Recognition .....	3202
<i>Noha Sarhan (Universität Hamburg, Germany) and Simone Frintrop (Universität Hamburg, Germany)</i>	
DiffHPE: Robust, Coherent 3D Human Pose Lifting with Diffusion .....	3212
<i>Cedric Rommel (Valeo.ai, Paris, France), Eduardo Valle (Valeo.ai, Paris, France; University of Campinas, Brazil), Mickael Chen (Valeo.ai, Paris, France), Souhaïel Khalfaoui (Valeo - Applied ML, Paris, France), Renaud Marlet (Valeo.ai, Paris, France; Univ Gustave Eiffel, CNRS, Marne-la-Vallee, France), Matthieu Cord (Valeo.ai, Paris, France; Sorbonne Université, Paris, France), and Patrick Perez (Valeo.ai, Paris, France)</i>	

## The second ROAD workshop & challenge: Event detection for situation awareness in autonomous driving (ROAD++)

STRIDE: Street View-Based Environmental Feature Detection and Pedestrian Collision Prediction .....	3222
<i>Cristina González (Universidad de los Andes, Colombia), Nicolás Ayobi (Universidad de los Andes, Colombia), Felipe Escallón (Universidad de los Andes, Colombia), Laura Baldovino-Chiquillo (Universidad de los Andes, Colombia), Maria Wilches-Mogollón (Universidad de los Andes, Colombia), Donny Pasos (Universidad de los Andes, Colombia), Nicole Ramírez (Universidad de los Andes, Colombia), Jose Pinzón (Pontificia Universidad Javeriana, Colombia; Universidad de los Andes, Colombia), Olga Sarmiento (Universidad de los Andes, Colombia), D Alex Quistberg (Drexel University, USA; Universidad de los Andes, Colombia), and Pablo Arbeláez (Universidad de los Andes, Colombia)</i>	
Surround-View Vision-Based 3D Detection for Autonomous Driving: A Survey .....	3235
<i>Apoorv Singh (Motional, Carnegie Mellon University)</i>	
Tracing the Influence of Predecessors on Trajectory Prediction .....	3245
<i>Mengmeng Liu (University of Twente, The Netherlands), Hao Cheng (University of Twente, The Netherlands), and Michael Ying Yang (University of Twente, The Netherlands)</i>	
Traffic Mirror Detection and Annotation Methods from Street Images of Open Data for Preventing Accidents at Intersections by Alert .....	3256
<i>Da Li (Fukuoka University, Japan), Hikaru Hagura (Kyoto Sangyo University, Japan), Taichi Miyabashira (Fukuoka University, Japan), Yukiko Kawai (Kyoto Sankyo University, Japan), and Shintaro Ono (Fukuoka University, Japan)</i>	
Efficient, Self-Supervised Human Pose Estimation with Inductive Prior Tuning .....	3263
<i>Nobline Yoo (Princeton University) and Olga Russakovsky (Princeton University)</i>	
Temporal DINO: A Self-Supervised Video Strategy to Enhance Action Prediction .....	3273
<i>Izzeddin Teeti (VAIL, Oxford Brookes University), Rongali Sai Bhargav (Indian Institute of Technology, Bombay), Vivek Singh (VAIL, Oxford Brookes University), Andrew Bradley (VAIL, Oxford Brookes University), Biplab Banerjee (Indian Institute of Technology, Bombay), and Fabio Cuzzolin (VAIL, Oxford Brookes University)</i>	

## Visual Continual Learning (VCL)

Progressive Feature Adjustment for Semi-Supervised Learning from Pretrained Models .....	3284
<i>Hai-Ming Xu (University of Adelaide), Lingqiao Liu (University of Adelaide), Hao Chen (Zhejiang University), Ehsan Abbasnejad (University of Adelaide), and Rafael Felix (University of Adelaide)</i>	
OpenIncrement: A Unified Framework for Open Set Recognition and Deep Class-Incremental Learning .....	3295
<i>Jiawen Xu (Technical University Berlin), Claas Grohnfeldt (Huawei Technologies), and Odej Kao (Technical University Berlin)</i>	
Transformer-Based Sensor Fusion for Autonomous Driving: A Survey .....	3304
<i>Apoorv Singh (Motional, Carnegie Mellon University)</i>	



Trajectory-Prediction with Vision: A Survey .....	3310
<i>Apoorv Singh (Motional, Carnegie Mellon University)</i>	
Memory-Augmented Variational Adaptation for Online Few-Shot Segmentation .....	3316
<i>Jie Liu (University of Amsterdam, Netherlands), Yingjun Du (University of Amsterdam, Netherlands), Zehao Xiao (University of Amsterdam, Netherlands), Cees G.M Snoek (University of Amsterdam, Netherlands), Jan-Jakob Sonke (University of Amsterdam, Netherlands), and Efstratios Gavves (University of Amsterdam, Netherlands)</i>	
Instant Continual Learning of Neural Radiance Fields .....	3326
<i>Ryan Po (Stanford University), Zhengyang Dong (Stanford University), Alexander W. Bergman (Stanford University), and Gordon Wetzstein (Stanford University)</i>	
ScrollNet: DynamicWeight Importance for Continual Learning .....	3337
<i>Fei Yang (Universitat Autònoma de Barcelona, Spain), Kai Wang (Universitat Autònoma de Barcelona, Spain), and Joost van de Weijer (Universitat Autònoma de Barcelona, Spain)</i>	
Identification of Novel Classes for Improving Few-Shot Object Detection .....	3348
<i>Zeyu Shangguan (University of Southern California) and Mohammad Rostami (University of Southern California)</i>	
SAM-Adapter: Adapting Segment Anything in Underperformed Scenes .....	3359
<i>Tianrun Chen (Zhejiang University; KOKONI, Moxin (Huzhou) Technology, China), Lanyun Zhu (Singapore University of Technology and Design, Singapore), Chaotao Deng (Huzhou University, China), Runlong Cao (Huzhou University, China), Yan Wang (Beihang University, China), Shangzhan Zhang (Zhejiang University, China), Zejian Li (Zhejiang University, China), Lingyun Sun (Zhejiang University, China), Ying Zang (Huzhou University, China), and Papa Mao (KOKONI, Moxin (Huzhou) Technology, China)</i>	
Experience Replay as an Effective Strategy for Optimizing Decentralized Federated Learning.....	3368
<i>Matteo Pennisi (University of Catania, Italy), Federica Proietto Salanitri (University of Catania, Italy), Giovanni Bellitto (University of Catania, Italy), Concetto Spampinato (University of Catania, Italy), Simone Palazzo (University of Catania, Italy), Bruno Casella (University of Turin, Italy), and Marco Aldinucci (University of Turin, Italy)</i>	
Clustering-Based Domain-Incremental Learning .....	3376
<i>Christiaan Lamers (Norce, Norway), René Vidal (University of Pennsylvania, USA), Nabil Belbachir (Norce, Norway), Niki van Stein (Leiden Institute of Advanced Computer Science, The Netherlands), Thomas Bäck (Leiden Institute of Advanced Computer Science, The Netherlands), and Paris Giampouras (Johns Hopkins University, USA)</i>	
Multimodal Parameter-Efficient Few-Shot Class Incremental Learning .....	3385
<i>Marco D’Alessandro (Neuraptic AI, Spain), Alberto Alonso (Neuraptic AI, Spain), Enrique Calabrés (Neuraptic AI, Spain), and Mikel Galar (Public University of Navarra, Spain)</i>	

