

Detailed Programme

Monday 21st November

10 am Registration Desk Opens

11:00 - 11:15 **Opening Ceremony**

11:15 - 12:15 **Keynote Speaker**

Philip Isola, Massachusetts Institute of Technology (MIT)

(Chair:)

12:15 - 12:45 🍽️ **Lunch Break**

12:45 - 14:00 **Oral Session 1: Segmentation (Chair:)**

45. Open-vocabulary Semantic Segmentation with Frozen Vision-Language Models. Chaofan Ma (Shanghai Jiao Tong University), Yuhuan Yang (Shanghai Jiao Tong University), Yan-Feng Wang (Cooperative medianet innovation center of Shanghai Jiao Tong University), Ya Zhang (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University), Weidi Xie (Shanghai Jiao Tong University)*

592 Semantics-Adding Flaw-Erasing Network for Semantic Human Matting Jiayu Sun (Dalian University of Technology)*, Zhanghan Ke (City University of Hong Kong), Ke Xu (City University of Hong Kong), Fan Shao (Wonxing Technology), Lihe Zhang (Dalian University of Technology), Huchuan Lu (Dalian University of Technology), Rynson W.H. Lau (City University of Hong Kong)

126 Few-shot Semantic Segmentation with Support-induced Graph Convolutional Network Jie Liu (University of Amsterdam)*, Yanqi Bao (Northeastern University), Wenzhe Yin (University of Amsterdam), haochen wang (UvA), Yang Gao (Nanjing University), Jan-Jakob Sonke (The Netherlands Cancer Institute), Efstratios

Gold Sponsors



Gaves (University of Amsterdam)

417 Masked Supervised Learning for Semantic Segmentation

Hasib Zunair (Concordia University),* Abdessamad Ben Hamza (Concordia University)

1002 Adversarial Vision Transformer for Medical Image Semantic Segmentation with Limited Annotations

Ziyang Wang (University of Oxford),* Will Zhao (Bucknell University), Zixuan Ni (CU Boulder), Yuchen Zheng (University of North Carolina at Chapel Hill)

14:00-15:30

Poster Session I (Chair:) & Tea Break

4. Enhancing Material Features Using Dynamic Backward Attention on Cross-Resolution Patches

Yuwen Heng (University of Southampton),* Yihong Wu (University of Southampton), Srinandan Dasmahapatra (University of Southampton), Hansung Kim (University Of Southampton)

5. MAC: Mask-Augmentation for Motion-Aware Video Representation Learning.

Arif Akar (Hacettepe University),* Ufuk Umut Senturk (Hacettepe University), Nazli Ikizler-Cinbis (Hacettepe University)

7. Self-distillation and Uncertainty Boosting Self-supervised Monocular Depth Estimation.

Hang Zhou (University of East Anglia),* Sarah Taylor (University of East Anglia), David Greenwood (University of East Anglia), Michal Mackiewicz (University of East Anglia)

8. TripleDNet: Exploring Depth Estimation with Self-Supervised Representation Learning

Ufuk Umut Senturk (Hacettepe University),* Arif Akar (Hacettepe University), Nazli Ikizler-Cinbis (Hacettepe University)

13. Domain Generalization Capability Enhancement for Binary Neural Networks

Jianming Ye (Peking University), Shunan Mao (Peking University), Shiliang Zhang (Peking University)*

18. Spatio-Temporal Learnable Proposals for End-to-End Video Object Detection

Khurram Azeem Hashmi (Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH),* Didier Stricker (DFKI), Muhammad Zeshan Afzal (Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH)

23. Deep Image Harmonization by Bridging the Reality Gap

Junyan Cao (Shanghai Jiao Tong University), Wenyan Cong (Shanghai Jiao Tong University), Li Niu (Shanghai Jiao Tong University),* Jianfu Zhang (Shanghai Jiao Tong University), Liqing Zhang (Shanghai Jiao Tong University)

Gold Sponsors



26. Category-Level Pose Retrieval with Contrastive Features Learnt with Occlusion Augmentation Georgios Kouros (KU Leuven),* Shubham Shrivastava (Ford Greenfield Labs), Cédric Picron (KU Leuven), Sushruth Nagesh (Ford Motor Company), Punarjay Chakravarty (Ford Motor Company), Tinne Tuytelaars (KU Leuven)

27. Approximating Continuous Convolutions for Deep Network Compression Theo W Costain (University Of Oxford),* Victor Adrian Prisacariu (University of Oxford)

30. EpipolarNVS: leveraging on Epipolar geometry for single-image Novel View Synthesis Gaétan Landreau (Meero-CEA),* Mohamed Tamaazousti (CEA Saclay)

31. Disentangling 3D Attributes from a Single 2D Image: Human Pose, Shape and Garment. Xue Hu (Imperial College London), Xinghui Li (University of Oxford), Benjamin Busam (Technical University of Munich), Yiren Zhou (Huawei Noah's Ark Lab), Ales Leonardis (Huawei Noah's Ark Lab), Shanxin Yuan (Huawei Noah's Ark Lab)*

33. TAG: Boosting Text-VQA via Text-aware Visual Question-answer Generation. Jun Wang (University of Maryland, College Park),* Mingfei Gao (Apple), Yuqian Hu (University of Maryland, College Park), Ramprasaath R. Selvaraju (Salesforce Research), Chetan Ramaiah (Google), Ran Xu (Salesforce Research), Joseph JaJa (University of Maryland, College Park), Larry Davis (University of Maryland)

34. Video Prediction at Multiple Spatio-Temporal Scales with Hierarchical Recurrent Networks. Angel Villar-Corrales (University of Bonn),* Ani Karapetyan (University of Bonn), Andreas Boltres (SAP), Sven Behnke (University of Bonn)

35. Training Binary Neural Networks the Easy Way Alasdair J Paren (University of Oxford),* Rudra Poudel (Toshiba Research)

37. LOCL: Learning Object-Attribute Composition using Localization Satish Kumar (University of California, Santa Barbara),* ASM Iftekhar (University of California, Santa Barbara), Ekta Prashnani (University of California, Santa Barbara), B. S. Manjunath (University of California Santa Barbara)

38. Positive Pair Distillation Considered Harmful: Continual Meta Metric Learning for Lifelong Object Re-Identification kai wang (Computer Vision Center),* Chenshen Wu (Computer Vision Center), Andy Bagdanov (University of Florence, Italy), Xialei Liu (Nankai University), Shiqi Yang (Computer Vision Center), Shangling Jui (Huawei Kirin Solution), Joost van de Weijer (Computer Vision Center)

Gold Sponsors



39. Deep Clustering by Semantic Contrastive Learning Jiabo Huang (Queen Mary University of London),* Shaogang Gong (Queen Mary University of London)

45. Open-vocabulary Semantic Segmentation with Frozen Vision-Language Models Chaofan Ma (Shanghai Jiao Tong University), Yuhuan Yang (Shanghai Jiao Tong University), Yan-Feng Wang (Cooperative medianet innovation center of Shanghai Jiao Tong University), Ya Zhang (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University), Weidi Xie (Shanghai Jiao Tong University)*

46. Pay Self-Attention to Audio-Visual Navigation Yinfeng Yu (Department of Computer Science and Technology, State Key Lab on Intelligent Technology and Systems, Tsinghua University, Beijing, China, Xinjiang University),* Lele Cao (EQT Group), Fuchun Sun (Tsinghua University), Xiaohong Liu (Tsinghua University), Liejun Wang (Xinjiang University)

48. Energy-Based Residual Latent Transport for Unsupervised Point Cloud Completion Ruikai Cui (Australian National University), Shi Qiu (ANU), Saeed Anwar (The Australian National University), Jing Zhang (Australian National University), Nick Barnes (ANU)*

52. ShowFace: Coordinated Face Inpainting with Memory-Disentangled Refinement Networks Zhuojie Wu (Beijing University of Posts and Telecommunications),* Xingqun Qi (University of Technology Sydney), Zijian Wang (Beijing university of Posts and Telecommunications), Wanting Zhou (Beijing University of Posts and Telecommunications), Kun Yuan (Kuaishou Technology), Muyi Sun (CRIPAC, Institute of Automation, Chinese Academy of Sciences), Zhenan Sun (Chinese of Academy of Sciences)

55. SearchTrack: Multiple Object Tracking with Object-Customized Search and Motion-Aware Features ZHONG MIN TSAI (National Taiwan University), YU-JU TSAI (National Taiwan University),* Chien-Yao Wang (Institute of Information Science, Academia Sinica), Hong-Yuan Mark Liao (Institute of Information Science, Academia Sinica, Taiwan), Youn-Long Lin (National Tsing Hua University), Yung-Yu Chuang (National Taiwan University)

56. MorphPool: Efficient Non-linear Pooling & Unpooling in CNNs Rick Groenendijk (University of Amsterdam),* Leo Dorst (University of Amsterdam), Theo Gevers (University of Amsterdam)

58. Detailed Annotations of Chest X-Rays via CT Projection for Report Understanding Constantin Marc Seibold (Karlsruhe Institute of Technology),* Simon Reiß (Karlsruhe Institute of Technology), M. Sauqib Sarfraz (KIT), Matthias A. Fink (University Hospital Heidelberg), Victoria Mayer (Mayer), Jan Sellner (German Cancer Research Center), Moon Sung Kim (University Hospital of Essen), Klaus H. Maier-Hein

Gold Sponsors



(German Cancer Research Center (DKFZ)), Jens Kleesiek (Institute for AI in Medicine (IKIM), University Hospital Essen), Rainer Stiefelhagen (Karlsruhe Institute of Technology)

60. Propagating Difference Flows for Efficient Video Super-Resolution Ruisheng Gao (University of Science and Technology of China), Zeyu Xiao (University of Science and Technology of China), Zhiwei Xiong (University of Science and Technology of China)*

61. One-Pot Multi-Frame Denoising Lujia Jin (Peking University), Shi Zhao (Peking University), Lei Zhu (Beijing University of Posts and Telecommunications), Qian Chen (University of Science and Technology of China), Yanye Lu (Peking University)*

63. Automatic universal taxonomies for multi-domain semantic segmentation Petra Bevandić (Faculty of Electrical Engineering and Computing),* Sinisa Segvic (UniZg-FER)

67. DiffSketching: Sketch Control Image Synthesis with Diffusion Models Qiang Wang (Beijing university of posts and telecommunications),* Di Kong (Beijing University of Posts and Telecommunications), Fengyin Lin (Beijing University of Posts and Telecommunications), Yonggang Qi (Beijing University of Posts and Telecommunications)

70. Re-Attention Transformer for Weakly Supervised Object Localization Hui Su (Zhejiang Lab), Yue Ye (Zhejiang Lab), Zhiwei Chen (Zhejiang Lab), Mingli Song (Zhejiang University), Lechao Cheng (Zhejiang Lab)*

72. SPARC: Sparse Render-and-Compare for CAD model alignment in a single RGB Image Florian Maximilian Langer (Department of Engineering, University of Cambridge),* Gwangbin Bae (University of Cambridge), Ignas Budvytis (Department of Engineering, University of Cambridge), Roberto Cipolla (University of Cambridge)

73. Hybrid Cost Volume Regularization for Memory-efficient Multi-view Stereo Networks Qingtian Zhu (Peking University),* Zizhuang Wei (Peking University), Zhongtao Wang (Wuhan University), Yisong Chen (Peking University), GUOPING WANG (PEKING UNIVERSITY)

76. Rethinking Graph Neural Networks for Unsupervised Video Object Segmentation Daizong Liu (Peking University),* Wei Hu (Peking University)

78. Event-based Non-Rigid Reconstruction from Contours Yuxuan Xue (Max Planck Institute for Intelligent Systems),* Haolong Li (Max Planck Institute for Intelligent Systems), Stefan Leutenegger (TU

Gold Sponsors



Munich), Joerg Stueckler (Max-Planck-Institute for Intelligent Systems)

80. Beyond the CLS Token: Image Reranking using Pretrained Vision Transformers Chao Zhang (Toshiba Europe Limited),* Stephan Liwicki (Toshiba Europe Limited), Roberto Cipolla (University of Cambridge)

82. Motion-Aware Graph Reasoning Hashing for Self-supervised Video Retrieval Ziyun Zeng (Tsinghua University),* Jinpeng Wang (Tsinghua University), Bin Chen (Harbin Institute of Technology, Shenzhen), Yuting Wang (Tsinghua University), Shu-Tao Xia (Tsinghua University)

88. Blind Removal of Facial Foreign Shadows Yaojie Liu (Google Research), Andrew Z Hou (Michigan State University),* Xinyu Huang (BOSCH Research North America), Liu Ren (BOSCH Research North America), Xiaoming Liu (Michigan State University)

89. StyleFaceUV: a 3D Face UV Map Generator for View-Consistent Face Image Synthesis Wei-Chieh Chung (National Taiwan University), Jian-Kai Zhu (National Taiwan University), I-Chao Shen (The University of Tokyo),* Yu-Ting Wu (National Taipei University), Yung-Yu Chuang (National Taiwan University)

90. Convolutional Sparse Coding Network Via Improved Proximal Gradient For Compressed Sensing Magnetic Resonance Imaging Xiaofan Wang (Yanshan University), Yali Zhang (Yanshan University), PENGYU Li (Yanshan University), Jinjia Wang (Yanshan University)*

91. Learning to Construct 3D Building Wireframes from 3D Line Clouds Yicheng Luo (Beijing University of Posts and Telecommunications),* Jing Ren (ETH Zurich), Xuefei Zhe (Tencent AI lab), Di Kang (Tencent), Yajing Xu (Beijing University of Posts and Telecommunications), Peter Wonka (KAUST), Linchao Bao (Tencent AI Lab)

92. OSM: An Open Set Matting Framework with OOD Detection and Few-Shot Matting Yuhongze Zhou (McGill University),* Issam Hadj Laradji (ServiceNow), Liguang Zhou (The Chinese University of Hong Kong, Shenzhen), Derek Nowrouzezahrai (McGill University)

94. Subtask-dominated Supervised Pretraining Transfer Learning for Person Search Chuang Liu (Shanghai Jiao Tong University),* Hua Yang (Shanghai Jiao Tong University), Shibao Zheng (SJTU)

96. XCon: Learning with Experts for Fine-grained Category Discovery Yixin Fei (Tongji University), Zhongkai Zhao (Tongji University), Siwei Yang (Tongji University), Bingchen Zhao (University of Edinburgh)*

Gold Sponsors



104. Style2NeRF: An Unsupervised One-Shot NeRF for Semantic 3D Reconstruction James Charles (Cambridge University),* Wim Abbeloos (Toyota Motor Europe), Daniel Olmeda Reino (Toyota Motor Europe), Roberto Cipolla (University of Cambridge)

106. Visible Watermark Removal with Dynamic Kernel and Semantic-aware Propagation Xing Zhao (Shanghai Jiao Tong University), Li Niu (Shanghai Jiao Tong University),* Liqing Zhang (Shanghai Jiao Tong University)

107. A Simple Plugin for Transforming Images to Arbitrary Scales Qinye Zhou (Shanghai Jiao Tong University), Ziyi Li (Shanghai Jiao Tong University), Weidi Xie (Shanghai Jiao Tong University),* Xiaoyun Zhang (Shanghai Jiao Tong University), Yan-Feng Wang (Cooperative medianet innovation center of Shanghai Jiao Tong University), Ya Zhang (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University)

