

October 20–24, 2021
Virtual Event, China



Association for
Computing Machinery

Advancing Computing as a Science & Profession



MM '21

Proceedings of the 29th ACM International Conference on
Multimedia

Sponsored by:

ACM SIGMM

General Chairs:

Heng Tao Shen, Yueting Zhuang & John R. Smith

Program Chairs:

Yang Yang, Pablo Cesar, Florian Metze & Balakrishnan Prabhakaran

Proceedings Chairs:

Liqiang Nie, Qianru Sun & Peng Cui



**Association for
Computing Machinery**

Advancing Computing as a Science & Profession

**The Association for Computing Machinery
1601 Broadway, 10th Floor
New York, NY 10019-7434**

Copyright © 2021 by the Association for Computing Machinery, Inc. (ACM). Permission to make digital or hard copies of portions of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permission to republish from: permissions@acm.org or Fax +1 (212) 869-0481.

For other copying of articles that carry a code at the bottom of the first or last page, copying is permitted provided that the per-copy fee indicated in the code is paid through www.copyright.com.

ISBN: 978-1-4503-8651-7

Additional copies may be ordered prepaid from:

ACM Order Department
PO Box 30777
New York, NY 10087-0777, USA

Phone: 1-800-342-6626 (USA and Canada)
+1-212-626-0500 (Global)
Fax: +1-212-944-1318
E-mail: acmhelp@acm.org
Hours of Operation: 8:30 am – 4:30 pm ET

Printed in the USA

Welcome Message from the General Chairs

Welcome all of you to join ACM Multimedia 2021, the 29th ACM International Conference on Multimedia. ACM Multimedia is the premier international conference series in the area of multimedia within the field of computer science. Since 1993, ACM Multimedia has been bringing together worldwide researchers and practitioners from academia and industry to present their innovative research and to discuss recent advancements in multimedia.

Due to the continuous influence of COVID-19 all over the world since the end of 2019, the traditional onsite international conferences have to move to online mode. Notably, the last year ACM Multimedia conference 2020, has been successfully held totally online for the first time in history. During the online conference in the last year, face-to-face gatherings had been replaced by all-hands live sessions, parallel live question-answering (Q&A) sessions, and presentations from the authors via pre-recorded videos. This would be a reasonable solution to hold such a large-scale meeting during the COVID-19 period. Fortunately, the good epidemic prevention effect in China allows us to consider a more appropriate hybrid form to hold the conference onsite and online jointly. Our vision is to enrich the interaction of attendees all over the world via both physically face-to-face and virtually online fashions.

The hybrid conference this year indeed brings greater challenges for organization. Firstly, we need to choose an appropriate conference venue that can accommodate hundreds of attendees. We finally find a good place in the downtown of the Chengdu city, where the attendees will definitely have a pleasant conference trip. Secondly, unlike the presentation form of papers on the purely online conference last year, all the conference papers this year are expected to be presented more informatively to both onsite and online attendees. Therefore, we adopt the traditional onsite conference form and divide the papers to oral and poster sessions with live streaming in each conference room, encouraging the onsite attendees to present their work physically and online attendees to provide live talks. Besides, we will also build a web portable platform (both Apps and websites) to collect the pre-recorded videos and posters of all papers, which will facilitate all attendees to view these resources before and during the conference. Moreover, free poster printing service will be available during the conference to enable the attendees onsite to view the printed posters of the authors who cannot physically attend the conference, which is expected to further enrich their exchange on the research work. Besides, we will also build remote interaction with authors in Q&A sessions via Zoom meetings. We really expect both onsite and online attendees to take advantage of the onsite oral/poster session and Q&A sessions after studying as many as possible related papers to interact directly with the paper authors.

During the paper submission period of this year conference, we have tackled various unprecedented challenges and extended paper submission deadlines to provide sufficient time for participated teams. Nonetheless, we still have received a record-breaking number of 1,942 valid submissions for the main conference tracks (1,698 for last year) with an overall acceptance rate of 27.9% (542 papers), covering all areas and subareas of multimedia. This year, the conference will also be organized with a large diversity of tracks. In addition to the main track, the conference has tracks devoted to brave new ideas, industrial track, reproducibility companion, multimedia grand challenge, open source competition, interactive arts, and to multimedia demos.

This year we are honored to welcome six distinguished keynote speakers from both academia and industry. Through the 3-day main conference period, two keynotes will be assigned in each day. In the first main conference day, the opening keynote will be delivered by Prof. Wen Gao, who is currently a Boya Chair Professor at Peking university, and the Director of Pengcheng Laboratory, China. He is an academician of Chinese Academy of Engineering, a fellow of ACM and IEEE. In his talk, Prof. Wen Gao will introduce the history of video coding for machining (VCM), list some early works on pattern analysis based on compressed data domain, some efforts from ISO/IEC MPEG group on MPEG-7 CDVS (compact descriptor for visual search) and CDVA (compact descriptors for visual analysis), some ongoing projects on AVS working group and MPEG working group, give the key techniques and challenges on VCM, and overview its future. The second keynote will be delivered by Professor H. V. Jagadish, is Bernard A Galler Collegiate Professor at the University of Michigan in Ann Arbor, and Director of the Michigan Institute for Data Science. He is a fellow of ACM and AAAS. With recent amazing progress in machine intelligence, it is becoming increasingly easy to “convert” information reliably from one medium to another. In his talk, Professor H. V. Jagadish will outline some possibilities with high quality semantic media conversion. Multimedia data, so converted, becomes easy to use, to aggregate, and to analyze, leading to new Data Science opportunities. The third and the fourth keynote speeches will be held in the second main conference day, at the beginning of the afternoon session. The third keynote will be given by Professor Cordelia Schmid, who is a research director at INRIA Rhone-Alpes and holds a joint appointment with Google research. She is a fellow of IEEE. In her talk, Professor Cordelia Schmid will present recent progress on large-scale learning of multimodal video representations. The talk will cover the VideoBert, a joint model for video and language, repurposing the Bert model for multimodal data, as well as a recent Video Vision Transformer (ViViT) which is fully transformer based. The fourth keynote will be delivered by Dr. Jingren Zhou, Senior Vice President at Ant Group. He is a fellow of IEEE and responsible for driving data intelligence infrastructure and several key data-driven businesses at Ant Group. In his talk, Dr. Jingren Zhou will present their experiences and applications of large-scale multi-modality pretrained models such as M6 (Multi-Modality to Multi-Modality Multitask Mega-transformer), developed at Alibaba and Ant Group. Finally, the last two keynotes will be performed on the third main conference day at the beginning of the morning session. The fifth keynote will be provided by Professor James Lester, Distinguished University Professor of Computer Science and Director of the Center for Educational Informatics at North Carolina State University. He is a fellow of AAAI and the Director of the National Science Foundation AI Institute for Engaged Learning. In his talk, Professor James Lester will consider how (and at what rate) AI technologies for education will evolve, discuss emerging innovations in AI-augmented learning environments for formal and informal contexts, and explore what competencies will be elevated in an AI-pervasive workforce. The last keynote will be given by Dr. Zhengyou Zhang, the Chief Scientist at Tencent, China, and is the Director of Tencent AI Lab and Tencent Robotics X. Dr. Zhang is a fellow of ACM and IEEE. With the rapid development of digital technologies such as VR, AR, XR, and more importantly the almost ubiquitous mobile broadband coverage, we are entering an Integrated Physical-Digital World (IPhD), the tight integration of virtual world with the physical world. In his talk, Dr. Zhengyou Zhang will describe their work on digital human for this IPhD world, including computer vision techniques for building digital humans, multimodal text-to-speech synthesis (voice and lip shapes), speech-driven face animation, neural-network-based body motion control, human-digital-human interaction, and an emotional video game anchor.

This year, we are also happy to have two panel sessions on the first and third main conference day at the afternoon session, respectively. The topic of the first panel session will be “The Next Generation of Multimedia Search and Recommendation”, which will be hosted by Professor Tat-Seng Chua, who is the KITHCT Chair Professor at the School of Computing, National University of Singapore. The theme of the second panel session will be “Social Signals and Multimedia: Past, Present, Future”, which will be hosted by Professor Hung. Let us brainstorm these in this panel.

This year, ACM Multimedia's great tradition of Multimedia Grand Challenges (MGCs) continues. In particular, we have organized nine MGCs, selected from dozens of proposals based on the strength of the proposals, their relevance to the community, significance of tasks and diversity of organizers. The continuity of the challenges has been also emphasized, in order to support sustained and substantial progress in the state of the art. The accepted MGCs will provide a timely reflection of the cutting-edge research and industrial applications of the multimedia computing technology. The tasks range from low-level multimedia data transmission, middle-level object detection and tracking, and high-level multi-modal video understanding and group behavior analysis, with organizers teamed up by top universities, research institutes and Internet giants worldwide. Furthermore, on the tutorial and workshop days before and after the 3-day main conference, eight tutorials provide the opportunity to learn about new directions in multimedia research. Besides, we will have thirteen workshops which cover a wide range of cutting-edge topics for the field of multimedia.

Like every year in the past, we would like to give our sincere thanks to everyone who have contributed to ACM Multimedia 2021. We are grateful to this year's outstanding organizing committee consisting of devoted volunteers, who have worked extremely hard to build such an impressive program. We also appreciate the Area Chairs for their hard work in guaranteeing the novelty and scientific rigor of the main-conference program. We would also like to give our special thanks to the committed reviewers for their great work in reviewing submissions across all tracks of the conference. We are also grateful for both the onsite and online conference “local chairs” and the finance chair, who have resolved numerous unexpected difficulties in making our conference program available to all attendees. We also greatly appreciate the support from all the sponsors from both academia and industry. Finally, we would like to thank all the authors who have submitted to ACM Multimedia 2021. The heart of the conference is your research discoveries, which is what drives our field forward.

We close by wishing all conference participants a dream onsite and online journey experience in the 5-day series of conference sessions. It will be a great opportunity for all attendees to meet together via the conference venue in the beautiful city Chengdu, China as well as the online platform, to witness the latest progress of multimedia areas. Enjoy a wonderful hybrid conference ACM Multimedia 2021!

Heng Tao Shen

*ACM Multimedia 2021
General Chair
University of Electronic Science
and Technology of China, China*

Yueting Zhuang

*ACM Multimedia 2021
General Chair
Zhejiang University, China*

John R. Smith

*ACM Multimedia 2021
General Chair
IBM, USA*

Message from the Technical Program Chairs

Welcome to the 29th ACM International Conference on Multimedia! We would like to take this opportunity to provide an overview of the exciting program of this year, the review process and to explain how the program has been crafted.

This year we received a record-breaking number of submissions: 2544. Two rounds of desk rejects resulted in 1942 valid submission, ready for the review process, which covered the different areas and subareas of the conference. From creation and production of multimedia content to content recommendation, from content delivery and transmission to user experience, and from multimedia content understanding to better engaging users with content. The conference was highly competitive this year, with 542 accepted (~27.9% acceptance rate) papers, representing all areas and subareas of multimedia. The paper review was conducted in a double-blind manner, with several aspects aimed to achieve a very rigorous review process, including an author rebuttal, review discussion, meta-reviews, and within-area and cross-area discussions.

Following the tradition of the previous years, papers were submitted to ACM Multimedia 2021 under the four major themes of multimedia, Engagement, Experience, Systems and Understanding. Each theme was further subdivided into three or more areas. The Engagement theme focuses on research about how users engage with multimedia material: What emotions or social signals are conveyed or induced? How to search and browse efficiently? How can we gain insight from large multimedia collections? The Experience theme focuses on the user experience when engaging with multimedia material: What interaction and quality of experience? How can we better understand art and cultural material? What applications can be designed and how? The System theme is devoted to the mechanisms and frameworks for building large-scale multimedia applications: How can we improve system components? How can we transport and deliver multimedia content efficiently? What mechanisms for data management and indexing? Finally, the Understanding theme is about machines analyzing and interpreting content with an emphasis on a multimodal perspective: How can we combine multiple modalities? How do vision and language interact? How to process different modalities for a new interpretation?

For each area, a number of Area Chairs (219) were nominated. The program chairs recruited 167 Area Chairs and other 52 (out of 175) were selected via an open call, trying to increase diversity and inclusiveness. Each paper was assigned to an Area Chair, who would steer and monitor the review process, acting as a meta-reviewer. The Area Chair controlled the review quality, initiated and coordinated the discussion among reviewers, and proposed a recommendation. The final decision was taken during the Technical Program committee meeting (TPC), held online (around 47 hours) via videoconference on 25th – 27th June 2021. We would like here to heartily thank the Area Chairs for their invaluable contribution in striving for a fair review process.

In total, we received 2544 submissions, from which 602 were rejected following two rounds of desk rejects. The first one, done by the program chairs, focused on the submission details; and the second one, coordinated by the Area Chairs, focused on the format and content of the article. We adhered strictly to the definition of the multimedia field that we adopted in the call and in the reviewer guidelines. Each valid submission was assigned to reviewers and ACs with the help of the Toronto Paper Matching System (TMS). The 1942 valid submissions were assigned to at least 3 reviewers under the supervision of one area and were given the opportunity of rebuttal to enable review adjustment and discussion among reviewers. Double-blind reviewing policies

were strictly maintained throughout the entire process. We did our utmost to ensure that ACs and reviewers did not know authors' identities, even though anonymity becomes difficult to maintain as more and more submissions appear concurrently on arXiv.org. The review process included a total of 1233 reviewers. We take this opportunity to thank them all for their commitment to the community and for their constructive comments.

As in previous years, the program (542 papers) is divided into oral (179 papers) and poster (363 posters) presentations. It includes the final selection of best paper candidates as well. A long list of initial best paper candidates was proposed by the area chairs on a per area basis and cross-checked against nominations made by reviewers. The final selection was made by the TPC chairs, in consultation with the general chairs and the best paper committee chair. All the accepted papers are excellent examples of multimedia research according to our conference scoping guidelines.

We wish to thank everyone involved for their time and dedication to making the ACM Multimedia 2021 program possible. The success of the conference entirely relies on the time and effort invested by the authors into producing high-quality research, on the care taken by the reviewers in writing thorough and professional reviews, and on the commitment by the area chairs to reconciling the reviews and writing detailed consolidation reports. We also wish to thank the general chairs and the other organizing committee members for their efforts into the organization of the event.

Pablo Cesar

*ACM Multimedia 2021
Technical Program Chair
CWI & TU Delft, The Netherlands*

Florian Metze

*ACM Multimedia 2021
Technical Program Chair
Facebook, Inc., USA*

Balakrishnan Prabhakaran

*ACM Multimedia 2021
Technical Program Chair
University of Texas at Dallas, USA*

Yang Yang

*ACM Multimedia 2021
Technical Program Chair
University of Electronic Science and
Technology of China, China*

Table of Contents

ACM Multimedia 2021 Organization	lvii
ACM Multimedia 2021 Area Chairs	lviii
ACM Multimedia 2021 Program Committee	lxv
ACM Multimedia 2021 Sponsor & Supporters	lxiii

Keynote Talks I & II

• Video Coding for Machine	1
<i>Weno Gao (Peking University & Peng-Cheng Laboratory)</i>	
• Semantic Media Conversion: Possibilities and Limits	2
<i>H. V. Jagadish (University of Michigan)</i>	

Session 1: Deep Learning for Multimedia-I

• Image Re-composition via Regional Content-Style Decoupling	3
<i>Rong Zhang (Zhejiang University), Wei Li (Inceptio), Yiqun Zhang (Zhejiang University), Hong Zhang (SenseTime Group Ltd.), Jinhui Yu (Zhejiang University), Ruigang Yang (Inceptio), Weiwei Xu (Zhejiang University)</i>	
• Deep Clustering based on Bi-Space Association Learning	12
<i>Hao Huang (GE Research), Shinjae Yoo (Brookhaven National Laboratory), Chenxiao Xu (Stony Brook University)</i>	
• Feature Stylization and Domain-aware Contrastive Learning for Domain Generalization	22
<i>Seogkyu Jeon, Kibeom Hong, Pilhyeon Lee, Jewook Lee, Hyeran Byun (Yonsei University)</i>	
• HDA-Net: Horizontal Deformable Attention Network for Stereo Matching	32
<i>Qi Zhang, Xuesong Zhang, Baoping Li, Yuzhong Chen, Anlong Ming (Beijing University of Posts and Telecommunications)</i>	
• MBRS: Enhancing Robustness of DNN-based Watermarking by Mini-Batch of Real and Simulated JPEG Compression	41
<i>Zhaoyang Jia, Han Fang, Weiming Zhang (University of Science and Technology of China)</i>	
• From Synthetic to Real: Image Dehazing Collaborating with Unlabeled Real Data	50
<i>Ye Liu (Tianjin University), Lei Zhu (University of Cambridge), Shunda Pei (Tianjin University), Huazhu Fu (Inception Institute of Artificial Intelligence), Jing Qin (Hong Kong Polytechnic University), Qing Zhang (Sun Yat-sen University), Liang Wan, Wei Feng (Tianjin University)</i>	

Session 2: Deep Learning for Multimedia-II

• Video Semantic Segmentation with Sparse Temporal Transformer	59
<i>Jiangtong Li, Wentao Wang, Junjie Chen, Li Niu (Shanghai Jiao Tong University), Jianlou Si, Chen Qian (SenseTime Research), Liqing Zhang (Shanghai Jiao Tong University)</i>	
• Diverse Image Inpainting with Bidirectional and Autoregressive Transformers	69
<i>Yingchen Yu (Nanyang Technological University & Alibaba Group), Fangneng Zhan, Rongliang Wu (Nanyang Technological University), Jianxiong Pan (DAMO Academy, Alibaba Group), Kaiwen Cui, Shijian Lu (Nanyang Technological University), Feiyang Ma, Xuansong Xie (DAMO Academy, Alibaba Group), Chunyan Miao (Nanyang Technological University)</i>	
• SSFlow: Style-guided Neural Spline Flows for Face Image Manipulation	79
<i>Hanbang Liang, Xianxu Hou, Linlin Shen (Shenzhen University & Shenzhen Institute of Artificial Intelligence of Robotics of Society)</i>	

- **Constrained Graphic Layout Generation via Latent Optimization** 88
Kotaro Kikuchi, Edgar Simo-Serra (*Waseda University*), Mayu Otani, Kota Yamaguchi (*CyberAgent*)
- **Transfer Vision Patterns for Multi-Task Pixel Learning**..... 97
Xiaoya Zhang, Ling Zhou, Yong Li, Zhen Cui, Jin Xie, Jian Yang (*Nanjing University of Science and Technology*)
- **Object-aware Long-short-range Spatial Alignment for Few-Shot Fine-Grained Image Classification** 107
Yike Wu, Bo Zhang (*Fudan University*), Gang Yu (*Tencent*),
Weixi Zhang, Bin Wang, Tao Chen, Jiayuan Fan (*Fudan University*)

Session 3: Brave New Idea

- **Recycling Discriminator: Towards Opinion-Unaware Image Quality Assessment Using Wasserstein GAN**..... 116
Yunan Zhu, Haichuan Ma, Jialun Peng, Dong Liu, Zhiwei Xiong (*University of Science and Technology of China*)
- **Learning Kinematic Formulas from Multiple View Videos**..... 126
Liangchen Song, Sheng Liu (*University at Buffalo*), Celong Liu (*InnoPeak Technology Inc.*),
Zhong Li (*InnoPeak Technology Inc.*), Yuqi Ding (*Louisiana State University*),
Yi Xu (*InnoPeak Technology Inc.*), Junsong Yuan (*University at Buffalo*)
- **DEPA: Self-Supervised Audio Embedding for Depression Detection** 135
Pingyue Zhang, Mengyue Wu, Heinrich Dinkel, Kai Yu (*Shanghai Jiao Tong University*)
- **Retinomorphic Sensing: A Novel Paradigm for Future Multimedia Computing** 144
Zhaodong Kang, Jianing Li, Lin Zhu, Yonghong Tian (*Peking University*)
- **Metaverse for Social Good: A University Campus Prototype**..... 153
Haihan Duan, Jiaye Li, Sizheng Fan, Zhonghao Lin (*The Chinese University of Hong Kong & Shenzhen Institute of Artificial Intelligence and Robotics for Society*),
Xiao Wu (*White Matrix Inc.*),
Wei Cai (*The Chinese University of Hong Kong & Shenzhen Institute of Artificial Intelligence and Robotics for Society*)

Session 4: Deep Learning for Multimedia-III

- **Enhanced Invertible Encoding for Learned Image Compression** 162
Yueqi Xie, Ka Leong Cheng, Qifeng Chen (*Hong Kong University of Science and Technology*)
- **DC-GNet: Deep Mesh Relation Capturing Graph Convolution Network for 3D Human Shape Reconstruction**..... 171
Shihao Zhou, Mengxi Jiang, Shanshan Cai, Yunqi Lei (*Xiamen University*)
- **Deep Marginal Fisher Analysis based CNN for Image Representation and Classification** 181
Xun Cai, Jiajing Chai, Yanbo Gao, Shuai Li, Bo Zhu (*Shandong University*)
- **Learning Structure Affinity for Video Depth Estimation**..... 190
Yuanzhouhan Cao, Yidong Li (*Beijing Jiaotong University*),
Haokui Zhang (*Northwestern Polytechnical University*), Chao Ren (*Sichuan University*),
Yifan Liu (*The University of Adelaide*)
- **X-GGM: Graph Generative Modeling for Out-of-distribution Generalization in Visual Question Answering**..... 199
Jingjing Jiang, Ziyi Liu, Yifan Liu, Zhixiong Nan, Nanning Zheng (*Xi'an Jiaotong University*)
- **DSSL: Deep Surroundings-person Separation Learning for Text-based Person Retrieval** 209
Aichun Zhu, Zijie Wang, Yifeng Li, Xili Wan, Jing Jin (*Nanjing Tech University*),
Tian Wang (*Beihang University*), Fangqiang Hu (*Nanjing Tech University*),
Gang Hua (*China University of Mining and Technology*)