Multi-Task Hypergraphs for Semi-Supervised Learning Using Earth Observations .....	3396
<i>Mihai Pirvu (University Politehnica of Bucharest; Institute of Mathematics "Simion Stoilow" of the Romanian Academy), Alina Marcu (University Politehnica of Bucharest; Institute of Mathematics "Simion Stoilow" of the Romanian Academy), Maria Alexandra Dobrescu (University Politehnica of Bucharest), Ahmed Nabil Belbachir (NORCE Norwegian Research Center), and Marius Leordeanu (University Politehnica of Bucharest; Institute of Mathematics of the Romanian Academy; NORCE Norwegian Research Center)</i>	
Multi-Task Consistency for Active Learning .....	3407
<i>Aral Hekimoglu (Technical University of Munich), Philipp Friedrich (BMW Group), Walter Zimmer (Technical University of Munich), Michael Schmidt (BMW Group), Alvaro Marcos-Ramiro (BMW Group), and Alois Knoll (Technical University of Munich)</i>	
Class-Incremental Learning Using Diffusion Model for Distillation and Replay .....	3417
<i>Quentin Jodelet (Tokyo Institute of Technology, Japan; Artificial Intelligence Research Center, Japan), Xin Liu (Artificial Intelligence Research Center, Japan), Yin Jun Phua (Tokyo Institute of Technology, Japan), and Tsuyoshi Murata (Tokyo Institute of Technology, Japan; Artificial Intelligence Research Center, Japan)</i>	
Flashback for Continual Learning .....	3426
<i>Leila Mahmoodi (Monash University, Data61, CSIRO), Mehrtash Harandi (Monash University, Data61, CSIRO), and Peyman Moghadam (Data61, CSIRO, Queensland University of Technology)</i>	
Continual Evidential Deep Learning for Out-of-Distribution Detection .....	3436
<i>Eduardo Aguilar (Universidad Católica del Norte, Chile; Universitat Autònoma de Barcelona; Universitat de Barcelona, Spain), Bogdan Raducanu (Universitat Autònoma de Barcelona, Spain), Petia Radeva (Universitat Autònoma de Barcelona; Universitat de Barcelona, Spain), and Joost Van de Weijer (Universitat Autònoma de Barcelona, Spain)</i>	
Continual Learning with Deep Streaming Regularized Discriminant Analysis .....	3447
<i>Joe Khawand (Ecole Polytechnique; Sony Computer Science Laboratories Paris), Peter Hanappe (Sony Computer Science Laboratories Paris), and David Colliaux (Sony Computer Science Laboratories Paris)</i>	
FedRCIL: Federated Knowledge Distillation for Representation based Contrastive Incremental Learning .....	3455
<i>Athanasios Psaltis (Centre for Research and Technology Hellas; University of West Attica, Greece), Christos Chatzikonstantinou (Centre for Research and Technology Hellas, Greece), Charalampos Z. Patrikakis (University of West Attica, Greece), and Petros Daras (Centre for Research and Technology Hellas, Greece)</i>	
SATHUR: Self Augmenting Task Hallucinical Unified Representation for Generalized Class Incremental Learning .....	3465
<i>Sathursan Kanagarajah (University of Moratuwa, Sri Lanka), Thanuja Ambegoda (University of Moratuwa, Sri Lanka), and Ranga Rodrigo (University of Moratuwa, Sri Lanka)</i>	

Memory Population in Continual Learning via Outlier Elimination .....	3473
<i>Julio Hurtado (University of Pisa), Alain Raymond-Sáez (Pontificia Universidad Católica de Chile), Vladimir Araujo (Pontificia Universidad Católica de Chile; KU Leuven), Vincenzo Lomonaco (University of Pisa), Alvaro Soto (Pontificia Universidad Católica de Chile), and Davide Bacciu (University of Pisa)</i>	
AR-TTA: A Simple Method for Real-World Continual Test-Time Adaptation .....	3483
<i>Damian Sójka (Poznań University of Technology; IDEAS NCBR), Sebastian Cygert (IDEAS NCBR; Gdańsk University of Technology), Bartłomiej Twardowski (IDEAS NCBR; Autonomous University of Barcelona; Computer Vision Center), and Tomasz Trzciniński (IDEAS NCBR; Warsaw University of Technology; Tooploox)</i>	
Looking Through the Past: Better Knowledge Retention for Generative Replay in Continual Learning .....	3488
<i>Valeriya Khan (IDEAS NCBR), Sebastian Cygert (IDEAS NCBR; Gdansk University of Technology), Bartłomiej Twardowski (IDEAS NCBR; Computer Vision Center; Universitat Autònoma Barcelona), and Tomasz Trzciniński (IDEAS NCBR; Warsaw University of Technology; Jagiellonian University; Tooploox)</i>	
Decision Boundary Optimization for Few-Shot Class-Incremental Learning .....	3493
<i>Chenxu Guo (Beihang University), Qi Zhao (Beihang University), Shuchang Lyu (Beihang University), Binghao Liu (Beihang University), Chunlei Wang (Beihang University), Lijiang Chen (Beihang University), and Guangliang Cheng (University of Liverpool)</i>	
Adapt Your Teacher: Improving Knowledge Distillation for Exemplar-Free Continual Learning ..	3504
<i>Filip Szatkowski (Warsaw University of Technology; IDEAS NCBR), Mateusz Pyla (IDEAS NCBR; Jagiellonian University), Marcin Przewięźlikowski (IDEAS NCBR; Jagiellonian University), Sebastian Cygert (IDEAS NCBR; Gdańsk University of Technology), Bartłomiej Twardowski (IDEAS NCBR; Autonomous University of Barcelona; Computer Vision Center), and Tomasz Trzciniński (Warsaw University of Technology; IDEAS NCBR; Tooploox)</i>	
A Comprehensive Empirical Evaluation on Online Continual Learning .....	3510
<i>Albin Soutif-Cormerais (Computer Vision Center, Universitat Autònoma de Barcelona, Spain), Antonio Carta (University of Pisa, Italy), Andrea Cossu (Scuola Normale Superiore, University of Pisa, Italy), Julio Hurtado (University of Pisa, Italy), Vincenzo Lomonaco (University of Pisa, Italy), Joost Van de Weijer (Computer Science Center, Universitat Autònoma de Barcelona, Spain), and Hamed Hemati (University of St. Gallen, Switzerland)</i>	
TKIL: Tangent Kernel Optimization for Class Balanced Incremental Learning .....	3521
<i>Jinlin Xiang (University of Washington, USA) and Eli Shlizerman (University of Washington, USA)</i>	
Improving Replay Sample Selection and Storage for Less Forgetting in Continual Learning .....	3532
<i>Daniel Brignac (University of Arizona), Niels Lobo (University of Central Florida), and Abhijit Mahalanobis (University of Arizona)</i>	

Selective Freezing for Efficient Continual Learning .....	3542
<i>Amelia Sorrenti (University of Catania, Italy), Giovanni Bellitto (University of Catania, Italy), Federica Proietto Salanitri (University of Catania, Italy), Matteo Pennisi (University of Catania, Italy), Concetto Spampinato (University of Catania, Italy), and Simone Palazzo (University of Catania, Italy)</i>	
Confusion Mixup Regularized Multimodal Fusion Network for Continual Egocentric Activity Recognition .....	3552
<i>Hanxin Wang (University of Electronic Science and Technology of China), Shuchang Zhou (University of Electronic Science and Technology of China), Qingbo Wu (University of Electronic Science and Technology of China), Hongliang Li (University of Electronic Science and Technology of China), Fanman Meng (University of Electronic Science and Technology of China), Linfeng Xu (University of Electronic Science and Technology of China), and Heqian Qiu (University of Electronic Science and Technology of China)</i>	
Margin Contrastive Learning with Learnable-Vector for Continual Learning .....	3562
<i>Kotaro Nagata (Meijo University, Japan) and Kazuhiro Hotta (Meijo University, Japan)</i>	
A Simple Signal for Domain Shift .....	3569
<i>Goirik Chakrabarty (IISER, Pune, India), Manogna Sreenivas (IISc, Bangalore, India), and Soma Biswas (IISc, Bangalore, India)</i>	
On the Effectiveness of LayerNorm Tuning for Continual Learning in Vision Transformers .....	3577
<i>Thomas De Min (University of Trento), Massimiliano Mancini (University of Trento), Karteek Alahari (Inria), Xavier Alameda-Pineda (Inria), and Elisa Ricci (University of Trento; Fondazione Bruno Kessler)</i>	
Comparative Study of Natural Replay and Experience Replay in Online Object Detection .....	3587
<i>Baptiste Wagner (Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, France), Denis Pellerin (Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, France), and Sylvain Huet (Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, France)</i>	
Unseen And Adverse Outdoor Scenes Recognition Through Event-Based Captions .....	3596
<i>Hidetomo Sakaino (Vision Recognition Group, Weather Transportation Lab., Weathernews Inc.)</i>	