108. ELDA: Using Edges to Have an Edge on Semantic Segmentation Based UDA Ting-Hsuan Liao (National Tsing Hua University), Huang-Ru Liao (National Tsing Hua University), Shan-Ya Yang (National Tsing Hua University), Jie-En Yao (National Tsing Hua University), Li-Yuan Tsao (National Tsing Hua University), Hsu-Shen Liu (National Tsing Hua University), Chen-Hao Chao (National Tsing Hua University), Bo-Wun Cheng (National Tsing Hua University), Chia-Che Chang (MediaTek Inc.), Yi-Chen Lo (MediaTek Inc.), Chun-Yi Lee (National Tsing Hua University)*

111. UV-Based 3D Hand-Object Reconstruction with Grasp Optimization Ziwei Yu (National University of Singapore),* Linlin Yang (University of Bonn), You Xie (TUM), Ping Chen (Bean Tech), Angela Yao (National University of Singapore)

120. ARCSC-Net: An Approximate Residual Convolutional Sparse Coding Network For Compressed Sensing MRI Qian Wang (Yanshan University), PENGYU Li (Yanshan University), Jinjia Wang (Yanshan University)*

121. Pro-DDPM: Progressive Growing of Variable Denoising Diffusion Probabilistic Models for Faster Convergence Rohit Gandikota (Northeastern University),* Nicholas Brown (Northeastern University)

126. Few-shot Semantic Segmentation with Support-induced Graph Convolutional Network Jie Liu (University of Amsterdam),* Yanqi Bao (Northeastern University), Wenzhe Yin (University of Amsterdam), haochen wang (UvA), Yang Gao (Nanjing University), Jan-Jakob Sonke (The Netherlands Cancer Institute), Efstratios Gavves (University of Amsterdam)

Gold Sponsors



127. Pose-graph via Adaptive Image Re-ordering Daniel Barath (ETH Zürich),* Jana Noskova (CMP CTU FEE), Iván Eichhardt (Eötvös Loránd University), Jiri Matas (Czech Technical University, Prague)

131. SiNeRF: Sinusoidal Neural Radiance Fields for Joint Pose Estimation and Scene Reconstruction Yitong Xia (ETH Zurich),* Hao Tang (ETH Zurich), Radu Timofte (University of Wurzburg & ETH Zurich), Luc Van Gool (ETH Zurich)

132. Multi-View Multi-Person 3D Pose Estimation with Uncalibrated Camera Networks Yan Xu (Carnegie Mellon University),* Kris Kitani (Carnegie Mellon University)

134. Can I see an Example? Active Learning the Long Tail of Attributes and Relations Tyler L Hayes (RIT),* Maximillian Nickel (Facebook AI Research), Christopher Kanan (University of Rochester), Ludovic Denoyer (No University), Arthur Szlam (Facebook)

139. Bootstrapping Human Optical Flow and Pose Aritro Roy Arko (The University of British Columbia),* Jim Little (University of British Columbia, Canada), Kwang Moo Yi (University of British Columbia)

148. DeepSpark: A Lightweight Model when ISP and Deep Learning Sparks Hongyang Chen (Xi'an Jiaotong University),* Kaisheng Ma (Tsinghua University)

149. Font Representation Learning via Paired-glyph Matching Junho Cho (Seoul National University),* Kyuewang Lee (Seoul National University), Jin Young Choi (Seoul National University)

155. Shape preserving facial landmarks with Graph Attention Networks Andrés Prados Torreblanca (Universidad Politécnica de Madrid), José M Buenaposada (Universidad Rey Juan Carlos),* Luis Baumela (Universidad Politecnica de Madrid)

163. KPE: Keypoint Pose Encoding for Transformer-based Image Generation Soon Yau Cheong (University of Surrey),* Armin Mustafa (University of Surrey), Andrew Gilbert (University of Surrey)

173. ISG: I can See Your Gene Expression Yan Yang (The Australian National University),* Liyuan Pan (The Australian National University), Liu liu (ANU (Australian National University)), Eric A Stone (The Australian National University)

176. Two-View Left Ventricular Segmentation and Ejection Fraction Estimation in 2D Echocardiograms Frank Cally A Tabuco (University of the Philippines - Diliman),* Jose Donato A Magno (Philippine General Hospital), Nathaniel Jr Orillaza (University of the Philippines Manila), Rani Ailyna V Domingo (UP College of Medicine), Prospero C. Naval (University of the Philippines - Diliman)

Gold Sponsors



181. Meta Transferring for Deblurring Po-Sheng Liu (National Yang Ming Chiao Tung University),* Fu-Jen Tsai (National Tsing Hua University), Yan-Tsung Peng (National Chengchi University), Chung-Chi Tsai (Qualcomm Technology), Chia-Wen Lin (National Tsing Hua University), Yen-Yu Lin (National Yang Ming Chiao Tung University)

182. Debiasing Image-to-Image Translation Model Md. Mehrab Tanjim (University of California, San Diego),* Krishna Kumar Singh (Adobe Research), Kushal Kafle (Adobe Research), Ritwik Sinha (Adobe Research), Garrison Cottrell (University of California, San Diego)

185. Learning Object-level Point Augmentor for Semi-supervised 3D Object Detection Cheng-Ju Ho (National Yang Ming Chiao Tung University),* Chen-Hsuan Tai (National Yang Ming Chiao Tung University), Yi-Hsuan Tsai (Google), Yen-Yu Lin (National Yang Ming Chiao Tung University), Ming-Hsuan Yang (University of California at Merced)

190. Dual Decision Improves Open-Set Panoptic Segmentation Haiming Xu (The University of Adelaide),* Hao Chen (Huawei Noah's Ark Lab), Lingqiao Liu (University of Adelaide), Yufei Yin (University of Science and Technology of China)

191. SeA: Selective Attention for Fine-grained Visual Categorization yajie Chen (Nanjing University of Science and Technology),* Huan Wang (Nanjing University of Science & Technology), PAN PEIWEN (Nanjing University of Science and Technology)

197. Inharmonious Region Localization with Auxiliary Style Feature Penghao Wu (Shanghai Jiao Tong University), Li Niu (Shanghai Jiao Tong University),* Liqing Zhang (Shanghai Jiao Tong University)

198. Inharmonious Region Localization via Recurrent Self-Reasoning Penghao Wu (Shanghai Jiao Tong University), Li Niu (Shanghai Jiao Tong University),* Jing Liang (Shanghai Jiao Tong University), Liqing Zhang (Shanghai Jiao Tong University)

200. End-to-End Learning of Multi-category 3D Pose and Shape Estimation Yigit Baran Can (ETH Zurich),* Alexander Liniger (ETH Zurich), Danda Pani Paudel (ETH Zürich), Luc Van Gool (ETH Zurich)

204. GameCodec: Neural Cloud Gaming Video Codec Hoang Le (Qualcomm AI),* Reza Pourreza (Qualcomm), Amir Said (Qualcomm Technologies, Inc.), Guillaume Sautiere (Qualcomm AI Research), Auke Wiggers (Qualcomm AI Research)

210. Learning ODIN Amir Jevnisek (Tel-Aviv University),* Shai Avidan

Gold Sponsors



(Tel Aviv University)

211. Weak-shot Semantic Segmentation by Transferring Semantic Affinity and Boundary Siyuan Zhou (Shanghai Jiao Tong University), Li Niu (Shanghai Jiao Tong University),* Jianlou Si (SenseTime), Chen Qian (SenseTime), Liqing Zhang (Shanghai Jiao Tong University)

220. Self-Supervised Robustifying Guidance for Monocular 3D Face Reconstruction Hitika Tiwari (Indian Institute of Technology Kanpur & National Yang Ming Chiao Tung University),* Min-Hung Chen (Microsoft), Yi-Min Tsai (MediaTek), Hsien-Kai Kuo (MediaTek), Hung-Jen Chen (MediaTek), Kevin Jou (MediaTek Inc.), K. S. Venkatesh (IIT Kanpur), Yong-Sheng Chen (National Yang Ming Chiao Tung University)

222. Ki-Pode: Keypoint-based Implicit Pose Distribution Estimation of Rigid Objects Thorbjorn M. Iversen (University of Southern Denmark),* Rasmus Haugaard (University of Southern Denmark), Anders G Buch (University of Southern Denmark)

224. Towards Unsupervised Sketch-based Image Retrieval Conghui Hu (National University of Singapore),* Yongxin Yang (University of Surrey), Yunpeng Li (University of Surrey), Timothy Hospedales (Edinburgh University), Yi-Zhe Song (University of Surrey)

225. Feature Embedding by Template Matching as a ResNet Block Ada Gorgun (Middle East Technical University),* Yeti Z. Gurbuz (Middle East Technical University), Aydin Alatan (Middle East Technical University, Turkey)

227. In the Eye of Transformer: Global-Local Correlation for Egocentric Gaze Estimation Bolin Lai (Georgia Institute of Technology),* Miao Liu (Georgia Institute of Technology), Fiona Ryan (Georgia Institute of Technology), James Rehg (Georgia Institute of Technology)

228. Selective Colour Restoration of Underwater Surfaces Chau Yi Li (Queen Mary University of London, UK),* ANDREA CAVALLARO (Queen Mary University of London, UK)

229. Semantic Segmentation with Active Semi-Supervised Representation Learning Aneesh Rangnekar (Rochester Institute of Technology),* Christopher Kanan (University of Rochester), Matthew J Hoffman (Rochester Institute of Technology)

230. Privacy Vulnerability of Split Computing to Data-Free Model Inversion Attacks Xin Dong (Harvard University),* Hongxu Yin (NVIDIA), Jose M. Alvarez (NVIDIA), Jan Kautz (NVIDIA), Pavlo Molchanov (NVIDIA), H.T. Kung (Harvard University)

Gold Sponsors



231. Hybrid-Learning Video Moment Retrieval across Multi-Domain Labels Weitong Cai (Queen Mary University of London),* Jiabo Huang (Queen Mary University of London), Shaogang Gong (Queen Mary University of London)

237. Unsupervised Domain Adaptive Fundus Image Segmentation with Few Labeled Source Data Qianbi Yu (University of Sydney),* Dongnan Liu (University of Sydney), Chaoyi Zhang (University of Sydney), XINWEN ZHANG (University of Sydney), Weidong Cai (University of Sydney)

238. You Only Need 90K Parameters to Adapt Light: a Light Weight Transformer for Image Enhancement and Exposure Correction Ziteng Cui (The University of Tokyo), Kunchang Li (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Lin Gu (RIKEN, AIP / The University of Tokyo),* Shenghan Su (Shanghai Jiao Tong University), Peng Gao (Chinese university of hong kong), ZhengKai Jiang (Tencent Youtu Lab), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Tatsuya Harada (The University of Tokyo / RIKEN)

239. Learnable Descriptive Convolutional Network for Face Anti-Spoofing Pei-Kai Huang (National Tsing Hua University),* Hui-Yu Ni (National Tsing Hua University), Yan-Qin Ni (National Central University), Chiou-Ting Hsu (National Tsing Hua University)

241. Trident Pyramid Networks for Object Detection Cédric Picon (KU Leuven),* Tinne Tuytelaars (KU Leuven)

243. CICC: Channel Pruning via the Concentration of Information and Contributions of Channels Yihao Chen (Zhejiang University),* Guanzhong Tian (Ningbo Research Institute, Zhejiang University), Zhishan Li (Zhejiang University), Yingqing Yang (Zhejiang University), Lei Xie (Zhejiang University), Yong Liu (Zhejiang University), longhua ma (NingboTech University), Shanqi Liu (Zhejiang University)

244. Content-Diverse Comparisons improve IQA William Thong (Sony AI),* Jose Costa Pereira (Huawei Noah's Ark Lab), Sarah Parisot (Huawei Noah's Ark Lab), Ales Leonardis (Huawei Noah's Ark Lab), Steven McDonagh (Huawei Noah's Ark Lab)

250. A Tri-Layer Plugin to Improve Occluded Detection Guanqi Zhan (University of Oxford),* Weidi Xie (Shanghai Jiao Tong University), Andrew Zisserman (University of Oxford)

251. Dress Well via Fashion Cognitive Learning Kaicheng PANG (Hong Kong Polytechnic University), Xingxing Zou (Laboratory for Artificial Intelligence in Design, The Hong Kong Polytechnic University), Waikeng Wong (Institute of Textiles and Clothing, The Hong Kong Polytechnic University)*

Gold Sponsors



258. EAPruning: Universal Evolutionary Pruning for Vision Transformers and CNNs Qingyuan Li (Meituan), Bo Zhang (Meituan),* Xiangxiang Chu (Meituan)

259. Check Your Other Door! Creating Backdoor Attacks in the Frequency Domain Hasan Abed Al Kader Hammoud (King Abdullah University of Science and Technology),* Bernard Ghanem (KAUST)

261. USB: Universal-Scale Object Detection Benchmark Yosuke Shinya (DENSO CORPORATION / independent researcher)*

266. Edge Detection of Motion-blurred Images based on GAN networks Feng Li (Donghua university), Jiyu Li (Donghua University)*

267. LDEdit: Towards Generalized Text Guided Image Manipulation via Latent Diffusion Models Paramanand Chandramouli (University of Siegen),* Kanchana Vaishnavi Gandikota (University of Siegen)

268. Exploring Localization for Self-supervised Fine-grained Contrastive Learning di wu (Westlake University), Siyuan Li (Westlake University),* Zelin Zang (Zhejiang University & Westlake University), Stan Z. Li (Westlake University)

272. Learning to Wear: Details-Preserved Virtual Try-on via Disentangling Clothes and Wearer Sangho Lee (Seoul National University), Seo Young Lee (Seoul National University), Joonseok Lee (Google Research & Seoul National University)*

279. Rethinking the Evaluation of Unbiased Scene Graph Generation Xingchen Li (Zhejiang University),* Long Chen (Columbia University), Jian Shao (Zhejiang University), Shaoning Xiao (Zhejiang University), Songyang Zhang (University of Rochester), Jun Xiao (Zhejiang University)

281. Improving Gradient Paths for Binary Convolutional Neural Networks Baozhou Zhu (Tudelft),* Peter Hofstee (IBM), Jinho Lee (Yonsei University), Zaid Alars (TUDelft)

285. Dual Pyramid Generative Adversarial Networks for Semantic Image Synthesis Shijie Li (Bonn University),* Ming-Ming Cheng (Nankai University), Jürgen Gall (University of Bonn)

288. Region-of-Interest Based Neural Video Compression Yura M. Perugachi-Diaz (Vrije Universiteit Amsterdam), Guillaume Sautiere (Qualcomm AI Research), Davide Abati (Qualcomm AI Research),* Yang Yang (Qualcomm Inc.), Amirhossein Habibian (Qualcomm AI Research), Taco S. Cohen (Qualcomm AI Research)

289. Compressing Video Calls using Synthetic Talking Heads Madhav Agarwal (IIIT-Hyderabad),* Anchit Gupta (IIIT Hyderabad), Rudrabha Mukhopadhyay (IIIT Hyderabad), Vinay Namboodiri

Gold Sponsors



(University of Bath), C.V. Jawahar (IIIT-Hyderabad)

290. Implicit texture mapping for multi-view video synthesis

Mohamed I Lakhali (Huawei)*, ANDREA CAVALLARO (Queen Mary University of London, UK), Oswald Lanz (Free University of Bozen-Bolzano)

293. TransResNet: Integrating the Strengths of ViTs and CNNs for High Resolution Medical Image Segmentation via Feature Grafting

Muhammad Hamza Sharif (MBZUAI), Dmitry Demidov (Mohamed bin Zayed University of Artificial Intelligence), Asif Hanif (Muhammad Bin Zayed University of Artificial Intelligence), Mohammad Yaqub (Mohamed Bin Zayed University of Artificial Intelligence), Min Xu (Carnegie Mellon University)*