Session 5: Emerging Multimedia Applications-I

- **Diverse Multimedia Layout Generation with Multi Choice Learning** 218
David D. Nguyen (*University of New South Wales*), Surya Nepal (*CSIRO, Data61*),
Salil S. Kanhere (*University of New South Wales Sydney*)

- **Viewing from Frequency Domain: A DCT-based Information Enhancement Network for Video Person Re-Identification** 227
Liangchen Liu, Xi Yang, Nannan Wang (*Xidian University*),
Xinbo Gao (*Chongqing University of Posts and Telecommunications*)
- **Unsupervised Portrait Shadow Removal via Generative Priors** 236
Yingqing He, Yazhou Xing, Tianjia Zhang, Qifeng Chen (*Hong Kong University of Science and Technology*)
- **Multimodal Global Relation Knowledge Distillation for Egocentric Action Anticipation**..... 245
Yi Huang (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Xiaoshan Yang (*Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; & Peng Cheng Lab*),
Changsheng Xu (*Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Peng Cheng Lab; & CASIA-LLVision Joint Lab*)
- **Exploring Pathologist Knowledge for Automatic Assessment of Breast Cancer Metastases in Whole-slide Image** 255
Liuwan Wang, Li Sun, Mingjie Zhang, Huigang Zhang, Wang Ping, Rong Zhou,
Jun Sun (*Fujitsu Research & Development Center CO., LTD*)
- **Towards Multiple Black-boxes Attack via Adversarial Example Generation Network** 264
Duan Mingxing, Kenli Li (*Hunan University*), Lingxi Xie, Qi Tian (*Huawei Inc*),
Bin Xiao (*Hong Kong Polytechnic University*)

Session 6: Emerging Multimedia Applications-II

- **DocTr: Document Image Transformer for Geometric Unwarping and Illumination Correction** 273
Hao Feng, Yuechen Wang (*University of Science and Technology of China*),
Wengang Zhou (*University of Science and Technology of China & Hefei Comprehensive National Science Center*),
Jiajun Deng (*University of Science and Technology of China*),
Houqiang Li (*University of Science and Technology of China & Hefei Comprehensive National Science Center*)
- **Self-supervised Multi-view Multi-Human Association and Tracking** 282
Yiyang Gan, Ruize Han, Liqiang Yin, Wei Feng (*Tianjin University*),
Song Wang (*University of South Carolina*)
- **Learning Fine-Grained Motion Embedding for Landscape Animation** 291
Hongwei Xue (*University of Science and Technology of China*),
Bei Liu, Huan Yang, Jianlong Fu (*Microsoft Research*),
Houqiang Li (*University of Science and Technology of China*), Jiebo Luo (*University of Rochester*)
- **Multi-label Pattern Image Retrieval via Attention Mechanism Driven Graph Convolutional Network** 300
Ying Li, Hongwei Zhou, Yeyu Yin, Jiaquan Gao (*Nanjing Normal University*)
- **Collocation and Try-on Network: Whether an Outfit is Compatible**..... 309
Na Zheng, Xuemeng Song, Qingying Niu, Xue Dong (*Shandong University*),
Yibing Zhan (*JD Explore Academy*), Liqiang Nie (*Shandong University*)
- **MeronymNet: A Hierarchical Model for Unified and Controllable Multi-Category Object Generation** 318
Rishabh Baghel, Abhishek Trivedi, Tejas Ravichandran, Ravi Kiran Sarvadevabhatla
(*International Institute of Information Technology, Hyderabad*)

Session 7: Emerging Multimedia Applications-III

- **Ada-VSR: Adaptive Video Super-Resolution with Meta-Learning** 327
Akash Gupta, Padmaja Jonnalagedda, Bir Bhanu, Amit K. Roy-Chowdhury (*University of California, Riverside*)
- **CoReD: Generalizing Fake Media Detection with Continual Representation using Distillation** 337
Minha Kim, Shahroz Tariq, Simon S. Woo (*Sungkyunkwan University*)
- **SRNet: Spatial Relation Network for Efficient Single-stage Instance Segmentation in Videos** 347
Xiaowen Ying, Xin Li, Mooi Choo Chuah (*Lehigh University*)

- **Personality Recognition by Modelling Person-specific Cognitive Processes using Graph Representation** 357
Zilong Shao (*Shenzhen University & Shenzhen Institute of Artificial Intelligence of Robotics of Society*), Siyang Song (*University of Cambridge*), Shashank Jaiswal (*University of Nottingham*), Linlin Shen (*Shenzhen University & Shenzhen Institute of Artificial Intelligence of Robotics of Society*), Michel Valstar (*University of Nottingham*), Hatice Gunes (*University of Cambridge*)
- **Enhancing Knowledge Tracing via Adversarial Training** 367
Xiaopeng Guo, Zhijie Huang, Jie Gao, Mingyu Shang, Maojing Shu, Jun Sun (*Peking University*)
- **Beyond OCR + VQA: Involving OCR into the Flow for Robust and Accurate TextVQA** 376
Gangyan Zeng, Yuan Zhang (*Communication University of China*), Yu Zhou (*Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Xiaomeng Yang (*Institute of Information Engineering, Chinese Academy of Sciences*)

Poster Session 1

- **JPGNet: Joint Predictive Filtering and Generative Network for Image Inpainting** 386
Qing Guo (*Tianjin University & Nanyang Technological University*), Xiaoguang Li (*University of South Carolina*), Felix Juefei-Xu (*Alibaba Group*), Hongkai Yu (*Cleveland State University*), Yang Liu (*Nanyang Technological University & Zhejiang Sci-Tech University*), Song Wang (*University of South Carolina*)
- **AdvFilter: Predictive Perturbation-aware Filtering against Adversarial Attack via Multi-domain Learning** 395
Yihao Huang (*East China Normal University*), Qing Guo (*Nanyang Technological University*), Felix Juefei-Xu (*Alibaba Group, USA*), Lei Ma (*University of Alberta*), Weikai Miao (*East China Normal University*), Yang Liu (*Nanyang Technology University*), Geguang Pu (*East China Normal University*)
- **Pixel-level Intra-domain Adaptation for Semantic Segmentation** 404
Zizheng Yan, Xianggang Yu (*CUHK-Shenzhen & Shenzhen Research Institute of Big Data*), Yipeng Qin (*Cardiff University*), Yushuang Wu, Xiaoguang Han, Shuguang Cui (*CUHK-Shenzhen & Shenzhen Research Institute of Big Data*)
- **Mask is All You Need: Rethinking Mask R-CNN for Dense and Arbitrary-Shaped Scene Text Detection** 414
Xugong Qin, Yu Zhou, Youhui Guo (*Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Dayan Wu (*Institute of Information Engineering, Chinese Academy of Sciences*), Zhihong Tian (*Guangzhou University*), Ning Jiang, Hongbin Wang (*Mashang Consumer Finance Co., Ltd.*), Weiping Wang (*Institute of Information Engineering, Chinese Academy of Sciences*)
- **Windowing Decomposition Convolutional Neural Network for Image Enhancement** 424
Chuanjun Zheng, Daming Shi, Yukun Liu (*Shenzhen University*)
- **Joint Optimization in Edge-Cloud Continuum for Federated Unsupervised Person Re-identification** 433
Weiming Zhuang (*Nanyang Technological University & SenseTime Research*), Yonggang Wen (*Nanyang Technological University*), Shuai Zhang (*SenseTime Research*)
- **Multi-view 3D Smooth Human Pose Estimation based on Heatmap Filtering and Spatio-temporal Information** 442
Zehai Niu (*University of Chinese Academy of Sciences*), Ke Lu (*University of Chinese Academy of Sciences & Peng Cheng Laboratory*), Jian Xue, Haifeng Ma, Runchen Wei (*University of Chinese Academy of Sciences*)
- **Imitative Learning for Multi-Person Action Forecasting** 451
Yu-Ke Li, Pin Wang (*University of California, Berkeley*), Mang Ye (*Wuhan University*), Ching-Yao Chan (*University of California, Berkeley*)
- **Stereo Video Super-Resolution via Exploiting View-Temporal Correlations** 460
Ruikang Xu, Zeyu Xiao, Mingde Yao, Yueyi Zhang, Zhiwei Xiong (*University of Science and Technology of China*)
- **M3TR: Multi-modal Multi-label Recognition with Transformer** 469
Jiawei Zhao, Yifan Zhao (*Beihang University*), Jia Li (*Beihang University & Peng Cheng Laboratory*)

• TACR-Net: Editing on Deep Video and Voice Portraits	478
Luchuan Song, Bin Liu, Guojun Yin, Xiaoyi Dong (<i>University of Science and Technology of China</i>), Yufei Zhang (<i>Meituan</i>), Jia-Xuan Bai (<i>University of Science and Technology of China</i>)	
• Annotation-Efficient Untrimmed Video Action Recognition	487
Yixiong Zou (<i>Peking University & Carnegie Mellon University</i>), Shanghang Zhang (<i>University of California, Berkeley</i>), Guangyao Chen, Yonghong Tian (<i>Peking University</i>), Kurt Keutzer (<i>University of California, Berkeley</i>), José M. F. Moura (<i>Carnegie Mellon University</i>)	
• Face-based Voice Conversion: Learning the Voice behind a Face	496
Hsiao-Han Lu, Shao-En Weng, Ya-Fan Yen, Hong-Han Shuai, Wen-Huang Cheng (<i>National Yang Ming Chiao Tung University</i>)	
• A Large-Scale Benchmark for Food Image Segmentation	506
Xiongwei Wu (<i>Singapore Management University</i>), Xin Fu (<i>Beijing Jiaotong University</i>), Ying Liu, Ee-Peng Lim (<i>Singapore Management University</i>), Steven C.H. Hoi (<i>Singapore Management University & Salesforce Research Asia</i>), Qianru Sun (<i>Singapore Management University</i>)	
• HAT: Hierarchical Aggregation Transformers for Person Re-identification	516
Guowen Zhang, Pingping Zhang, Jinqing Qi (<i>Dalian University of Technology</i>), Huchuan Lu (<i>Dalian University of Technology & Pengcheng Lab</i>)	
• Long-Range Feature Propagating for Natural Image Matting	526
Qinglin Liu, Haozhe Xie, Shengping Zhang (<i>Harbin Institute of Technology</i>), Bineng Zhong (<i>Guangxi Normal University</i>), Rongrong Ji (<i>Xiamen University</i>)	
• Towards Controllable and Photorealistic Region-wise Image Manipulation	535
Ansheng You (<i>Peking University</i>), Chenglin Zhou, Qixuan Zhang, Lan Xu (<i>ShanghaiTech University</i>)	
• Information-Growth Attention Network for Image Super-Resolution	544
Zhuangzi Li (<i>SECE, Shenzhen Graduate School, Peking University & Peng Cheng Laboratory</i>), Ge Li (<i>SECE, Shenzhen Graduate School, Peking University</i>), Thomas Li (<i>AIT of Peking University & ITRDIT of Peking University</i>), Shan Liu (<i>Tencent America</i>), Wei Gao (<i>SECE, Shenzhen Graduate School, Peking University & Peng Cheng Laboratory</i>)	
• Anchor-free 3D Single Stage Detector with Mask-Guided Attention for Point Cloud	553
Jiale Li (<i>Zhejiang University</i>), Hang Dai (<i>Mohamed bin Zayed University of Artificial Intelligence</i>), Ling Shao (<i>Inception Institute of Artificial Intelligence</i>), Yong Ding (<i>Zhejiang University</i>)	
• Shape Controllable Virtual Try-on for Underwear Models	563
Xin Gao, Zhenjiang Liu (<i>Alibaba Group</i>), Zunlei Feng, Chengji Shen (<i>Zhejiang University</i>), Kairi Ou, Haihong Tang (<i>Alibaba Group</i>), Mingli Song (<i>Zhejiang University</i>)	
• E²Net: Excitative-Expansile Learning for Weakly Supervised Object Localization	573
Zhiwei Chen, Liujuan Cao (<i>Xiamen University</i>), Yunhang Shen (<i>Tencent Youtu Lab</i>), Feihong Lian (<i>Xiamen University</i>), Yongjian Wu (<i>Tencent Youtu Lab</i>), Rongrong Ji (<i>Xiamen University</i>)	
• Few-shot Fine-Grained Action Recognition via Bidirectional Attention and Contrastive Meta-Learning	582
Jiahao Wang, Yunhong Wang, Sheng Liu, Annan Li (<i>Beihang University</i>)	
• Selective Dependency Aggregation for Action Classification	592
Yi Tan, Yanbin Hao, Xiangnan He (<i>University of Science and Technology of China</i>), Yinwei Wei, Xun Yang (<i>National University of Singapore</i>)	
• Conditional Directed Graph Convolution for 3D Human Pose Estimation	602
Wenbo Hu (<i>The Chinese University of Hong Kong & DAMO Academy, Alibaba Group</i>), Changgong Zhang (<i>DAMO Academy, Alibaba Group</i>), Fangneng Zhan (<i>Nanyang Technological University</i>), Lei Zhang (<i>The Hong Kong Polytechnic University & DAMO Academy, Alibaba Group</i>), Tien-Tsin Wong (<i>The Chinese University of Hong Kong</i>)	
• Cross Chest Graph for Disease Diagnosis with Structural Relational Reasoning	612
Gangming Zhao (<i>University of Chinese Academy of Sciences & University of Hong Kong</i>)	
• ZiGAN: Fine-grained Chinese Calligraphy Font Generation via a Few-shot Style Transfer Approach	621
Qi Wen (<i>NetEase Fuxi AI Lab</i>), Shuang Li, Bingfeng Han (<i>Beijing Institute of Technology</i>), Yi Yuan (<i>NetEase Fuxi AI Lab</i>)	

• Cycle-Consistent Inverse GAN for Text-to-Image Synthesis	630
Hao Wang, Guosheng Lin (<i>Nanyang Technological University</i>), Steven C. H. Hoi (<i>Singapore Management University</i>), Chunyan Miao (<i>Nanyang Technological University</i>)	
• Fully Quantized Image Super-Resolution Networks	639
Hu Wang, Peng Chen (<i>The University of Adelaide</i>), Bohan Zhuang (<i>Monash University</i>), Chunhua Shen (<i>Amazon Australia</i>)	
• AKECP: Adaptive Knowledge Extraction from Feature Maps for Fast and Efficient Channel Pruning	648
Haonan Zhang, Longjun Liu, Hengyi Zhou, Wenxuan Hou, Hongbin Sun, Nanning Zheng (<i>Xi'an Jiaotong University</i>)	
• Dynamic Momentum Adaptation for Zero-Shot Cross-Domain Crowd Counting	658
Qiangqiang Wu, Jia Wan, Antoni B. Chan (<i>City University of Hong Kong</i>)	
• Auto-MSFNet: Search Multi-scale Fusion Network for Salient Object Detection	667
Miao Zhang, Tingwei Liu, Yongri Piao, Shunyu Yao (<i>Dalian University of Technology</i>), Huchuan Lu (<i>Dalian University of Technology & Pengcheng Lab</i>)	
• Few-shot Unsupervised Domain Adaptation with Image-to-Class Sparse Similarity Encoding	677
Shengqi Huang, Wanqi Yang (<i>Nanjing Normal University</i>), Lei Wang (<i>University of Wollongong</i>), Luping Zhou (<i>University of Sydney</i>), Ming Yang (<i>Nanjing Normal University</i>)	
• Semantic-aware Transfer with Instance-adaptive Parsing for Crowded Scenes Pose Estimation	686
Xuanhan Wang, Lianli Gao, Yan Dai, Yixuan Zhou, Jingkuan Song (<i>University of Electronic Science and Technology of China</i>)	
• Multimodal Dialog System: Relational Graph-based Context-aware Question Understanding	695
Haoyu Zhang (<i>Shandong University</i>), Meng Liu (<i>Shandong Jianzhu University</i>), Zan Gao (<i>Shandong Artificial Intelligence Institute</i>), Xiaoqiang Lei (<i>Kuaishou Technology</i>), Yinglong Wang (<i>Shandong Artificial Intelligence Institute</i>), Liqiang Nie (<i>Shandong University</i>)	
• Shadow Detection via Predicting the Confidence Maps of Shadow Detection Methods	704
Jingwei Liao, Yanli Liu, Guanyu Xing, Housheng Wei, Jueyu Chen (<i>Sichuan University</i>), Songhua Xu (<i>University of South Carolina</i>)	
• Motion Prediction via Joint Dependency Modeling in Phase Space	713
Pengxiang Su (<i>Jilin University</i>), Zhenguang Liu (<i>Zhejiang University</i>), Shuang Wu (<i>Nanyang Technological University</i>), Lei Zhu (<i>Shandong Normal University</i>), Yifang Yin (<i>National University of Singapore</i>), Xuanjing Shen (<i>Jilin University</i>)	
• Q-Art Code: Generating Scanning-robust Art-style QR Codes by Deformable Convolution	722
Hao Su (<i>Beihang University</i>), Jianwei Niu (<i>Beihang University & Zhengzhou University</i>), Xuefeng Liu, Qingfeng Li, Ji Wan (<i>Beihang University</i>), Mingliang Xu (<i>Zhengzhou University</i>)	
• Depth Quality-Inspired Feature Manipulation for Efficient RGB-D Salient Object Detection	731
Wenbo Zhang (<i>Sichuan University</i>), Ge-Peng Ji (<i>Inception Institute of Artificial Intelligence</i>), Zhuo Wang, Keren Fu, Qijun Zhao (<i>Sichuan University</i>)	
• Revisiting Mid-Level Patterns for Cross-Domain Few-Shot Recognition	741
Yixiong Zou (<i>Peking University & Carnegie Mellon University</i>), Shanghang Zhang (<i>University of California, Berkeley</i>), Jianpeng Yu (<i>Carnegie Mellon University</i>), Yonghong Tian (<i>Peking University</i>), José M. F. Moura (<i>Carnegie Mellon University</i>)	
• Space-Angle Super-Resolution for Multi-View Images	750
Yuqi Sun, Ri Cheng, Bo Yan, Shili Zhou (<i>Fudan University</i>)	
• Weakly-Supervised Video Object Grounding via Stable Context Learning	760
Wei Wang, Junyu Gao, Changsheng Xu (<i>Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; & Peng Cheng Laboratory</i>)	

• Modeling the Uncertainty for Self-supervised 3D Skeleton Action Representation Learning	769
Yukun Su (<i>South China University of Technology</i>), Guosheng Lin (<i>Nanyang Technological University</i>), Ruizhou Sun, Yun Hao, Qingyao Wu (<i>South China University of Technology</i>)	
• D³Net: Dual-Branch Disturbance Disentangling Network for Facial Expression Recognition	779
Rongyun Mo, Yan Yan (<i>Xiamen University</i>), Jing-Hao Xue (<i>University College London</i>), Si Chen (<i>Xiamen University of Technology</i>), Hanzi Wang (<i>Xiamen University</i>)	
• Towards a Unified Middle Modality Learning for Visible-Infrared Person Re-Identification	788
Yukang Zhang, Yan Yan, Yang Lu, Hanzi Wang (<i>Xiamen University</i>)	
• ROSITA: Enhancing Vision-and-Language Semantic Alignments via Cross-and Intra-modal Knowledge Integration	797
Yuhao Cui, Zhou Yu (<i>Hangzhou Dianzi University</i>), Chunqi Wang, Zhongzhou Zhao, Ji Zhang (<i>Alibaba Group</i>), Meng Wang (<i>Hefei University of Technology</i>), Jun Yu (<i>Hangzhou Dianzi University</i>)	
• Object Point Cloud Classification via Poly-Convolutional Architecture Search	807
Xuanxiang Lin, Ke Chen, Kui Jia (<i>South China University of Technology & Peng Cheng Laboratory</i>)	
• Semantic-Guided Relation Propagation Network for Few-shot Action Recognition	816
Xiao Wang, Weirong Ye (<i>Xiamen University</i>), Zhongang Qi, Xun Zhao (<i>Tencent PCG</i>), Guangge Wang (<i>Xiamen University</i>), Ying Shan (<i>Tencent PCG</i>), Hanzi Wang (<i>Xiamen University</i>)	
• Anti-Distillation Backdoor Attacks: Backdoors Can Really Survive in Knowledge Distillation	826
Yunjie Ge, Qian Wang, Baolin Zheng, Xinlu Zhuang (<i>Wuhan University</i>), Qi Li (<i>Tsinghua University</i>), Chao Shen (<i>Xi'an Jiaotong University</i>), Cong Wang (<i>City University of Hong Kong</i>)	
• One-stage Context and Identity Hallucination Network	835
Yinglu Liu (<i>JD AI Research</i>), Mingcan Xiang (<i>JD AI Research & University of Massachusetts Amherst</i>), Hailin Shi, Tao Mei (<i>JD AI Research</i>)	
• Mitigating Generation Shifts for Generalized Zero-Shot Learning	844
Zhi Chen, Yadan Luo, Sen Wang, Ruihong Qiu (<i>The University of Queensland</i>), Jingjing Li (<i>University of Electronic Science and Technology of China</i>), Zi Huang (<i>The University of Queensland</i>)	
• Weakly-Supervised Temporal Action Localization via Cross-Stream Collaborative Learning	853
Yuan Ji, Xu Jia (<i>Dalian University of Technology</i>), Huchuan Lu (<i>Dalian University of Technology & Peng Cheng Laboratory</i>), Xiang Ruan (<i>Tiwaki Co.Ltd.</i>)	
• Deep Interactive Video Inpainting: An Invisibility Cloak for Harry Potter	862
Cheng Chen (<i>Alibaba Group & Huazhong University of Science and Technology</i>), Jiayin Cai (<i>Tsinghua University</i>), Yao Hu, Xu Tang (<i>Alibaba Group</i>), Xinggang Wang (<i>Huazhong University of Science and Technology</i>), Chun Yuan (<i>Tsinghua University</i>), Xiang Bai (<i>Huazhong University of Science and Technology</i>), Song Bai (<i>University of Oxford</i>)	
• Searching Motion Graphs for Human Motion Synthesis	871
Chenchen Liu, Yadong Mu (<i>Peking University</i>)	
• When Video Classification Meets Incremental Classes	880
Hanbin Zhao, Xin Qin, Shihao Su, Yongjian Fu, Zibo Lin, Xi Li (<i>Zhejiang University</i>)	
• Fast and Accurate Lane Detection via Frequency Domain Learning	890
Yulin He, Wei Chen (<i>National University of Defense Technology</i>), Zhengfa Liang (<i>Defense Innovation Institute</i>), Dan Chen (<i>Wuhan University</i>), Yusong Tan, Xin Luo, Chen Li, Yulan Guo (<i>National University of Defense Technology</i>)	
• Learning Multi-context Aware Location Representations from Large-scale Geotagged Images	899
Yifang Yin (<i>National University of Singapore</i>), Ying Zhang (<i>Northwestern Polytechnical University</i>), Zhenguang Liu (<i>Zhejiang Gongshang University</i>), Yuxuan Liang (<i>National University of Singapore</i>), Sheng Wang (<i>Alibaba Group</i>), Rajiv Ratn Shah (<i>Indraprastha Institute of Information Technology Delhi</i>), Roger Zimmermann (<i>National University of Singapore</i>)	