## Workshop on Adversarial Robustness in the Real World (AROW)

Adversarial Examples with Specular Highlights .....	3604
<i>Vanshika Vats (University of California Santa Cruz, USA) and Koteswar Rao Jerripothula (Indraprastha Institute of Information Technology-Delhi, India)</i>	
IPCert: Provably Robust Intellectual Property Protection for Machine Learning .....	3614
<i>Zhengyuan Jiang (Duke University), Minghong Fang (Duke University), and Neil Zhenqiang Gong (Duke University)</i>	

Fair Robust Active Learning by Joint Inconsistency .....	3624
<i>Tsung-Han Wu (National Taiwan University, Taiwan), Hung-Ting Su (National Taiwan University, Taiwan), Shang-Tse Chen (National Taiwan University, Taiwan), and Winston H. Hsu (National Taiwan University, Taiwan; Mobile Drive Technology, Taiwan)</i>	
Classification Robustness to Common Optical Aberrations .....	3634
<i>Patrick Müller (Hochschule Düsseldorf - University of Applied Sciences; University of Siegen), Alexander Braun (Hochschule Düsseldorf - University of Applied Sciences), and Margret Keuper (University of Siegen; Max-Planck-Institute for Informatics, Saarland Informatics Campus)</i>	
Defense-Prefix for Preventing Typographic Attacks on CLIP .....	3646
<i>Hiroki Azuma (The University of Tokyo, Japan) and Yusuke Matsui (The University of Tokyo, Japan)</i>	
Semantically Enhanced Scene Captions with Physical and Weather Condition Changes .....	3656
<i>Hidetomo Sakaino (Visual Recognition Group, Weather Transportation Lab., Weathernews Inc.)</i>	
PRAT: PProfiling Adversarial Attacks .....	3669
<i>Rahul Ambati (University of Central Florida, USA), Naveed Akhtar (The University of Western Australia), Ajmal Mian (The University of Western Australia), and Yogesh S Rawat (University of Central Florida, USA)</i>	
On the Adversarial Robustness of Multi-Modal Foundation Models .....	3679
<i>Christian Schlarmann (University of Tübingen, Germany) and Matthias Hein (University of Tübingen, Germany)</i>	
Black-Box Attacks on Image Activity Prediction and its Natural Language Explanations .....	3688
<i>Alina Elena Baia (Idiap Research Institute, Switzerland), Valentina Poggioni (University of Perugia, Italy), and Andrea Cavallaro (Idiap Research Institute, Switzerland; École Polytechnique Fédérale de Lausanne, Switzerland; Queen Mary University of London, U.K.)</i>	
OMG-ATTACK: Self-Supervised On-Manifold Generation of Transferable Evasion Attacks .....	3698
<i>Ofir Bar Tal (Tel Aviv University, Israel), Adi Haviv (Tel Aviv University, Israel), and Amit H. Bermano (Tel Aviv University, Israel)</i>	
On the Unreasonable Vulnerability of Transformers for Image Restoration - and an easy fix .....	3709
<i>Shashank Agnihotri (University of Siegen), Kanchana Vaishnavi Gandikota (University of Siegen), Julia Grabinski (University of Siegen; University of Offenburg), Paramanand Chandramouli (University of Siegen), and Margret Keuper (University of Siegen; Max-Planck-Institute for Informatics, Saarland Informatics Campus)</i>	
Targeted Adversarial Attacks on Generalizable Neural Radiance Fields .....	3720
<i>András Horváth (Pázmány Péter Catholic University, Hungary) and Csaba M. Józsa (Nokia Bell Labs, Hungary)</i>	

# Artificial Intelligence for Humanitarian Assistance and Disaster Response (AIHADR)

Guarding the Guardians: Automated Analysis of Online Child Sexual Abuse .....	3730
<i>Juanita Puentes (Universidad de los Andes, Colombia), Angela Castillo (Universidad de los Andes, Colombia), Wilmar Osejo (Aulas en Paz, Colombia), Yuly Calderón (Universidad de los Andes, Colombia), Viviana Quintero (Aulas en Paz, Colombia), Lina Saldarriaga (Aulas en Paz, Colombia), Diana Agudelo (Universidad de los Andes, Colombia), and Pablo Arbeláez (Universidad de los Andes, Colombia)</i>	
Rapid Flood Inundation Forecast Using Fourier Neural Operator .....	3735
<i>Alexander Y. Sun (The University of Texas, USA), Zhi Li (University of Oklahoma, USA), Wonhyun Lee (The University of Texas at Austin, USA), Qixing Huang (The University of Texas at Austin, USA), Bridget R. Scanlon (The University of Texas at Austin, USA), and Clint Dawson (The University of Texas at Austin, USA)</i>	
Fusing VHR Post-Disaster Aerial Imagery and LiDAR Data for Roof Classification in the Caribbean .....	3742
<i>Isabelle Tingzon (The World Bank Group, GFDRR, USA), Nuala Margaret Cowan (The World Bank Group, GFDRR, USA), and Pierre Chrzanowski (The World Bank Group, GFDRR, USA)</i>	
Deep Learning Driven Detection of Tsunami Related Internal GravityWaves: A Path Towards open-Ocean Natural Hazards Detection .....	3750
<i>Valentino Constantinou (Terran Orbital Corporation, California), Michela Ravanelli (Institut de Physique du Globe de Paris, France), Hamlin Liu (University of California, California), and Jacob Bortnik (University of California, California)</i>	
TeleViT: Teleconnection-Driven Transformers Improve Subseasonal to Seasonal Wildfire Forecasting .....	3756
<i>Ioannis Prapas (National Observatory of Athens; Universitat de València), Nikolaos-Ioannis Bountos (National Observatory of Athens; Harokopio University of Athens), Spyros Kondylatos (National Observatory of Athens; Universitat de València), Dimitrios Michail (Harokopio University of Athens), Gustau Camps-Valls (Universitat de València), and Ioannis Papoutsis (National Observatory of Athens)</i>	
Rapid Building Damage Assessment Workflow: An Implementation for the 2023 Rolling Fork, Mississippi Tornado Event .....	3762
<i>Caleb Robinson (Microsoft AI for Good Research Lab), Simone Fobi Nsutezo (Microsoft AI for Good Research Lab), Anthony Ortiz (Microsoft AI for Good Research Lab), Tina Sederholm (Microsoft AI for Good Research Lab), Rahul Dodhia (Microsoft AI for Good Research Lab), Cameron Birge (Microsoft Philanthropies), Kasie Richards (American Red Cross), Kris Pitcher (American Red Cross), Paulo Duarte (American Red Cross), and Juan M. Lavista Ferres (Microsoft AI for Good Research Lab)</i>	