296. Exemplar Learning for Medical Image Segmentation Qing En (Carleton University)*, Yuhong Guo (Carleton University)

298. APSNet: Attention Based Point Cloud Sampling Yang Ye (Georgia State University), Xiulong Yang (Georgia State University), Shihao Ji (Georgia State University)*

299. Rethinking Prototypical Contrastive Learning through Alignment, Uniformity and Correlation Shentong Mo (Carnegie Mellon University), Zhun Sun (Tohoku University)*, Chao Li (RIKEN)

300. Learning visual representations for transfer learning by suppressing texture Shlok Kumar Mishra (University of Maryland College Park)*, Anshul Shah (Johns Hopkins University), Ankan Bansal (Amazon.com), Janit Anjaria (University of Maryland, College Park), Jonghyun Choi (Yonsei University), Abhinav Shrivastava (University of Maryland), Abhishek Sharma (Cruise LLC), David Jacobs (University of Maryland)

302. Towards Efficient Neural Scene Graphs by Learning Consistency Fields Yeji Song (Seoul National University), Chaerin Kong (Seoul National University), Seo Young Lee (Seoul National University), Nojun Kwak (Seoul National University)*, Joonseok Lee (Google Research & Seoul National University)

307. Towards Scalable Spectral Clustering via Spectrum-Preserving Sparsification Yongyu Wang (Michigan Technological University)*, Zhuo Feng (Stevens Institute of Technology)

309. G2Net: Generic Game-Theoretic Network for Partial-Label Image Classification Rabab Abdelfattah (University of South Carolina)*, Xin Zhang (University of South Carolina), Mostafa M. Fouda (Idaho State University), XIAOFENG Wang (University of South Carolina), Song Wang (University of South Carolina)

Gold Sponsors



311. Track Targets by Dense Spatio-Temporal Position Encoding

Jinkun Cao (Carnegie Mellon University),* Hao Wu (Bytedance Inc.),
Kris Kitani (Carnegie Mellon University)

313. Scale-Prior Deformable Convolution for Exemplar-Guided Class-Agnostic Counting

Wei Lin (City University of Hong Kong),* Kunlin Yang (Sensetime Group Limited), Xinzhu Ma (The University of Sydney), Junyu Gao (Northwestern Polytechnical University, Center for OPTical IMagery Analysis and Learning), Lingbo Liu (Hong Kong Polytechnic University), Shinan Liu (SenseTime Group Limited), Jun Hou (SenseTime Group Limited), Shuai Yi (SenseTime Group Limited), Antoni Chan (City University of Hong Kong, Hong, Kong)

314. Spatio-Temporal Fusion-based Monocular 3D Lane Detection

Yin Wang (JiLin University),* Qiuyi Guo (SenseTime Group Limited), Peiwen Lin (SenseTime Group Limited), Guangliang Cheng (Sensetime Group Limited), Jian Wu (Jilin university)

315. Task Generalizable Spatial and Texture Aware Image Downsizing Network

Lin Ma (Samsung),* Weiming Li (Samsung Research China – Beijing (SRC-B)), Hongsheng Li (The Chinese University of Hong Kong), Qiang Wang (Samsung Research China, Beijing), Ji-Yeon Kim (Samsung Advanced Institute of Technology)

320. Zero-shot Visual Commonsense Immorality Prediction

Yujin Jeong (Korea University), seongbeom Park (Korea University), Suhong Moon (UC Berkeley), Jinkyu Kim (Korea University)*

322. Signing Outside the Studio: Benchmarking Background Robustness for Continuous Sign Language Recognition

Youngjoon Jang (KAIST),* Youngtaek Oh (KAIST), Jae Won Cho (KAIST), Dong-Jin Kim (Hanyang University), Joon Son Chung (KAIST), In So Kweon (KAIST)

323. Dist2: Distribution-Guided Distillation for Object Detection

Tianchu Guo (Artificial Intelligence Center, DAMO Academy, Alibaba Group, Hangzhou, China),* Pengyu Li (Alibaba Group), Wei Liu (Alibaba), Bin Luo (DAMO Academy, Alibaba Group), biao wang (Alibaba)

886. SVS: Adversarial refinement for sparse novel view synthesis

Violeta Menéndez González (University of Surrey),* Andrew Gilbert (University of Surrey), Graeme Phillipson (BBC), Stephen Jolly (BBC), Simon Hadfield (University of Surrey)

896. Class-Balanced Loss Based on Class Volume for Long Tailed Object Recognition

ZhiJian Zheng (National University of Singapore),* Teck Khim Ng (National University of Singapore)

Gold Sponsors

909. Maximizing Mutual Shape Information Md Amirul Islam (Ryerson University),* Matthew Kowal (York University), Patrick Esser (Heidelberg University), Bjorn Ommer (University of Munich), Konstantinos G Derpanis (York University), Neil Bruce (University of Guelph)

939. FitCLIP: Refining Large-Scale Pretrained Image-Text Models for Zero-Shot Video Understanding Tasks Santiago Castro (University of Michigan),* Fabian Caba (Adobe Research)

968. Multi-Task Edge Prediction in Temporally-Dynamic Video Graphs Osman Ülger (University of Amsterdam),* Julian Wiederer (Ulm University), Mohsen Ghafourian (TomTom), Vasileios Belagiannis (Otto von Guericke University Magdeburg), Pascal Mettes (University of Amsterdam)

972. Quantitative Metrics for Evaluating Explanations of Video DeepFake Detectors Federico Baldassarre (KTH - Royal Institute of Technology),* Quentin Debar (Huawei Ireland Research Center), Gonzalo Fiz Pontiveros (Huawei), Tri Kurniawan Wijaya (Huawei Ireland Research Centre)

977. Semantic Segmentation under Adverse Conditions: A Weather and Nighttime-aware Synthetic Data-based Approach Abdulrahman Kerim (Lancaster University),* Felipe C Chamone (Universidade Federal de Minas Gerais), Washington L. S. Ramos (Universidade Federal de Minas Gerais), Leandro Soriano Marcolino (Lancaster University), Erickson R. Nascimento (UFMG), Richard Jiang (Associate Professor, Lancaster University)

990. Adapting branched networks to realise progressive intelligence Jack Dymond (University of Southampton),* Sebastian Stein (University of Southampton), Steve R Gunn (University of Southampton)

1009. Analysing Training-Data Leakage from Gradients through Linear Systems and Gradient Matching Cangxiong Chen (University of Bath),* Neill Campbell (University of Bath)

1039. Program Generation from Diverse Video Demonstrations Anthony Manchin (University of Adelaide),* Jamie Sherrah (AIML), Qi Wu (University of Adelaide), Anton van den Hengel (University of Adelaide)

1055. Continuous Hand Gesture Recognition using Deep Coarse and Fine Hand Features Hazem Wannous (University of Lille),* Jean-Philippe Vandeborre (IMT Nord Europe)

1063. Joint Reconstruction and Super Resolution of Hyper-Spectral CTIS Images Mazen Mel (University of Padua),*

Gold Sponsors



Alexander Gatto (Sony Europe B.V.), Pietro Zanuttigh (University of Padova)

1083. Sampling Based On Natural Image Statistics Improves Local Surrogate Explainers Ricardo Kleinlein (Universidad Politécnica de Madrid),* Alexander Hepburn (University of Bristol), Raul Santos Rodriguez (University of Bristol), Fernando Fernández-Martínez (Universidad Politécnica de Madrid)

15:30-16:45

Orals Session 2: 3D Analysis Part I (Chair:)

78 Event-based Non-Rigid Reconstruction from Contours Yuxuan Xue (Max Planck Institute for Intelligent Systems),* Haolong Li (Max Planck Institute for Intelligent Systems), Stefan Leutenegger (TU Munich), Joerg Stueckler (Max-Planck-Institute for Intelligent Systems)

365 TetGAN: A Convolutional Neural Network for Tetrahedral Mesh Generation William M Gao (University of Chicago),* April Wang (Threedle), Gal Metzer (Tel Aviv University), Raymond A Yeh (Toyota Technological Institute at Chicago), Rana Hanocka (University of Chicago)

708 VoRF: Volumetric Relightable Faces Pramod Rao (Max-Planck-Institut für Informatik),* Mallikarjun B R (Max Planck Institute for Informatics), Gereon Fox (Max Planck Institute for Informatics), Tim Weyrich (Friedrich-Alexander Universität Erlangen-Nürnberg), Bernd Bickel (IST Austria), Hanspeter Pfister (Harvard University), Wojciech Matusik (MIT), Ayush Tewari (MIT), Christian Theobalt (MPI Informatik), Mohamed Elgharib (Max Planck Institute for Informatics)

17:00 – 19:00

Welcome Drinks Reception in Duchy sponsored by Intel

Gold Sponsors

ROKE

vivo

woven planet

intel.

IET

Tuesday 22nd November

8:00 **Registration Opens**

09:00-10:00 **Keynote Speaker**
Dacheng Tao, University of Sydney
(Chair:)

10:00-11:15 **Spotlights 1 (Chair:)**

32. Where are my Neighbors? Exploiting Patches Relations in Self-Supervised Vision Transformer. Guglielmo Camporese (University of Padova), Elena Izzo (University of Padua), Lamberto Ballan (University of Padova)*

92. OSM: An Open Set Matting Framework with OOD Detection and Few-Shot Matting Yuhongze Zhou (McGill University),* Issam Hadj Laradji (ServiceNow), Liguang Zhou (The Chinese University of Hong Kong, Shenzhen), Derek Nowrouzezahrai (McGill University)

111. UV-Based 3D Hand-Object Reconstruction with Grasp Optimization Ziwei Yu (National University of Singapore),* Linlin Yang (University of Bonn), You Xie (TUM), Ping Chen (Bean Tech), Angela Yao (National University of Singapore)

223. Geometric priors for data-efficient learning of 3D-mirror symmetry from a single image Yancong Lin (Delft University of Technology),* Silvia-Laura L Pintea (TU Delft), Jan C van Gemert (Delft University of Technology)

307. Towards Scalable Spectral Clustering via Spectrum-Preserving Sparsification Yongyu Wang (Michigan Technological University),* Zhuo Feng (Stevens Institute of Technology)

585. DUDA: Online-Offline Dual Domain Adaption for Semantic Segmentation Antao Pan (Zhejiang University),* Yawei Luo (Zhejiang University), Yi Yang (Zhejiang University), Jun Xiao (Zhejiang University)

899. CASAPose: Class-Adaptive and Semantic-Aware Multi-Object Pose Estimation Niklas Gard (Fraunhofer HHI),* Anna Hilsmann (Fraunhofer HHI), Peter Eisert (Fraunhofer Heinrich Hertz Institute)

13. Domain Generalization Capability Enhancement for Binary Neural Networks Jianming Ye (Peking University), Shunan Mao

Gold Sponsors



(Peking University), Shiliang Zhang (Peking University)*

61. One-Pot Multi-Frame Denoising Lujia Jin (Peking University), Shi Zhao (Peking University), Lei Zhu (Beijing University of Posts and Telecommunications), Qian Chen (University of Science and Technology of China), Yanye Lu (Peking University)*

116. Learning Fine-Grained Visual Understanding for Video Question Answering via Decoupling Spatial-Temporal Modeling Hsin-Ying Lee (National Taiwan University),* Hung-Ting Su (National Taiwan University), Bing-Chen Tsai (National Taiwan University), Tsung-Han Wu (National Taiwan University), Jia-Fong Yeh (National Taiwan University), Winston H. Hsu (National Taiwan University)

207. Multi-hop Modulated Graph Convolutional Networks for 3D Human Pose Estimation Jae Yung Lee (KT (Korea Telecom)),* IGIL KIM (KT)

227. In the Eye of Transformer: Global-Local Correlation for Egocentric Gaze Estimation Bolin Lai (Georgia Institute of Technology),* Miao Liu (Georgia Institute of Technology), Fiona Ryan (Georgia Institute of Technology), James Rehg (Georgia Institute of Technology)

237. Unsupervised Domain Adaptive Fundus Image Segmentation with Few Labeled Source Data Qianbi Yu (University of Sydney),* Dongnan Liu (University of Sydney), Chaoyi Zhang (University of Sydney), XINWEN ZHANG (University of Sydney), Weidong Cai (University of Sydney)

268. Exploring Localization for Self-supervised Fine-grained Contrastive Learning di wu (Westlake University), Siyuan Li (Westlake University),* Zelin Zang (Zhejiang University & Westlake University), Stan Z. Li (Westlake University)

298. APSNet: Attention Based Point Cloud Sampling Yang Ye (Georgia State University), Xiulong Yang (Georgia State University), Shihao Ji (Georgia State University)*

11:15-11:30

 **Tea Break**

11:30-13:00

Orals Session 3: 3D Analysis Part II (Chair:)

1049. Data Augmentation-free Unsupervised Learning for 3D Point Cloud Understanding (Representation learning) Guofeng Mei (UTS),* Cristiano Saltori (University of Trento), Fabio Poiesi (Fondazione Bruno Kessler), Jian Zhang (UTS), Elisa Ricci (University of Trento), Nicu Sebe (University of Trento), Qiang Wu (University of Trento)

Gold Sponsors

ROKE

vivo

 **woven planet**

intel.

IET

Technology Sydney)

429. STPLS3D: A Large-Scale Synthetic and Real Aerial Photogrammetry 3D Point Cloud Dataset Meida Chen (the University of Southern California - Institute for Creative Technologies), Qingyong Hu (University of Oxford),* Zifan Yu (Arizona State University), Hugues THOMAS (University of Toronto), Andrew Feng (USC ICT), Yu Hou (Carnegie Mellon University), Kyle McCullough (USC ICT), Fengbo Ren (Arizona State University), Lucio Soibelman (University of Southern California)

578. Robustifying the Multi-Scale Representation of Neural Radiance Fields Nishant Jain (iit roorkee), Suryansh Kumar (ETH Zurich),* Luc Van Gool (ETH Zurich)

383. Finding Directions in GAN's Latent Space for Neural Face Reenactment Stella Bounareli (Kingston University of London),* Vasileios Argyriou (Kingston University London), Georgios Tzimiropoulos (Queen Mary University of London)

13:00-14:00

 **Lunch Break**

BMVA AGM and Q & A Session

14:00- 15:30

Poster Session II (Chair:)

32. Where are my Neighbors? Exploiting Patches Relations in Self-Supervised Vision Transformer Guglielmo Camporese (University of Padova), Elena Izzo (University of Padua), Lamberto Ballan (University of Padova)*

116. Learning Fine-Grained Visual Understanding for Video Question Answering via Decoupling Spatial-Temporal Modeling Hsin-Ying Lee (National Taiwan University),* Hung-Ting Su (National Taiwan University), Bing-Chen Tsai (National Taiwan University), Tsung-Han Wu (National Taiwan University), Jia-Fong Yeh (National Taiwan University), Winston H. Hsu (National Taiwan University)

207. Multi-hop Modulated Graph Convolutional Networks for 3D Human Pose Estimation Jae Yung Lee (KT (Korea Telecom)),* IGIL KIM (KT)

223. Geometric priors for data-efficient learning of 3D-mirror symmetry from a single image Yancong Lin (Delft University of Technology),* Silvia-Laura L Pintea (TU Delft), Jan C van Gemert (Delft University of Technology)

329. Humans need not label more humans: Occlusion Copy &

Gold Sponsors

ROKE

vivo

 **woven planet**

intel.