- **MV-TON: Memory-based Video Virtual Try-on network** 908
Xiaojing Zhong (*South China University of Technology*), Zhonghua Wu (*Nanyang Technological University*),
Taizhe Tan (*Guangdong University of Technology*), Guosheng Lin (*Nanyang Technological University*),
Qingyao Wu (*South China University of Technology*)
- **Token Shift Transformer for Video Classification**..... 917
Hao Zhang (*City University of Hong Kong*), Yanbin Hao (*University of Science and Technology of China*),
Chong-Wah Ngo (*Singapore Management University*)
- **Attribute-specific Control Units in StyleGAN for Fine-grained Image Manipulation**..... 926
Rui Wang (*Huazhong University of Science and Technology*), Jian Chen, Gang Yu (*Tencent*),
Li Sun (*East China Normal University*),
Changqian Yu, Changxin Gao, Nong Sang (*Huazhong University of Science and Technology*)
- **Attention-driven Graph Clustering Network** 935
Zhihao Peng, Hui Liu (*City University of Hong Kong*), Yuheng Jia (*Southeast University*),
Junhui Hou (*City University of Hong Kong*)
- **Lifting the Veil of Frequency in Joint Segmentation and Depth Estimation** 944
Tianhao Fu, Yingying Li, Xiaoqing Ye, Xiao Tan, Hao Sun (*Baidu*),
Fumin Shen (*University of Electronic Science and Technology of China*), Errui Ding (*Baidu*)

Panel 1

- **The Next Generation Multimodal Conversational Search and Recommendation** 953
Joao Magalhaes (*Universidade NOVA Lisboa*), Tat-Seng Chua (*National University of Singapore*),
Tao Mei (*AI Research of JD.com*), Alan Smeaton (*Dublin City University*)

Session 8: Emerging Multimedia Applications-IV

- **VoteHMR: Occlusion-Aware Voting Network for Robust 3D Human Mesh Recovery from Partial Point Clouds** 955
Guanze Liu (*Beihang University*), Yu Rong (*The Chinese University of Hong Kong*), Lu Sheng (*Beihang University*)
- **MageAdd: Real-Time Interaction Simulation for Scene Synthesis** 965
Shao-Kui Zhang, Yi-Xiao Li, Yu He (*Tsinghua University*), Yong-Liang Yang (*University of Bath*),
Song-Hai Zhang (*Tsinghua University & BNRist*)
- **Cross-View Exocentric to Egocentric Video Synthesis**..... 974
Gaowen Liu (*Cisco Systems*), Hao Tang (*ETH*), Hugo M. Latapie (*Cisco Systems*),
Jason J. Corso (*Stevens Institute for Artificial Intelligence*), Yan Yan (*Illinois Institute of Technology*)
- **EVRNet: Efficient Video Restoration on Edge Devices** 983
Sachin Mehta (*University of Washington*), Amit Kumar (*Facebook Inc.*), Fitsum Reda (*Google*),
Varun Nasery, Vikram Mulukutla, Rakesh Ranjan, Vikas Chandra (*Facebook Inc.*)
- **Multimodal Entity Linking: A New Dataset and A Baseline**..... 993
Jingru Gan (*Institute of Computing Technology, Chinese Academic of Science & University of Chinese Academy of Sciences*),
Jinchang Luo, Haiwei Wang (*Baidu Inc.*),
Shuhui Wang (*Institute of Computing Technology, Chinese Academic of Science*), Wei He (*Baidu Inc.*),
Qingming Huang (*University of Chinese Academy of Sciences & Institute of Computing Technology, Chinese Academy of Sciences*)
- **AI-Lyricist: Generating Music and Vocabulary Constrained Lyrics** 1002
Xichu Ma, Ye Wang, Min-Yen Kan, Wee Sun Lee (*National University of Singapore*)

Session 9: Emotional and Social Signals in Multimedia

- **CaFGraph: Context-aware Facial Multi-graph Representation for Facial Action Unit Recognition** 1029
Yingjie Chen (*Peking University*),
Diqi Chen (*Intelligent Computing Research Center, Advanced Institute of Information Technology (AIIT), Peking University*), Yizhou Wang, Tao Wang, Yun Liang (*Peking University*)

- **Self-Supervised Regional and Temporal Auxiliary Tasks for Facial Action Unit Recognition** 1038
Jingwei Yan, Jingjing Wang, Qiang Li, Chunmao Wang, Shiliang Pu (*Hikvision Research Institute*)
- **HetEmotionNet: Two-Stream Heterogeneous Graph Recurrent Neural Network for Multi-modal Emotion Recognition** 1047
Ziyu Jia, Youfang Lin (*Beijing Jiaotong University & Beijing Key Laboratory of Traffic Data Analysis and Mining*),
Jing Wang (*Beijing Jiaotong University & Beijing Key Laboratory of Traffic Data Analysis and Mining*),
Zhiyang Feng, Xiangheng Xie, Caijie Chen (*Beijing Jiaotong University*)
- **Simplifying Multimodal Emotion Recognition with Single Eye Movement Modality** 1057
Xu Yan (*University of Washington*), Li-Ming Zhao, Bao-Liang Lu (*Shanghai Jiao Tong University*)
- **Learning What and When to Drop: Adaptive Multimodal and Contextual Dynamics for Emotion Recognition in Conversation** 1064
Feiyu Chen (*University of Electronic Science and Technology of China & Sichuan Artificial Intelligence Research Institute*),
Zhengxiao Sun (*University of Electronic Science and Technology of China*),
Deqiang Ouyang (*Chongqing University*), Xueliang Liu (*Hefei University of Technology*), Jie Shao (*University of Electronic Science and Technology of China & Sichuan Artificial Intelligence Research Institute*)
- **Zero-shot Video Emotion Recognition via Multimodal Protagonist-aware Transformer Network** 1074
Fan Qi (*Hefei University of Technology & Peng Cheng Laboratory*),
Xiaoshan Yang, Changsheng Xu (*Institute of Automation, Chinese Academy of Sciences, University of Chinese Academy of Sciences, & Peng Cheng Laboratory*)

Session 10: Industrial Track

- **Show, Read and Reason: Table Structure Recognition with Flexible Context Aggregator** 1084
Hao Liu, Xin Li, Bing Liu, Deqiang Jiang, Yinsong Liu, Bo Ren (*Tencent YouTu Lab*),
Rongrong Ji (*Xiamen University*),
- **TRANSFUSION: Multi-Modal Fusion for Video Tag Inference via Translation-based Knowledge Embedding** 1093
Di Jin (*University of Michigan*), Zhongang Qi, Yingmin Luo, Ying Shan (*Tencent*)
- **RecycleNet: An Overlapped Text Instance Recovery Approach** 1102
Yiqing Hu, Yan Zheng, Xinghua Jiang, Hao Liu, Deqiang Jiang, Yinsong Liu, Bo Ren (*Tencent YouTu Lab*),
Rongrong Ji (*Xiamen University*)
- **ARShoe: Real-Time Augmented Reality Shoe Try-on System on Smartphones** 1111
Shan An (*JD.COM Inc. & Beihang University*), Guangfu Che, Jinghao Guo (*JD.COM Inc.*),
Haogang Zhu (*Beihang University & Beijing Advanced Innovation Center for Big Data-Based Precision Medicine*),
Junjie Ye, Fangru Zhou, Zhaoqi Zhu, Dong Wei (*JD.COM Inc.*), Aishan Liu (*Beihang University*),
Wei Zhang (*JD.COM Inc.*)
- **Inferring the Importance of Product Appearance with Semi-supervised Multi-modal Enhancement: A Step Towards the Screenless Retailing** 1120
Yongshun Gong (*Shandong University*), Jinfeng Yi, Dong-Dong Chen (*JD AI Research*),
Jian Zhang (*University of Technology Sydney*), Jiayu Zhou (*Michigan State University*),
Zhihua Zhou (*Nanjing University*)
- **AsyNCE: Disentangling False-Positives for Weakly-Supervised Video Grounding** 1129
Cheng Da, Yanhao Zhang, Yun Zheng, Pan Pan, Yinghui Xu (*Alibaba Group*),
Chunhong Pan (*Institute of Automation, Chinese Academy of Sciences*)
- **Unifying Multimodal Transformer for Bi-directional Image and Text Generation** 1138
Yupan Huang (*Sun Yat-sen University*), Hongwei Xue (*University of Science and Technology of China*),
Bei Liu (*Microsoft Research Asia*), Yutong Lu (*Sun Yat-sen University*)
- **Once and for All: Self-supervised Multi-modal Co-training on One-billion Videos at Alibaba** 1148
Lianghua Huang, Yu Liu, Xiangzeng Zhou, Ansheng You, Ming Li, Bin Wang, Yingya Zhang,
Pan Pan, Xu Yinghui (*Alibaba Group*)

- **L2RS: A Learning-to-Rescore Mechanism for Hybrid Speech Recognition** 1157
Yuanfeng Song (*The Hong Kong University of Science and Technology & WeBank Co., Ltd*),
Di Jiang, Xuefang Zhao, Qian Xu (*WeBank Co., Ltd*),
Raymond Chi-Wing Wong (*Hong Kong University of Science and Technology*), Lixin Fan (*WeBank Co., Ltd*),
Qiang Yang (*The Hong Kong University of Science and Technology & WeBank Co., Ltd*)
- **Distantly Supervised Semantic Text Detection and Recognition
for Broadcast Sports Videos Understanding** 1167
Avijit Shah, Topojoy Biswas, Sathish Ramadoss (*Yahoo! Research*), Deven Santosh Shah (*Microsoft*)
- **Focusing on Persons: Colorizing Old Images Learning from Modern Historical Movies** 1176
Xin Jin, Zhonglan Li, Ke Liu (*Beijing Electronic Science and Technology Institute*),
Dongqing Zou (*SenseTime Research and Tetras.AI; Qing Yuan Research Institute, Shanghai Jiao Tong University*),
Xiaodong Li, Xingfan Zhu, Ziyin Zhou, Qilong Sun, Qingyu Liu (*Beijing Electronic Science and Technology
Institute*)
- **Personalized Multi-modal Video Retrieval on Mobile Devices** 1185
Haotian Zhang, Allan D. Jepson, Iqbal Mohamed, Konstantinos G. Derpanis, Ran Zhang, Afsaneh Fazly
(*Samsung AI Center – Toronto*)
- **Boosting End-to-end Multi-Object Tracking and Person Search
via Knowledge Distillation**..... 1192
Wei Zhang (*CRIPAC & NLPR, CASIA and University of Chinese Academy of Sciences*),
Lingxiao He, Peng Chen, Xingyu Liao, Wu Liu (*JD AI Research*), Qi Li, Zhenan Sun (*CRIPAC & NLPR, CASIA*)
- **A Virtual Character Generation and Animation System for E-Commerce Live Streaming ...** 1202
Li Hu, Bang Zhang, Peng Zhang, Jinwei Qi, Jian Cao, Daiheng Gao, Haiming Zhao,
Xiaoduan Feng, Qi Wang, Lian Zhuo, Pan Pan, Yinghui Xu (*Alibaba Group*)
- **Improving Fake News Detection by Using an Entity-enhanced Framework
to Fuse Diverse Multimodal Clues**..... 1212
Peng Qi (*Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of
Sciences; & Zhengzhou University*),
Juan Cao (*Institute of Computing Technology, Chinese Academy of Sciences & University of Chinese Academy of
Sciences*), Xirong Li (*Renmin University of China*),
Huan Liu (*Zhengzhou University*), Qiang Sheng (*Institute of Computing Technology, Chinese Academy of Sciences
& University of Chinese Academy of Sciences*),
Xiaoyue Mi (*Institute of Computing Technology, Chinese Academy of Sciences & University of Chinese Academy of
Sciences*), Qin He, Yongbiao Lv, Chenyang Guo, Yingchao Yu (*Hangzhou ZhongkeRuijian Technology Co., Ltd.*)

Session 11: Multimedia HCI and Quality of Experience

- **Fast Video Visual Quality and Resolution Improvement using SR-UNet**..... 1221
Federico Vaccaro, Marco Bertini, Tiberio Uricchio, Alberto Del Bimbo (*Università degli Studi di Firenze*)
- **MS-GraphSIM: Inferring Point Cloud Quality via Multiscale Graph Similarity**..... 1230
Yujie Zhang, Qi Yang, Yiling Xu (*Shanghai Jiao Tong University*)
- **I Know Your Keyboard Input: A Robust Keystroke Eavesdropper
Based-on Acoustic Signals**..... 1239
Jia-Xuan Bai, Bin Liu, Luchuan Song (*University of Science and Technology of China*)
- **Perceptual Quality Assessment of Internet Videos**..... 1248
Jiahua Xu, Jing Li, Xingguang Zhou (*Alibaba Group*), Wei Zhou (*University of Science and Technology of China*),
Baichao Wang (*Alibaba Group*), Zhibo Chen (*University of Science and Technology of China*)
- **Using Interaction Data to Predict Engagement with Interactive Media**..... 1258
Jonathan Carlton (*University of Manchester and BBC R&D*), Andy Brown (*BBC R&D*),
Caroline Jay, John Keane (*University of Manchester*)
- **Air-Text: Air-Writing and Recognition System**..... 1267
Sun-Kyung Lee, Jong-Hwan Kim (*Korea Advanced Institute of Science and Technology*)

Session 12: Multimodal Analysis and Description-I

- **How to Learn a Domain-Adaptive Event Simulator?** 1275
Daxin Gu (*Beihang University*), Jia Li (*Beihang University & Peng Cheng Laboratory*),
Yu Zhang (*Beihang University*), Yonghong Tian (*Peking University & Peng Cheng Laboratory*)

- **A Stepwise Matching Method for Multi-modal Image based on Cascaded Network**..... 1284
Jinming Mu, Shuiping Gou, Shasha Mao (*Xidian University*),
Shankui Zheng (*Beijing Interstellar Glory Space Technology Co., Ltd.*)
- **SINGA-Easy: An Easy-to-Use Framework for MultiModal Analysis**..... 1293
Naili Xing, Chris Yeung (*National University of Singapore*),
Cheng-Hao Cai (*National University of Singapore & National University of Singapore (Suzhou) Research Institute*),
Teck Khim Ng, Wei Wang, Kaiyuan Yang, Nan Yang (*National University of Singapore*),
Meihui Zhang (*Beijing Institute of Technology*), Gang Chen (*Zhejiang University*),
Beng Chin Ooi (*National University of Singapore*)
- **Informative Class-Conditioned Feature Alignment for Unsupervised Domain Adaptation..** 1303
Wanxia Deng (*National University of Defense Technology*), Yawen Cui (*University of Oulu*),
Zhen Liu, Gangyao Kuang, Dewen Hu (*National University of Defense Technology*),
Matti Pietikäinen (*University of Oulu*), Li Liu (*National University of Defense Technology*)
- **Hierarchical Multi-Task Learning for Diagram Question Answering
with Multi-Modal Transformer** 1313
Zhaoquan Yuan Xiao Peng, Xiao Wu (*Southwest Jiaotong University*),
Changsheng Xu (*Institute of Automation, Chinese Academy of Sciences*)
- **Differentiated Learning for Multi-Modal Domain Adaptation** 1322
Jianming Lv, Kaijie Liu, Shengfeng He (*South China University of Technology*)

Session 13: Multimodal Analysis and Description-II

- **Two-stage Visual Cues Enhancement Network for Referring Image Segmentation** 1331
Yang Jiao (*Fudan University*), Zequn Jie, Weixin Luo (*Meituan*),
Jingjing Chen, Yu-Gang Jiang (*Fudan University*), Xiaolin Wei, Lin Ma (*Meituan*)
- **Partial Tubal Nuclear Norm Regularized Multi-view Learning** 1341
Yongyong Chen (*Harbin Institute of Technology, Shenzhen*), Shuqin Wang (*Shandong University of Science and
Technology*), Chong Peng (*Qingdao University*), Guangming Lu (*Harbin Institute of Technology, Shenzhen*),
Yicong Zhou (*University of Macau*)
- **Deep Unsupervised 3D SfM Face Reconstruction Based on Massive
Landmark Bundle Adjustment**..... 1350
Yuxing Wang, Yawen Lu (*Rochester Institute of Technology*), Zhihua Xie (*Jiangxi Sciences and Technology
Normal University*), Guoyu Lu (*Rochester Institute of Technology*)
- **SimuLLR: Simultaneous Lip Reading Transducer with Attention-Guided
Adaptive Memory** 1359
Zhijie Lin, Zhou Zhao, Haoyuan Li, Jinglin Liu (*Zhejiang University*),
Meng Zhang, Xingshan Zeng (*Huawei Noah's Ark Lab*), Xiaofei He (*Zhejiang University*)
- **Dense Semantic Contrast for Self-Supervised Visual Representation Learning**..... 1368
Xiaoni Li, Yu Zhou, Yifei Zhang (*Institute of Information Engineering, Chinese Academy of Sciences & University
of Chinese Academy of Sciences*),
Aoting Zhang (*Institute of Information Engineering, Chinese Academy of Sciences*),
Wei Wang (*Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy
of Sciences*), Ning Jiang, Haiying Wu (*Mashang Consumer Finance Co., Ltd.*),
Weiping Wang (*Institute of Information Engineering, Chinese Academy of Sciences*)
- **Multiple Object Tracking by Trajectory Map Regression
with Temporal Priors Embedding** 1377
Xingyu Wan, Sanping Zhou, Jinjun Wang, Rongye Meng (*Xi'an Jiaotong University*)

Session 14: Multimedia Cloud, Edge and Device Computing

- **DeepGame: Efficient Video Encoding for Cloud Gaming** 1387
Omar Mossad, Khaled Diab (*Simon Fraser University*), Ihab Amer (*AMD*),
Mohamed Hefeeda (*Simon Fraser University*)
- **ChartPointFlow for Topology-Aware 3D Point Cloud Generation**..... 1396
Takumi Kimura (*Kobe University*), Takashi Matsubara (*Osaka University*),
Kuniaki Uehara (*Osaka Gakuin University*)

- **Co-learning: Learning from Noisy Labels with Self-supervision** 1405
Cheng Tan Jun Xia, Lirong Wu, Stan Z. Li (*Westlake University & Westlake Institute for Advanced Study*)
- **Graph Convolutional Multi-modal Hashing for Flexible Multimedia Retrieval** 1414
Xu Lu, Lei Zhu (*Shandong Normal University*), Li Liu, Liqiang Nie (*Shandong University*),
Huaxiang Zhang (*Shandong Normal University*)
- **Hybrid Network Compression via Meta-Learning** 1423
Jianming Ye, Shiliang Zhang (*Peking University*), Jingdong Wang (*Microsoft Research Asia*)
- **Two-pronged Strategy: Lightweight Augmented Graph Network Hashing for Scalable Image Retrieval** 1432
Hui Cui, Lei Zhu (*Shandong Normal University*),
Jingjing Li (*University of Electronic Science and Technology of China*),
Zhiyong Cheng (*Shandong Artificial Intelligence Institute*),
Zheng Zhang (*Harbin Institute of Technology, Shenzhen*)