FireFly: A Synthetic Dataset for Ember Detection in Wildfire .....	3767
<i>Yue Hu (University of Southern California), Xinan Ye (University of Southern California), Yifei Liu (University of California, Irvine), Souvik Kundu (Intel Labs, USA), Gourav Datta (University of Southern California), Srikar Mutnuri (University of Southern California), Namu Asavisanu (University of Southern California), Nora Ayanian (Brown University), Konstantinos Psounis (University of Southern California), and Peter Beerel (University of Southern California)</i>	
Drones4Good: Supporting Disaster Relief Through Remote Sensing and AI .....	3772
<i>Nina Merkle (German Aerospace Center (DLR), Remote Sensing Technology Institute), Reza Bahmanyar (German Aerospace Center (DLR), Remote Sensing Technology Institute), Corentin Henry (German Aerospace Center (DLR), Remote Sensing Technology Institute), Seyed Majid Azimi (German Aerospace Center (DLR), Remote Sensing Technology Institute), Xiangtian Yuan (German Aerospace Center (DLR), Remote Sensing Technology Institute), Simon Schopferer (German Aerospace Center (DLR), Institute of Flight Systems), Veronika Gstaiger (German Aerospace Center (DLR), Remote Sensing Technology Institute), Stefan Auer (German Aerospace Center (DLR), Remote Sensing Technology Institute), Anne Schneibel (German Aerospace Center (DLR), Remote Sensing Data Center), Marc Wieland (German Aerospace Center (DLR), Remote Sensing Data Center), and Thomas Kraft (German Aerospace Center (DLR), Institute of Optical Sensor Systems)</i>	
Estimation of Human Condition at Disaster Site Using Aerial Drone Images .....	3777
<i>Tomoki Arai (National Institute of Advanced Industrial Science and Technology (AIST), Japan; University of Tsukuba, Japan), Kenji Iwata (National Institute of Advanced Industrial Science and Technology (AIST), Japan), Kensho Hara (National Institute of Advanced Industrial Science and Technology (AIST), Japan), and Yutaka Satoh (National Institute of Advanced Industrial Science and Technology (AIST), Japan; University of Tsukuba, Japan)</i>	
Open Problems in Computer Vision for Wilderness SAR and The Search for Patricia Wu-Murad .	3786
<i>Thomas Manzini (Texas A&amp;M University) and Robin Murphy (Texas A&amp;M University)</i>	

## BioImage Computing Workshop (BIC)

ACTIS: Improving Data Efficiency by Leveraging Semi-Supervised Augmentation Consistency Training for Instance Segmentation .....	3792
<i>Josef Lorenz Rumberger (Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association, Germany; Charité University Medicine, Germany; Humboldt-Universität zu Berlin, Faculty of Mathematics and Natural Sciences, Germany; Helmholtz Imaging), Jannik Franzen (Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association, Germany; Charité University Medicine, Germany; Helmholtz Imaging), Peter Hirsch (Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association, Germany; Humboldt-Universität zu Berlin, Faculty of Mathematics and Natural Sciences, Germany; Helmholtz Imaging), Jan-Philipp Albrecht (Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association, Germany; Humboldt-Universität zu Berlin, Faculty of Mathematics and Natural Sciences, Germany; Helmholtz Imaging), and Dagmar Kainmueller (Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association, Germany; Potsdam University, Digital Engineering Faculty, Germany; Helmholtz Imaging)</i>	
Class-Guided Image-to-Image Diffusion: Cell Painting from Brightfield Images with Class Labels .....	3802
<i>Jan Oscar Cross-Zamirski (DAMTP, University of Cambridge; Discovery Sciences, R&amp;D, AstraZeneca), Praveen Anand (Discovery Sciences, R&amp;D, AstraZeneca), Guy Williams (Discovery Sciences, R&amp;D, AstraZeneca), Elizabeth Mouchet (Discovery Sciences, R&amp;D, AstraZeneca), Yin Hai Wang (Discovery Sciences, R&amp;D, AstraZeneca), and Carola-Bibiane Schönlieb (DAMTP, University of Cambridge)</i>	
Complex-Valued Retrievals from Noisy Images Using Diffusion Models .....	3812
<i>Nadav Torem (Technion - Israel Institute of Technology, Israel), Roi Ronen (Technion - Israel Institute of Technology, Israel), Yoav Y. Schechner (Technion - Israel Institute of Technology, Israel), and Michael Elad (Technion - Israel Institute of Technology, Israel)</i>	
Deep Learning Framework Using Sparse Diffusion MRI for Diagnosis of Frontotemporal Dementia .....	3823
<i>Abhishek Tiwari (Shiv Nadar Institution of Eminence, India), Ananya Singhal (Shiv Nadar Institution of Eminence, India), Saurabh J. Shigwan (Shiv Nadar Institution of Eminence, India), and Rajeev Kumar Singh (Shiv Nadar Institution of Eminence, India)</i>	
DeepContrast: Deep Tissue Contrast Enhancement Using Synthetic Data Degradations and OOD Model Predictions .....	3830
<i>Nuno Pimpão Martins (MPI-CBG, Germany), Yannis Kalaidzidis (MPI-CBG, Germany), Marino Zerial (MPI-CBG, Germany), and Florian Jug (MPI-CBG, Germany; Fondazione Human Technopole, Italy)</i>	
Direct Unsupervised Denoising .....	3840
<i>Benjamin Salmon (University of Birmingham, United Kingdom) and Alexander Krull (University of Birmingham, United Kingdom)</i>	
Discrete Representation Learning for Modeling Imaging-Based Spatial Transcriptomics Data .....	3848
<i>Dig Vijay Kumar Yarlagadda (Cornell University, USA), Joan Massagué (Memorial Sloan Kettering Cancer Center, USA), and Christina Leslie (Memorial Sloan Kettering Cancer Center, USA)</i>	



Focus on Content not Noise: Improving Image Generation for Nuclei Segmentation by Suppressing Steganography in CycleGAN .....	3858
<i>Jonas Utz (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Tobias Weise (Friedrich-Alexander-Universität Erlangen-Nürnberg; FAU Erlangen-Nürnberg, Germany), Maja Schlereth (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Fabian Wagner (FAU Erlangen-Nürnberg, Germany), Mareike Thies (FAU Erlangen-Nürnberg, Germany), Mingxuan Gu (FAU Erlangen-Nürnberg, Germany), Stefan Uderhardt (FAU Erlangen-Nürnberg, Germany), and Katharina Breininger (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)</i>	
Generating Synthetic Computed Tomography (CT) Images to Improve the Performance of Machine Learning Model for Pediatric Abdominal Anomaly Detection .....	3867
<i>Samayan Bhattacharya (Jadavpur University, India), Avigyan Bhattacharya (Jadavpur University, India), and Sk Shahnawaz (Jadavpur University, India)</i>	
Leveraging Classic Deconvolution and Feature Extraction in Zero-Shot Image Restoration .....	3876
<i>Tomáš Chobola (Technical University of Munich; Helmholtz Munich - German Research Center for Environmental Health, Germany), Gesine Müller (Georg-August-University Göttingen, Germany), Veit Dausmann (GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany), Anton Theileis (GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany), Jan Taucher (GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany), Jan Huisken (Georg-August-University Göttingen, Germany), and Tingying Peng (Helmholtz Munich - German Research Center for Environmental Health, Germany)</i>	
NU-Net: A Self-Supervised Smart Filter for Enhancing Blobs in Bioimages .....	3886
<i>Seongbin Lim (École Polytechnique, Institut Polytechnique de Paris, France), Emmanuel Beaurepaire (Institut Polytechnique de Paris, France), and Anatole Chessel (Institut Polytechnique de Paris, France)</i>	
On the Risk of Manual Annotations in 3D Confocal Microscopy Image Segmentation .....	3896
<i>Justin Sonneck (Leibniz-Institut für Analytische Wissenschaften - ISAS - e.V.), Shuo Zhao (Leibniz-Institut für Analytische Wissenschaften - ISAS - e.V.), and Jianxu Chen (Leibniz-Institut für Analytische Wissenschaften - ISAS - e.V.)</i>	
PCTrans: Position-Guided Transformer with Query Contrast for Biological Instance Segmentation .....	3905
<i>Qi Chen (University of Science and Technology of China, China), Wei Huang (University of Science and Technology of China, China), Xiaoyu Liu (University of Science and Technology of China, China), Jiacheng Li (University of Science and Technology of China, China), and Zhiwei Xiong (University of Science and Technology of China, China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, China)</i>	
Reinforcement Learning for Instance Segmentation with high-Level Priors .....	3915
<i>Paul Hilt (University Heidelberg), Maedeh Zarvandi (EMBL Heidelberg), Edgar Kaziakhmedov (EMBL Heidelberg), Sourabh Bhide (EMBL Heidelberg), Maria Leptin (EMBL Heidelberg), Constantin Pape (University Goettingen), and Anna Kreshuk (EMBL Heidelberg)</i>	