IET

Paste for Occluded Human Instance Segmentation Evan Ling (Hyundai Motor Group Innovation Center in Singapore),* dezha HUANG (Hyundai Motor Group Innovation Center in Singapore (HMGICS)), Minhoe Hur (AIRS Company, Hyundai Motor Group)

337. Learning Clothes-irrelevant Cues for Clothes-Changing Person Re-identification Jingyi Mu (Nanjing University of Science and Technology),* Yong Li (Nanjing University of Science and Technology), Jun Li (Nanjing University of Science and Technology), Jian Yang (Nanjing University of Science and Technology)

340. DisPositioNet: Disentangled Pose and Identity in Semantic Image Manipulation Azade Farshad (Technical University of Munich),* Yousef Yeganeh (Technical University of Munich), Helisa Dharmo (Huawei Noah's Ark Lab), Federico Tombari (Google, TU Munich), Nassir Navab ("TU Munich, Germany")

342. VID-Trans-ReID: Enhanced Video Transformers for Person Re-identification Aishah Alsehaim (Durham university),* Toby P Breckon (Durham University)

344. Non-uniform Sampling Strategies for NeRF on 360° images Takashi Otonari (The University of Tokyo),* Satoshi Ikehata (National Institute of Informatics), Kiyoharu Aizawa (The University of Tokyo)

345. ScannerNet: A Deep Network for Scanner-Quality Document Images under Complex Illumination Chih-Jou Hsu (National Taiwan University), Yu-Ting Wu (National Taipei University),* Ming-Sui Lee (National Taiwan University), Yung-Yu Chuang (National Taiwan University)

347. Resolving Semantic Confusions for Improved Zero-Shot Detection Sandipan Sarma (Indian Institute of Technology Guwahati),* SUSHIL KUMAR (Indian Institute of Technology Guwahati), Arijit Sur (IIT Guwahati)

348. T4DT: Tensorizing Time for Learning Temporal 3D Visual Data Mikhail Usvyatsov (ETH Zürich),* Rafael Ballester (IE University), Lina Bashaeva (Skolkovo institute of science and technology), Konrad Schindler (ETH), Gonzalo Ferrer (Skolkovo Institute of Science and Technology), Ivan Oseledets (Skolkovo Institute of Science and Technology)

351. Unsupervised Flow Refinement near Motion Boundaries Shuzhi Yu (Duke University),* Hannah H Kim (Duke University), Shuai Yuan (Duke University), Carlo Tomasi (Duke University)

353. Efficient Vision-Language Pretraining with Visual Concepts and Hierarchical Alignment Mustafa Shukor (Sorbonne University),* Guillaume Couairon (Facebook AI Research), Matthieu Cord (Sorbonne University)

Gold Sponsors



355. Capturing Temporal Information in a Single Frame: Channel Sampling Strategies for Action Recognition Kiyoon Kim (The University of Edinburgh),* Shreyank N Gowda (University of Edinburgh), Oisin Mac Aodha (University of Edinburgh), Laura Sevilla-Lara (University of Edinburgh)

357. Geometry Driven Progressive Warping for One-Shot Face Animation Yatao Zhong (Microsoft),* Faezeh Amjadi (Microsoft), Ilya Zharkov (Microsoft)

358. Biologically Plausible Variational Policy Gradient with Spiking Recurrent Winner-Take-All Networks Zhile Yang (University of Leeds),* Shangqi Guo (Tsinghua University), Ying Fang (Fujian Normal University), Jian Liu (University of Leeds)

359. Dual-lens Reference Image Super-Resolution Jing Zhu (Samsung Research America), Wenbo Li (Samsung Research America),* Hongxia Jin (Samsung Research America)

361. MaterialNet: Multi-scale Texture Hierarchy and Multi-view Surface Reflectance for Material Type Recognition DONGJIN LEE (Kyung Hee University), Seungkyu Lee (Kyung Hee University)*

364. LcT: Locally-Enhanced Cross-Window Vision Transformer Canhui Wei (Southwest University), Huiwei Wang (Southwest University)*

365. TetGAN: A Convolutional Neural Network for Tetrahedral Mesh Generation William M Gao (University of Chicago),* April Wang (Threedle), Gal Metzger (Tel Aviv University), Raymond A Yeh (Toyota Technological Institute at Chicago), Rana Hanocka (University of Chicago)

367. Boosting Adversarial Robustness From The Perspective of Effective Margin Regularization Ziquan Liu (City University of Hong Kong),* Antoni Chan (City University of Hong Kong, Hong, Kong)

369. Less is More: Facial Landmarks can Recognize a Spontaneous Smile Md. Tahrim Faroque Tushar (North South University), Yan Yang (The Australian National University), Md Zakir Hossain (The Australian National University), Sheikh Motahar Naim (Amazon), Nabeel Mohammed (North South University), Shafin Rahman (North South University)*

370. CounTR: Transformer-based Generalised Visual Counting Chang Liu (Shanghai Jiao Tong University), Yujie Zhong (University of Oxford), Andrew Zisserman (University of Oxford), Weidi Xie (Shanghai Jiao Tong University)*

371. BOAT: Bilateral Local Attention Vision Transformer Tan Yu (Baidu Research),* Gangming Zhao (The University of Hong Kong),

Gold Sponsors



Ping Li (Baidu), Yizhou Yu (The University of Hong Kong)

372. SSR: An Efficient and Robust Framework for Learning with Unknown Label Noise Chen Feng (Queen Mary University of London),* Georgios Tzimiropoulos (Queen Mary University of London), Ioannis Patras (Queen Mary University of London)

373. Unsupervised Low Light Image Enhancement Transformer Based on Dual Contrastive Learning Fengji Ma (Beihang University),* Jinping Sun (Beihang University)

378. Free-form 3D Scene Inpainting with Dual-stream GAN Ru-Fen Jheng (National Taiwan University),* Tsung-Han Wu (National Taiwan University), Jia-Fong Yeh (National Taiwan University), Winston H. Hsu (National Taiwan University)

381. PaRK-Detect: Towards Efficient Multi-Task Satellite Imagery Road Extraction via Patch-Wise Keypoints Detection Shenwei Xie (Beijing University of Posts and Telecommunications),* Wanfeng Zheng (Beijing University of Posts and Telecommunications), Zhenglin Xian (Beijing University of Posts and Telecommunications), Junli Yang (Beijing University of Posts and Telecommunications), Chuang Zhang (Beijing University of Posts and Telecommunications), Ming Wu (Beijing University of Posts and Telecommunications)

383. Finding Directions in GAN's Latent Space for Neural Face Reenactment Stella Bounareli (Kingston University of London),* Vasileios Argyriou (Kingston University London), Georgios Tzimiropoulos (Queen Mary University of London)

385. Information Theoretic Representation Distillation Roy V Miles (Imperial College London),* Adrian Lopez (Imperial College London), Krystian Mikolajczyk (Imperial College London)

386. Ranking Aggregation with Interactive Feedback for Collaborative Person Re-identification Ji Huang (Wuhan University), Chao Liang (Wuhan University),* Yue Zhang (Wuhan University), Zhongyuan Wang (Wuhan University), Chunjie Zhang (Beijing Jiaotong University)

388. A Memory Transformer Network for Incremental Learning Thomas Bird (UCL),* Ahmet Iscen (Google), Mathilde Caron (Facebook Artificial Intelligence Research), Alireza Fathi (Google), Cordelia Schmid (Google)

389. Taylor Swift: Taylor Driven Temporal Modeling for Swift Future Frame Prediction Emad Bahrami (University of Bonn),* Mohammad Saber Pourheydari (University of Bonn), Mohsen Fayyaz (Microsoft), Gianpiero Francesca (Toyota-Europe), Mehdi Noroozi (Bosch Gmb), Jürgen Gall (University of Bonn)

390. CroCPS: Addressing Photometric Challenges in

Gold Sponsors



Self-Supervised Category-Level 6D Object Poses with Cross-Modal Learning

Pengyuan Wang (TUM),* Lorenzo Garattoni (Toyota-Europe), Sven Meier (Toyota Motor Europe), Nassir Navab ("TU Munich, Germany"), Benjamin Busam (Technical University of Munich)

392. Robust Action Segmentation from Timestamp Supervision

Yaser Souri (Microsoft), Yazan Abu Farha (Birzeit University), Emad Bahrami (University of Bonn),* Gianpiero Francesca (Toyota-Europe), Jürgen Gall (University of Bonn)

394. Variational Simultaneous Stereo Matching and Defogging in Low Visibility

Yining Ding (Heriot-Watt University),* Andrew Wallace (Heriot Watt University), Sen Wang (Imperial College London)

395. Sparse in Space and Time: Audio-visual Synchronisation with Trainable Selectors

Vladimir Iashin (Tampere University),* Weidi Xie (Shanghai Jiao Tong University), Esa Rahtu (Tampere University), Andrew Zisserman (University of Oxford)

397. Segmentation Assisted U-shaped Multi-scale Transformer for Crowd Counting

Yifei Qian (University of St Andrews),* Liangfei Zhang (University of St Andrews), Xiaopeng Hong (Harbin Institute of Technology), Carl Donovan (University of St Andrews), Ognjen Arandjelovic (University of St Andrews)

398. MUAD: Multiple Uncertainties for Autonomous Driving benchmark for multiple uncertainty types and tasks

Gianni Franchi (ENSTA Paris),* Xuanlong Yu (ENSTA Paris), Andrei Bursuc (valeo.ai), Angel Tena (Next Limit Technologies), Rémi Kazmierczak (ENSTA Paris), Severine Dubuisson (Aix-Marseille University), Emanuel Aldea (Paris-Saclay University), David Filliat (ENSTA Paris)

405. Semi-Supervised Object Detection with Object-wise Contrastive Learning and Regression Uncertainty

honggyu choi (KAIST),* Zhixiang Chen (University of Sheffield), Xuepeng Shi (Imperial College London), Tae-Kyun (T-K) Kim (KAIST/Imperial College London)

406. Revisiting Self-Supervised Contrastive Learning for Facial Expression Recognition

Yuxuan Shu (Imperial College London),* Xiao Gu (Imperial College London), Guang-Zhong Yang (SJTU), Benny P L Lo (Imperial College London)

407. Flynet: Max it, Excite it, Quantize it

Luis Guerra (Monash University),* Tom Drummond (University of Melbourne)

411. HSPA: Hough Space Pattern Analysis as an Answer to Local Description Ambiguities for 3D Pose Estimation

Fabrice Mayran de Chamisso (CEA, LIST),* Boris Meden (CEA List), Mohamed Tamaazousti (CEA Saclay)

Gold Sponsors

417. Masked Supervised Learning for Semantic Segmentation Hasib Zunair (Concordia University),* Abdessamad Ben Hamza (Concordia University)

418. Fill in Fabrics: Body-Aware Self-Supervised Inpainting for Image-Based Virtual Try-On Hasib Zunair (Concordia University),* Yan Gobeil (Decathlon Canada), Samuel Mercier (Decathlon Canada), Abdessamad Ben Hamza (Concordia University)

420. Selective Partial Domain Adaptation Pengxin Guo (Southern University of Science and Technology),* Jinjing Zhu (Southern University of Science and Technology), Yu Zhang (Southern University of Science and Technology)

421. Unified Negative Pair Generation toward Well-discriminative Feature Space for Face Recognition Jun-uk Jung (KOREATECH),* seon-hoon SH Lee (koreatech), Heung-Seon Oh (KOREATECH), Yong-Jun Park (KOREATECH), Sung-Bin Son (KOREATECH), Joo-chan Park (KOREATECH)

424. Pyramid Region-based Slot Attention Network for Temporal Action Proposal Generation Shuaicheng Li (Sensetime Research),* Feng Zhang (Fudan University), Rui-Wei Zhao (Fudan University), Kunlin Yang (Sensetime Group Limited), Lingbo Liu (Hong Kong Polytechnic University), Rui Feng (Fudan University), Jun Hou (SenseTime Group Limited)

427. RGB-T Multi-Modal Crowd Counting Based on Transformer Zhengyi Liu (Anhui University),* Wei Wu (Anhui University), Yacheng Tan (Anhui University), Guanghui Zhang (Anhui University)

428. Disentangling based Environment-Robust Feature Learning for Person ReID Yifan Liu (Tsinghua University), Ya-Li Li (Tsinghua University), Shengjin Wang (Tsinghua University)*

429. STPLS3D: A Large-Scale Synthetic and Real Aerial Photogrammetry 3D Point Cloud Dataset Meida Chen (the University of Southern California - Institute for Creative Technologies), Qingyong Hu (University of Oxford),* Zifan Yu (Arizona State University), Hugues THOMAS (University of Toronto), Andrew Feng (USC ICT), Yu Hou (Carnegie Mellon University), Kyle McCullough (USC ICT), Fengbo Ren (Arizona State University), Lucio Soibelman (University of Southern California)

439. Dual-Pixel Raindrop Removal Yizhou Li (Tokyo Institute of Technology),* Yusuke Monno (Tokyo Institute of Technology), Masatoshi Okutomi (Tokyo Institute of Technology)

443. Disentangling Content and Motion for Text-Based Neural Video Manipulation Levent Karacan (Iskenderun Technical University),* Tolga Kerimoğlu (Boğaziçi University), İsmail Ata İnan (Boğaziçi University), Tolga Birdal (TU Munich), Erkut Erdem

Gold Sponsors



(Hacettepe University), Aykut Erdem (Koc University)

444. MagFormer: Hybrid Video Motion Magnification Transformer from Eulerian and Lagrangian Perspectives Sicheng Gao (Beihang University), Yutang Feng (Beihang University), Xuhui Liu (Beihang University), Baochang Zhang (Beihang University), Zichen Zhu (Harbin Institute of Technology), David Doermann (University at Buffalo), Linlin Yang (University of Bonn)*

445. Defect Transfer GAN: Diverse Defect Synthesis for Data Augmentation Ruyu Wang (Robert Bosch GmbH),* Sabrina Hoppe (Bosch), Eduardo Monari (Robert Bosch Corporate Research), Marco Huber (University of Stuttgart)

446. Towards Robust In-domain and Out-of-Domain Generalization: Contrastive Learning with Prototype Alignment and Collaborative Attention Yuan-Jhe Kuo (National Tsing Hua University),* Cheng-Yu Yang (National Tsing-Hua University), Chiou-Ting Hsu (National Tsing Hua University)

447. A Unified Mixture-View Framework for Unsupervised Representation Learning Xiangxiang Chu (Meituan), Xiaohang Zhan (The Chinese University of Hong Kong), Bo Zhang (Meituan)*

449. CLAD: A Contrastive Learning based Approach for Background Debiasing Ke Wang (EPFL),* Harshitha Machiraju (EPFL), Oh-Hyeon Choung (Epfl), Michael Herzog (EPFL), Pascal Frossard (EPFL)

458. Fractional Optimization Model for Infrared and Visible Image Fusion Kang Zhang (Nanjing University of Science and Technology), Shiwei Wu (Nanjing University of Science and Technology), Zhiliang Wu (Nanjing University of Science and Technology), Xia Yuan (Nanjing University of Science and Technology),* ChunXia Zhao (Nanjing university of science and technology)

460. Doubly Contrastive End-to-End Semantic Segmentation for Autonomous Driving under Adverse Weather Jongoh Jeong (KAIST), Jong-Hwan Kim (KAIST)*

469. Object Tracking Network Based on Deformable Attention Mechanism Kexin Chen (Nanjing University of Posts and Telecommunications),* baojie fan (njupt), xiaobin Guo (Nanjing University of Posts and Telecommunications)