Interactive Arts

- **Reconstruction: A Motion Driven Interactive Artwork Inspired by Chinese Shadow Puppet** 1441
Wenli Jiang, Chong Cao (*Beihang University*)
- **Syntropic Counterpoints: Metaphysics of The Machines** 1443
Predrag K. Nikolic, Ruiyang Liu (*ShanghaiTech University*), Shengcheng Luo (*University of Aizu*)
- **Kandinsky Mobile: Abstract Art-Inspired Interactive Visualization of Social Discussions on Mobile Devices** 1446
Castillo Clarence Fitzgerald Gumtang (*Nanyang Technological University*),
Sourav S. Bhowmick (*Nanyang Technological University*)
- **Sand Scope: An Interactive Installation for Revealing the Connection Between Mental Space and Life Space in a Microcosm of the World** 1449
Lyn Chao-ling Chen (*Fu Jen Catholic University*)
- **Heraclitus's Forest: An Interactive Artwork for Oral History** 1452
Lin Wang (*Guangzhou Academy of Fine Arts*),
Zhonghao Lin, Wei Cai (*The Chinese University of Hong Kong*)
- **Affective Color Fields: Reimagining Rothkoesque Artwork as an Interactive Companion for Artistic Self-Expression** 1454
Aiden Kang, Liang Wang, Ziyu Zhou, Zhe Huang, Robert J.K. Jacob (*Tufts University*)
- **Apercevoir: Bio Internet of Things Interactive System** 1456
You-Yang Hu, Chiao-Chi Chou, Chia-Wei Li (*National Tsing Hua University*)

Poster Session 2

- **Visual Co-Occurrence Alignment Learning for Weakly-Supervised Video Moment Retrieval** 1459
Zheng Wang, Jingjing Chen, Yu-Gang Jiang (*Fudan University*)
- **Adaptive Normalized Representation Learning for Generalizable Face Anti-Spoofing** 1469
ShuBao Liu (*East China Normal University*),
Ke-Yue Zhang, Taiping Yao, Mingwei Bi, Shouhong Ding, Jilin Li, Feiyue Huang (*Tencent*),
Lizhuang Ma (*East China Normal University*)
- **Imitating Arbitrary Talking Style for Realistic Audio-Driven Talking Face Synthesis** 1478
Haozhe Wu, Jia Jia, Haoyu Wang (*Tsinghua University*),
Yishun Dou, Chao Duan, Qingshan Deng (*HiSilicon Company*)
- **Pose-guided Inter- and Intra-part Relational Transformer for Occluded Person Re-Identification** 1487
Zhongxing Ma, Yifan Zhao (*Beihang University*), Jia Li (*Beihang University & Peng Cheng Laboratory*)

- **VLAD-VSA: Cross-Domain Face Presentation Attack Detection with Vocabulary Separation and Adaptation** 1497
 Jiong Wang, Zhou Zhao, Weike Jin (*Zhejiang University*), Xinyu Duan (*Huawei Cloud*),
 Zhen Lei (*Institute of Automation, Chinese Academy of Sciences*), Baoxing Huai (*Huawei Cloud*),
 Yiling Wu (*Huawei Technology Co., Ltd*), Xiaofei He (*Zhejiang University*)
- **End-to-End Video Object Detection with Spatial-Temporal Transformers** 1507
 Lu He, Qianyu Zhou (*Shanghai Jiao Tong University*), Xiangtai Li (*Peking University*),
 Li Niu (*Shanghai Jiao Tong University*), Guangliang Cheng, Xiao Li (*Sensetime Research*),
 Wenxuan Liu (*University of California, Los Angeles*), Yunhai Tong (*Peking University*),
 Lizhuang Ma, Liqing Zhang (*Shanghai Jiao Tong University*)
- **Joint-teaching: Learning to Refine Knowledge for Resource-constrained Unsupervised Cross-modal Retrieval**..... 1517
 Peng-Fei Zhang, Jiasheng Duan, Zi Huang, Hongzhi Yin (*University of Queensland*)
- **AggNet for Self-supervised Monocular Depth Estimation: Go An Aggressive Step Further**..... 1526
 Zhi Chen (*University of Science and Technology of China*), Xiaoqing Ye (*Baidu, Inc.*),
 Liang Du (*Fudan University*), Wei Yang, Liusheng Huang (*University of Science and Technology of China*),
 Xiao Tan (*Baidu, Inc.*), Zhenbo Shi (*University of Science and Technology of China*),
 Fumin Shen (*University of Electronic Science and Technology of China*), Errui Ding (*Baidu, Inc.*)
- **Boosting Lightweight Single Image Super-resolution via Joint-distillation** 1535
 Xiaotong Luo, Qiuyuan Liang (*Xiamen University*), Ding Liu (*Bytedance Inc.*),
 Yanyun Qu (*Xiamen University*)
- **Discriminator-free Generative Adversarial Attack** 1544
 Shaohao Lu, Yuqiao Xian (*Sun Yat-Sen University*), Ke Yan, Yi Hu, Xing Sun, Xiaowei Guo, Feiyue Huang
 (*Tencent Youtu Lab*), Wei-Shi Zheng (*Sun Yat-Sen University*)
- **Former-DFER: Dynamic Facial Expression Recognition Transformer**..... 1553
 Zengqun Zhao, Qingshan Liu (*Nanjing University of Information Science & Technology*)
- **Discovering Density-Preserving Latent Space Walks in GANs for Semantic Image Transformations**..... 1562
 Guanyue Li, Yi Liu, Xiwen Wei (*South China University of Technology*),
 Yang Zhang (*City University of Hong Kong*), Si Wu, Yong Xu (*South China University of Technology*),
 Hau-San Wong (*City University of Hong Kong*)
- **MGH: Metadata Guided Hypergraph Modeling for Unsupervised Person Re-identification** 1571
 Yiming Wu, Xintian Wu, Xi Li, Jian Tian (*Zhejiang University*)
- **Recovering the Unbiased Scene Graphs from the Biased Ones**..... 1581
 Meng-Jiun Chiou (*National University of Singapore*), Henghui Ding (*ByteDance AI Lab*),
 Hanshu Yan (*National University of Singapore*), Changhu Wang (*ByteDance AI Lab*),
 Roger Zimmermann, Jiashi Feng (*National University of Singapore*)
- **Cross-modal Consensus Network for Weakly Supervised Temporal Action Localization** 1591
 Fa-Ting Hong (*Sun Yat-sen University, Tencent PCG, Ministry of Education, & Pazhou Lab*),
 Jia-Chang Feng (*Sun Yat-sen University, Tencent PCG, & Ministry of Education*),
 Dan Xu (*Hong Kong University of Science and Technology*), Ying Shan (*Tencent PCG*),
 Wei-Shi Zheng (*Sun Yat-sen University, Peng Cheng Laboratory, & Ministry of Education*)
- **Searching a Hierarchically Aggregated Fusion Architecture for Fast Multi-Modality Image Fusion** 1600
 Risheng Liu, Zhu Liu, Jinyuan Liu, Xin Fan (*Dalian University of Technology*)
- **SuperFront: From Low-resolution to High-resolution Frontal Face Synthesis** 1609
 Yu Yin (*Northeastern University*), Joseph P. Robinson (*Vicarious Surgical*),
 Songyao Jiang, Yue Bai, Can Qin, Yun Fu (*Northeastern University*)
- **Learning Segment Similarity and Alignment in Large-Scale Content Based Video Retrieval** 1618
 Chen Jiang, Kaiming Huang, Sifeng He, Xudong Yang, Wei Zhang, Xiaobo Zhang, Yuan Cheng, Lei Yang,
 Qing Wang, Furong Xu, Tan Pan, Wei Chu (*Ant Group*)

- **Cut-Thumbnail: A Novel Data Augmentation for Convolutional Neural Network**..... 1627
Tianshu Xie, Xuan Cheng, Xiaomin Wang, Minghui Liu, Jiali Deng, Tao Zhou, Ming Liu
(*University of Electronic Science and Technology of China*)
- **Diffusing the Liveness Cues for Face Anti-spoofing** 1636
Sheng Li, Xun Zhu (*Fudan University*), Guorui Feng (*Shanghai University*),
Xinpeng Zhang, Zhenxing Qian (*Fudan University*)
- **Co-Transport for Class-Incremental Learning** 1645
Da-Wei Zhou, Han-Jia Ye, De-Chuan Zhan (*Nanjing University*)
- **Skeleton-Contrastive 3D Action Representation Learning** 1655
Fida Mohammad Thoker, Hazel Doughty, Cees G. M. Snoek (*University of Amsterdam*)
- **Fast-forwarding, Rewinding, and Path Exploration in Interactive Branched Video Streaming**..... 1664
Albin Vogel, Erik Kronberg, Niklas Carlsson (*Linköping University*),
- **Multiview Detection with Shadow Transformer (and View-Coherent Data Augmentation)**..... 1673
Yunzhong Hou, Liang Zheng (*Australian National University*)
- **Domain Generalization via Feature Variation Decorrelation** 1683
Chang Liu, Lichen Wang, Kai Li, Yun Fu (*Northeastern University*)
- **Occlusion-aware Bi-directional Guided Network for Light Field Salient Object Detection**.. 1692
Dong Jing (*Beijing Jiaotong University & Beijing Key Lab of Traffic Data Analysis and Mining, Beijing*),
Shuo Zhang (*Beijing Jiaotong University; Beijing Key Lab of Traffic Data Analysis and Mining, Beijing & CAAC Key Laboratory of Intelligent Passenger Service of Civil Aviation*),
Runmin Cong (*Beijing Jiaotong University & Beijing Key Laboratory of Advanced Information Science and Network Technology*),
Youfang Lin (*Beijing Jiaotong University; Beijing Key Lab of Traffic Data Analysis and Mining, Beijing & CAAC Key Laboratory of Intelligent Passenger Service of Civil Aviation*)
- **One-Stage Visual Grounding via Semantic-Aware Feature Filter** 1702
Jiabo Ye (*East China Normal University*),
Xin Lin, Liang He (*East China Normal University & Shanghai Key Laboratory of Multidimensional Information Processing*), Dingbang Li, Qin Chen (*East China Normal University*)
- **Few-Shot Multi-Agent Perception** 1712
Chenyou Fan, Junjie Hu (*Shenzhen Institute of Artificial Intelligence and Robotics for Society*),
Jianwei Huang (*The Chinese University of Hong Kong, Shenzhen*)
- **SI3DP: Source Identification Challenges and Benchmark for Consumer-Level 3D Printer Forensics** 1721
Bo Seok Shim, Yoo Seung Shin, Seong Wook Park, Jong-Uk Hou (*Hallym University*)
- **Exploring Sequence Feature Alignment for Domain Adaptive Detection Transformers** 1730
Wen Wang (*University of Science and Technology of China*),
Yang Cao (*University of Science and Technology of China & Hefei Comprehensive National Science Center*),
Jing Zhang (*The University of Sydney*), Fengxiang He (*JD Explore Academy*),
Zheng-Jun Zha (*University of Science and Technology of China*),
Yonggang Wen (*Nanyang Technological University*), Dacheng Tao (*JD Explore Academy*)
- **Towards Realistic Visual Dubbing with Heterogeneous Sources**..... 1739
Tianyi Xie (*Shanghai Jiao Tong University*), Liucheng Liao (*University of Electronic Science and Technology of China*),
Cheng Bi, Benlai Tang, Xiang Yin (*ByteDance AI Lab*),
Jianfei Yang (*Nanyang Technological University*),
Mingjie Wang (*University of Guelph & Memorial University of Newfoundland*),
Jiali Yao, Yang Zhang, Zejun Ma (*ByteDance AI Lab*)
- **Deep Self-Supervised t-SNE for Multi-modal Subspace Clustering**..... 1748
Qianqian Wang, Wei Xia (*Xidian University*), Zhiqiang Tao (*Santa Clara University*),
Quanxue Gao (*Xidian University*), Xiaochun Cao (*Chinese Academy of Sciences*)
- **Multimodal Video Summarization via Time-Aware Transformers** 1756
Kindi Shang (*National University of Singapore*), Zehuan Yuan, Anran Wang, Changhu Wang (*ByteDance Inc*)

- **State-aware Video Procedural Captioning** 1766
Taichi Nishimura (*Kyoto University*), Atsushi Hashimoto, Yoshitaka Ushiku (*OMRON SINIC X Corporation*), Hiroataka Kameko, Shinsuke Mori (*Kyoto University*)
- **AMSS-Net: Audio Manipulation on User-Specified Sources with Textual Queries** 1775
Woosung Choi, Minseok Kim (*Korea University*),
Marco A. Martínez Ramírez (*Queen Mary University of London*),
Jaehwa Chung (*Korea National Open University*), Soonyoung Jung (*Korea University*)
- **Fully Functional Image Manipulation Using Scene Graphs in A Bounding-Box Free Way** 1784
Sitong Su, Lianli Gao, Junchen Zhu, Jie Shao, Jingkuan Song
(*University of Electronic Science and Technology of China*)
- **Multi-Level Counterfactual Contrast for Visual Commonsense Reasoning** 1793
Xi Zhang (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Feifei Zhang (*Institute of Automation, Chinese Academy of Sciences*),
Changsheng Xu (*Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; & Peng Cheng Laboratory*)
- **Data-Free Ensemble Knowledge Distillation for Privacy-conscious Multimedia Model Compression** 1803
Zhiwei Hao (*Beijing Institute of Technology*), Yong Luo (*Wuhan University*),
Han Hu, Jianping An (*Beijing Institute of Technology*), Yonggang Wen (*Nanyang Technological University*)
- **SM-SGE: A Self-Supervised Multi-Scale Skeleton Graph Encoding Framework for Person Re-Identification** 1812
Haocong Rao (*Shenzhen Institute of Advanced Technology & The Chinese University of Hong Kong, Hong Kong*),
Xiping Hu (*Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences & The Chinese University of Hong Kong, Hong Kong & Lanzhou University*),
Jun Cheng (*Shenzhen Institutes of Advanced Technology & The Chinese University of Hong Kong, Hong Kong*),
Bin Hu (*Beijing Institute of Technology & Lanzhou University*)
- **Video Transformer for Deepfake Detection with Incremental Learning** 1821
Sohail Ahmed Khan, Hang Dai (*Mohamed bin Zayed University of Artificial Intelligence*)
- **Chinese Character inpainting with Contextual Semantic Constraints** 1829
Jiahao Wang, Gang Pan (*Tianjin University*), Di Sun (*Tianjin University of Science and Technology*),
Jiawan Zhang (*Tianjin University*)
- **Curriculum-Based Meta-learning** 1838
Ji Zhang (*Center for Future Media, University of Electronic Science and Technology of China*),
Jingkuan Song (*Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China*),
Yazhou Yao (*Nanjing University of Science and Technology*),
Lianli Gao (*Center for Future Media, University of Electronic Science and Technology of China*)
- **Ego-Deliver: A Large-Scale Dataset For Egocentric Video Analysis** 1847
Haonan Qiu (*East China Normal University*), Pan He (*University of Florida*), Shuchun Liu, Weiyuan Shao,
Feiyun Zhang, Jiajun Wang (*Alibaba Group*), Liang He, Feng Wang (*East China Normal University*)
- **Adversarial Pixel Masking: A Defense against Physical Attacks for Pre-trained Object Detectors** 1856
Ping-Han Chiang, Chi-Shen Chan, Shan-Hung Wu (*National Tsing Hua University*)
- **Knowledge-Supervised Learning: Knowledge Consensus Constraints for Person Re-Identification** 1866
Li Wang, Baoyu Fan, Zhenhua Guo, Yaqian Zhao, Runze Zhang, Rengang Li, Weifeng Gong, Endong Wang
(*Inspur Electronic Information Industry Co., Ltd. & State Key Laboratory of High-end Server & Storage Technology*)
- **View-normalized Skeleton Generation for Action Recognition** 1875
Qingzhe Pan, Zhifu Zhao, Xuemei Xie, Jianan Li, Yuhan Cao, Guangming Shi (*Xidian University*)
- **Learning Hierarchical Embedding for Video Instance Segmentation** 1884
Zheyun Qin, Xiankai Lu (*Shandong University*), Xiushan Nie (*Shandong Jianzhu University*),
Xiantong Zhen (*University of Amsterdam*), Yilong Yin (*Shandong University*)
- **Text as Neural Operator: Image Manipulation by Text Instruction** 1893
Tianhao Zhang (*Google Research*), Hung-Yu Tseng (*University of California, Merced*), Lu Jiang (*Google Research*),
Weilong Yang (*Waymo*), Honglak Lee (*University of Michigan*), Irfan Essa (*Google Research*)

- **DSANet: Dynamic Segment Aggregation Network for Video-Level Representation Learning** 1903
Wenhao Wu (*Baidu Inc.*), Yuxiang Zhao (*Baidu Inc. & Shenzhen Institute of Advanced Technology, CAS*), Yanwu Xu (*University of Pittsburgh*), Xiao Tan, Dongliang He, Zhikang Zou, Jin Ye, Yingying Li, Mingde Yao, Zichao Dong, Yifeng Shi (*Baidu, Inc.*)
- **StrucTexT: Structured Text Understanding with Multi-Modal Transformers** 1912
Yulin Li (*Baidu Inc.*), Yuxi Qian (*Beijing University of Posts and Telecommunications*), Yuechen Yu, Xiameng Qin, Chengquan Zhang (*Baidu Inc.*), Yan Liu (*Taikang Insurance Group*), Kun Yao, Junyu Han, Jingtuo Liu, Errui Ding (*Baidu Inc.*)
- **Local Graph Convolutional Networks for Cross-Modal Hashing** 1921
Yudong Chen, Sen Wang (*The University of Queensland*), Jianglin Lu (*Shenzhen University*), Zhi Chen (*The University of Queensland*), Zheng Zhang (*Harbin Institute of Technology & Peng Cheng Laboratory*), Zi Huang (*The University of Queensland*)
- **Metric Learning for Anti-Compression Facial Forgery Detection** 1929
Shenhao Cao, Qin Zou (*Wuhan University*), Xiuqing Mao (*Information Engineering University*), Dengpan Ye, Zhongyuan Wang (*Wuhan University*)
- **ASFM-Net: Asymmetrical Siamese Feature Matching Network for Point Completion** 1938
Yaqi Xia (*Xidian University*), Yan Xia (*Technical University of Munich*), Wei Li (*Inceptio*), Rui Song, Kailang Cao (*Xidian University*), Uwe Stilla (*Technical University of Munich*)
- **Capsule-based Object Tracking with Natural Language Specification** 1948
Ding Ma, Xiangqian Wu (*Harbin Institute of Technology*)
- **Faster-PPN: Towards Real-Time Semantic Segmentation with Dual Mutual Learning for Ultra-High Resolution Images** 1957
Bicheng Dai, Kaisheng Wu, Tong Wu (*Xiamen University*), Kai Li (*Northeastern University*), Yanyun Qu (*Xiamen University*), Yuan Xie (*East China Normal University*), Yun Fu (*Northeastern University*)
- **Distributed Attention for Grounded Image Captioning** 1966
Nenglu Chen (*The University of Hong Kong*), Xingjia Pan (*Youtu Lab, Tencent*), Runnan Chen, Lei Yang (*The University of Hong Kong*), Zhiwen Lin, Yuqiang Ren, Haolei Yuan, Xiaowei Guo, Feiyue Huang (*Youtu Lab, Tencent*), Wenping Wang (*University of Hong Kong*)
- **Multi-initialization Optimization Network for Accurate 3D Human Pose and Shape Estimation** 1976
Zhiwei Liu (*Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; & ObjectEye, Inc.*), Xiangyu Zhu (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Lu Yang (*Beijing University of Posts and Telecommunications*), Xiang Yan (*Dilusense Technology Corporation.*), Ming Tang (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Zhen Lei (*Institute of Automation Chinese Academy of Sciences; University of Chinese Academy of Sciences; & Hong Kong Institute of Science and Innovation, Chinese Academy of Sciences*), Guibo Zhu (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Xuetao Feng, Yan Wang (*Alibaba Group*), Jinqiao Wang (*Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; & ObjectEye, Inc.*)
- **Feedback Network for Mutually Boosted Stereo Image Super-Resolution and Disparity Estimation** 1985
Qinyan Dai, Juncheng Li, Qiaosi Yi, Faming Fang, Guixu Zhang (*East China Normal University*)
- **Merging Multiple Template Matching Predictions in Intra Coding with Attentive Convolutional Neural Network** 1994
Qijun Wang, Guodong Zheng (*Anhui University*)
- **Camera-Agnostic Person Re-Identification via Adversarial Disentangling Learning** 2002
Hao Ni, Jingkuan Song, Xiaosu Zhu (*University of Electronic Science and Technology of China*), Feng Zheng (*Southern University of Science and Technology*), Lianli Gao (*University of Electronic Science and Technology of China*)