SortedAP: Rethinking Evaluation Metrics for Instance Segmentation .....	3925
<i>Long Chen (RWTH Aachen University, Germany), Yuli Wu (RWTH Aachen University, Germany), Johannes Stegmaier (RWTH Aachen University, Germany), and Dorit Merhof (University of Regensburg, Germany)</i>	
Spatio-Temporal Analysis of Patient-Derived Organoid Videos Using Deep Learning for the Prediction of Drug Efficacy .....	3932
<i>Leo Fillioux (CentraleSupélec, Université Paris-Saclay), Emilie Gontran (Gustave Roussy), Jérôme Cartry (Gustave Roussy), Jacques RR Mathieu (Gustave Roussy), Sabrina Bedja (Gustave Roussy), Alice Boilève (Gustave Roussy), Paul-Henry Cournède (CentraleSupélec, Université Paris-Saclay), Fanny Jaulin (Gustave Roussy), Stergios Christodoulidis (CentraleSupélec, Université Paris-Saclay), and Maria Vakalopoulou (CentraleSupélec, Université Paris-Saclay)</i>	
The TYC Dataset for Understanding Instance-Level Semantics and Motions of Cells in Microstructures .....	3942
<i>Christoph Reich (Technische Universität Darmstadt), Tim Prangemeier (Technische Universität Darmstadt), and Heinz Koepl (Technische Universität Darmstadt)</i>	
Towards Hierarchical Regional Transformer-Based Multiple Instance Learning .....	3954
<i>Josef Cersovsky (Bayer AG, Germany), Sadegh Mohammadi (Bayer AG, Germany), Dagmar Kainmueller (Max-Delbrueck-Center for Molecular Medicine, Germany), and Johannes Hoehne (Bayer AG, Germany)</i>	
Transformer-Based Detection of Microorganisms on High-Resolution Petri Dish Images .....	3963
<i>Nikolas Ebert (Mannheim University for Applied Science; RPTU Kaiserslautern-Landau, Germany), Didier Stricker (RPTU Kaiserslautern-Landau, Germany), and Oliver Wasenmüller (Mannheim University for Applied Science, Germany)</i>	
Virtual Perturbations to Assess Explainability of Deep-Learning Based Cell Fate Predictors.....	3973
<i>Christopher J. Soelistyo (University College London, UK; The Alan Turing Institute, UK), Guillaume Charras (University College London, UK), and Alan R. Lowe (University College London, UK; The Alan Turing Institute, UK)</i>	

## **Robustness and Reliability of Autonomous Vehicles in the Open-world (BRAVO)**

On the Interplay of Convolutional Padding and Adversarial Robustness .....	3983
<i>Paul Gavrikov (Offenburg University) and Janis Keuper (Offenburg University; Fraunhofer ITWM)</i>	

What Does Really Count? Estimating Relevance of Corner Cases for Semantic Segmentation in Automated Driving .....	3993
<i>Jasmin Breitenstein (Technische Universität Braunschweig, Germany), Florian Heidecker (University of Kassel, Germany), Maria Lyssenko (Robert Bosch GmbH, Germany), Daniel Bogdoll (KIT; FZI Research Center for Information Technology, Germany), Maarten Bieshaar (Bosch Center for Artificial Intelligence, Germany), J. Marius Zöllner (KIT; FZI Research Center for Information Technology, Germany), Bernhard Sick (University of Kassel, Germany), and Tim Fingscheidt (Technische Universität Braunschweig, Germany)</i>	
GPS-GLASS: Learning Nighttime Semantic Segmentation Using Daytime Video and GPS Data ....	4003
<i>Hongjae Lee (Korea University, Republic of Korea), Changwoo Han (Korea University, Republic of Korea), Jun-Sang Yoo (Korea University, Republic of Korea), and Seung-Won Jung (Korea University, Republic of Korea)</i>	
Camera-Based Road Snow Coverage Estimation .....	4013
<i>Kai Cordes (VISCODA GmbH) and Hellward Broszjo (VISCODA GmbH)</i>	
You can have your ensemble and run it too – Deep Ensembles Spread Over Time .....	4022
<i>Isak Meding (Zenseact), Alexander Bodin (Zenseact), Adam Tonderski (Zenseact, Lund University), Joakim Johnander (Zenseact, Linköping University), Christoffer Petersson (Zenseact, Chalmers University of Technology), and Lennart Svensson (Chalmers University of Technology)</i>	
T-FFTRadNet: Object Detection with Swin Vision Transformers from Raw ADC Radar Signals ....	4032
<i>James Giroux (University of Ottawa, Canada), Martin Bouchard (University of Ottawa, Canada), and Robert Laganiere (University of Ottawa, Canada; Sensor Cortek Inc., Canada)</i>	
Unsupervised Domain Adaptation for Self-Driving from Past Traversal Features .....	4042
<i>Travis Zhang (Cornell University, USA), Katie Luo (Cornell University, USA), Cheng Perng Phoo (Cornell University, USA), Yurong You (Cornell University, USA), Wei-Lun Chao (The Ohio State University, USA), Bharath Hariharan (Cornell University, USA), Mark Campbell (Cornell University, USA), and Kilian Q. Weinberger (Cornell University, USA)</i>	
Introspection of 2D Object Detection Using Processed Neural Activation Patterns in Automated Driving Systems .....	4049
<i>Hakan Yekta Yatbaz (University of Warwick), Mehrdad Dianati (University of Warwick; Queen’s University of Belfast), Konstantinos Koufos (University of Warwick), and Roger Woodman (University of Warwick)</i>	
Synthetic Dataset Acquisition for a Specific Target Domain .....	4057
<i>Joshua Niemeijer (German Aerospace Center (DLR), Germany), Sudhanshu Mittal (University of Freiburg, Germany), and Thomas Brox (University of Freiburg, Germany)</i>	
Anomaly-Aware Semantic Segmentation via Style-Aligned OoD Augmentation .....	4067
<i>Dan Zhang (Bosch Center for Artificial Intelligence), Kaspar Sakmann (Bosch Center for Artificial Intelligence), William Beluch (Bosch Center for Artificial Intelligence), Robin Huttmacher (Bosch Center for Artificial Intelligence), and Yumeng Li (Bosch Center for Artificial Intelligence)</i>	

An Empirical Analysis of Range for 3D Object Detection .....	4076
<i>Neehar Peri (Carnegie Mellon University), Mengtian Li (Carnegie Mellon University), Benjamin Wilson (Georgia Institute of Technology), Yu-Xiong Wang (University of Illinois Urbana-Champaign), James Hays (Georgia Institute of Technology), and Deva Ramanan (Carnegie Mellon University)</i>	
On Offline Evaluation of 3D Object Detection for Autonomous Driving .....	4086
<i>Tim Schreier (University of Tübingen, Germany), Katrin Renz (University of Tübingen, Germany), Andreas Geiger (University of Tübingen, Germany), and Kashyap Chitta (University of Tübingen, Germany)</i>	
Identifying Systematic Errors in Object Detectors with the SCROD Pipeline .....	4092
<i>Valentyn Boreiko (Bosch Center for Artificial Intelligence, Robert Bosch GmbH, Germany; University of Tübingen, Germany), Matthias Hein (University of Tübingen, Germany), and Jan Hendrik Metzen (Bosch Center for Artificial Intelligence, Robert Bosch GmbH, Germany)</i>	
Sensitivity Analysis of AI-Based Algorithms for Autonomous Driving on Optical Wavefront Aberrations Induced by the Windshield .....	4102
<i>Dominik Werner Wolf (Volkswagen Group), Markus Ulrich (Karlsruhe Institute of Technology), and Nikhil Kapoor (CARIAD SE)</i>	