470. Trans2k: Unlocking the Power of Deep Models for Transparent Object Tracking Alan Lukezic (University of Ljubljana),* Žiga Trojer (University of Ljubljana), Jiri Matas (Czech Technical University, Prague), Matej Kristan (University of Ljubljana)

471. Multi-body Self-Calibration Andrea Porfiri Dal Cin (Politecnico di Milano),* Giacomo Boracchi (Politecnico di Milano), Luca Magri

Gold Sponsors



(Politecnico di Milano)

472. Anomaly Detection and Localization Using Attention-Guided Synthetic Anomaly and Test-Time Adaptation

Behzad Bozorgtabar (EPFL),* Dwarikanath Mahapatra (Inception Institute of Artificial Intelligence), Jean-Philippe Thiran (École Polytechnique Fédérale de Lausanne)

473. K-Space Transformer for Fast MRI Reconstruction

Ziheng Zhao (Shanghai Jiao Tong University),* Tianjiao Zhang (Shanghai Jiao Tong University), Weidi Xie (Shanghai Jiao Tong University), Yan-Feng Wang (Cooperative medianet innovation center of Shanghai Jiao Tong University), Ya Zhang (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University)

474. SGENet: Spatial Guided Enhancement Network for Image Motion Deblurring

Chia-Hung Yeh (National Taiwan Normal University),* Yu-Chieh Wang (National Taiwan Normal University)

475. On the Importance of Image Encoding in Automated Chest X-Ray Report Generation

Otabek N Nazarov (Mohamed Bin Zayed University of Artificial Intelligence),* Mohammad Yaqub (Mohamed Bin Zayed University of Artificial Intelligence), Karthik Nandakumar (Mohamed Bin Zayed University of Artificial Intelligence)

476. IronDepth: Iterative Refinement of Single-View Depth using Surface Normal and its Uncertainty

Gwangbin Bae (University of Cambridge),* Ignas Budvytis (Department of Engineering, University of Cambridge), Roberto Cipolla (University of Cambridge)

479. Neighbor Regularized Bayesian Optimization for Hyperparameter Optimization

Lei Cui (Tsinghua University), Yangguang Li (SenseTime Group Limited),* Xin Lu (SenseTime Research), Fenggang Liu (SenseTime), Dong An (Institute of Automation, Chinese Academy of Sciences)

480. Fixed Point Layers for Geodesic Morphological Operations

Santiago Velasco-Forero (MINES ParisTech),* Ayoub Rhim (Ecole Nationale des Ponts et Chaussées), Jesus Angulo (Mines Paris Tech)

481. Unleashing the Potential of Vision-Language Models for Long-Tailed Visual Recognition

Teli Ma (Shanghai Artificial Intelligence Laboratory),* Shijie Geng (Rutgers University), Mengmeng Wang (Zhejiang University), Sheng Xu (Beihang University), Hongsheng Li (The Chinese University of Hong Kong), Baochang Zhang (Beihang University), Peng Gao (Chinese university of hong kong), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences)

482. BaseTransformers: Attention over base data-points for One Shot Learning

Mayug Maniparambil (ML Labs),* Kevin McGuinness

Gold Sponsors



(DCU), Noel O Connor (Home)

484. SAGE: Saliency-Guided Mixup with Optimal Rearrangements

Avery Ma (University of Toronto and Vector Institute),* Nikita Dvornik (Samsung AI Center Toronto), Ran Zhang (Samsung AI Center Toronto), Leila Pishdad (Borealis AI), Konstantinos G Derpanis (York University, Vector Institute and Samsung AI Center Toronto), Afsaneh Fazly (Samsung AI Center Toronto)

486. Cluster-level pseudo-labelling for source-free cross-domain facial expression recognition

Alessandro Conti (University of Trento),* Paolo Rota (University of Trento), Yiming Wang (Fondazione Bruno Kessler), Elisa Ricci (University of Trento)

488. Information Removal at the bottleneck in Deep Neural Networks

Enzo Tartaglione (Télécom Paris - Institut Polytechnique de Paris)*

491. Enhancing Person Synthesis in Complex Scenes via Intrinsic and Contextual Structure Modeling

Xi Tian (University of Bath),* Yongliang Yang (University of Bath), Qi Wu (University of Adelaide)

497. Classification of Biomedical Journal Images using Retargeting-Based Data Augmentation and Visually Explainable Attention Priors

Vinit Veerendraveer Singh (University Of Delaware),* Chandra Kambhamettu (University of Delaware)

499. Cross-Modal Fusion Distillation for Fine-Grained Sketch-Based Image Retrieval

Abhra Chaudhuri (University of Exeter),* Massimiliano Mancini (University of Tübingen), Yanbei Chen (University of Tübingen), Zeynep Akata (University of Tübingen), Anjan Dutta (University of Surrey)

501. Beyond Deterministic Translation for Unsupervised Domain Adaptation

Eleni Chiou (University College London),* Eleftheria Panagiotaki (University College London), Iasonas Kokkinos (Snap / University College London)

506. Stating Comparison Score Uncertainty and Verification Decision Confidence Towards Transparent Face Recognition

Marco Huber (Fraunhofer IGD),* Philipp Terhörst (Paderborn University), Florian Kirchbuchner (Fraunhofer Institute for Computer Graphics Research IGD), Naser Damer (Fraunhofer IGD), Arjan Kuijper (Fraunhofer Institute for Computer Graphics Research IGD and Mathematical and Applied Visual Computing group, TU Darmstadt)

509. Why Do Self-Supervised Models Transfer? On the Impact of Invariance on Downstream Tasks

Linus Ericsson (University of Edinburgh),* Henry Gouk (University of Edinburgh), Timothy Hospedales (Edinburgh University)

510. CNeRV: Content-adaptive Neural Representation for Visual

Gold Sponsors



Data Hao Chen (University of Maryland),* Matthew A Gwilliam (University of Maryland), Bo He (University of Maryland), Ser-Nam Lim (Meta AI), Abhinav Shrivastava (University of Maryland)

512. Teaching StyleGAN to Read: Improving Text-to-image Synthesis with U2C Transfer Learning Vinicius G Pereira (Puc Rio), JONATAS WEHRMANN (PUCRS)*

522. Face editing using a regression-based approach in the StyleGAN latent space Saeid Motiiian (Adobe),* Siavash Khodadadeh (University of Central Florida), Shabnam Ghadar (Adobe), Baldo Faieta (Adobe), Ladislau Boloni (University of Central Florida)

526. ORA3D: Overlap Region Aware Multi-view 3D Object Detection WONSEOK ROH (Korea University), Gyusam Chang (Korea university), Giljoo Nam (Facebook Inc.), Seokha Moon (Yonsei University), Chanyoung Kim (Korea University), Younghyun Kim (Hyundai), Sangpil Kim (Korea University), Jinkyu Kim (Korea University)*

532. Robust normalizing flows using Bernstein-type polynomials Sameera Ramasinghe (University of Adelaide),* Kasun Fernando (University of Toronto), Salman Khan (MBZUAI), Nick Barnes (ANU)

533. Unconditional Image-Text Pair Generation with Multimodal Cross Quantizer Hyungyung Lee (KAIST GSAI),* Sungjin Park (Korea Advanced Institute of Science and Technology), Joonseok Lee (Google Research & Seoul National University), Edward Choi (KAIST)

535. Distilling Representational Similarity using Centered Kernel Alignment (CKA) Aninda Saha (The University of Queensland),* Alina N Bialkowski (The University of Queensland), Sara Khalifa (CSIRO)

538. Centered Symmetric Quantization for Hardware-Efficient Low-Bit Neural Networks Faaiz Asim (Ulsan National Institute of Science and Technology (UNIST)),* Jaewoo Park (UNIST), Azat Azamat (UNIST), Jongeun Lee (UNIST)

541. On Temporal Granularity in Self-Supervised Video Representation Learning Rui Qian (Cornell University), Yeqing Li (Google), Liangzhe Yuan (Google Research), Boqing Gong (Google), Ting Liu (Google Research), Matthew Brown (Google), Serge Belongie (University of Copenhagen), Ming-Hsuan Yang (Google Research), Hartwig Adam (Google), Yin Cui (Google)*

542. RORD: A Real-world Object Removal Dataset Min-Cheol Sagong (Korea Univ.),* Yoon-Jae Yeo (Korea University), Seung-Won Jung (Korea University), Sung-Jea Ko (Korea University)

543. CLIPFont: Text Guided Vector WordArt Generation Yiren Song (Shanghai Jiaotong University),* Yuxuan Zhang (Shanghai jiaotong

Gold Sponsors



University)

544. Learning to Segment Object Affordances on Synthetic Data for Task-oriented Robotic Handovers Albert D Christensen (Aalborg University),* Daniel Lehotsky (Aalborg University), Marius W Jørgensen (Aalborg University), Dimitris Chrysostomou (Aalborg University)

546. Adversarial Pixel Restoration as a Pretext Task for Transferable Perturbations Hashmat Shadab Malik (MBZUAI),* Shahina Kunhimon (Mohamed bin Zayed University of Artificial Intelligence), Muzammal Naseer (MBZUAI), Salman Khan (MBZUAI/ANU), Fahad Shahbaz Khan (MBZUAI)

549. Towards Self-Supervised Gaze Estimation Arya Farkhondeh (Sapienza University of Rome),* Cristina Palmero (Universitat de Barcelona), Simone Scardapane (Sapienza University), Sergio Escalera (Computer Vision Center (UAB) & University of Barcelona,)

550. Multi-View Neural Surface Reconstruction with Structured Light Chunyu Li (Preferred Networks, Inc.),* Taisuke Hashimoto (Preferred Networks, Inc.), Eiichi Matsumoto (Preferred Networks, Inc.), Hiroharu Kato (Preferred Networks)

551. Are we pruning the correct channels in image-to-image translation models? Yiyong Li (BIGO Ltd.), Zhun Sun (Tohoku University),* Chao Li (RIKEN)

552. Local Feature Extraction from Salient Regions by Feature Map Transformation Yerim Jung (Inha Univ.), NUR SURIZA SYAZWANY BINTI AHMAD NIZAM (Inha University), Sang-Chul Lee (Inha univ)*

555. Masked Vision-Language Transformers for Scene Text Recognition Jie Wu (Westone Information Industry INC.),* Ying Peng (Westone Information Industry INC.), Shengming Zhang (Westone Information Industry INC), Weigang Qi (Westone Information Industry INC.), zhang jian (Westone Information Industry INC)

564. SP-ViT: Learning 2D Spatial Priors for Vision Transformers Yuxuan Zhou (University of Mannheim),* Wangmeng Xiang (The Hong Kong Polytechnic University), Chao Li (Alibaba), Biao Wang (Alibaba), Xihan Wei (Alibaba), Lei Zhang ("Hong Kong Polytechnic University, Hong Kong, China"), Margret Keuper (University of Mannheim), Xian-Sheng Hua (Damo Academy, Alibaba Group)

569. Flow-based GAN for 3D Point Cloud Generation from a Single Image Yao Wei (University of Twente), George Vosselman ("University of Twente, the Netherlands"), Michael Ying Yang (University of Twente)*

576. Membership Privacy-Preserving GAN Heonseok Ha (Seoul

Gold Sponsors



National University), Uiwon Hwang (Seoul National University), Jaehee Jang (Seoul National University), Ho Bae (Ewha Womans University), Sungroh Yoon (Seoul National University)*

577. Event Transformer FlowNet for optical flow estimation Yi Tian (Institut de Robòtica i Informàtica Industrial CSIC-UPC),* Juan Andrade-Cetto (Institut de Robòtica i Informàtica Industrial CSIC-UPC)

578. Robustifying the Multi-Scale Representation of Neural Radiance Fields Nishant Jain (iit roorkee), Suryansh Kumar (ETH Zurich),* Luc Van Gool (ETH Zurich)

580. SVL-Adapter: Self-Supervised Adapter for Vision-Language Pretrained Models Omiros Pantazis (University College London),* Gabriel Brostow (University College London), Kate Jones (University College London), Oisín Mac Aodha (University of Edinburgh)

581. Image-to-Image Translation with Text Guidance Bowen Li (University of Oxford),* Philip Torr (University of Oxford), Thomas Lukasiewicz (University of Oxford)

584. SalLiDAR: Saliency Knowledge Transfer Learning for 3D Point Cloud Understanding Guanqun Ding (University of Tsukuba),* Nevrez Imamoglu (AIST), Ali Caglayan (National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan), Masahiro Murakawa (National Institute of Advanced Industrial Science and Technology (AIST)), Ryosuke Nakamura (National Institute of Advanced Industrial Science and Technology)

585. DUDA: Online-Offline Dual Domain Adaption for Semantic Segmentation Antao Pan (Zhejiang University),* Yawei Luo (Zhejiang University), Yi Yang (Zhejiang University), Jun Xiao (Zhejiang University)

587. Towards Unified Multi-Excitation for Unsupervised Video Prediction Junyan Wang (UNSW Sydney),* Qin Likun (Institute of Microelectronics of Chinese Academy of Science), Peng Zhang (Durham University), Yang Long (Durham University), Bingzhang Hu (Hefei CAS Dihuge Automation CO., LTD), Maurice Pagnucco (UNSW), Shizheng Wang (Chinese Academy of Sciences), Yang Song (University of New South Wales)

589. Class-Prototypes for Contrastive Learning in Weakly-Supervised 3D Point Cloud Segmentation Rong Li (South China University of Technology),* Anh-Quan CAO (Inria), Raoul de Charette (Inria)

592. Semantics-Adding Flaw-Erasing Network for Semantic Human Matting Jiayu Sun (Dalian University of Technology),* Zhanghan Ke (City University of Hong Kong), Ke Xu (City University of Hong Kong), Fan Shao (Wonxing Technology), Lihe Zhang (Dalian

Gold Sponsors



University of Technology), Huchuan Lu (Dalian University of Technology), Rynson W.H. Lau (City University of Hong Kong)

594. clip2latent: Text driven sampling of a pre-trained StyleGAN using denoising diffusion and CLIP Justin N.M. Pinkney (Lambda, Inc),* Chuan Li (Lambda, Inc)

601. Polishing Network for Decoding of Higher-Quality Diverse Image Captions Yue Zheng (Tsinghua University),* Ya-Li Li (Tsinghua University), Shengjin Wang (Tsinghua University)

603. HDR Reconstruction from Bracketed Exposures and Events Richard Shaw (Huawei London Research Centre),* Sibi Catley-Chandar (Huawei Noah's Ark Lab), Ales Leonardis (Huawei Noah's Ark Lab), Eduardo Pérez Pellitero (Huawei Noah's Ark Lab)

605. Improving Interpretability by Information Bottleneck Saliency Guided Localization Hao Zhou (Jiangsu university), keyang Cheng (Jiangsu University),* Yu Si (Jiangsu University), Liuyang Yan (Jiangsu University)

606. Distilling and Refining Domain-Specific Knowledge for Semi-Supervised Domain Adaptation Ju Hyun Kim (Dongguk University),* Ba Hung Ngo (Dongguk University), Jae Hyeon Park (Dongguk University), Jung Eun Kwon (Dongguk university), Ho Sub Lee (Daegu University), Sung In Cho (Dongguk University)

607. GLAMI-1M: A Multilingual Image-Text Fashion Dataset Vaclav Kosar (GLAMI), Antonín Hoskovec (GLAMI), Milan Šulc (Rossum.ai),* Radek Bartyzal (GLAMI)

609. Weakly-supervised Fingerspelling Recognition in British Sign Language Videos Prajwal K R (VGG, Oxford),* Hannah Bull (LISN (CNRS)), Liliane Momeni (University of Oxford), Samuel Albanie (University of Cambridge), Gul Varol (Ecole des Ponts ParisTech), Andrew Zisserman (University of Oxford)

899. CASAPose: Class-Adaptive and Semantic-Aware Multi-Object Pose Estimation Niklas Gard (Fraunhofer HHI),* Anna Hilsmann (Fraunhofer HHI), Peter Eisert (Fraunhofer Heinrich Hertz Institute)

900. Revisiting Deep Fisher Vectors: Using Fisher Information to Improve Object Classification SARAH AHMED (Center of Artificial Intelligence in Health Sciences, ICCBS), Tayyaba Azim (University of Southampton),* Joseph Early (University of Southampton), Sarvapali Ramchurn (University of Southampton)

15:30 - 16:30

Keynote Speaker

Dima Damen, University of Bristol

Gold Sponsors

ROKE

vivo

woven planet

intel.