Session 15: Best Paper Session

- **Speech2AffectiveGestures: Synthesizing Co-Speech Gestures with Generative Adversarial Affective Expression Learning** 2027
Uttaran Bhattacharya, Elizabeth Childs, Nicholas Rewkowski, Dinesh Manocha
(*University of Maryland, College Park*)
- **Video Background Music Generation with Controllable Music Transformer** 2037
Shangzhe Di, Zeren Jiang, Si Liu, Zhaokai Wang, Leyan Zhu, Zexin He (*Beihang University*),
Hongming Liu (*Charterhouse School*), Shuicheng Yan (*Sea AI Lab*)
- **PIMNet: A Parallel, Iterative and Mimicking Network for Scene Text Recognition** 2046
Zhi Qiao, Yu Zhou (*Institute of Information Engineering, Chinese Academy of Sciences & School of Cyber Security, University of Chinese Academy of Sciences*), Jin Wei (*Communication University of China*),
Wei Wang (*Institute of Information Engineering, Chinese Academy of Sciences & School of Cyber Security, University of Chinese Academy of Sciences*), Yuan Zhang (*Communication University of China*),
Ning Jiang, Hongbin Wang (*Mashang Consumer Finance Co., Ltd.*),
Weiping Wang (*Institute of Information Engineering, Chinese Academy of Sciences*)
- **Theophany: Multimodal Speech Augmentation in Instantaneous Privacy Channels** 2056
Abhishek Kumar (*University of Helsinki*), Tristan Braud (*Hong Kong University of Science and Technology*),
Lik Hang Lee (*Korea Advanced Institute of Science and Technology*), Pan Hui (*Hong Kong University of Science and Technology & University of Helsinki*)
- **aBio: Active Bi-Olfactory Display Using Subwoofers for Virtual Reality** 2065
You-Yang Hu, Yao-Fu Jan, Kuan-Wei Tseng, You-Shin Tsai, Hung-Ming Sung (*National Taiwan University*),
Jin-Yao Lin (*Tainan National University of the Arts*), Yi-Ping Hung (*National Taiwan University*)

Session 16: ACM SIGMM Awards

Poster Session 3

- **Learning to Understand Traffic Signs**..... 2076
Yunfei Guo, Wei Feng, Fei Yin (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Tao Xue, Shuqi Mei (*Tencent Technology (Beijing) Co., Ltd.*),
Cheng-Lin Liu (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences & CAS Center for Excellence of Brain Science and Intelligence Technology*),
- **R-GAN: Exploring Human-like Way for Reasonable Text-to-Image Synthesis via Generative Adversarial Networks**..... 2085
Yanyuan Qiao, Qi Chen, Chaorui Deng (*University of Adelaide*),
Ning Ding (*South China University of Technology*),
Yuankai Qi (*University of Adelaide*), Mingkui Tan (*South China University of Technology*),
Xincheng Ren (*Yanan University*), Qi Wu (*University of Adelaide*)
- **Cross-modality Discrepant Interaction Network for RGB-D Salient Object Detection** 2094
Chen Zhang (*Beijing Jiaotong University & Beijing Key Laboratory of Advanced Information Science and Network Technology*),
Runmin Cong (*Beijing Jiaotong University & Beijing Key Laboratory of Advanced Information Science and Network Technology & City Univeristy of Hong Kong*),
Qinwei Lin (*Beijing Jiaotong University*), Lin Ma (*Meituan*),
Feng Li, Yao Zhao (*Beijing Jiaotong University & Beijing Key Laboratory of Advanced Information Science and Network Technology*),
Sam Kwong (*City Univeristy of Hong Kong*)
- **Deconfounded and Explainable Interactive Vision-Language Retrieval of Complex Scenes** 2103
Junda Wu (*New York University*), Tong Yu (*Carnegie Mellon University*),
Shuai Li (*Shanghai Jiao Tong University*)
- **Long Short-term Convolutional Transformer for No-Reference Video Quality Assessment**..... 2112
Junyong You (*Norwegian Research Centre*)

- **Automatic Channel Pruning with Hyper-parameter Search and Dynamic Masking** 2121
Baopu Li (*Baidu USA LLC*), Yanwen Fan (*Baidu Inc.*), Zhihong Pan (*Baidu USA LLC*),
Yuchen Bian (*Baidu Research*), Gang Zhang (*Baidu Inc.*)
- **SVHAN: Sequential View Based Hierarchical Attention Network
for 3D Shape Recognition** 2130
Yue Zhao, Weizhi Nie (*Tianjin University*), An-An Liu (*Tianjin University & Chinese Academy of Sciences*),
Zan Gao (*Shandong Artificial Intelligence Institute*), Yuting Su (*Tianjin University*)
- **ASFD: Automatic and Scalable Face Detector** 2139
Jian Li (*Tencent Youtu Lab*), Bin Zhang (*Southeast University*),
Yabiao Wang, Ying Tai, Zhenyu Zhang, Chengjie Wang, Jilin Li, Xiaoming Huang (*Tencent Youtu Lab*),
Yili Xia (*Southeast University*)
- **BridgeNet: A Joint Learning Network of Depth Map Super-Resolution
and Monocular Depth Estimation** 2148
Qi Tang, Runmin Cong, Ronghui Sheng, Lingzhi He (*Institute of Information Science, Beijing Jiaotong University
& Beijing Key Laboratory of Advanced Information Science and Network Technology*),
Dan Zhang (*UISEE Technology (Beijing) Co., Ltd.*), Yao Zhao (*Institute of Information Science, Beijing Jiaotong
University & Beijing Key Laboratory of Advanced Information Science and Network Technology*),
Sam Kwong (*City Univeristy of Hong Kong*)
- **LSTC: Boosting Atomic Action Detection with Long-Short-Term Context** 2158
Yuxi Li, Boshen Zhang, Jian Li, Yabiao Wang (*Tencent Youtu Lab*),
Weiyao Lin (*Shanghai Jiao Tong University*), Chengjie Wang, Jilin Li, Feiyue Huang (*Tencent Youtu Lab*)
- **UACANet: Uncertainty Augmented Context Attention for Polyp Segmentation** 2167
Taehun Kim, Hyemin Lee, Daijin Kim (*Pohang University of Science and Technology*)
- **Weight Evolution: Improving Deep Neural Networks Training
through Evolving Inferior Weight Values** 2176
Zhenquan Lin, Kailing Guo (*South China University of Technology*),
Xiaofen Xing (*South China University of Technology and UBTECH-SCUT Union Laboratory*),
Xiangmin Xu (*South China University of Technology and Institute of Modern Industrial Technology of SCUT in
Zhongshan*)
- **Coarse to Fine: Domain Adaptive Crowd Counting via Adversarial Scoring Network** 2185
Zhikang Zou (*Huazhong University of Science and Technology & Baidu Inc.*),
Xiaoye Qu, Pan Zhou (*Huazhong University of Science and Technology*),
Shuangjie Xu (*Hong Kong University of Science and Technology*),
Xiaoqing Ye, Wenhao Wu, Jin Ye (*Baidu Inc.*)
- **Towards Adversarial Patch Analysis and Certified Defense against Crowd Counting** 2195
Qiming Wu (*Huazhong University of Science and Technology*), Zhikang Zou (*Baidu Inc.*),
Pan Zhou (*Huazhong University of Science and Technology*), Xiaoqing Ye (*Baidu Inc.*),
Binghui Wang (*Illinois Institute of Technology*), Ang Li (*Duke University*)
- **Conceptual and Syntactical Cross-modal Alignment with Cross-level Consistency
for Image-Text Matching** 2205
Pengpeng Zeng, Lianli Gao, Xinyu Lyu, Shuaiqi Jing, Jingkuan Song
(*University of Electronic Science and Technology of China*)
- **SSPU-Net: Self-Supervised Point Cloud Upsampling via Differentiable Rendering** 2214
Yifan Zhao, Le Hui, Jin Xie (*Nanjing University of Science and Technology*)
- **VmAP: A Fair Metric for Video Object Detection** 2224
Anupam Sobti, Vaibhav Mavi, M Balakrishnan, Chetan Arora (*Indian Institute of Technology Delhi*)
- **Source Data-free Unsupervised Domain Adaptation for Semantic Segmentation** 2233
Mucong Ye, Jing Zhang, Jinpeng Ouyang, Ding Yuan (*Beihang University*)
- **“Yes, Attention Is All You Need” for Exemplar based Colorization** 2243
Wang Yin, Peng Lu, Zhaoran Zhao (*Beijing University of Posts and Telecommunications*),
Xujun Peng (*Information Sciences Institute, University of Southern California*)

- **Heuristic Depth Estimation via Recurrent Prediction with Confidence-Aware Loss** 2252
 Jiehua Zhang (*Hangzhou Dianzi University*),
 Liang Li (*Institute of Computing Technology, Chinese Academy of Sciences*),
 Chenggang Yan, Yaoqi Sun (*Hangzhou Dianzi University*),
 Tao Shen (*Kunming University of Science and Technology*), Jiyong Zhang (*Hangzhou Dianzi University*),
 Zhan Wang (*Moreal Pte. Ltd*)
- **Unsupervised Cross-Modal Distillation for Thermal Infrared Tracking** 2262
 Jingxian Sun (*Northwestern Polytechnical University*), Lichao Zhang (*Air Force Engineering University*),
 Yufei Zha (*Northwestern Polytechnical University*), Abel Gonzalez-Garcia (*wrnch*),
 Peng Zhang (*Northwestern Polytechnical University*), Wei Huang (*Nanchang University*),
 Yanning Zhang (*Northwestern Polytechnical University*)
- **ABPNet: Adaptive Background Modeling for Generalized Few Shot Segmentation** 2271
 Kaiqi Dong, Wei Yang, Zhenbo Xu, Liusheng Huang, Zhidong Yu
 (*University of Science and Technology of China*)
- **Towards Reasoning Ability in Scene Text Visual Question Answering** 2281
 Qingqing Wang, Liqiang Xiao (*Shanghai Jiao Tong University*), Yue Lu (*East China Normal University*),
 Yaohui Jin, Hao He (*Shanghai Jiao Tong University*)
- **Multi-caption Text-to-Face Synthesis: Dataset and Algorithm** 2290
 Jianxin Sun, Qi Li (*Center for Research on Intelligent Perception and Computing, NLPR, CASIA & School of Artificial Intelligence, University of Chinese Academy of Sciences (UCAS)*),
 Weining Wang (*Center for Research on Intelligent Perception and Computing, NLPR, CASIA*),
 Jian Zhao (*Institute of North Electronic Equipment*),
 Zhenan Sun (*Center for Research on Intelligent Perception and Computing, NLPR, CASIA & School of Artificial Intelligence, University of Chinese Academy of Sciences (UCAS)*)
- **Multimodal Compatibility Modeling via Exploring the Consistent and Complementary Correlations** 2299
 Weili Guan (*Monash University*), Haokun Wen, Xuemeng Song (*Shandong University*),
 Chung-Hsing Yeh (*Monash University*), Xiaojun Chang (*MIT University*), Liqiang Nie (*Shandong University*)
- **CDD: Multi-view Subspace Clustering via Cross-view Diversity Detection** 2308
 Shudong Huang (*Sichuan University*), Ivor W. Tsang (*University of Technology Sydney*),
 Zenglin Xu (*Harbin Institute of Technology*), Jiancheng Lv, Quanhui Liu (*Sichuan University*)
- **Learning Spatio-temporal Representation by Channel Aliasing Video Perception** 2317
 Yiqi Lin, Jimpeng Wang, Manlin Zhang, Andy J. Ma (*Sun Yat-sen University*)
- **Efficient Sparse Attacks on Videos using Reinforcement Learning** 2326
 Huanqian Yan, Xingxing Wei (*Beihang University*)
- **AdvHash: Set-to-set Targeted Attack on Deep Hashing with One Single Adversarial Patch** 2335
 Shengshan Hu, Yechao Zhang, Xiaogeng Liu (*Huazhong University of Science and Technology*),
 Leo Yu Zhang (*Deakin University*), Minghui Li (*Huazhong University of Science and Technology*),
 Hai Jin (*Huazhong University of Science and Technology*)
- **TransRefer3D: Entity-and-Relation Aware Transformer for Fine-Grained 3D Visual Grounding** 2344
 Dailan He, Yusheng Zhao, Junyu Luo (*Beihang University*),
 Tianrui Hui, Shaofei Huang (*Institute of Information Engineering, Chinese Academy of Sciences*),
 Aixi Zhang (*Alibaba Group*), Si Liu (*Institute of Artificial Intelligence*)
- **Single Image 3D Object Estimation with Primitive Graph Networks** 2353
 Qian He (*School of Information Science and Technology, ShanghaiTech University & Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
 Desen Zhou (*Department of Computer Vision Technology (VIS), Baidu Inc.*),
 Bo Wan (*Department of Electrical Engineering (ESAT), KU Leuven*),
 Xuming He (*School of Information Science and Technology, ShanghaiTech University & Shanghai Engineering Research Center of Intelligent Vision and Imaging*)

- **Boosting Mobile CNN Inference through Semantic Memory** 2362
 Yun Li (*University of Science and Technology of China*), Chen Zhang (*Damo Academy, Alibaba Group*),
 Shihao Han (*Rose-Hulman Institute of Technology*), Li Lina Zhang (*Microsoft Research*),
 Baoqun Yin (*University of Science and Technology of China*),
 Yunxin Liu (*Institute for AI Industry Research (AIR), Tsinghua University*),
 Mengwei Xu (*State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and
 Telecommunications*)
- **Knowing When to Quit: Selective Cascaded Regression with Patch Attention
 for Real-Time Face Alignment** 2372
 Gil Shapira (*Samsung Semiconductor Israel R&D Center (SIRC) & Bar-Ilan University*),
 Noga Levy, Ishay Goldin, Roy J. Jevnisek (*Samsung Semiconductor Israel R&D Center (SIRC)*)
- **End-to-end Boundary Exploration for Weakly-supervised Semantic Segmentation** 2381
 Jianjun Chen (*Institute of Information Engineering, Chinese Academy of Sciences & School of Cyber Security,
 University of Chinese Academy of Sciences*),
 Shancheng Fang, Hongtao Xie, Zheng-Jun Zha (*University of Science and Technology of China*),
 Yue Hu (*Institute of Information Engineering, Chinese Academy of Sciences & School of Cyber Security, University
 of Chinese Academy of Sciences*),
 Jianlong Tan (*Institute of Information Engineering, Chinese Academy of Sciences & School of Cyber Security,
 University of Chinese Academy of Sciences*)
- **SFE-Net: EEG-based Emotion Recognition with Symmetrical Spatial Feature Extraction** 2391
 Xiangwen Deng, Junlin Zhu, Shangming Yang (*University of Electronic Science and Technology of China*)
- **Bridging the Gap between Low-Light Scenes: Bilevel Learning for Fast Adaptation**..... 2401
 Dian Jin, Long Ma, Risheng Liu, Xin Fan (*Dalian University of Technology*)
- **Handling Difficult Labels for Multi-label Image Classification
 via Uncertainty Distillation**..... 2410
 Liangchen Song, Jialian Wu (*University at Buffalo*),
 Ming Yang, Qian Zhang (*Horizon Robotics*), Yuan Li (*Google*), Junsong Yuan (*University at Buffalo*)
- **Perception-Oriented Stereo Image Super-Resolution**..... 2420
 Chenxi Ma, Bo Yan, Weimin Tan, Xuhao Jiang (*Fudan University*)
- **ReLLIE: Deep Reinforcement Learning for Customized Low-Light Image Enhancement**..... 2429
 Rongkai Zhang, Lanqing Guo, Siyu Huang, Bihan Wen (*Nanyang Technological University*)
- **Intrinsic Temporal Regularization for High-resolution Human Video Synthesis**..... 2438
 Lingbo Yang (*Peking University & Alibaba Group*), Zhanning Gao (*Alibaba Group*),
 Siwei Ma, Wen Gao (*Peking University*)
- **A2W: Context-Aware Recommendation System for Mobile Augmented
 Reality Web Browser**..... 2447
 Kit Yung Lam (*Hong Kong University of Science and Technology*),
 Lik Hang Lee (*Korea Advanced Institute of Science and Technology*),
 Pan Hui (*Hong Kong University of Science and Technology*)
- **Cross-modal Self-Supervised Learning for Lip Reading: When Contrastive
 Learning meets Adversarial Training** 2456
 Changchong Sheng, Matti Pietikäinen (*University of Oulu*), Qi Tian (*Xidian University*),
 Li Liu (*University of Oulu*)
- **OsGG-Net: One-step Graph Generation Network for Unbiased Head Pose Estimation**..... 2465
 Shentong Mo (*Carnegie Mellon University*), Xin Miao (*Institute of Automation, Chinese Academy of Sciences*)
- **Multi-Modal Multi-Instance Learning for Retinal Disease Recognition** 2474
 Xirong Li (*Renmin University of China*), Yang Zhou (*Beijing Visionary Intelligence Ltd.*),
 Jie Wang, Hailan Lin (*Renmin University of China*),
 Jianchun Zhao, Dayong Ding (*Beijing Visionary Intelligence Ltd.*),
 Weihong Yu, Youxin Chen (*Peking Union Medical College Hospital*)
- **Locally Adaptive Structure and Texture Similarity for Image Quality Assessment**..... 2483
 Keyan Ding (*City University of Hong Kong*), Yi Liu, Xueyi Zou (*Huawei Technologies*),
 Shiqi Wang, Kede Ma (*City University of Hong Kong*)
- **CALLip: Lipreading using Contrastive and Attribute Learning** 2492
 Yiyang Huang, Xuefeng Liang, Chaowei Fang (*Xidian University*)

- **Cross-Modal Recipe Embeddings by Disentangling Recipe Contents and Dish Styles**..... 2501
Yu Sugiyama, Keiji Yanai (*The University of Electro-Communications*)
- **TDI TextSpotter: Taking Data Imbalance into Account in Scene Text Spotting** 2510
Yu Zhou, Hongtao Xie, Shancheng Fang (*University of Science and Technology of China*),
Jing Wang (*Huawei Cloud & AI*),
Zhengjun Zha, Yongdong Zhang (*University of Science and Technology of China*)
- **Position-Augmented Transformers with Entity-Aligned Mesh for TextVQA** 2519
Xuanyu Zhang, Qing Yang (*Du Xiaoman Financial*)
- **Learning Contextual Transformer Network for Image Inpainting** 2529
Ye Deng, Siqi Hui (*Xi'an Jiaotong University*),
Sanping Zhou (*Xi'an Jiaotong University & Shunan Academy of Artificial Intelligence*),
Deyu Meng, Jinjun Wang (*Xi'an Jiaotong University*)
- **Milliseconds Color Stippling** 2539
Lei Ma (*Peking University & Beijing Academy of Artificial Intelligence*),
Jian Shi (*Institute of Automation, Chinese Academy of Sciences*),
Yanyun Chen (*Institute of Software, Chinese Academy of Sciences*)
- **AFD-Net: Adaptive Fully-Dual Network for Few-Shot Object Detection** 2549
Longyao Liu, Bo Ma, Yulin Zhang, Xin Yi, Haozhi Li (*Beijing Institute of Technology*)
- **Missing Data Imputation for Solar Yield Prediction using Temporal Multi-Modal Variational Auto-Encoder**..... 2558
Meng Shen, Huaizheng Zhang, Yixin Cao, Fan Yang, Yonggang Wen (*Nanyang Technological University*)
- **Understanding Chinese Video and Language via Contrastive Multimodal Pre-Training** 2567
Chenyi Lei (*University of Science and Technology of China & Alibaba Group*), Shixian Luo (*Alibaba Group*),
Yong Liu (*Nanyang Technological University*),
Wanggui He, Jiamang Wang, Guoxin Wang, Haihong Tang (*Alibaba Group*),
Chunyan Miao (*Nanyang Technological University*),
Houqiang Li (*University of Science and Technology of China*)
- **DehazeFlow: Multi-scale Conditional Flow Network for Single Image Dehazing**..... 2577
Hongyu Li (*Beihang University*), Jia Li (*Beihang University & National Astronomical Observatories, Chinese Academy of Sciences & Peng Cheng Laboratory*),
Dong Zhao (*Beihang University*),
Long Xu (*National Astronomical Observatories, Chinese Academy of Sciences & Peng Cheng Laboratory*)
- **GCM-Net: Towards Effective Global Context Modeling for Image Inpainting** 2586
Huan Zheng, Zhao Zhang, Yang Wang (*Hefei University of Technology*),
Zheng Zhang (*Harbin Institute of Technology, Shenzhen*), Mingliang Xu (*Zhengzhou University*),
Yi Yang (*University of Technology Sydney*), Meng Wang (*Hefei University of Technology*)
- **Embracing the Dark Knowledge: Domain Generalization Using Regularized Knowledge Distillation**..... 2595
Yufei Wang (*Nanyang Technological University*), Haoqiang Li (*City University of Hong Kong*),
Lap-pui Chau, Alex C. Kot (*Nanyang Technological University*)
- **Cluster and Scatter: A Multi-grained Active Semi-supervised Learning Framework for Scalable Person Re-identification**..... 2605
Bingyu Hu, Zheng-Jun Zha, Jiawei Liu, Xierong Zhu, Hongtao Xie (*Univ. of Science and Technology of China*)
- **Dual Graph Convolutional Networks with Transformer and Curriculum Learning for Image Captioning** 2615
Xinzhi Dong (*Wuhan University*), Chengjiang Long (*JD Finance America Corporation*),
Wenju Xu (*InnoPeak Technology Inc.*), Chunxia Xiao (*Wuhan University*)
- **Build Your Own Bundle - A Neural Combinatorial Optimization Method** 2625
Qilin Deng, Kai Wang, Minghao Zhao, Runze Wu, Yu Ding, Zhene Zou, Yue Shang, Jianrong Tao,
Changjie Fan (*NetEase Games*)
- **Unsupervised Image Deraining: Optimization Model Driven Deep CNN** 2634
Changfeng Yu (*Huazhong University of Science and Technology*), Yi Chang (*Peng Cheng Laboratory*),
Yi Li (*Huazhong University of Science and Technology*),
Xi-Le Zhao (*University of Electronic Science and Technology of China*),
Luxin Yan (*Huazhong University of Science and Technology*),