## LatinX in AI Research Workshop @ ICCV 2023 (LXCV)

Gaussian Image Anomaly Detection with Greedy Eigencomponent Selection .....	4112
<i>Tetiana Gula (PSL University, France) and João P.C. Bertoldo (PSL University, France)</i>	
Sub-Ensembles for Fast Uncertainty Estimation in Neural Networks .....	4121
<i>Matias Valdenegro-Toro (University of Groningen, Netherlands)</i>	
A Simple and Robust Framework for Cross-Modality Medical Image Segmentation Applied to Vision Transformers .....	4130
<i>Matteo Bastico (Mines Paris, Université PSL, France), David Ryckelynck (Mines Paris, Université PSL, France), Laurent Corté (Mines Paris, Université PSL, France), Yannick Tillier (Mines Paris, Université PSL, France), and Etienne Decencière (Mines Paris, Université PSL, France)</i>	
An Experimental Protocol for Neural Architecture Search in Super-Resolution .....	4141
<i>Jesús Leopoldo Llano García (Tecnologico de Monterrey), Raúl Monroy (Tecnologico de Monterrey), and Víctor Adrián Sosa Hernández (Tecnologico de Monterrey)</i>	
Exploring Image Classification Robustness and Interpretability with Right for the Right Reasons Data Augmentation .....	4149
<i>Flávio Arthur Oliveira Santos (Universidade Federal de Pernambuco, Cidade Universitária, Brazil) and Cleber Zanchettin (Universidade Federal de Pernambuco, Cidade Universitária, Brazil)</i>	

Optical Solutions for Spectral Imaging Inverse Problems with a Shift-Variant System. .... 4159  
*Sergio Urrea (Universidad Industrial de Santander, Colombia), Roman Jacome (Universidad Industrial de Santander, Colombia), M. Salman Asif (University of California Riverside, USA), Henry Arguello (Universidad Industrial de Santander, Colombia), and Hans Garcia (Universidad Industrial de Santander, Colombia)*

Improving Automatic Endoscopic Stone Recognition Using a Multi-view Fusion Approach  
 Enhanced with Two-Step Transfer Learning ..... 4167  
*Francisco Lopez-Tiro (Tecnologico de Monterrey, Mexico; Université de Lorraine, France), Elias Villalvazo-Avila (Tecnologico de Monterrey, Mexico), Juan Pablo Betancur-Rengifo (Tecnologico de Monterrey, Mexico), Ivan Reyes-Amezcuca (CINVESTAV, Mexico), Jacques Hubert (Université de Lorraine, France), Gilberto Ochoa-Ruiz (Tecnologico de Monterrey, Mexico), and Christian Daul (Université de Lorraine, France)*

## **The 2nd computer vision for Metaverse workshop (CV4Metaverse)**

UPGPT: Universal Diffusion Model for Person Image Generation, Editing and Pose Transfer ..... 4175  
*Soon Yau Cheong (University of Surrey, United Kingdom), Armin Mustafa (University of Surrey, United Kingdom), and Andrew Gilbert (University of Surrey, United Kingdom)*

Generative Approach for Probabilistic Human Mesh Recovery Using Diffusion Models ..... 4185  
*Hanbyel Cho (KAIST, South Korea) and Junmo Kim (KAIST, South Korea)*

Personalized 3D Human Pose and Shape Refinement ..... 4191  
*Tom Wehrbein (Leibniz University Hannover, Germany; Epic Games, USA), Bodo Rosenhahn (Leibniz University Hannover, Germany), Iain Matthews (Epic Games, USA), and Carsten Stoll (Epic Games, USA)*

Extract-and-Adaptation Network for 3D Interacting Hand Mesh Recovery ..... 4202  
*Joonkyu Park (Dept. of ECE&ASRI), Daniel Sungho Jung (Seoul National University, Korea; SNU-LG AI Research Center), Gyeongsik Moon (Meta Reality Labs Research), and Kyoung Mu Lee (Dept. of ECE&ASRI; IPAI, Seoul National University, Korea; SNU-LG AI Research Center)*

Effective Whole-Body Pose Estimation with Two-Stages Distillation ..... 4212  
*Zhendong Yang (Tsinghua Shenzhen International Graduate School), Ailing Zeng (International Digital Economy Academy (IDEA)), Chun Yuan (Tsinghua Shenzhen International Graduate School), and Yu Li (International Digital Economy Academy (IDEA))*

BoDiffusion: Diffusing Sparse Observations for Full-Body Human Motion Synthesis ..... 4223  
*Angela Castillo (Universidad de Los Andes, Colombia), Maria Escobar (Universidad de Los Andes, Colombia), Guillaume Jeanneret (University of Caen Normandie, ENSICAEN, CNRS, France), Albert Pumarola (Meta AI), Pablo Arbeláez (Universidad de Los Andes, Colombia), Ali Thabet (Meta AI), and Artsiom Sanakoyeu (Meta AI)*

Intrinsic Appearance Decomposition Using Point Cloud Representation .....	4234
<i>Xiaoyan Xing (UvA-Bosch Delta Lab, University of Amsterdam), Konrad Groh (Bosch Center of Artificial Intelligence), Sezer Karaoglu (UvA-Bosch Delta Lab, University of Amsterdam), and Theo Gevers (UvA-Bosch Delta Lab, University of Amsterdam)</i>	
Noise-in, Bias-out: Balanced and Real-Time MoCap Solving .....	4239
<i>Georgios Albanis (Moverse; University of Thessaly, Greece), Nikolaos Zioulis (Moverse), Spyridon Thermos (Moverse), Anargyros Chatzitofis (Moverse), and Kostas Kolomvatsos (University of Thessaly, Greece)</i>	
Temporally Consistent Semantic Segmentation Using Spatially Aware Multi-view Semantic Fusion for Indoor RGB-D Videos .....	4250
<i>Fengyuan Sun (University of Amsterdam &amp; 3DUniversum, The Netherlands), Sezer Karaoglu (University of Amsterdam &amp; 3DUniversum, The Netherlands), and Theo Gevers (University of Amsterdam &amp; 3DUniversum, The Netherlands)</i>	
Efficient 3D Reconstruction, Streaming and Visualization of Static and Dynamic Scene Parts for Multi-Client Live-Telepresence in Large-Scale Environments .....	4260
<i>Leif Van Holland (University of Bonn, Germany), Patrick Stotko (University of Bonn, Germany), Stefan Krumpfen (University of Bonn, Germany), Reinhard Klein (University of Bonn, Germany), and Michael Weinmann (University of Bonn, Germany; Delft University of Technology, The Netherlands)</i>	
On-Device Real-Time Custom Hand Gesture Recognition .....	4275
<i>Esha Uboweja (Google LLC, USA), David Tian (Google LLC, USA), Qifei Wang (Google LLC, USA), Yi-Chun Kuo (Google LLC, USA), Joe Zou (Google LLC, USA), Lu Wang (Google LLC, USA), George Sung (Google LLC, USA), and Matthias Grundmann (Google LLC, USA)</i>	
MAMMOS: MAPPING Multiple Human MOTION with Scene Understanding and Natural Interactions ..	4280
<i>Donggeun Lim (Seoul National University, Korea), Cheongi Jeong (Samsung Research, Korea), and Young Min Kim (Seoul National University, Korea)</i>	
NOVA: NOvel View Augmentation for Neural Composition of Dynamic Objects .....	4290
<i>Dakshit Agrawal (Carnegie Mellon University, USA), Jiajie Xu (Carnegie Mellon University, USA), Siva Karthik Mustikovela (Cruise LLC, USA), Ioannis Gkioulekas (Carnegie Mellon University, USA), Ashish Shrivastava (Cruise LLC, USA), and Yuning Chai (Cruise LLC, USA)</i>	
FARMARe: a Furniture-Aware Multi-Task Methodology for Recommending Apartments Based on the user Interests .....	4295
<i>Ali Abdari (University of Udine, University of Naples Federico II, Italy), Alex Falcon (University of Udine, Italy), and Giuseppe Serra (University of Udine, Italy)</i>	