IET

(Chair:)

16:30 - 17:30

Orals Session 4: Image Restoration (Chair:)

439. Dual-Pixel Raindrop Removal Yizhou Li (Tokyo Institute of Technology),* Yusuke Monno (Tokyo Institute of Technology), Masatoshi Okutomi (Tokyo Institute of Technology)

974. Contrastive Learning for Controllable Blind Video Restoration Givi Meishvili (Computer Vision Group - Computer Science Department - University of Bern), Abdelaziz Djelouah (Disney Research),* Shinobu Hattori (Disney), Christopher Schroers (DisneyResearch|Studios)

288. Region-of-Interest Based Neural Video Compression Yura M. Perugachi-Diaz (Vrije Universiteit Amsterdam), Guillaume Sautiere (Qualcomm AI Research), Davide Abati (Qualcomm AI Research),* Yang Yang (Qualcomm Inc.), Amirhossein Habibian (Qualcomm AI Research), Taco S. Cohen (Qualcomm AI Research)

17:30-18:30

Drinks reception to start in the Ashes suite

18:30-23:45

Dinner in the England suite/Welcome Reception**Gold Sponsors****ROKE****vivo** **woven
planet****intel.****IET**

Wednesday 23rd November

8:00 **Registration Opens**

9:00-10:00 **Keynote Speaker**
Siyu Tang, ETH Zürich (ETH)
(Chair:)

10:00-11:15 **Spotlights 2 (Chair)**

305. Casual Indoor HDR Radiance Capture from Omnidirectional Images. Pulkit Gera (CVIT,IIT),* Mohammad Reza Karimi Dastjerdi (Université Laval), Charles Renaud (Université Laval), P. J. Narayanan (IIT-Hyderabad), Jean-Francois Lalonde (Laval University)

362. Multiple Object Tracking from appearance by hierarchically clustering tracklets. Andreu Girbau (National Institute of Informatics),* Ferran Marques (Universitat Politecnica de Catalunya), Shin'ichi Satoh (National Institute of Informatics)

377. iiTransformer: A Unified Approach to Exploiting Local and Non-local Information for Image Restoration. Soo Min Kang (Samsung Research),* YoungChan Song (Samsung Research), Hanul Shin (Samsung Research), Tammy Lee (Samsung Research)

395. Sparse in Space and Time: Audio-visual Synchronisation with Trainable Selectors. Vladimir Iashin (Tampere University),* Weidi Xie (Shanghai Jiao Tong University), Esa Rahtu (Tampere University), Andrew Zisserman (University of Oxford)

400. Instance Segmentation of Dense and Overlapping Objects via Layering. Long Chen (LFB, RWTH Aachen),* Yuli Wu (RWTH Aachen University), Dorit Merhof (RWTH Aachen University)

517. An Empirical Verification of Wide Networks Theory. Dario Balboni (Scuola Normale Superiore),* Davide Bacciu (Univeristy of Pisa)

542. RORD: A Real-world Object Removal Dataset. Min-Cheol Sagong (Korea Univ.),* Yoon-Jae Yeo (Korea Univiersity), Seung-Won Jung (Korea University), Sung-Jea Ko (Korea University)

Gold Sponsors



550. Multi-View Neural Surface Reconstruction with Structured Light. Chunyu Li (Preferred Networks, Inc.),* Taisuke Hashimoto (Preferred Networks, Inc.), Eiichi Matsumoto (Preferred Networks, Inc.), Hiroharu Kato (Preferred Networks)

554. Guess What Moves: Unsupervised Video and Image Segmentation by Anticipating Motion. Subhabrata Choudhury (University of Oxford),* Laurynas Karazija (University of Oxford), Iro Laina (University of Oxford), Andrea Vedaldi (Oxford University), Christian Rupprecht (University of Oxford)

603. HDR Reconstruction from Bracketed Exposures and Events. Richard Shaw (Huawei London Research Centre),* Sibi Catley-Chandar (Huawei Noah's Ark Lab), Ales Leonardis (Huawei Noah's Ark Lab), Eduardo Pérez Pellitero (Huawei Noah's Ark Lab)

654. Self-Improving SLAM in Dynamic Environments: Learning When to Mask. Adrian Bojko (CEA - Commissariat à l'Energie Atomique et aux Energies Alternatives),* Romain Dupont (DRT-LIST-DIASI-SIALV-LVML / CEA SACLAY), Mohamed Tamaazousti (CEA Saclay), Herve Le Borgne (CEA LIST)

768. Dual Space Multiple Instance Representative Learning for Medical Image Classification. Xiaoxian Zhang (Chongqing University), Sheng Huang (Chongqing University),* Yi Zhang (Chongqing University), Xiaohong Zhang (Chongqing University), Mingchen Gao (University at Buffalo, SUNY), Liu Chen (The First Affiliated Hospital of Army Medical University)

785. Self-Supervised Learning of Inlier Events for Event-based Optical Flow. Jun Nagata (Keio University),* Yoshimitsu Aoki (Keio University)

916. DA-CIL: Towards Domain Adaptive Class-Incremental 3D Object Detection. Ziyuan Zhao (Nanyang Technological University),* Mingxi Xu (Nanyang Technological University), Peisheng Qian (I2R, A*STAR), Ramanpreet Pahwa (I2R), richard chang (Institute for Infocomm Research)

1053. COAT: Correspondence-driven Object Appearance Transfer. Sangryul Jeon (UC Berkeley / ICSI),* Zhifei Zhang (Adobe Research), Zhe Lin (Adobe Research), Scott Cohen (Adobe Research), Zhihong Ding (Adobe Research), Kwanghoon Sohn (Yonsei Univ.)

Gold Sponsors



11:15-11:30

Tea Break

11:30-13:00

**Oral Session 5:
Recognition/Identification/Tracking/Detection (Part I)**

486. Cluster-level pseudo-labelling for source-free cross-domain facial expression recognition Alessandro Conti (University of Trento),* Paolo Rota (University of Trento), Yiming Wang (Fondazione Bruno Kessler), Elisa Ricci (University of Trento)

342. VID-Trans-ReID: Enhanced Video Transformers for Person Re-identification Aishah Alsehaim (Durham university),* Toby P Breckon (Durham University)

470. Trans2k: Unlocking the Power of Deep Models for Transparent Object Tracking Alan Lukezic (University of Ljubljana),* Žiga Trojer (University of Ljubljana), Jiri Matas (Czech Technical University, Prague), Matej Kristan (University of Ljubljana)

311. Track Targets by Dense Spatio-Temporal Position Encoding Jinkun Cao (Carnegie Mellon University),* Hao Wu (Bytedance Inc.), Kris Kitani (Carnegie Mellon University)

13:00-14:00

Lunch Break

14:00-15:30

Poster Session & Industrial Booth Exhibition (Chair)

305. Casual Indoor HDR Radiance Capture from Omnidirectional Images Pulkit Gera (CVIT,IIT),* Mohammad Reza Karimi Dastjerdi (Université Laval), Charles Renaud (Université Laval), P. J. Narayanan (IIT-Hyderabad), Jean-Francois Lalonde (Laval University)

356. An Action Is Worth Multiple Words: Handling Ambiguity in Action Recognition Kiyoon Kim (The University of Edinburgh),* Davide Moltisanti (University of Edinburgh), Oisín Mac Aodha (University of Edinburgh), Laura Sevilla-Lara (University of Edinburgh)

362. Multiple Object Tracking from appearance by hierarchically clustering tracklets Andreu Girbau (National Institute of Informatics),* Ferran Marques (Universitat Politècnica de Catalunya), Shin'ichi Satoh (National Institute of Informatics)

377. iiTransformer: A Unified Approach to Exploiting Local and Non-local Information for Image Restoration Soo Min Kang (Samsung Research),* YoungChan Song (Samsung Research), Hanul Shin (Samsung Research), Tammy Lee (Samsung Research)

400. Instance Segmentation of Dense and Overlapping Objects via Layering Long Chen (LfB, RWTH Aachen),* Yuli Wu (RWTH Aachen University), Dorit Merhof (RWTH Aachen University)

Gold Sponsors

485. Polycentric Clustering and Structural Regularization for Source-free Unsupervised Domain Adaptation Guan Xinyu (Nanjing University of Aeronautics and Astronautics), Han Sun (NUAA)*, Ningzhong Liu (Nanjing University of Aeronautics and Astronautics), Huiyu Zhou (University of Leicester)

517. An Empirical Verification of Wide Networks Theory Dario Balboni (Scuola Normale Superiore)*, Davide Bacciu (Univeristy of Pisa)

554. Guess What Moves: Unsupervised Video and Image Segmentation by Anticipating Motion Subhabrata Choudhury (University of Oxford)*, Laurynas Karazija (University of Oxford), Iro Laina (University of Oxford), Andrea Vedaldi (Oxford University), Christian Rupprecht (University of Oxford)

610. VL4Pose: Active Learning Through Out-Of-Distribution Detection For Pose Estimation Megh Shukla (EPFL)*, Pankaj Singh (Mercedes-Benz Research and Development India), Roshan Roy (Lockheed Martin), Shuaib Ahmed (Mercedes-Benz R&D India), Alexandre Alahi (EPFL)

611. Part-based Face Recognition with Vision Transformers Zhonglin Sun (Queen Mary University of London)*, Georgios Tzimiropoulos (Queen Mary University of London)

612. Layer Folding: Neural Network Depth Reduction using Activation Linearization Amir Ben Dror (Samsung)*, Niv Zehngut (Samsung), Avraham Raviv (SIRC), Evgeny Artyomov (Samsung), Ran Vitek (Samsung Israel R&D Center)

615. Wide Feature Projection with Fast and Memory-Economic Attention for Efficient Image Super-Resolutio Minghao Fu (University of Electronic Science and Technology of China), Dongyang Zhang (University of Electronic Science and Technology of China), Min Lei (University of Electronic Science and Technology of China), Kun He (University of Electronic Science and Technology of China), Changyu Li (University of Electronic Science and Technology of China), Jie Shao (University of Electronic Science and Technology of China)*

618. FoGMesh: 3D Human Mesh Recovery in Videos with Focal Transformer and GRU Yihao He (Jiangnan University)*, Xiaoning Song (Jiangnan University), Tianyang Xu (Jiangnan University), Hua Yang (Jiangnan University), Xiao-Jun Wu (Jiangnan University)

619. AssocFormer: Association Transformer on Multi-label Classification Xin Xing (University of Kentucky)*, Chong Peng (Qingdao University), Yu Zhang (University of Kentucky), Ai-Ling Lin (MU-Radiology), Nathan Jacobs (Washington University in St. Louis)

Gold Sponsors



622. Turbo Training with Token Dropout Tengda Han (University of Oxford),* Weidi Xie (Shanghai Jiao Tong University), Andrew Zisserman (University of Oxford)

630. FIND: An Unsupervised Implicit 3D Model of Articulated Human Feet Oliver S Boyne (University of Cambridge),* James Charles (Cambridge University), Roberto Cipolla (University of Cambridge)

632. D-STEP: Dynamic Spatio-Temporal Pruning Avraham Raviv (SIRC),* Yonatan Dinai (Samsung Israel R&D Center), Igor Drozdov (Samsung), Niv Zehngut (Samsung), Ishay Goldin (Samsung)

636. Hierarchical Residual Learning Based Vector Quantized Variational Autoencoder for Image Reconstruction and Generation Mohammad Adiban (NTNU),* Kalin Stefanov (Monash University), Sabato M Siniscalchi (Norwegian University of Science and Technology), Giampiero Salvi (NTNU)

639. Personalised CLIP or: how to find your vacation videos Bruno Korbar (University of Oxford),* Andrew Zisserman (University of Oxford)

644. Consistency-CAM: Towards Improved Weakly Supervised Semantic Segmentation Sai Rajeswar (University of Montreal),* Issam Hadj Laradji (ServiceNow), Pau Rodriguez (ServiceNow), David Vazquez (Element AI), Aaron Courville (MILA, Université de Montréal)

648. Scaling up Instance Segmentation using Approximately Localized Phrases Karan Desai (University of Michigan),* Ishan Misra (Facebook AI Research), Justin Johnson (University of Michigan), Laurens van der Maaten (Facebook)

649. Partially-Supervised Novel Object Captioning Using Context from Paired Data Shashank Bujimalla (Intel),* Mahesh Subedar (Intel), Omesh Tickoo (Intel)

654. Self-Improving SLAM in Dynamic Environments: Learning When to Mask Adrian Bojko (CEA - Commissariat à l'Energie Atomique et aux Energies Alternatives),* Romain Dupont (DRT-LIST-DIASI-SIALV-LVML / CEA SACLAY), Mohamed Tamaazousti (CEA Saclay), Herve Le Borgne (CEA LIST)

660. Two-Stream Transformer Architecture for Long Form Video Understanding Edward Fish (University of Surrey),* Jon Weinbren (University of Surrey), Andrew Gilbert (University of Surrey)

666. Attention Distillation: self-supervised vision transformer students need more guidance kai wang (Computer Vision Center),* Fei Yang (Computer Vision Center, UAB, Barcelona), Joost van de Weijer (Computer Vision Center)

Gold Sponsors



669. A Closer Look at Temporal Ordering in the Segmentation of Instructional Videos Anil Batra (University of Edinburgh),* Shreyank N Gowda (University of Edinburgh), Laura Sevilla-Lara (University of Edinburgh), Frank Keller (University of Edinburgh)

674. MoBYv2: Self-supervised Active Learning for Image Classification Razvan Caramalau (Imperial College),* Binod Bhattarai (University College London), Danail Stoyanov (UCL), Tae-Kyun Kim (Imperial College London)

679. Multi-task Transformation Learning for Robust Out-of-Distribution Detection Sina Mohseni (NVIDIA),* Arash Vahdat (NVIDIA), Jay Yadawa (NVIDIA)

682. Imagining Hidden Supporting Objects using Volumetric Conditional GANs and Differentiable Stability Scores Hector Basevi (University of Birmingham),* Ales Leonardis (University of Birmingham)

689. Towards Device Efficient Conditional Image Generation Nisarg A Shah (AI Foundation),* Gaurav Bharaj (AI Foundation)

693. Rethinking Group Fisher Pruning for Efficient Label-Free Network Compression Jong-Ryul Lee (ETRI),* Yong-Hyuk Moon (ETRI)

699. Improving Dense Representation Learning by Superpixelization and Contrasting Cluster Assignment Robin Karlsson (Nagoya University),* Tomoki Hayashi (Nagoya University), Keisuke Fujii (Nagoya University / RIKEN), Alexander Carballo (Nagoya University), Kento Ohtani (Nagoya University), Kazuya Takeda (Nagoya University)