Keynote Talks III & IV

- **Do you see what I see? Large-scale Learning from Multimodal Videos** 2643
Cordelia Schmid (*INRIA/Google*)
- **Large-scale Multi-Modality Pretrained Models: Applications and Experiences** 2644
Jingren Zhou (*Ant Group*)

Session 17: Multimodal Fusion and Embedding-I

- **Multi-Source Fusion and Automatic Predictor Selection for Zero-Shot Video Object Segmentation** 2645
Xiaoqi Zhao, Youwei Pang, Jiaying Yang, Lihe Zhang (*Dalian University of Technology*),
Huchuan Lu (*Dalian University of Technology & Pengcheng Lab*)
- **Self-supervised Consensus Representation Learning for Attributed Graph** 2654
Changshu Liu (*University of Electronic Science and Technology of China*),
Liangjian Wen (*Huawei Technologies Company Limited*),
Zhao Kang, Guangchun Luo, Ling Tian (*University of Electronic Science and Technology of China*)
- **Efficient Multi-Modal Fusion with Diversity Analysis** 2663
Shuhui Qu, Yan Kang, Janghwan Lee (*Samsung Display America Lab*)
- **GCCN: Geometric Constraint Co-attention Network for 6D Object Pose Estimation** 2671
Yongming Wen, Yiquan Fang, Junhao Cai, Kimwa Tung, Hui Cheng (*Sun Yat-Sen University*)
- **Cross-Modal Generalization: Learning in Low Resource Modalities via Meta-Alignment** 2680
Paul Pu Liang, Peter Wu (*Carnegie Mellon University*), Liu Ziyin (*University of Tokyo*),
Louis-Philippe Morency, Ruslan Salakhutdinov (*Carnegie Mellon University*)
- **Elastic Tactile Simulation Towards Tactile-Visual Perception** 2690
Yikai Wang (*Tsinghua University*), Wenbing Huang (*Tsinghua University & Pazhou Laboratory*),
Bin Fang, Fuchun Sun (*Tsinghua University*), Chang Li (*JD Explore Academy*)

Session 18: Multimodal Fusion and Embedding-II

- **A Novel Patch Convolutional Neural Network for View-based 3D Model Retrieval** 2699
Zan Gao, Yuxiang Shao (*Qilu University of Technology*), Weili Guan (*Monash University*),
Meng Liu (*Shandong Jianzhu University*), Zhiyong Cheng (*Shandong AI Institute*),
Shengyong Chen (*Tianjin University of Technology*)
- **Semi-Autoregressive Image Captioning** 2708
Xu Yan, Zhengcong Fei, Zekang Li, Shuhui Wang, Qingming Huang (*Key Lab of Intell. Info. Process.,
Institute of Comput. Tech., CAS, Beijing, China*), Qi Tian (*Cloud BU, Huawei Technologies, Shenzhen, China*)
- **One-Stage Incomplete Multi-view Clustering via Late Fusion** 2717
Yi Zhang, Xinwang Liu, Siwei Wang, Jiyuan Liu, Sisi Dai, En Zhu (*National University of Defense Technology*)
- **Self-Representation Subspace Clustering for Incomplete Multi-view Data** 2726
Jiyuan Liu, Xinwang Liu, Yi Zhang, Pei Zhang, Wenxuan Tu, Siwei Wang, Sihang Zhou, Weixuan Liang,
Siqi Wang, Yuexiang Yang (*National University of Defense Technology*)
- **Is Visual Context Really Helpful for Knowledge Graph? A Representation Learning Perspective** 2735
Meng Wang (*Southeast University*), Sen Wang (*The University of Queensland*), Han Yang (*Peking University*),
Zheng Zhang (*Harbin Institute of Technology, Shenzhen*), Xi Chen (*Tencent*), Guilin Qi (*Southeast University*)
- **Knowledge Perceived Multi-modal Pretraining in E-commerce** 2744
Yushan Zhu (*Zhejiang University*), Huaixiao Zhao (*Alibaba Group*), Wen Zhang, Ganqiang Ye (*Zhejiang
University*), Hui Chen (*Alibaba Group*), Ningyu Zhang, Huajun Chen (*Zhejiang University*)

Session 19: Video Program and Demo Session

- **Text2Video: Automatic Video Generation Based on Text Scripts** 2753
Yipeng Yu, Zirui Tu, Longyu Lu, Xiao Chen, Hui Zhan, Zixun Sun (*Tencent*)

- **A System for Interactive and Intelligent AD Auxiliary Screening** 2756
Sen Yang, Qike Zhao, Lanxin Miao, Min Chen, Lianli Gao (*University of Electronic Science and Technology of China*), Jingkuan Song (*University of Electronic Science and Technology of China & Institute of Neurology, Sichuan Academy of Medical Sciences & Provincial People's Hospital*), Weidong Le (*Institute of Neurology, Sichuan Academy of Medical Sciences & Provincial People's Hospital*)
- **Move As You Like: Image Animation in E-Commerce Scenario** 2759
Borun Xu (*University of Electronic Science and Technology of China*), Biao Wang (*Alibaba Group*), Jiale Tao (*University of Electronic Science and Technology of China*), Tiezheng Ge, Yuning Jiang (*Alibaba Group*), Wen Li, Lixin Duan (*University of Electronic Science and Technology of China*)
- **MDMS: Music Data Matching System for Query Variant Retrieval** 2762
Rinita Roy, Ruben Mayer (*Technical University of Munich*), Hans-Arno Jacobsen (*University of Toronto*)
- **Community Generated VR Painting using Eye Gaze** 2765
Mu Mu, Murtada Dohan (*University of Northampton*)
- **Sync Glass: Virtual Pouring and Toasting Experience with Multimodal Presentation** 2768
Yuki Tajima, Toshiharu Horiuchi, Gen Hattori (*KDDI Research, Inc.*)
- **VideoDiscovery: An Automatic Short-Video Generation System for E-commerce Live-streaming** 2771
Yanhao Zhang, Qiang Wang, Yun Zheng, Pan Pan, Yinghui Xu (*Alibaba Group*)
- **SmartSales: An AI-Powered Telemarketing Coaching System in FinTech** 2774
Yuanfeng Song (*The Hong Kong University of Science and Technology & WeBank Co., Ltd*), Xuefang Zhao, Di Jiang, Xiaoling Huang, Weiwei Zhao, Qian Xu (*WeBank Co., Ltd*), Raymond Chi-Wing Wong (*The Hong Kong University of Science and Technology*), Qiang Yang (*The Hong Kong University of Science and Technology & WeBank Co., Ltd*)
- **SmartMeeting: Automatic Meeting Transcription and Summarization for In-Person Conversations** 2777
Yuanfeng Song (*The Hong Kong University of Science and Technology & WeBank Co., Ltd*), Di Jiang, Xuefang Zhao, Xiaoling Huang, Qian Xu (*WeBank Co., Ltd*), Raymond Chi-Wing Wong (*The Hong Kong University of Science and Technology*), Qiang Yang (*The Hong Kong University of Science and Technology & WeBank Co., Ltd*)
- **Aesthetic Evaluation and Guidance for Mobile Photography** 2780
Hao Lou, Heng Huang, Chaoen Xiao, Xin Jin (*Beijing Electronic Science and Technology Institute*)
- **A Question Answering System for Unstructured Table Images** 2783
Wenyuan Xue, Siqi Cai, Wen Wang, Qingyong Li (*Beijing Jiaotong University*), Baosheng Yu (*The University of Sydney*), Yibing Zhan (*JD Explore Academy*), Dacheng Tao (*JD Explore Academy & The University of Sydney*)
- **Post2Story: Automatically Generating Storylines from Microblogging Platforms** 2786
Xujian Zhao, Chongwei Wang (*Southwest University of Science and Technology*), Peiquan Jin (*University of Science and Technology of China*), Hui Zhang, Chunming Yang, Bo Li (*Southwest University of Science and Technology*)
- **ViDA-MAN: Visual Dialog with Digital Humans** 2789
Tong Shen, Jiawei Zuo, Fan Shi (*JD AI Research*), Jin Zhang, Liqin Jiang (*Migu Culture Technology*), Meng Chen, Zhengchen Zhang, Wei Zhang, Xiaodong He, Tao Mei (*JD AI Research*)
- **A Picture is Worth a Thousand Words: A Unified System for Diverse Captions and Rich Images Generation** 2792
Yupan Huang (*Sun Yat-sen University*), Bei Liu, Jianlong Fu (*Microsoft Research Asia*), Yutong Lu (*Sun Yat-sen University*)
- **Softly: Simulated Empathic Touch between an Agent and a Human** 2795
Maxime Grandidier, Fabien Boucaud, Indira Thouvenin (*Université de Technologie de Compiègne*), Catherine Pelachaud (*Sorbonne Université*)
- **RecipeLog: Recipe Authoring App for Accurate Food Recording** 2798
Akihisa Ishino, Yoko Yamakata (*The University of Tokyo*), Hiroaki Karasawa (*Hongo Software Development*), Kiyoharu Aizawa (*The University of Tokyo*)

- **iART: A Search Engine for Art-Historical Images to Support Research in the Humanities ...** 2801
Matthias Springstein (*TIB – Leibniz Information Centre for Science and Technology*),
Stefanie Schneider (*Ludwig Maximilian University of Munich*), Javad Rahnama (*University Paderborn*),
Eyke Hüllermeier, Hubertus Kohle (*Ludwig Maximilian University of Munich*),
Ralph Ewerth (*TIB – Leibniz Information Center for Science and Technology*)
- **ArtiVisual: A Platform to Generate and Compare Art** 2804
Jardenna Mohazzab, Abe Vos, Jonathan van Westendorp, Lucas Lageweg, Dylan Prins, Aritra Bhowmik
(*University of Amsterdam*)
- **GCNIllustrator: Illustrating the Effect of Hyperparameters
on Graph Convolutional Networks** 2807
Ivona Najdenkoska, Jeroen den Boef, Thomas Schneider, Justo van der Werf, Reinier de Ridder,
Fajar Fathurrahman, Marcel Worring (*University of Amsterdam*)
- **On-demand Action Detection System using Pose Information** 2810
Noboru Yoshida, Jianquan Liu (*NEC Corporation*)
- **APF: An Adversarial Privacy-preserving Filter to Protect Portrait Information** 2813
Xian Zhao, Jiaming Zhang, Xiaowen Huang (*Beijing Jiaotong University*)
- **Text-driven 3D Avatar Animation with Emotional and Expressive Behaviors** 2816
Li Hu, Jinwei Qi, Bang Zhang, Pan Pan, Yinghui Xu (*Alibaba Group*)
- **Text to Scene: A System of Configurable 3D Indoor Scene Synthesis** 2819
Xinyan Yang (*Ministry of Education*),
Fei Hu, Long Ye (*Communication University of China*)
- **MovieREP: A New Movie Reproduction Framework for Film Soundtrack** 2822
Ruiqi Wang, Long Ye, Qin Zhang (*Communication University of China*)

Session 20: Multimodal Fusion and Embedding-III

- **DSP: Dual Soft-Paste for Unsupervised Domain Adaptive Semantic Segmentation** 2825
Li Gao (*Wuhan University*), Jing Zhang (*The University of Sydney*),
Lefei Zhang (*Wuhan University*), Dacheng Tao (*JD Explore Academy*)
- **Generating Point Cloud from Single Image in The Few Shot Scenario** 2834
Yu Lin, Jinghui Guo, Yang Gao, Yi-fan Li, Zhuoyi Wang, Latifur Khan (*University of Texas at Dallas*)
- **Product-oriented Machine Translation with Cross-modal Cross-lingual Pre-training** 2843
Yuqing Song (*Renmin University of China*), Shizhe Chen (*INRIA*), Qin Jin (*Renmin University of China*),
Wei Luo, Jun Xie, Fei Huang (*Alibaba Damo Academy*)
- **Pre-training Graph Transformer with Multimodal Side Information
for Recommendation** 2853
Yong Liu (*Nanyang Technological University*), Susen Yang (*Alibaba Group*),
Chenyi Lei (*Alibaba Group & University of Science and Technology of China*),
Guoxin Wang (*Alibaba Group & Zhejiang University*),
Haihong Tang (*Alibaba Group*), Juyong Zhang (*University of Science and Technology of China*),
Aixin Sun, Chunyan Miao (*Nanyang Technological University*)
- **Learning Disentangled Factors from Paired Data in Cross-Modal Retrieval:
An Implicit Identifiable VAE Approach** 2862
Minyoung Kim, Ricardo Guerrero, Vladimir Pavlovic (*Samsung AI Center Cambridge*),
- **Progressive Graph Attention Network for Video Question Answering** 2871
Liang Peng, Shuangji Yang, Yi Bin, Guoqing Wang (*University of Electronic Science and Technology of China*)

Session 21: Media Interpretation-I

- **Mix-order Attention Networks for Image Restoration** 2880
Tao Dai (*Shenzhen University*), Yalei Lv, Bin Chen, Zhi Wang (*Tsinghua University*),
Zexuan Zhu (*Shenzhen University*), Shu-Tao Xia (*Tsinghua University*)
- **Vehicle Counting Network with Attention-based Mask Refinement
and Spatial-awareness Block Loss** 2889
Ji Zhang, Jian-Jun Qiao, Xiao Wu, Wei Li (*Southwest Jiaotong University*)

- **DPT: Deformable Patch-based Transformer for Visual Recognition** 2899
Zhiyang Chen (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Yousong Zhu, Chaoyang Zhao (*Institute of Automation, Chinese Academy of Sciences*), Guosheng Hu (*AnyVision*), Wei Zeng (*Peking University*), Jinqiao Wang (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Ming Tang (*Institute of Automation, Chinese Academy of Sciences*)
- **Scene Text Image Super-Resolution via Parallely Contextual Attention Network**..... 2908
Cairong Zhao, Shuyang Feng (*Tongji University*), Brian Nlong Zhao (*University of Southern California*), Zhijun Ding (*Tongji University*), Jun Wu (*Fudan University*), Fumin Shen, Heng Tao Shen (*University of Electronic Science and Technology of China*)
- **Improving Pedestrian Detection from a Long-tailed Domain Perspective** 2918
Mengyuan Ding, Shanshan Zhang, Jian Yang (*Nanjing University of Science and Technology*)
- **Robust Shadow Detection by Exploring Effective Shadow Contexts**..... 2927
Xianyong Fang, Xiaohao He, Linbo Wang (*Anhui University*), Jianbing Shen (*University of Macau*)

Session 22: Doctoral Symposium

- **End-to-end Quality of Experience Evaluation for HTTP Adaptive Streaming** 2936
Babak Taraghi (*Alpen-Adria-Universität Klagenfurt*)
- **Generative Adversarial Network for Text-to-Face Synthesis and Manipulation**..... 2940
Yutong Zhou (*Ritsumeikan University*)
- **GAN-aided Serial Dependence Study in Medical Image Perception** 2945
Zhihang Ren (*University of California, Berkeley*)
- **Image Style Transfer with Generative Adversarial Networks** 2950
Ru Li (*University of Electronic Science and Technology of China*)
- **Annotation-Efficient Semantic Segmentation with Shape Prior Knowledge**..... 2955
Yuhang Lu (*University of South Carolina*)
- **Neural-based Rendering and Application** 2960
Peng Dai (*The University of Hong Kong*)
- **Towards Bridging Video and Language by Caption Generation and Sentence Localization** 2964
Shaoliang Chen (*Fudan University*)
- **Situational Anomaly Detection in Multimedia Data under Concept Drift** 2969
Pratibha Kumari (*Indian Institute of Technology Ropar*)
- **Dynamic Knowledge Distillation with Cross-Modality Knowledge Transfer** 2974
Guangzhi Wang (*National University of Singapore*)

Session 23: Media Interpretation-II

- **WeClick: Weakly-Supervised Video Semantic Segmentation with Click Annotations** 2995
Peidong Liu, Zibin He, Xiyu Yan (*Tsinghua University*), Yong Jiang, Shu-Tao Xia (*Tsinghua University & Peng Cheng Laboratory*), Feng Zheng (*Southern University of Science and Technology*), Hu Maowei (*Tsinghua University & Shenzhen Rejoice Sport Tech. Co., LTD*)
- **Towards Cross-Granularity Few-Shot Learning: Coarse-to-Fine Pseudo-Labeling with Visual-Semantic Meta-Embedding**..... 3005
Jinhai Yang, Hua Yang, Lin Chen (*Shanghai Jiao Tong University & Shanghai Key Lab of Digital Media Processing and Transmission*)
- **Disentangled Representation Learning and Enhancement Network for Single Image De-Raining** 3015
Guoqing Wang (*University of Electronic Science and Technology of China*), Changming Sun (*CSIRO Data61*), Xing Xu, Jingjing Li, Zheng Wang, Zeyu Ma (*University of Electronic Science and Technology of China*)

- **Towards Robust Cross-domain Image Understanding with Unsupervised Noise Removal ..** 3024
Lei Zhu, Zhaojing Luo, Wei Wang (*National University of Singapore*),
Meihui Zhang (*Beijing Institute of Technology*), Gang Chen (*Zhejiang University*),
Kaiping Zheng (*National University of Singapore*)
- **Exploiting BERT for Multimodal Target Sentiment Classification through Input Space Translation** 3034
Zaid Khan, Yun Fu (*Northeastern University*)
- **Video Representation Learning with Graph Contrastive Augmentation** 3043
Jingran Zhang, Xing Xu, Fumin Shen (*University of Electronic Science and Technology of China*),
Yazhou Yao (*Nanjing University of Science and Technology*),
Jie Shao, Xiaofeng Zhu (*University of Electronic Science and Technology of China*)

Poster Session 4

- **An EM Framework for Online Incremental Learning of Semantic Segmentation** 3052
Shipeng Yan (*ShanghaiTech University; Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences; & University of Chinese Academy of Sciences*),
Jiale Zhou, Jiangwei Xie, Songyang Zhang (*ShanghaiTech University*),
Xuming He (*ShanghaiTech University & Shanghai Engineering Research Center of Intelligent Vision and Imaging*)
- **I2V-GAN: Unpaired Infrared-to-Visible Video Translation** 3061
Shuang Li, Bingfeng Han, Zhenjie Yu, Chi Harold Liu (*Beijing Institute of Technology*),
Kai Chen, Shuigen Wang (*Yantai IRay Technologies Lt. Co.*)
- **Implicit Feedbacks are Not Always Favorable: Iterative Relabeled One-Class Collaborative Filtering against Noisy Interactions** 3070
Zitai Wang (*Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Qianqian Xu (*Institute of Computing Technology, Chinese Academy of Sciences*),
Zhiyong Yang (*Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Xiaochun Cao (*Institute of Information Engineering, Chinese Academy of Sciences & Sun Yat-sen University*),
Qingming Huang (*University of Chinese Academy of Sciences & Institute of Computing Technology*)
- **InsPose: Instance-Aware Networks for Single-Stage Multi-Person Pose Estimation.....** 3079
Dahu Shi (*Hikvision Research Institute*), Xing Wei (*Xi'an Jiaotong University*),
Xiaodong Yu, Wenming Tan, Ye Ren, Shiliang Pu (*Hikvision Research Institute*)
- **Implicit Feature Refinement for Instance Segmentation** 3088
Lufan Ma (*Tsinghua University*), Tiancai Wang, Bin Dong (*MEGVII Technology*),
Jiangpeng Yan, Xiu Li (*Tsinghua University*), Xiangyu Zhang (*MEGVII Technology*)
- **Question-controlled Text-aware Image Captioning** 3097
Anwen Hu (*Renming University of China*), Shizhe Chen (*INRIA*), Qin Jin (*Renmin University of China*)
- **Style-Aware Image Recommendation for Social Media Marketing** 3106
Yiwei Zhang, Toshihiko Yamasaki (*The University of Tokyo*)
- **WePerson: Learning a Generalized Re-identification Model from All-weather Virtual Data** 3115
He Li, Mang Ye, Bo Du (*Wuhan University*)
- **Polar Ray: A Single-stage Angle-free Detector for Oriented Object Detection in Aerial Images.....** 3124
Shuai Liu, Lu Zhang, Shuai Hao (*Dalian University of Technology*),
Huchuan Lu (*Dalian University of Technology & Peng Cheng Laboratory*),
You He (*Naval Aeronautical University*)
- **Self-Contrastive Learning with Hard Negative Sampling for Self-supervised Point Cloud Learning** 3133
Bi'an Du, Xiang Gao, Wei Hu (*Peking University*), Xin Li (*West Virginia University*)
- **Generally Boosting Few-Shot Learning with HandCrafted Features.....** 3143
Yi Zhang (*Chongqing University*),
Sheng Huang (*Chongqing University & Ministry of Education Key Laboratory of Dependable Service Computing in Cyber Physical Society*), Fengtao Zhou (*Chongqing University*)