## The 2nd Workshop and Challenges for Out-of-Distribution Generalization in Computer Vision (OODCV)

Confusing Large Models by Confusing Small Models .....	4306
<i>Vitor Albiero (Meta AI), Raghav Mehta (McGill University), Ivan Evtimov (Meta AI), Samuel Bell (Meta AI), Levent Sagun (Meta AI), and Aram Markosyan (Meta AI)</i>	
Misalignment-Free Relation Aggregation for Multi-Source-Free Domain Adaptation .....	4315
<i>Hao-Wei Yeh (The University of Tokyo, Japan), Qier Meng (The University of Tokyo, Japan), and Tatsuya Harada (The University of Tokyo, Japan; RIKEN, Japan)</i>	
Consistency Regularization for Generalizable Source-Free Domain Adaptation .....	4325
<i>Longxiang Tang (Tsinghua University), Kai Li (NEC Laboratories America), Chunming He (Tsinghua University), Yulun Zhang (ETH Zurich, Switzerland), and Xiu Li (Tsinghua University)</i>	
Unsupervised Camouflaged Object Segmentation as Domain Adaptation .....	4336
<i>Yi Zhang (LIVIA, École de Technologie Supérieure, Canada) and Chengyi Wu (Henan Polytechnic University, China)</i>	
Class-Aware Memory Guided Unbiased Weighting for Universal Domain Adaptive Object Detection .....	4347
<i>Qinghai Lang (Chongqing University, China), Zhenwei He (Chongqing University of Technology, China), Xiaowei Fu (Chongqing University, China), and Lei Zhang (Chongqing University, China)</i>	
AD-CLIP: Adapting Domains in Prompt Space Using CLIP .....	4357
<i>Mainak Singha (Indian Institute of Technology Bombay, India), Harsh Pal (Indian Institute of Technology Bombay, India), Ankit Jha (Indian Institute of Technology Bombay, India), and Biplab Banerjee (Indian Institute of Technology Bombay, India)</i>	
Raising the Bar on the Evaluation of Out-of-Distribution Detection .....	4367
<i>Jishnu Mukhoti (University of Oxford; Meta AI), Tsung-Yu Lin (Meta AI), Bor-Chun Chen (Meta AI), Ashish Shah (Meta AI), Philip H.S. Torr (University of Oxford), Puneet K. Dokania (University of Oxford), and Ser-Nam Lim (Meta AI)</i>	
A Re-Parameterized Vision Transformer (ReVT) for Domain-Generalized Semantic Segmentation .....	4378
<i>Jan-Aike Termöhlen (Technische Universität Braunschweig, Germany), Timo Bartels (Technische Universität Braunschweig, Germany), and Tim Fingscheidt (Technische Universität Braunschweig, Germany)</i>	
LORD: Leveraging Open-Set Recognition with Unknown Data .....	4388
<i>Tobias Koch (e.solutions GmbH; Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), Christian Riess (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany), and Thomas Köhler (e.solutions GmbH, Germany)</i>	
Masking Strategies for Background Bias Removal in Computer Vision Models .....	4399
<i>Ananthu Aniraj (Inria; University of Montpellier; LIRMM; UMR-Tetis), Cassio F. Dantas (Inrae; University of Montpellier; UMR-Tetis), Dino Ienco (Inrae; University of Montpellier; UMR-Tetis), and Diego Marcos (Inria; University of Montpellier; LIRMM; UMR-Tetis)</i>	

Assessing the Impact of Diversity on the Resilience of Deep Learning Ensembles: A Comparative Study on Model Architecture, Output, Activation, and Attribution .....	4408
<i>Rafael Rosales (Intel Labs, Intel Corporation), Pablo Munoz (Intel Labs, Intel Corporation), and Michael Paulitsch (Intel Labs, Intel Corporation)</i>	
DatasetEquity: Are All Samples Created Equal? In The Quest For Equity Within Datasets .....	4419
<i>Shubham Shrivastava (Ford Motor Company), Xianling Zhang (Ford Motor Company), Sushruth Nagesh (Ford Motor Company), and Armin Parchami (Ford Motor Company)</i>	
Benchmarking Image Classifiers for Physical Out-of-Distribution Examples Detection .....	4429
<i>Ojaswee Ojaswee (IISER Bhopal, India), Akshay Agarwal (IISER Bhopal, India), and Nalini Ratha (University at Buffalo, USA)</i>	
Gradient Estimation for Unseen Domain Risk Minimization with Pre-Trained Models .....	4438
<i>Byounggyu Lew (Hyperconnect; Seoul National University), Donghyun Son (Hyperconnect; Seoul National University), and Buru Chang (Hyperconnect; Sogang University)</i>	
Leveraging Visual Attention for out-of-Distribution Detection .....	4449
<i>Luca Cultrera (University of Florence, MICC), Lorenzo Seidenari (University of Florence, MICC), and Alberto Del Bimbo (University of Florence, MICC)</i>	
SC2GAN: Rethinking Entanglement by Self-Correcting Correlated GAN Space .....	4459
<i>Zikun Chen (ModiFace Inc.), Han Zhao (University of Illinois at Urbana-Champaign), Parham Aarabi (University of Toronto), and Ruowei Jiang (ModiFace Inc.)</i>	
Can Self-Supervised Representation Learning Methods Withstand Distribution Shifts and Corruptions? .....	4469
<i>Prakash Chandra Chhipa (Luleå Tekniska Universitet, Sweden), Johan Rodahl Holmgren (Luleå Tekniska Universitet, Sweden), Kanjar De (Luleå Tekniska Universitet, Sweden; Fraunhofer Heinrich-Hertz-Institut, Germany), Rajkumar Saini (Luleå Tekniska Universitet, Sweden), and Marcus Liwicki (Luleå Tekniska Universitet, Sweden)</i>	

## Uncertainty Estimation for Computer Vision (UnCV)

Far Away in the Deep Space: Dense Nearest-Neighbor-Based Out-of-Distribution Detection .....	4479
<i>Silvio Galesso (University of Freiburg), Max Argus (University of Freiburg), and Thomas Brox (University of Freiburg)</i>	
Gaussian Latent Representations for Uncertainty Estimation Using Mahalanobis Distance in Deep Classifiers .....	4490
<i>Aishwarya Venkataramanan (Université de Lorraine, CNRS, LIEC; Georgia Tech Europe; LTSER-“Zone Atelier Moselle”, France), Assia Benbihi (Czech Technical University in Prague), Martin Laviale (Université de Lorraine, CNRS, LIEC; LTSER-“Zone Atelier Moselle”, France), and Cédric Pradalier (Georgia Tech Europe; LTSER-“Zone Atelier Moselle”, France)</i>	



Probabilistic MIMO U-Net: Efficient and Accurate Uncertainty Estimation for Pixel-Wise Regression .....	4500
<i>Anton Baumann (University of the Bundeswehr Munich \ Germany), Thomas Roßberg (University of the Bundeswehr Munich, Germany), and Michael Schmitt (University of the Bundeswehr Munich, Germany)</i>	
Calibrated Out-of-Distribution Detection with a Generic Representation .....	4509
<i>Tomáš Vojíš (Czech Technical University in Prague), Jan Šochman (Czech Technical University in Prague), Rahaf Aljundi (Toyota Motor Europe, Belgium), and Jiří Matas (Czech Technical University in Prague)</i>	
DELO: Deep Evidential LiDAR Odometry Using Partial Optimal Transport .....	4519
<i>Sk Aziz Ali (German Research Center for Artificial Intelligence (DFKI), Germany; University of Luxembourg, Luxembourg), Djamila Aouada (University of Luxembourg, Luxembourg), Gerd Reis (German Research Center for Artificial Intelligence (DFKI), Germany), and Didier Stricker (German Research Center for Artificial Intelligence (DFKI), Germany)</i>	
Dual-Level Interaction for Domain Adaptive Semantic Segmentation .....	4529
<i>Dongyu Yao (Wuhan University, China) and Boheng Li (Wuhan University, China)</i>	
UncLe-SLAM: Uncertainty Learning for Dense Neural SLAM .....	4539
<i>Erik Sandström (ETH Zürich, Switzerland), Kevin Ta (ETH Zürich, Switzerland), Luc Van Gool (ETH Zürich, Switzerland; KU Leuven, Belgium), and Martin R. Oswald (ETH Zürich, Switzerland; University of Amsterdam, Netherlands)</i>	
Distance Matters For Improving Performance Estimation Under Covariate Shift .....	4551
<i>Mélanie Roschewitz (Imperial College London) and Ben Glocker (Imperial College London)</i>	
Identifying Out-of-Domain Objects with Dirichlet Deep Neural Networks .....	4562
<i>Ahmed Hammam (Stellantis, Opel Automobile GmbH, Institute of Measurement and Control Systems, Karlsruhe Institute of Technology, Germany), Frank Bonarens (Stellantis, Opel Automobile GmbH, Germany), Seyed Eghbal Ghobadi (Technische Hochschule Mittelhessen, Germany), and Christoph Stiller (Karlsruhe Institute of Technology, Germany)</i>	
A Simple and Explainable Method for Uncertainty Estimation Using Attribute Prototype Networks .....	4572
<i>Claudius Zelenka (Kiel University, Germany), Andrea Göhring (Kiel University, Germany), Daniyal Kazempour (Kiel University, Germany), Maximilian Hünemörder (Kiel University, Germany), Lars Schmarje (Kiel University, Germany), and Peer Kröger (Kiel University, Germany)</i>	
Biased Class disagreement: detection of out of distribution instances by using differently biased semantic segmentation models. ....	4582
<i>Roberto Alcover-Couso (Autonomous University of Madrid (UAM), Madrid, Spain), Juan C. SanMiguel (Autonomous University of Madrid (UAM), Madrid, Spain), and Marcos Escudero-Viñolo (Autonomous University of Madrid (UAM), Madrid, Spain)</i>	