703. LIIF-GAN: Learning Representation With Local Implicit Image Function and GAN for Realistic Images on a Continuous Scale Jun Seok Kang (University of Science and Technology), Sang Chul Ahn (KIST)*

705. Multi-task Curriculum Learning based on Gradient Similarity Hiroaki Igarashi (DENSO Corporation),* Kenichi Yoneji (DENSO), Kohta Ishikawa (Denso IT Laboratory, Inc.), Rei Kawakami (Tokyo Institute of Technology), Teppei Suzuki (Denso IT Laboratory), Shingo Yashima (Denso IT Laboratory), Ikuro Sato (Tokyo Institute of Technology / Denso IT Laboratory)

708. VoRF: Volumetric Relightable Face Pramod Rao (Max-Planck-Institut für Informatik),* Mallikarjun B R (Max Planck Institute for Informatics), Gereon Fox (Max Planck Institute for Informatics), Tim Weyrich (Friedrich-Alexander Universität Erlangen-Nürnberg), Bernd Bickel (IST Austria), Hanspeter Pfister

Gold Sponsors



(Harvard University), Wojciech Matusik (MIT), Ayush Tewari (MIT), Christian Theobalt (MPI Informatik), Mohamed Elgharib (Max Planck Institute for Informatics)

709. Probing Visual-Audio Representation for Video Highlight Detection via Hard-Pairs Guided Contrastive Learning Shuaicheng Li (SenseTime Research), Feng Zhang (Fudan University),* Kunlin Yang (SenseTime Group Limited), Lingbo Liu (Hong Kong Polytechnic University), Shinan Liu (SenseTime Group Limited), Jun Hou (SenseTime Group Limited), Shuai Yi (SenseTime Group Limited)

712. AISFormer: Amodal Instance Segmentation with Transformer Minh Q Tran (University of Arkansas),* Khoa HV Vo (University of Arkansas), Kashu Yamazaki (University of Arkansas), Arthur Fernandes (Cobb-Vantress / Tyson Foods, Inc), Michael T Kidd (University of Arkansas), Ngan Le (University of Arkansas)

715. One-shot Network Pruning at Initialization with Discriminative Image Patches Yinan Yang (Ritsumeikan University),* Ying Ji (Nagoya University), Yu Wang (Hitotsubashi University), Heng Qi (Dalian University of Technology), Jien Kato (Ritsumeikan University)

719. Spatio-temporal tendency reasoning for human body pose and shape estimation from videos Boyang Zhang (Ningxia University), Hu Cao (Ningxia University), Kehua Ma (Ningxia University), Pan Li (NingXia University), Lei Lin (Ningxia university), Suping Wu (Ningxia University)*

723. PPL: Pairwise Prototype Learning for Masked Face Recognition MINSOO KIM (KIST),* GIPYO NAM (KIST), Yu-Jin Hong (Hoseo university), Ig-Jae Kim (KIST)

725. Dual consistency assisted multi-confident learning for the hepatic vessel segmentation using noisy label Nam Phuong Nguyen (VinBrain),* Tuan Van Vo (VinBrain), Soan T. M. Duong (Le Quy Don Technical University), Chanh D Tr Nguyen (VinBrain), Trung Bui (Individual), QUOC HUNG TRUONG (VINBRAIN)

726. Memory-Driven Text-to-Image Generation Bowen Li (University of Oxford),* Philip Torr (University of Oxford), Thomas Lukasiewicz (University of Oxford)

728. Handling Class-Imbalance for Improved Zero-Shot Domain Generalization Ahmad Arfeen (IISc bangalore), Titir Dutta (Indian Institute of Science, Bangalore), Soma Biswas (Indian Institute of Science, Bangalore)*

729. Improving Local Features with Relevant Spatial Information by Vision Transformer for Crowd Counting Nguyen Hoang Tran (VinBrain),* Ta Duc Huy (Vinbrain), Soan T. M. Duong (Le Quy Don

Gold Sponsors



Technical University), Phan Nguyen (VinBrain), Dao Huu Hung (VinBrain), Chanh D Tr Nguyen (VinBrain), Trung Bui (Individual), QUOC HUNG TRUONG (VINBRAIN)

731. How to Train Vision Transformer on Small-scale Datasets?

Hanan Gani (Mohamed Bin Zayed University of Artificial Intelligence),* Muzammal Naseer (MBZUAI), Mohammad Yaqub (Mohamed Bin Zayed University of Artificial Intelligence)

736. Revisiting single-gated Mixtures of Experts

Amelie Royer (Qualcomm Research),* Iliia Karmanov (Qualcomm Research), Andrii Skliar (Qualcomm AI Research), Babak Ehteshami Bejnordi (Qualcomm AI Reseach), Tijmen Blankevoort (Qualcomm)

737. PAUMER: Patch Pausing Transformer for Semantic Segmentation

Evann Courdier (Idiap Research Institute),* Prabhu Teja Sivaprasad (Idiap Research Institute), François Fleuret (University of Geneva)

740. ViewNeRF: Unsupervised Viewpoint Estimation Using Category-Level Neural Radiance Fields

Octave Mariotti (University of Edinburgh),* Oisin Mac Aodha (University of Edinburgh), Hakan Bilen (University of Edinburgh)

743. Low Light Video Enhancement by Learning on Static Videos with Cross-Frame Attention

Shivam Chhriolya (Indian Institute of Science),* Sameer Malik (Indian Institute Of Science), Rajiv Soundararajan (Indian Institute of Science)

746. Explorable Data Consistent CT Reconstruction

Hannah Dröge (University of Siegen),* Yuval Bahat (Princeton University), Felix Heide (Princeton / Algolux), Michael Moeller (University of Siegen)

752. Siamese U-Net for Image Anomaly Detection and Segmentation with Contrastive Learning

Chia Ying Lin (National Tsing Hua University),* Shang-Hong Lai (National Tsing Hua University)

758. Face Pyramid Vision Transformer

Khawar Islam (floppydisk.ai),* Muhammad Zaigham Zaheer (UST), Arif Mahmood (Information Technology University)

763. Scale-Equivariant U-Net

Mateus Sangalli (Mines Paris Tech),* Samy Blusseau (Mines Paris Tech), Santiago Velasco-Forero (MINES ParisTech), Jesus Angulo (Mines Paris Tech)

768. Dual Space Multiple Instance Representative Learning for Medical Image Classification

Xiaoxian Zhang (Chongqing University), Sheng Huang (Chongqing University),* Yi Zhang (Chongqing University), Xiaohong Zhang (Chongqing University), Mingchen Gao (University at Buffalo, SUNY), Liu Chen (The First

Gold Sponsors



Affiliated Hospital of Army Medical University)

770. Parallel and Robust Text Rectifier for Scene Text Recognition

Bingcong Li (Ping An property&casualty insurance company of China.LTD.), xin tang (Ping An property&casualty insurance company of China.LTD.), Jun Wang (Ping An Technology (Shenzhen) Co. Ltd.), Liang Diao (Ping An property&casualty insurance company of China.LTD.), RUI FANG (Ping An Property & Casualty Insurance Company of China), Guotong Xie (Ping An Technology (Shenzhen) Co. Ltd.), Weifu Chen (Guangzhou Maritime University)*

772. Visual-Semantic Transformer for Scene Text Recognition

Liang Diao (Ping An property&casualty insurance company of China.LTD.), xin tang (Ping An property&casualty insurance company of China.LTD.), Jun Wang (Ping An Technology (Shenzhen) Co. Ltd.), RUI FANG (Ping An Property & Casualty Insurance Company of China), Guotong Xie (Ping An Technology (Shenzhen) Co. Ltd.), Weifu Chen (Guangzhou Maritime University)*

776. Anatomically constrained CT image translation for heterogeneous blood vessel segmentation

Giammarco La Barbera (Télécom Paris),* Haithem Boussaid (Philips Research Paris), Francesco Maso (Télécom Paris), Sabine Sarnacki (IMAG2, Imagine Institute), Rouet Laurence (Philips Research Paris), Pietro Gori (Télécom Paris), Isabelle Bloch (Télécom Paris)

778. Robust Target Training for Multi-Source Domain Adaptation

Zhongying Deng (University of Surrey),* Da Li (Samsung), Yi-Zhe Song (University of Surrey), Tao Xiang (University of Surrey)

779. Morphological Network: How Far Can We Go with Morphological Neurons?

Sanchayan Santra (Institute for Dataability Sciences, Osaka University),* Ranjan Mondal (Samsung Research, Bangalore), Soumendu Sundar Mukherjee (Indian Statistical Institute, Kolkata), Bhabatosh Chanda (Indian Statistical Institute)

782. XDGAN: Multi-Modal 3D Shape Generation in 2D Space

Hassan Abu Alhaija (NVIDIA),* Alara Dirik (Bogazici University), André Knörig (PCH Innovations), Sanja Fidler (University of Toronto, NVIDIA), Maria Shugrina (NVIDIA)

785. Self-Supervised Learning of Inlier Events for Event-based Optical Flow

Jun Nagata (Keio University),* Yoshimitsu Aoki (Keio University)

790. Universal Perturbation Attack on Differentiable No-Reference Image- and Video-Quality Metrics

Ekaterina Shumitskaya (Lomonosov Moscow State University),* Anastasia

Gold Sponsors



Antsiferova (Lomonosov Moscow State University), Dmitriy S Vatolin (Lomonosov Moscow State University)

797. Efficient Spatial Reasoning for Human Pose Estimation YING HUANG (Hangzhou Normal University),* Shanfeng Hu (Northumbria University), zike zhang (Hangzhou Normal University)

798. Knowledge Diversification in Ensembles of Identical Neural Networks Bishshoy Das (IIT Delhi),* Sumantra Dutta Roy (Indian Institute of Technology Delhi)

807. Re-examining Distillation for Continual Object Detection Eli Verwimp (KU Leuven),* Kuo Yang (Huawei Noah's Ark Lab), Sarah Parisot (Huawei Noah's Ark Lab), Lanqing Hong (Huawei Noah's Ark Lab), Steven McDonagh (Huawei Noah's Ark Lab), Eduardo Pérez Pellitero (Huawei Noah's Ark Lab), Matthias De Lange (KU Leuven), Tinne Tuytelaars (KU Leuven)

810. Search for Concepts: Learning Visual Concepts Using Direct Optimization Pradyumna Reddy (University College London),* Paul Guerrero (Adobe), Niloy Mitra (University College London)

817. AVisT: A Benchmark for Visual Object Tracking in Adverse Visibility Mubashir Noman (MBZUAI),* Wafa H Al Ghallabi (MBZUAI), Daniya Kareem (MBZUAI), Christoph Mayer (ETH Zurich), Akshay Dudhane (Mohamed bin Zayed University of Artificial Intelligence), Martin Danelljan (ETH Zurich), Hisham Cholakkal (MBZUAI), Salman Khan (MBZUAI/ANU), Luc Van Gool (ETH Zurich), Fahad Shahbaz Khan (MBZUAI)

821. $\mathcal{S}^2\text{-Flow}$: Joint Semantic and Style Editing of Facial Images

Krishnakant Singh (TU Darmstadt),* Simone Schaub-Meyer (TU Darmstadt), Stefan Roth (TU Darmstadt)

825. Efficient Feature Extraction for High-resolution Video Frame Interpolation Moritz Nottebaum (TU Darmstadt), Stefan Roth (TU Darmstadt), Simone Schaub-Meyer (TU Darmstadt)*

826. HiFECap: Monocular High-Fidelity and Expressive Capture of Human Performances Yue Jiang (Max Planck Institute for Informatics), Marc Habermann (Max Planck Institute for Informatics), Vladislav Golyanik (MPI for Informatics),* Christian Theobalt (MPI Informatik)

829. Pseudo-Label Denoising Techniques for Semi-Supervised Semantic Segmentation Sebastian Scherer (University of Augsburg),* Robin Schön (University of Augsburg), Rainer Lienhart ("Universität Augsburg, Germany")

Gold Sponsors



833. Analysis of Training Object Detection Models with Synthetic Data Bram Vanherle (Hasselt University),* Steven Moonen (Uhassey), Nick Michiels (Hasselt University), Frank Van Reeth (Hasselt University)

839. Unifying the Visual Perception of Humans and Machines on Fine-Grained Texture Similarity Weibo Wang (Ocean University of China), Xinghui Dong (Ocean University of China)*

843. A Cascade Dense Connection Fusion Network for Depth Completion Rizhao Fan (University of Bologna),* Zhigen Li (Ping An Technology), Matteo Poggi (University of Bologna), Stefano Mattocchia (University of Bologna)

844. Correlation between Alignment-and-Uniformity and Performance of Dense Contrastive Representations Jong Hak Moon (KAIST),* Wonjae Kim (NAVER AI Lab), Edward Choi (KAIST)

846. Wide-Range MRI Artifact Removal with Transformers Lennart Alexander Van der Goten (KTH Royal Institute of Technology),* Kevin Smith (KTH Royal Institute of Technology)

847. Progressive Multi-stage Interactive Training in Mobile Network for Fine-grained Classification Zhenxin Wu (Department of Computer Science, Jinan University),* Qingliang Chen (Jinan University), Yongjian Huang (Guangzhou Xuanyuan Research Institute Co., Ltd.)