- **ROECS: A Robust Semi-direct Pipeline Towards Online Extrinsic Correction of the Surround-view System** 3153
Tianjun Zhang (*Tongji University*), Nlong Zhao (*University of Southern California*), Ying Shen, Xuan Shao, Lin Zhang (*Tongji University*), Yicong Zhou (*University of Macau*)
- **Pseudo Graph Convolutional Network for Vehicle ReID** 3162
Wen Qian (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Zhiqun He (*SenseTime Group Limited*), Silong Peng, Chen Chen (*Institute of Automation, Chinese Academy of Sciences*), Wei Wu (*SenseTime Group Limited*)
- **Towards Fast and High-Quality Sign Language Production** 3172
Wencan Huang, Wenwen Pan, Zhou Zhao (*Zhejiang University*), Qi Tian (*Huawei Cloud & AI*)
- **Effective De-identification Generative Adversarial Network for Face Anonymization** 3182
Zhenzhong Kuang, Huigui Liu, Jun Yu (*Hangzhou Dianzi University*), Aikui Tian, Lei Wang (*Shandong University of Technology*), Jianping Fan (*AI Lab at Lenovo Research*), Noboru Babaguchi (*Osaka University*)
- **Cross-modal Retrieval and Synthesis (X-MRS): Closing the Modality Gap in Shared Subspace Learning** 3192
Ricardo Guerrero, Hai X. Pham, Vladimir Pavlovic (*Samsung AI Center*)
- **When Face Completion Meets Irregular Holes: An Attributes Guided Deep Inpainting Network** 3202
Jie Xiao, Dandan Zhan, Haoran Qi (*Sun Yat-Sen University*), Zhi Jin (*Sun Yat-Sen University & Guangdong Provincial Key Laboratory of Fire Science and Technology*)
- **Non-Linear Fusion for Self-Paced Multi-View Clustering** 3211
Zongmo Huang (*University of Electronic Science and Technology of China*), Yazhou Ren (*University of Electronic Science and Technology of China & Institute of Electronic and Info, Engineering of UESTC in Guangdong*), Xiaorong Pu (*University of Electronic Science and Technology of China*), Lifang He (*Lehigh University*)
- **Counterfactual Debiasing Inference for Compositional Action Recognition** 3220
Pengzhan Sun (*University of Electronic Science and Technology of China*), Bo Wu (*MIT-IBM Watson AI Lab*), Xunsong Li, Wen Li, Lixin Duan (*University of Electronic Science and Technology of China*), Chuang Gan (*MIT-IBM Watson AI Lab*)
- **STST: Spatial-Temporal Specialized Transformer for Skeleton-based Action Recognition ...** 3229
Yuhan Zhang (*University of Electronic Science and Technology of China*), Bo Wu (*MIT-IBM Watson AI Lab*), Wen Li, Lixin Duan (*University of Electronic Science and Technology of China*), Chuang Gan (*MIT-IBM Watson AI Lab*)
- **Exploring Gradient Flow Based Saliency for DNN Model Compression**..... 3238
Xinyu Liu (*City University of Hong Kong*), Baopu Li (*Baidu USA LLC*), Zhen Chen, Yixuan Yuan (*City University of Hong Kong*)
- **An Adaptive Iterative Inpainting Method with More Information Exploration**..... 3247
Shengjie Chen, Zhenhua Guo, Bo Yuan (*Tsinghua University*)
- **Assisting News Media Editors with Cohesive Visual Storylines** 3257
Gonçalo Marcelino, David Semedo, André Mourão (*Universidade NOVA de Lisboa*), Saverio Blasi (*BBC Research and Development*), João Magalhães (*Universidade NOVA de Lisboa*), Marta Mrak (*BBC Research and Development*)
- **MM-Flow: Multi-modal Flow Network for Point Cloud Completion**..... 3266
Yiqiang Zhao, Yiyao Zhou, Rui Chen, Bin Hu, Xiding Ai (*Tianjin University*)
- **Long-tailed Distribution Adaptation** 3275
Zhiliang Peng, Wei Huang, Zonghao Guo, Xiaosong Zhang, Jianbin Jiao, Qixiang Ye (*University of Chinese Academy of Sciences*)
- **Lesion-Inspired Denoising Network: Connecting Medical Image Denoising and Lesion Detection** 3283
Kecheng Chen, Kun Long, Yazhou Ren (*University of Electronic Science and Technology of China*), Jiayu Sun (*West China Hospital of SiChuan University*), Xiaorong Pu (*University of Electronic Science and Technology of China*)

- **Domain Adaptive Semantic Segmentation without Source Data**..... 3293
Fuming You, Jingjing Li (*University of Electronic Science and Technology of China*),
Lei Zhu (*Shandong Normal University*), Zhi Chen, Zi Huang (*University of Queensland*)
- **Cross-modal Joint Prediction and Alignment for Composed Query Image Retrieval** 3303
Yuchen Yang (*University of Science and Technology of China*),
Min Wang (*Institute of Artificial Intelligence, Hefei Comprehensive National Science Center*),
Wengang Zhou, Houqiang Li (*University of Science and Technology of China & Institute of Artificial Intelligence, Hefei Comprehensive National Science Center*)
- **JDMAN: Joint Discriminative and Mutual Adaptation Networks for Cross-Domain Facial Expression Recognition** 3312
Yingjian Li, Yingnan Gao, Bingzhi Chen (*Harbin Institute of Technology, Shenzhen*),
Zheng Zhang (*Harbin Institute of Technology, Shenzhen & Peng Cheng Laboratory*),
Lei Zhu (*Shandong Normal University*), Guangming Lu (*Harbin Institute of Technology, Shenzhen*)
- **Improving Weakly Supervised Object Localization via Causal Intervention** 3321
Feifei Shao (*Zhejiang University*), Yawei Luo (*Zhejiang University & Baidu Research*),
Li Zhang (*Zhejiang Insigma Digital Technology Co., Ltd.*),
Lu Ye (*Zhejiang University of Science and Technology*),
Siliang Tang, Yi Yang, Jun Xiao (*Zhejiang University*)
- **Imbalanced Source-free Domain Adaptation** 3330
Xinhao Li, Jingjing Li (*University of Electronic Science and Technology of China*),
Lei Zhu (*Shandong Normal University*),
Guoqing Wang (*University of Electronic Science and Technology of China*),
Zi Huang (*University of Queensland*)
- **Learning Transferrable and Interpretable Representations for Domain Generalization** 3340
Zhekai Du, Jingjing Li, Ke Lu (*University of Electronic Science and Technology of China*),
Lei Zhu (*Shandong Normal University*), Zi Huang (*University of Queensland*)
- **WAS-VTON: Warping Architecture Search for Virtual Try-on Network**..... 3350
Zhenyu Xie, Xujie Zhang, Fuwei Zhao (*Shenzhen Campus of Sun Yat-Sen University*),
Haoye Dong (*Sun Yat-sen University*), Michael C. Kampffmeyer (*UiT The Arctic University of Norway*),
Haonan Yan (*Momo Technology Company Limited*),
Xiaodan Liang (*Shenzhen Campus of Sun Yat-Sen University & DarkMatter AI Research*)
- **DFR-Net: A Novel Multi-Task Learning Network for Real-Time Multi-Instrument Segmentation** 3360
Yan-Jie Zhou (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Shi-Qi Liu (*Institute of Automation, Chinese Academy of Sciences*), Xiao-Liang Xie (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Zeng-Guang Hou (*Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; CAS Center for Excellence in Brain Science and Intelligence Technology; & MUST*)
- **From Superficial to Deep: Language Bias driven Curriculum Learning for Visual Question Answering** 3370
Mingrui Lao (*Leiden University*), Yanming Guo (*National University of Defense Technology*),
Yu Liu (*Dalian University of Technology*), Wei Chen, Nan Pu, Michael S. Lew (*Leiden University*)
- **Pairwise Emotional Relationship Recognition in Drama Videos: Dataset and Benchmark**..... 3380
Xun Gao, Yin Zhao, Jie Zhang, Longjun Cai (*Alibaba Group*)
- **Block Popularity Prediction for Multimedia Storage Systems Using Spatial-Temporal-Sequential Neural Networks** 3390
Yingying Cheng, Fan Zhang, Gang Hu, Yiwen Wang, Hanhui Yang,
Gong Zhang (*Huawei Technologies Co., Ltd.*),
Zhuo Cheng (*Tsinghua University, Huawei Technologies Co., Ltd.*)
- **Transferrable Contrastive Learning for Visual Domain Adaptation** 3399
Yang Chen (*University of Science and Technology of China*),
Yingwei Pan, Yu Wang, Ting Yao (*JD AI Research*),
Xinmei Tian (*University of Science and Technology of China*), Tao Mei (*JD AI Research*)

- **Weighted Gaussian Loss based Hamming Hashing** 3409
Rong-Cheng Tu, Xian-Ling Mao, Cihang Kong, Zihang Shao, Ze-Lin Li (*Beijing Institute of Technology*),
Wei Wei (*Huazhong University of Science and Technology*), Heyan Huang (*Beijing Institute of Technology*)
- **Domain-Aware SE Network for Sketch-based Image Retrieval with Multiplicative
Euclidean Margin Softmax** 3418
Peng Lu, Gao Huang (*Tsinghua University*), Hangyu Lin (*Fudan University*),
Wenming Yang (*Tsinghua University*),
Guodong Guo (*IDL, Baidu Research & National Engineering Lab for Deep Learning Technology and Application*),
Yanwei Fu (*Fudan University*)
- **FTAFace: Context-enhanced Face Detector with Fine-grained Task Attention** 3427
Deyu Wang, Dongchao Wen, Wei Tao, Lingxiao Yin (*Canon Innovative Solution (Beijing) Co., Ltd.*),
Tse-Wei Chen, Tadayuki Ito, Kinya Osa, Masami Kato (*Canon Inc.*)
- **Identity-aware Graph Memory Network for Action Detection** 3437
Jingcheng Ni (*Beihang University*), Jie Qin (*Nanjing University of Aeronautics and Astronautics*),
Di Huang (*Beihang University*)
- **Improving Robustness and Accuracy via Relative Information Encoding
in 3D Human Pose Estimation** 3446
Wenkang Shan (*Peking University*), Haopeng Lu (*Shanghai Jiao Tong University*),
Shanshe Wang (*Peking University*), Xinfeng Zhang (*University of Chinese Academy of Sciences*),
Wen Gao (*Peking University*)
- **Deep Neural Network Retrieval** 3455
Nan Zhong, Zhenxing Qian, Xinpeng Zhang (*Fudan University*)
- **Adversarial Learning with Mask Reconstruction for Text-Guided Image Inpainting** 3464
Xingcai Wu, Yucheng Xie, Jiaqi Zeng, Zhenguo Yang (*Guangdong University of Technology*),
Yi Yu (*National Institute of Informatics*), Qing Li (*The Hong Kong Polytechnic University*),
Wenyin Liu (*Guangdong University of Technology & Peng Cheng Laboratory*)
- **Spatiotemporal Inconsistency Learning for DeepFake Video Detection** 3473
Zhihao Gu (*Shanghai Jiao Tong University*),
Yang Chen, Taiping Yao, Shouhong Ding, Jilin Li, Feiyue Huang (*Tencent Youtu Lab*),
Lizhuang Ma (*Shanghai Jiao Tong University*)
- **VeloCity: Using Voice Assistants for Cyclists to Provide Traffic Reports** 3482
Gian-Luca Savino (*University of Bremen & University of St. Gallen*), Jessé Moraes Braga (*University of Bremen*),
Johannes Schöning (*University of Bremen & University of St. Gallen*)
- **Edit Like A Designer: Modeling Design Workflows for Unaligned Fashion Editing** 3492
Qiyu Dai, Shuai Yang, Wenjing Wang (*Peking University*), Wei Xiang (*Bigo*),
Jiaying Liu (*Peking University*)
- **Privacy-Preserving Portrait Matting** 3501
Jizhizi Li, Sihan Ma, Jing Zhang (*The University of Sydney*),
Dacheng Tao (*JD Explore Academy & The University of Sydney*)
- **A Transformer based Approach for Image Manipulation Chain Detection** 3510
Jiaxiang You, Yuanman Li (*Shenzhen University*), Jiantao Zhou (*University of Macau*),
Zhongyun Hua (*Harbin Institute of Technology, Shenzhen*),
Weiwei Sun (*University of Macau*), Xia Li (*Shenzhen University*)
- **HANet: Hierarchical Alignment Networks for Video-Text Retrieval** 3518
Peng Wu (*Xidian University*), Xiangteng He, Mingqian Tang, Yiliang Lv (*Alibaba Group*),
Jing Liu (*Xidian University*)
- **Scalable Multi-view Subspace Clustering with Unified Anchors** 3528
Mengjing Sun, Pei Zhang, Siwei Wang, Sihang Zhou, Wenxuan Tu, Xinwang Liu,
En Zhu, Changjian Wang (*National University of Defense Technology*)
- **PRNet: A Progressive Recovery Network for Revealing Perceptually Encrypted Images** 3537
Tao Xiang, Ying Yang, Shangwei Guo, Hangcheng Liu (*Chongqing University*),
Hantao Liu (*Cardiff University*)

- **FakeTagger: Robust Safeguards against DeepFake Dissemination via Provenance Tracking** 3546
Run Wang (*Wuhan University & Ministry of Education*), Felix Juefei-Xu (*Alibaba Group, USA*), Meng Luo (*Northeastern University*), Yang Liu (*Nanyang Technological University*), Lina Wang (*Wuhan University & Ministry of Education*)
- **Discriminative Latent Semantic Graph for Video Captioning** 3556
Yang Bai (*Newcastle University*), Junyan Wang (*University of New South Wales*), Yang Long (*Durham University*), Bingzhang Hu (*Hefei CAS Dihuge Automation Co., LTD*), Yang Song, Maurice Pagnucco (*University of New South Wales*), Yu Guan (*Newcastle University*)
- **From Image to Imuge: Immunized Image Generation** 3565
Qichao Ying, Zhenxing Qian (*Fudan University*), Hang Zhou (*Simon Fraser University*), Haisheng Xu (*NVIDIA*), Xinpeng Zhang (*Fudan University*), Siyi Li (*NVIDIA*)
- **Wisdom of (Binned) Crowds: A Bayesian Stratification Paradigm for Crowd Counting** 3574
Sravya Vardhani Shivapuja, Mansi Pradeep Khamkar, Divij Bajaj (*International Institute of Information Technology, Hyderabad*), Ganesh Ramakrishnan (*Indian Institute of Technology Bombay*), Ravi Kiran Sarvadevabhatla (*International Institute of Information Technology, Hyderabad*)
- **Demystifying Commercial Video Conferencing Applications** 3583
Insoo Lee, Jinsung Lee (*University of Colorado Boulder*), Kyunghan Lee (*Seoul National University*), Dirk Grunwald, Sangtae Ha (*University of Colorado Boulder*)
- **LightFEC: Network Adaptive FEC with a Lightweight Deep-Learning Approach** 3592
Han Hu, Sheng Cheng, Xinggong Zhang, Zongming Guo (*Peking University*)
- **SOGAN: 3D-Aware Shadow and Occlusion Robust GAN for Makeup Transfer** 3601
Yueming Lyu (*Institute of Automation, Chinese Academy of Sciences & Univ. of Chinese Academy of Sciences*), Jing Dong, Bo Peng, Wei Wang, Tieniu Tan (*Institute of Automation, Chinese Academy of Sciences*)

Reproducibility Papers

- **Reproducibility Companion Paper: Campus3D: A Photogrammetry Point Cloud Benchmark for Outdoor Scene Hierarchical Understanding** 3610
Yuqing Liao (*Wuhan University*), Xinke Li (*National University of Singapore*), Zekun Tong, Yabang Zhao (*National University of Singapore*), Andrew Lim (*Southwest Jiaotong University*), Zhenzhong Kuang (*Hangzhou Dianzi University & Osaka University*), Cise Midoglu (*Simula Research Laboratory (SRL)*)
- **Reproducibility Companion Paper: Norm-in-Norm Loss with Faster Convergence and Better Performance for Image Quality Assessment** 3615
Dingquan Li (*Peking University & Peng Cheng Lab.*), Tingting Jiang, Ming Jiang (*Peking University*), Vajira Lasantha Thambawita (*SimulaMet*), Haoliang Wang (*Adobe Research*)
- **Reproducibility Companion Paper: Kalman Filter-Based Head Motion Prediction for Cloud-Based Mixed Reality** 3619
Serhan Gül, Sebastian Bosse, Dimitri Podborski, Thomas Schierl, Cornelius Hellge (*Fraunhofer HHI*), Marc A. Kastner (*National Institute of Informatics*), Jan Zahálka (*Czech Technical University in Prague*)
- **Reproducibility Companion Paper: Blind Natural Video Quality Prediction via Statistical Temporal Features and Deep Spatial Features** 3622
Jari Korhonen, Yicheng Su (*Shenzhen University*), Junyong You (*Norwegian Research Centre*), Steven Hicks, Cise Midoglu (*Simula Metropolitan Center for Digital Engineering (SimulaMet)*)
- **Reproducibility Companion Paper: Describing Subjective Experiment Consistency by p-Value P-P Plot** 3627
Jakub Nawala, Lucjan Janowski, Bogdan Cmiel, Krzysztof Rusek (*AGH University of Science and Technology*), Marc A. Kastner (*National Institute of Informatics*), Jan Zahálka (*Czech Technical University in Prague*)
- **Reproducibility Companion Paper: Self-supervised Video Representation Learning Using Inter-intra Contrastive Framework** 3630
Li Tao, Xueting Wang, Toshihiko Yamasaki (*The University of Tokyo*), Jingjing Chen (*Fudan University*), Steven Hicks (*SimulaMet & Oslo Metropolitan University*)

- **Reproducibility Companion Paper: Visual Relation of Interest Detection** 3633
Fan Yu (*Nanjing University & Shenzhen Research Institute of Nanjing University*),
Haonan Wang (*Nanjing University*),
Tongwei Ren (*Nanjing University & Shenzhen Research Institute of Nanjing University*),
Jinhui Tang (*Nanjing University of Science and Technology*), Gangshan Wu (*Nanjing University*),
Jingjing Chen (*Fudan University*), Zhenzhong Kuang (*Hangzhou Dianzi University*)

- **Reproducibility Companion Paper: On Learning Disentangled Representation for Acoustic Event Detection** 3638
Lijian Gao (*Jiangsu University*), Qirong Mao (*Jiangsu University & Jiangsu Engineering Research Center of Big Data Ubiquitous Perception and Intelligent Agricultural Applications*),
Jingjing Chen (*Jiangsu University*), Ming Dong, Ratna Chinnam (*Wayne State University*),
Lucile Sassatelli, Miguel Romero Rondon (*Université Côte d'Azur*),
Ujjwal Sharma (*University of Amsterdam*)

Keynote Talk V & VI

- **AI and the Future of Education** 3642
James Lester (*North Carolina State University*)
- **Digital Human in an Integrated Physical-Digital World (IPhD)** 3643
Zhengyou Zhang (*Tencent AI Lab & Tencent Robotics X*)

Session 24: Media Interpretation-III

- **Cross-Camera Feature Prediction for Intra-Camera Supervised Person Re-identification across Distant Scenes**..... 3644
Wenhong Ge (*Sun Yat-sen University & Pazhou Lab*), Chunyan Pan (*Sun Yat-sen University*),
Ancong Wu (*Sun Yat-sen University & Pazhou Lab*),
Hongwei Zheng (*University of Chinese Academy of Sciences*),
Wei-Shi Zheng (*Sun Yat-sen University & Pazhou Lab & Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education*),
- **Video Visual Relation Detection via Iterative Inference** 3654
Xindi Shang, Yicong Li, Junbin Xiao (*National University of Singapore*),
Wei Ji (*National University of Singapore & Sea-NExT Joint Lab*),
Tat-Seng Chua (*National University of Singapore & Sea-NExT Joint Lab*)
- **Instance-wise or Class-wise? A Tale of Neighbor Shapley for Concept-based Explanation** .. 3664
Jiahui Li, Kun Kuang, Lin Li (*Zhejiang University*), Long Chen (*Columbia University*),
Songyang Zhang (*University of Rochester*), Jian Shao, Jun Xiao (*Zhejiang University*)
- **Multifocal Attention-Based Cross-Scale Network for Image De-raining**..... 3673
Zheyu Zhang, Yurui Zhu, Xueyang Fu, Zhiwei Xiong, Zheng-Jun Zha, Feng Wu
(*University of Science and Technology of China*)
- **PFFN: Progressive Feature Fusion Network for Lightweight Image Super-Resolution** 3682
Dongyang Zhang (*University of Electronic Science and Technology of China & Sichuan Artificial Intelligence Research Institute*),
Changyu Li, Ning Xie, Guoqing Wang (*University of Electronic Science and Technology of China*),
Jie Shao (*University of Electronic Science and Technology of China & Sichuan Artificial Intelligence Research Institute*)
- **InterBN: Channel Fusion for Adversarial Unsupervised Domain Adaptation** 3691
Mengzhu Wang, Wei Wang (*Dalian University of Technology*), Baopu Li (*Baidu Research*),
Xiang Zhang, Long Lan, Huibin Tan, Tianyi Liang, Wei Yu, Zhigang Luo (*National University of Defense Technology*)

Session 25: Multimedia Art, Entertainment and Culture

- **Learning to Compose Stylistic Calligraphy Artwork with Emotions**..... 3701
Shaoyu Yuan, Ruixue Liu, Meng Chen (*JD AI*),
Baoyang Chen, Zhijie Qiu (*Central Academy of Fine Arts*), Xiaodong He (*JD AI*)
- **Graph Neural Networks for Knowledge Enhanced Visual Representation of Paintings**..... 3710
Athanasios Efthymiou, Stevan Rudinac, Monika Kackovic, Marcel Worring,
Nachoem Wijnberg (*University of Amsterdam*)