Exploring Inlier and Outlier Specification for Improved Medical OOD Detection .....	4591
<i>Vivek Narayanaswamy (Lawrence Livermore National Labs, USA), Yamen Mubarka (Lawrence Livermore National Labs, USA), Rushil Anirudh (Lawrence Livermore National Labs, USA), Deepta Rajan (Microsoft, USA), and Jayaraman J. Thiagarajan (Lawrence Livermore National Labs, USA)</i>	
Adversarial Attacks Against Uncertainty Quantification .....	4601
<i>Emanuele Ledda (Sapienza University of Rome, Italy; University of Genova, Italy), Daniele Angioni (University of Cagliari, Italy; Sapienza University of Rome, Italy), Giorgio Piras (University of Cagliari, Italy; Sapienza University of Rome, Italy), Giorgio Fumera (University of Cagliari, Italy), Battista Biggio (University of Cagliari, Italy), and Fabio Roli (University of Genova, Italy)</i>	
Unsupervised Confidence Approximation: Trustworthy Learning from Noisy Labelled Data .....	4611
<i>Navid Rabbani (Université Clermont Auvergne; CHU Clermont-Ferrand, France) and Adrien Bartoli (Université Clermont Auvergne; CHU Clermont-Ferrand, France, France)</i>	

The Robust Semantic Segmentation UNCV2023 Challenge Results ..... 4620

Xuanlong Yu (SATIE, Paris-Saclay University and U2IS, ENSTA Paris, Institut Polytechnique de Paris, France), Yi Zuo (Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, China), Zitao Wang (Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, China), Xiaowen Zhang (Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, China), Jiaxuan Zhao (Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, China), Yuting Yang (Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, China), Licheng Jiao (Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, China), Rui Peng (Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, China), Xinyi Wang (Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, China), Junpei Zhang (Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, China), Kexin Zhang (Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, China), Fang Liu (Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education, Xidian University, China), Roberto Alcover-Couso (VPU-Lab, Autonomous University of Madrid (UAM), Spain), Juan C. SanMiguel (VPU-Lab, Autonomous University of Madrid (UAM), Spain), Marcos Escudero-Viñolo (VPU-Lab, Autonomous University of Madrid (UAM), Spain), Hanlin Tian (Imperial College London, UK), Kenta Matsui (Imperial College London, UK), Tianhao Wang (The University of Texas at Dallas, USA), Fahmy Adan (Imperial College London, UK), Zhitong Gao (ShanghaiTech University, China), Xuming He (ShanghaiTech University, China), Quentin Bouniot (LTCI, Télécom Paris, Institut Polytechnique de Paris, France), Hossein Moghaddam (LTCI, Télécom Paris, Institut Polytechnique de Paris, France), Shyam Nandan Rai (Politecnico di Torino, Italy), Fabio Cermelli (Politecnico di Torino, Italy), Carlo Masone (Politecnico di Torino, Italy), Andrea Pilzer (NVIDIA AI Technology Center, Italy), Elisa Ricci (University of Trento, Italy), Andrei Bursuc (valeo.ai, France), Arno Solin (Aalto University, Finland), Martin Trapp (Aalto University, Finland), Rui Li (Aalto University, Finland), Angela Yao (National University of Singapore, Singapore), Wenlong Chen (Imperial College London, UK), Ivor Simpson (University of Sussex, UK), Neill D. F. Campbell (University of Bath, UK), and Gianni Franchi (U2IS, ENSTA Paris, Institut Polytechnique de Paris, France)

## Vision-and-Language Algorithmic Reasoning Workshop (VLAR)

What If the TV Was Off? Examining Counterfactual Reasoning Abilities of Multi-Modal Language Models .....	4631
<i>Letian Zhang (Tongji University), Xiaotong Zhai (University of Warwick), Zhongkai Zhao (LunarAI), Xin Wen (LunarAI), and Bingchen Zhao (LunarAI)</i>	
Uni-NLX: Unifying Textual Explanations for Vision and Vision-Language Tasks .....	4636
<i>Fawaz Sammani (Vrije Universiteit Brussel, Belgium) and Nikos Deligiannis (Vrije Universiteit Brussel, Belgium)</i>	
SelfGraphVQA: A Self-Supervised Graph Neural Network for Scene-Based Question Answering	4642
<i>Bruno Cesar de Oliveira Souza (Universidade Estadual de Campinas, Brazil), Marius Aasan (University of Oslo, Norway), Helio Pedrini (Universidade Estadual de Campinas, Brazil), and Adin Ramirez Rivera (University of Oslo, Norway)</i>	
Understanding Video Scenes Through Text: Insights from Text-Based Video Question Answering.....	4648
<i>Soumya Jahagirdar (IIIT Hyderabad, India), Minesh Mathew (Wadhvani AI), Dimosthenis Karatzas (UAB, Spain), and C. V. Jawahar (IIIT Hyderabad, India)</i>	
Iterative Robust Visual Grounding with Masked Reference Based Centerpoint Supervision .....	4653
<i>Menghao Li (Beihang University, China), Chunlei Wang (Beihang University, China), Wenquan Feng (Beihang University, China), Shuchang Lyu (Beihang University, China), Guangliang Cheng (University of Liverpool, The United Kingdom), Xiangtai Li (S-Lab, Nanyang Technological University, Singapore), Binghao Liu (Beihang University, China), and Qi Zhao (Beihang University, China)</i>	
MMTF: Multi-Modal Temporal Fusion for Commonsense Video Question Answering .....	4659
<i>Mobeen Ahmad (PYLER CO., LTD., South Korea), Geontwoo Park (PYLER CO., LTD., South Korea), Dongchan Park (PYLER CO., LTD., South Korea), and Sanguk Park (PYLER CO., LTD., South Korea)</i>	
Pointing out Human Answer Mistakes in a Goal-Oriented Visual Dialogue .....	4665
<i>Ryosuke Oshima (Waseda University), Seitaro Shinagawa (Nara Institute of Science and Technology), Hideki Tsunashima (Waseda University), Qi Feng (Waseda University), and Shigeo Morishima (Waseda Research Institute for Science and Engineering)</i>	
Language-Enhanced RNR-Map: Querying Renderable Neural Radiance Field Maps with Natural Language .....	4671
<i>Francesco Taioli (University of Verona, Italy), Federico Cunico (University of Verona, Italy), Federico Girella (University of Verona, Italy), Riccardo Bologna (University of Verona, Italy), Alessandro Farinelli (University of Verona, Italy), and Marco Cristani (University of Verona, Italy)</i>	
CLIP-Decoder : ZeroShot Multilabel Classification Using Multimodal CLIP Aligned Representations .....	4677
<i>Muhammad Ali (Mohamed Bin Zayed University of AI, UAE) and Salman Khan (Mohamed Bin Zayed University of AI, UAE)</i>	

## Author Index