848. Animal Pose Refinement in 2D Images with 3D Constraints Xiaowei Dai (Sichuan University),* Shuiwang Li (Guilin University of Technology), Qijun Zhao (Sichuan University), hongyu yang (sichuan university)

854. G-CMP: Graph-enhanced Contextual Matrix Profile for unsupervised anomaly detection in sensor-based remote health monitoring Nivedita Bijlani (University of Surrey),* Oscar Mendez (University of Surrey), Samaneh Kouchaki (University of Surrey)

861. Global Filter Pruning with Self-Attention for Real-Time UAV Tracking Mengyuan Liu (Guilin University of Technology), Yuelong Wang (Guilin University of Technology), Qiangyu Sun (Hubei Enshi College), Shuiwang Li (Guilin University of Technology)*

864. Self-adversarial Multi-scale Contrastive Learning for Semantic Segmentation of Thermal Facial Images Jitesh N Joshi (University College London),* Nadia Berthouze (University College London), Youngjun Cho (University College London)

868. Prior-Aware Synthetic Data to the Rescue: Animal Pose Estimation with Very Limited Real Data Le Jiang (Northeastern

Gold Sponsors



University), Shuangjun Liu (Northeastern University), Xiangyu Bai (Northeastern University), Sarah Ostadabbas (Northeastern University)*

869. Adaptive-TTA: accuracy-consistent weighted test time augmentation method for the uncertainty calibration of deep learning classifiers Pedro Conde (Institute of Systems and Robotics, University of Coimbra),* Cristiano Premebida (University of Coimbra)

872. Dual-Curriculum Teacher for Domain-Inconsistent Object Detection in Autonomous Driving Longhui Yu (Peking University),* Yifan Zhang (National University of Singapore), Lanqing Hong (Huawei Noah's Ark Lab), Fei Chen (Huawei Noah's Ark Lab), Zhenguo Li (Huawei Noah's Ark Lab)

876. Adaptive Task Sampling and Variance Reduction for Gradient-Based Meta-Learning Zhuoqun Liu (Shanghai Jiao Tong University),* Yuankun Jiang (Shanghai Jiao Tong University), Chenglin Li (Shanghai Jiao Tong University), Wenrui Dai (Shanghai Jiao Tong University), Junni Zou (Shanghai Jiao Tong University), Hongkai Xiong (Shanghai Jiao Tong University)

877. Anatomy-Aware Self-Supervised Learning for Aligned Multi-Modal Medical Data Hongyu Hu (Shanghai Jiao Tong University), Tiancheng Lin (Shanghai Jiao Tong University), Yuanfan Guo (SJTU), Chunxiao Li (Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine), Rong Wu (Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine), Yi Xu (Shanghai Jiao Tong University)*

879. Multi-Scale Adversarial Learning and Difficult Supervision for Kidney and Kidney Tumor Segmentation Shenhai Zheng (Chongqing University of Posts and Telecommunications),* Qiuyu Sun (Chongqing University of Posts and Telecommunications), 鑫叶 (xinye), Weisheng Li (Chongqing University of Posts and Telecommunications), Laquan Li (Chongqing University of posts and telecommunications)

880. Estimating water turbidity from a smartphone camera Lina M Lozano Wilches (Imperial College London),* Chotiawat Jantarakasem (Imperial College London), Laure Sioné (Imperial College London), Michael Templeton (Imperial College London), Krystian Mikolajczyk (Imperial College London)

883. Performance Limiting Factors of Deep Neural Networks for Pedestrian Detection in Automated Driving Systems Yasin Bayzidi (Volkswagen AG),* Alen Smajic (Volkswagen AG), Jan David Schneider (Volkswagen AG), Fabian Hüger (Volkswagen AG), Ruby Moritz (Volkswagen), Alois C. Knoll (Robotics and Embedded Systems)

Gold Sponsors



888. Learning to Augment via Implicit Differentiation for Domain Generalization Tingwei Wang (University of Surrey),* Da Li (Samsung), Kaiyang Zhou (Nanyang Technological University), Tao Xiang (University of Surrey), Yi-Zhe Song (University of Surrey)

892. Efficient Self-Ensemble for Semantic Segmentation Walid Bousselham (Oregon Health and Science University),* Guillaume Thibault (Oregon Health and Science University), Lucas Pagano (Oregon Health and Science University), Archana Machireddy (Oregon Health and Science University), Joe Gray (Oregon Health and Science University), Young Hwan Chang (Oregon Health and Science University), Xubo Song (Oregon Health and Science University)

893. Copy-Pasting Coherent Depth Regions Improves Contrastive Learning for Urban-Scene Segmentation Liang Zeng (Delft University of Technology), Attila Lengyel (Delft University of Technology), Nergis Tomen (Delft University of Technology), Jan C van Gemert (Delft University of Technology)*

916. DA-CIL: Towards Domain Adaptive Class-Incremental 3D Object Detection Ziyuan Zhao (Nanyang Technological University),* Mingxi Xu (Nanyang Technological University), Peisheng Qian (I2R, A*STAR), Ramanpreet Pahwa (I2R), richard chang (Institute for Infocomm Research)

919. Global Contextual Complementary Network for Multi-View Stereo Yongrong Cao (Ningxia University), Xing Zheng (Ningxia University), Bin Wang (Ningxia University),* Suping Wu (Ningxia University), Pan Li (NingXia University), Zhixiang Yuan (Ningxia University), Lei Lin (Ningxia university), Yuxin Peng (Ningxia University)

938. Towards more efficient few-shot learning based human gesture recognition via dynamic vision sensors Linglin Jing (Loughborough University), Yifan Wang (Loughborough University), Tailin Chen (Newcastle University), Shirin Dora (Loughborough University), ZHIGANG JI (Liverpool John Moores University), Hui Fang (Loughborough University)*

946. GLPose: Global-Local Attention Network with Feature Interpolation Regularization for Head Pose Estimation of People Wearing Facial Masks Hsueh-Wei Chen (National Taiwan University),* Yi Chen (National Taiwan University), Pei-Yung Hsiao (National University of Kaohsiung), Li-Chen Fu (National Taiwan University), ZI-RONG DING (The Automotive Research & Testing Center)

Gold Sponsors



958. Global Proxy-based Hard Mining for Visual Place Recognition Amar Ali-Bey (Université Laval),* Brahim Chaib-draa (Laval University), Philippe Giguere (Laval University)

960. BIO-CC: Biologically inspired color constancy Oguzhan Ulucan (Universität Greifswald),* Diclehan Ulucan (Universität Greifswald), Marc Ebner (Universität Greifswald)

965. Dual Moving Average Pseudo-Labeling for Source-Free Inductive Domain Adaptation Hao Yan (Carleton University),* Yuhong Guo (Carleton University)

969. Reading Chinese in Natural Scenes with a Bag-of-Radicals Prior Yongbin Liu (Beihang University), Qingjie Liu (State Key Laboratory of Virtual Reality Technology and System, Beihang University, Beijing 100191, China),* Jiaxin Chen (Beihang University), Yunhong Wang (State Key Laboratory of Virtual Reality Technology and System, Beihang University, Beijing 100191, China)

973. Distilling Knowledge from Self-Supervised Teacher by Embedding Graph Alignment Yuchen Ma (Heidelberg University), Yanbei Chen (University of Tübingen),* Zeynep Akata (University of Tübingen)

974. Contrastive Learning for Controllable Blind Video Restoration Givi Meishvili (Computer Vision Group - Computer Science Department - University of Bern), Abdelaziz Djelouah (Disney Research),* Shinobu Hattori (Disney), Christopher Schroers (DisneyResearch|Studios)

996. PatchSwap: A Regularization Technique for Vision Transformers Sachin Chhabra (Arizona State University),* Hemanth Venkateswara (Arizona State University), baixin Li (Arizona State University)

1002. Adversarial Vision Transformer for Medical Image Semantic Segmentation with Limited Annotations Ziyang Wang (University of Oxford),* Will Zhao (Bucknell University), Zixuan Ni (CU Boulder), Yuchen Zheng (University of North Carolina at Chapel Hill)

1007. Domain Adaptation for the Segmentation of Confidential Medical Images Serban Stan (University of Southern California),* Mohammad Rostami (University of Southern California)

1011. Overcoming Catastrophic Forgetting for Continual Learning via Feature Propagation Xuejun Han (Carleton University), Yuhong Guo (Carleton University)*

Gold Sponsors



1019. Group Graph Convolutional Networks for 3D Human Pose Estimation zijian z zhang (Beijing University of Posts and Telecommunications)*

1035. Hugs Are Better Than Handshakes: Unsupervised Cross-Modal Transformer Hashing with Multi-granularity Alignment Jinpeng Wang (Tsinghua University),* Ziyun Zeng (Tsinghua University), Bin Chen (Harbin Institute of Technology, Shenzhen), Yuting Wang (Tsinghua University), Dongliang Liao (Data Quality Team, WeChat, Tencent Inc., China), Gongfu Li (Tencent Inc.), Yiru Wang (Tencent Inc.), Shu-Tao Xia (Tsinghua University)

1049. Data Augmentation-free Unsupervised Learning for 3D Point Cloud Understanding Guofeng Mei (UTS),* Cristiano Saltori (University of Trento), Fabio Poiesi (Fondazione Bruno Kessler), Jian Zhang (UTS), Elisa Ricci (University of Trento), Nicu Sebe (University of Trento), Qiang Wu (University of Technology Sydney)

1052. Mutual Conditional Probability for Self-Supervised Learning Takumi Kobayashi (National Institute of Advanced Industrial Science and Technology)*

1053. COAT: Correspondence-driven Object Appearance Transfer Sangryul Jeon (UC Berkeley / ICSI),* Zhifei Zhang (Adobe Research), Zhe Lin (Adobe Research), Scott Cohen (Adobe Research), Zhihong Ding (Adobe Research), Kwanghoon Sohn (Yonsei Univ.)

1054. Anatomical prior-inspired label refinement for weakly supervised liver tumor segmentation with volume-level labels Fei Lyu (Department of Computer Science, Hong Kong Baptist University),* Andy J Ma (Sun Yat-sen University), PongChi Yuen (Department of Computer Science, Hong Kong Baptist University)

1062. Dense Contrastive Loss for Instance Segmentation Hang Chen (Tsinghua University),* Chufeng Tang (Tsinghua University), Xiaolin Hu (Tsinghua University)

1071. Mutual Contrastive Low-rank Learning to Disentangle Whole Slide Image Representations for Glioma Grading Lipei Zhang (University of Cambridge), Yiran Wei (University of Cambridge), Ying Fu (Beijing Institute of Technology), Stephen J. Price (University of Cambridge), Carola-Bibiane B Schönlieb (Cambridge University), Chao Li (University of Cambridge)*

15:00 - 15:30 **Tea Break**

15:30 - 16:45 **Oral Session 6:**

Gold Sponsors

ROKE

vivo

woven planet

intel.

IET

Recognition/Identification/Tracking/Dection (Part II)
(Chair:)**469. Object Tracking Network Based on Deformable Attention**

Mechanism Kexin Chen (Nanjing University of Posts and Telecommunications),* baojie fan (njupt), xiaobin Guo (Nanjing University of Posts and Telecommunications)

347. Resolving Semantic Confusions for Improved Zero-Shot Detection Sandipan Sarma (Indian Institute of Technology Guwahati),* SUSHIL KUMAR (Indian Institute of Technology Guwahati), Arijit Sur (IIT Guwahati)

381. PaRK-Detect: Towards Efficient Multi-Task Satellite Imagery Road Extraction via Patch-Wise Keypoints Detection Shenwei Xie (Beijing University of Posts and Telecommunications),* Wanfeng Zheng (Beijing University of Posts and Telecommunications), Zhenglin Xian (Beijing University of Posts and Telecommunications), Junli Yang (Beijing University of Posts and Telecommunications), Chuang Zhang (Beijing University of Posts and Telecommunications), Ming Wu (Beijing University of Posts and Telecommunications)

16:45-18:00

Oral Session 7: Image and Video Understanding**Chair:**

279. Rethinking the Evaluation of Unbiased Scene Graph Generation Xingchen Li (Zhejiang University),* Long Chen (Columbia University), Jian Shao (Zhejiang University), Shaoning Xiao (Zhejiang University), Songyang Zhang (University of Rochester), Jun Xiao (Zhejiang University)

1035. Hugs Are Better Than Handshakes: Unsupervised Cross-Modal Transformer Hashing with Multi-granularity Alignment Jinpeng Wang (Tsinghua University),* Ziyun Zeng (Tsinghua University), Bin Chen (Harbin Institute of Technology, Shenzhen), Yuting Wang (Tsinghua University), Dongliang Liao (Data Quality Team, WeChat, Tencent Inc., China), Gongfu Li (Tencent Inc.), Yiru Wang (Tencent Inc.), Shu-Tao Xia (Tsinghua University)

639. Personalised CLIP or: how to find your vacation videos Bruno Korbar (University of Oxford),* Andrew Zisserman (University of Oxford)

96. XCon: Learning with Experts for Fine-grained Category Discovery Yixin Fei (Tongji University), Zhongkai Zhao (Tongji University), Siwei Yang (Tongji University), Bingchen Zhao (University of Edinburgh)*

Gold Sponsors

Gold Sponsors

ROKE vivo



intel. **IET**

Thursday 24th November

9:00-10:00 **Keynote Speaker**
Pascal Fua, École Polytechnique Fédérale de Lausanne (EPFL)

10:00-11:00 **Oral Session 8: Domain generalisation/Domain adaptation**
Chair:

888. Learning to Augment via Implicit Differentiation for Domain Generalization Tingwei Wang (University of Surrey),* Da Li (Samsung), Kaiyang Zhou (Nanyang Technological University), Tao Xiang (University of Surrey), Yi-Zhe Song (University of Surrey)

778. Robust Target Training for Multi-Source Domain Adaptation Zhongying Deng (University of Surrey),* Da Li (Samsung), Yi-Zhe Song (University of Surrey), Tao Xiang (University of Surrey)

546. Adversarial Pixel Restoration as a Pretext Task for Transferable Perturbations Hashmat Shadab Malik (MBZUAI),* Shahina Kunhimon (Mohamed bin Zayed University of Artificial Intelligence), Muzammal Naseer (MBZUAI), Salman Khan (MBZUAI/ANU), Fahad Shahbaz Khan (MBZUAI)

9:30-12:00 **Workshop: Multimodal Video Search by Examples (MVSE) (running in parallel at India Room)**

10:00-12:00 **Workshop: Universal Representations for Computer Vision (URCV) (running in parallel at Jardine Suite)**

11:00-11:30 **Tea Break**

11:30-12:30 **Oral Session 9: Network Architecture /Learning strategy**

779. Morphological Network: How Far Can We Go with Morphological Neurons? Sanchayan Santra (Institute for Dataability Sciences, Osaka University),* Ranjan Mondal (Samsung Research, Bangalore), Soumendu Sundar Mukherjee (Indian Statistical Institute, Kolkata), Bhabatosh Chanda (Indian Statistical Institute)

371. BOAT: Bilateral Local Attention Vision Transformer. Tan Yu (Baidu Research),* Gangming Zhao (The University of Hong Kong),

Gold Sponsors



Ping Li (Baidu), Yizhou Yu (The University of Hong Kong)

127. Pose-graph via Adaptive Image Re-ordering Daniel Barath (ETH Zürich),* Jana Noskova (CMP CTU FEE), Iván Eichhardt (Eötvös Loránd University), Jiri Matas (Czech Technical University, Prague)

471. Multi-body Self-Calibration Andrea Porfiri Dal Cin (Politecnico di Milano),* Giacomo Boracchi (Politecnico di Milano), Luca Magri (Politecnico di Milano)

12:30-13:30

Lunch Break

13:00-15:00

Workshop: Multimodal Video Search by Examples
(running in parallel at India Room)

13:00-15:00

Workshop: Universal Representations for Computer Vision
(running in parallel at Jardine Suite)

15:00-15:30

Closing Ceremony

Gold Sponsors

ROKE

vivo

 **woven planet**

intel.

IET

Thursday 24st November: Workshop URCV

- 10:00-10:15 **Welcome to workshop**
- 10:15-11:00 **Invited Talk by Vittorio Ferrari**
- 11:00-11:15 **Poster session (Coffee break)**
- 11:15-12:00 **Invited Talk by Neil Houlsby**
- 12:00-13:00 **Lunch**
- 13:00-13:45 **Invited Talk by Dima Damen**
- 13:45-14:00 **Oral papers**
- 14:00-14:45 **Invited Talk by Timothy Hospedales**
- 14:45-14:55 **Prizes and Conclusion**

Gold Sponsors



Thursday 24st November: Workshop MVSE

- 09:00-09:30 **Welcome and MVSE Project**
- 09:50-10:30 **Search Keynote (Alan Smeaton)**
- 10:30-10:45 **MVSE incremental hashing (Wing and Xing)**
- 10:45-11:00 **IRM2R: Image retrieval method based on two models re-ranking**
- 11:00-11:15 **Tea break**
- 11:15-11:55 **Vision Keynote (Miroslaw Bober)**
- 11:55-12:10 **MixProp: Towards High-Performance Image Recognition via Dual Batch Normalisation**
- 12:10-12:50 **Audio Keynote (Doug Oard)**
- 12:50-13:30 **Lunch break**
- 13:30-13:45 **Scene paper**
- 13:45-14:00 **Multimodal learning and representations (Nishant Rai)**
- 14:00-14:15 **MVSE Demo (Ivor Spence and Guanfeng Wu)**
- 14:15-14:30 **MVSE Interface design (Raymond Bond and Maurice Mulvenne)**
- 14:30-15:00 **Video retrieval assessment “in the wild”**

Gold Sponsors



Gold Sponsors

ROKE vivo

 woven
planet

intel. **IET**