- **ArtScience and the ICECUBE LED Display [ILDm^{^3}]** 3720
Mark-David Hosale, Robert Allison (*York University*),
Jim Madsen (*University of Wisconsin-Madison*), Marcus Gordon (*York University*)
- **PUGCQ: A Large Scale Dataset for Quality Assessment of Professional User-Generated Content** 3728
Guo Li (*Kingsoft Cloud*), Baoliang Chen, Lingyu Zhu (*City University of Hong Kong*),
Qinwen He, Hongfei Fan (*Kingsoft Cloud*), Shiqi Wang (*City University of Hong Kong*)
- **Combining Attention with Flow for Person Image Synthesis**..... 3737
Yurui Ren, Yubo Wu, Thomas H. Li (*Peking University*), Shan Liu (*Tencent America*),
Ge Li (*Peking University*)
- **Dual Learning Music Composition and Dance Choreography** 3746
Shuang Wu (*Nanyang Technological University*), Zhenguang Liu (*Zhejiang University*),
Shijian Lu (*Nanyang Technological University*), Li Cheng (*University of Alberta*)

Session 26: Open Source Competition

- **MMFashion: An Open-Source Toolbox for Visual Fashion Analysis** 3755
Xin Liu (*Nanyang Technology University*), Jiancheng Li (*Alibaba Group*),
Jiaqi Wang (*The Chinese University of Hong Kong*), Ziwei Liu (*Nanyang Technology University*)
- **Efficient Reinforcement Learning Development with RLzoo** 3759
Zihan Ding (*Imperial College London*), Tianyang Yu (*Nanchang University*),
Hongming Zhang (*Peking University*), Yanhua Huang (*Xiaohongshu Technology Co.*),
Guo Li (*Imperial College London*), Quancheng Guo, Luo Mai (*University of Edinburgh*),
Hao Dong (*Peking University*)
- **Fast and Flexible Human Pose Estimation with HyperPose** 3763
Yixiao Guo (*Peking University*), Jiawei Liu (*Tongji University*),
Guo Li (*Imperial College London*), Luo Mai (*University of Edinburgh*), Hao Dong (*Peking University*)
- **SmartEye: An Open Source Framework for Real-Time Video Analytics with Edge-Cloud Collaboration** 3767
Xuezhi Wang, Guanyu Gao (*Nanjing University of Science and Technology*)
- **ZoomSense: A Scalable Infrastructure for Augmenting Zoom** 3771
Tom Bartindale, Xiuzhi Chen, Harrison Marshall, Stanislav Pozdniakov,
Dan Richardson (*Monash University*)
- **Efficient Graph Deep Learning in TensorFlow with tf_geometric**..... 3775
Jun Hu (*Institute of Automation, Chinese Academy of Sciences*),
Shengsheng Qian, Quan Fang (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Youze Wang, Quan Zhao (*Hefei University of Technology*),
Huaiwen Zhang (*Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; & Ministry of Industry and Information Technology*),
Changsheng Xu (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*)
- **FaceX-Zoo: A PyTorch Toolbox for Face Recognition** 3779
Jun Wang, Yinglu Liu, Yibo Hu, Hailin Shi, Tao Mei (*JD AI Research*)
- **PyTorchVideo: A Deep Learning Library for Video Understanding**..... 3783
Haoqi Fan, Tullie Murrell, Heng Wang, Kalyan Vasudev Alwala, Yanghao Li, Yilei Li,
Bo Xiong, Nikhila Ravi, Meng Li, Haichuan Yang, Jitendra Malik, Ross Girshick,
Matt Feiszli, Aaron Adcock, Wan-Yen Lo, Christoph Feichtenhofer (*Facebook AI*)
- **AICoach: A System Framework for Online Realtime Workout Coach** 3787
Haocong Ying, Tie Liu, Mingxin Ai, Jiali Ding, Yuanyuan Shang (*Capital Normal University*)

- **MMOCR: A Comprehensive Toolbox for Text Detection, Recognition and Understanding** 3791
Zhanghui Kuang, Hongbin Sun, Zhizhong Li (*SenseTime Research*),
Xiaoyu Yue (*Center for Perceptual and Interactive Intelligence*), Tsui Hin Lin (*SenseTime Research*),
Jianyong Chen (*South China University of Technology*), Huaqiang Wei (*SenseTime Research*),
Yiqin Zhu (*South China University of Technology*), Tong Gao (*SenseTime Research*),
Wenwei Zhang (*Nanyang Technological University*),
Kai Chen (*SenseTime Research & Shanghai AI Laboratory*), Wayne Zhang (*SenseTime Research*),
Dahua Lin (*The Chinese University of Hong Kong*)
- **A Complete End to End Open Source Toolchain for the Versatile Video Coding (VVC) Standard** 3795
Adam Wiecekowski, Christian Lehmann, Benjamin Bross, Detlev Marpe (*Fraunhofer HHI*),
Thibaud Biatek, Mikael Raulet (*ATEME*), Jean Le Feuvre (*Telecom Paris, Institut Polytechnique de Paris*)
- **X-modaler: A Versatile and High-performance Codebase for Cross-modal Analytics** 3799
Yehua Li, Yingwei Pan, Jingwen Chen, Ting Yao, Tao Mei (*JD AI Research*)
- **Interpreting Super-Resolution CNNs for Sub-Pixel Motion Compensation in Video Coding** 3803
Luka Murn (*British Broadcasting Corporation & Dublin City University*),
Alan F. Smeaton (*Dublin City University*), Marta Mrak (*British Broadcasting Corporation*)

Session 27: Multimedia Search and Recommendation-I

- **Towards Accurate Localization by Instance Search** 3807
Yi-Geng Hong, Hui-Chu Xiao, Wan-Lei Zhao (*Xiamen University*)
- **Database-adaptive Re-ranking for Enhancing Cross-modal Image Retrieval** 3816
Rintaro Yanagi, Ren Togo, Takahiro Ogawa, Miki Haseyama (*Hokkaido University*)
- **Fine-grained Cross-modal Alignment Network for Text-Video Retrieval** 3826
Ning Han (*Hunan University*), Jingjing Chen (*Fudan University*), Guangyi Xiao (*Hunan University*), Hao Zhang (*City University of Hong Kong*), Yawen Zeng, Hao Chen (*Hunan University*)
- **Meta Self-Paced Learning for Cross-Modal Matching** 3835
Jiwei Wei, Xing Xu, Zheng Wang, Guoqing Wang (*University of Electronic Science and Technology of China*)
- **CausalRec: Causal Inference for Visual Debiasing in Visually-Aware Recommendation** 3844
Ruihong Qiu, Sen Wang, Zhi Chen, Hongzhi Yin, Zi Huang (*University of Queensland*)
- **Semi-supervised Domain Adaptive Retrieval via Discriminative Hashing Learning** 3853
Haifeng Xia, Taotao Jing (*Tulane University*), Chen Chen (*University of Central Florida*),
Zhengming Ding (*Tulane University*)

Session 28: Multimedia Search and Recommendation-II

- **Hierarchical View Predictor: Unsupervised 3D Global Feature Learning through Hierarchical Prediction among Unordered Views** 3862
Zhizhong Han (*Tsinghua University & Wayne State University*),
Xiyang Wang, Yu-Shen Liu (*Tsinghua University*), Matthias Zwicker (*University of Maryland, College Park*)
- **Mining Latent Structures for Multimedia Recommendation** 3872
Jinghao Zhang (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Yanqiao Zhu, Qiang Liu, Shu Wu, Shuhui Wang (*Institute of Computing Technology, Chinese Academy of Sciences*), Liang Wang (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
- **Why Do We Click: Visual Impression-aware News Recommendation** 3881
Jiahao Xun, Shengyu Zhang, Zhou Zhao, Jieming Zhu, Qi Zhang, Jingjie Li,
Xiuqiang He (*Huawei Noah's Ark Lab*), Xiaofei He (*Zhejiang University*),
Tat-Seng Chua (*National University of Singapore*), Fei Wu (*Zhejiang University*)

- **Identity-Preserving Face Anonymization via Adaptively Facial Attributes Obfuscation** 3891
Jingzhi Li, Lutong Han, Ruoyu Chen, Hua Zhang (*Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Bing Han (*Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Lili Wang (*Beihang University*), Xiaochun Cao (*Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences*)
- **CONQUER: Contextual Query-aware Ranking for Video Corpus Moment Retrieval** 3900
Zhijian Hou (*City University of Hong Kong*), Chong-Wah Ngo (*Singapore Management University*),
W. K. Chan (*City University of Hong Kong*)
- **Learning Unified Embeddings for Recommendation via Meta-path Semantics** 3909
Qianxiu Hao (*Institute of Computing Technology, Chinese Academy of Sciences, University of Chinese Academy of Science*),
Qianqian Xu (*Institute of Computing Technology, Chinese Academy of Sciences*),
Zhiyong Yang (*University of Chinese Academy of Sciences*),
Qingming Huang (*Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Science, Peng Cheng Laboratory*)

Session 29: Music, Speech and Audio Processing in Multimedia

- **ReconVAT: A Semi-Supervised Automatic Music Transcription Framework for Low-Resource Real-World Data** 3918
Kin Wai Cheuk, Dorien Herremans (*Singapore University of Technology and Design*),
Li Su (*Academia Sinica*)
- **Is Someone Speaking? Exploring Long-term Temporal Features for Audio-visual Active Speaker Detection** 3927
Ruijie Tao, Zexu Pan, Rohan Kumar Das, Xinyuan Qian, Mike Zheng Shou, Haizhou Li
(*National University of Singapore*)
- **Actions Speak Louder than Listening: Evaluating Music Style Transfer based on Editing Experience** 3936
Wei-Tsung Lu, Meng-Hsuan Wu (*Academia Sinica*), Yuh-Ming Chiu (*KKBOX, Inc.*),
Li Su (*Academia Sinica*)
- **Multi-Singer: Fast Multi-Singer Singing Voice Vocoder With A Large-Scale Corpus** 3945
Rongjie Huang, Feiyang Chen, Yi Ren, Jinglin Liu, Chenye Cui, Zhou Zhao (*Zhejiang University*)
- **MusicBERT: A Self-supervised Learning of Music Representation** 3955
Hongyuan Zhu, Ye Niu, Di Fu, Hao Wang (*Alibaba Cloud, Alibaba Group*)
- **UniCon: Unified Context Network for Robust Active Speaker Detection** 3964
Yuanhang Zhang, Susan Liang, Shuang Yang (*Institute of Computing Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Xiao Liu (*Tomorrow Advancing Life*), Zhongqin Wu (*Tomorrow Advancing Life*),
Shiguang Shan (*Institute of Computing Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Xilin Chen (*Institute of Computing Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences*)

Session 30: Multimedia Transport and Delivery

- **AITransfer: Progressive AI-powered Transmission for Real-Time Point Cloud Video Streaming** 3989
Yakun Huang, Yuanwei Zhu, Xiuquan Qiao, Zhijie Tan (*Beijing University of Posts and Telecommunications*),
Boyuan Bai (*University of Birmingham*)
- **Game Theory-driven Rate Control for 360-Degree Video Coding** 3998
Tiesong Zhao (*Fuzhou University & Peng Cheng Laboratory*),
Jielian Lin, Yanjie Song (*Fuzhou University*), Xu Wang (*Shenzhen University*),
Yuzhen Niu (*Fuzhou University*)

- **TBRA: Tiling and Bitrate Adaptation for Mobile 360-Degree Video Streaming** 4007
 Lei Zhang, Yanyan Suo, Ximing Wu (*Shenzhen University*),
 Feng Wang (*University of Mississippi*), Yuchi Chen (*Simon Fraser University*),
 Laizhong Cui (*Shenzhen University*), Jiangchuan Liu (*Simon Fraser University*),
 Zhong Ming (*Shenzhen University*)
- **QoE Ready to Respond: A QoE-aware MEC Selection Scheme for DASH-based Adaptive Video Streaming to Mobile Users** 4016
 Wanxin Shi (*Tsinghua University & Peng Cheng Laboratory*),
 Qing Li (*Peng Cheng Laboratory & Southern University of Science and Technology*),
 Ruishan Zhang, Gengbiao Shen (*Tsinghua University*),
 Yong Jiang (*Tsinghua University & Peng Cheng Laboratory*),
 Zhenhui Yuan (*Northumbria University*), Gabriel-Miro Muntean (*Dublin City University*)
- **Hierarchical Fusion for Practical Ghost-free High Dynamic Range Imaging** 4025
 Pengfei Xiong, Yu Chen (*Tencent*)
- **Edge-oriented Convolution Block for Real-time Super Resolution on Mobile Devices** 4034
 Xindong Zhang, Hui Zeng, Lei Zhang (*The Hong Kong Polytechnic University & Alibaba Group*)

Poster Session 5

- **Semantic Scalable Image Compression with Cross-Layer Priors** 4044
 Hanyue Tu, Li Li, Wengang Zhou, Houqiang Li (*University of Science and Technology of China*)
- **Cascade Cross-modal Attention Network for Video Actor and Action Segmentation from a Sentence** 4053
 Weidong Chen (*University of Chinese Academy of Sciences*),
 Guorong Li, Xinfeng Zhang (*University of Chinese Academy of Sciences*),
 Hongyang Yu (*Peng Cheng Laboratory*),
 Shuhui Wang (*Institute of Computing Technology Chinese Academy of Sciences*),
 Qingming Huang (*University of Chinese Academy of Sciences & Institute of Computing Technology Chinese Academy of Sciences*)
- **Extracting Useful Knowledge from Noisy Web Images via Data Purification for Fine-Grained Recognition** 4063
 Chuanyi Zhang, Yazhou Yao (*Nanjing University of Science and Technology*),
 Xing Xu, Jie Shao, Jingkuan Song (*University of Electronic Science and Technology of China*),
 Zechao Li, Zhenmin Tang (*Nanjing University of Science and Technology*)
- **Complementary Factorization towards Outfit Compatibility Modeling** 4073
 Tianyu Su, Xuemeng Song, Na Zheng (*Shandong University*), Weili Guan (*Monash University*),
 Yan Li (*Kuaishou Technology*), Liqiang Nie (*Shandong University*)
- **Open Set Face Anti-Spoofing in Unseen Attacks** 4082
 Xin Dong, Hao Liu, Weiwei Cai, Pengyuan Lv (*Ningxia University*), Zekuan Yu (*Fudan University*)
- **Interventional Video Relation Detection** 4091
 Yicong Li, Xun Yang, Xindi Shang, Tat-Seng Chua (*National University of Singapore*)
- **CanvasEmb: Learning Layout Representation with Large-scale Pre-training for Graphic Design** 4100
 Yuxi Xie (*National University of Singapore*), Danqing Huang (*Microsoft Research Asia*),
 Jinpeng Wang (*Meituan*), Chin-Yew Lin (*Microsoft Research Asia*)
- **Augmenting TV Shows via Uncalibrated Camera Small Motion Tracking in Dynamic Scene** 4109
 Yizhen Lao (*Hunan University*), Jie Yang, Xinying Wang (*MGTV*),
 Jianxin Lin, Yu Cao (*Hunan University*), Shien Song (*MGTV*)
- **SimuSLT: End-to-End Simultaneous Sign Language Translation** 4118
 Aoxiong Yin, Zhou Zhao, Jinglin Liu, Weike Jin (*Zhejiang University*),
 Meng Zhang, Xingshan Zeng (*Huawei Noah's Ark Lab*), Xiaofei He (*Zhejiang University*)

- **Mask and Predict: Multi-step Reasoning for Scene Graph Generation** 4128
 Hongshuo Tian (*Tianjin University & People's Daily Online*),
 Ning Xu, An-An Liu (*Tianjin University*), Chenggang Yan (*Hangzhou Dianzi University*),
 Zhendong Mao (*University of Science and Technology of China*), Quan Zhang (*Peking University*),
 Yongdong Zhang (*University of Science and Technology of China*)
- **Heterogeneous Face Recognition with Attention-guided Feature Disentangling** 4137
 Shanmin Yang, Xiao Yang, Yi Lin, Peng Cheng, Yi Zhang, Jianwei Zhang (*Sichuan University*)
- **Exploring the Quality of GAN Generated Images for Person Re-Identification** 4146
 Yiqi Jiang, Weihua Chen, Xiuyu Sun, Xiaoyu Shi, Fan Wang, Hao Li (*Alibaba Group*)
- **Multi-view Clustering via Deep Matrix Factorization and Partition Alignment** 4156
 Chen Zhang, Siwei Wang, Jiyuan Liu, Sihang Zhou, Pei Zhang, Xinwang Liu,
 En Zhu (*National University of Defense Technology*), Changwang Zhang (*Tencent Technology*)
- **Video Similarity and Alignment Learning on Partial Video Copy Detection** 4165
 Zhen Han, Xiangteng He, Mingqian Tang, Yiliang Lv (*Alibaba Group*)
- **No-Reference Video Quality Assessment with Heterogeneous Knowledge Ensemble** 4174
 Jinjian Wu, Yongxu Liu, Leida Li, Weisheng Dong, Guangming Shi (*Xidian University*)
- **Seeing is Believing? Effects of Visualization on Smart Device Privacy Perceptions** 4183
 Carlos Bermejo Fernandez (*Hong Kong University of Science and Technology*),
 Petteri Nurmi (*University of Helsinki*),
 Pan Hui (*Hong Kong University of Science and Technology & University of Helsinki*)
- **MHFC: Multi-Head Feature Collaboration for Few-Shot Learning** 4193
 Shuai Shao, Lei Xing (*China University of Petroleum (East China)*), Yan Wang (*Beihang University*),
 Rui Xu (*China University of Petroleum (East China)*), Chunyan Zhao (*Suzhou Centennial College*),
 Yanjiang Wang, Baodi Liu (*China University of Petroleum (East China)*)
- **Vision-guided Music Source Separation via a Fine-grained Cycle-Separation Network** 4202
 Shuo Ma, Yanli Ji, Xing Xu, Xiaofeng Zhu (*University of Electronic Science and Technology of China*)
- **GLM-Net : Global and Local Motion Estimation via Task-Oriented Encoder-Decoder Structure** 4211
 Yuchen Yang, Ye Xiang (*Beijing University of Technology*),
 Shuaicheng Liu (*University of Electronic Science and Technology of China*),
 Lifang Wu, Boxuan Zhao (*Beijing University of Technology*),
 Bing Zeng (*University of Electronic Science and Technology of China*)
- **Sensor-Augmented Egocentric-Video Captioning with Dynamic Modal Attention** 4220
 Katsuyuki Nakamura, Hiroki Ohashi, Mitsuhiro Okada (*Hitachi, Ltd.*)
- **Cross Modal Compression: Towards Human-comprehensible Semantic Compression** 4230
 Jiguo Li (*Institute of Computing Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
 Chuanmin Jia (*Peking University*), Xinfeng Zhang (*University of Chinese Academy of Sciences*),
 Siwei Ma (*Peking University & Information Technology R&D Innovation Center of Peking University*),
 Wen Gao (*Peking University*)
- **RAMS-Trans: Recurrent Attention Multi-scale Transformer for Fine-grained Image Recognition** 4239
 Yunqing Hu (*Zhejiang University*), Xuan Jin (*Alibaba Group*),
 Yin Zhang, Haiwen Hong, Jingfeng Zhang (*Zhejiang University*),
 Yuan He, Hui Xue (*Alibaba Group*)
- **Memory-Augmented Deep Unfolding Network for Compressive Sensing** 4249
 Jiechong Song, Bin Chen (*Peking University*),
 Jian Zhang (*Peking University & Peng Cheng Laboratory*)
- **Underwater Species Detection using Channel Sharpening Attention** 4259
 Lihao Jiang, Yi Wang, Qi Jia (*Dalian University of Technology*), Shengwei Xu (*ByteDance Inc*),
 Yu Liu, Xin Fan, Haojie Li, Risheng Liu, Xinwei Xue (*Dalian University of Technology*),
 Ruili Wang (*Massey University*)

- **Self-Supervised Pre-training on the Target Domain for Cross-Domain Person Re-identification** 4268
Junyin Zhang, Yongxin Ge (*Chongqing University*), Xinqian Gu (*University of Chinese Academy of Sciences*),
Boyu Hua, Tao Xiang (*Chongqing University*)
- **Exploring Graph-Structured Semantics for Cross-Modal Retrieval** 4277
Lei Zhang, Leiting Chen, Chuan Zhou (*University of Electronic Science and Technology of China*),
Fan Yang, Xin Li (*AIQ*)
- **Text is NOT Enough: Integrating Visual Impressions into Open-domain Dialogue Generation** 4287
Lei Shen (*Institute of Computing Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Haolan Zhan (*University of Chinese Academy of Sciences*), Xin Shen (*Australian National University*),
Yonghao Song, Xiaofang Zhao (*Institute of Computing Technology, Chinese Academy of Sciences*)
- **Quality Assessment of End-to-End Learned Image Compression: The Benchmark and Objective Measure**..... 4297
Yang Li (*Peking University*), Shiqi Wang (*City University of Hong Kong*),
Xinfeng Zhang (*University of Chinese Academy of Sciences*), Shanshe Wang, Siwei Ma (*Peking University*),
Yue Wang (*Bytedance Inc.*),
- **A Statistical Approach to Mining Semantic Similarity for Deep Unsupervised Hashing** 4306
Xiao Luo, Daqing Wu (*Peking University*), Zeyu Ma (*Harbin Institute of Technology*),
Chong Chen (*Alibaba Group*), Minghua Deng (*Peking University*),
Jianqiang Huang, Xian-Sheng Hua (*Alibaba Group*)
- **BAM: Bilateral Activation Mechanism for Image Fusion** 4315
Zi-Rong Jin, Liang-Jian Deng, Tian-Jing Zhang, Xiao-Xu Jin
(*University of Electronic Science and Technology of China*)
- **Self-supervising Action Recognition by Statistical Moment and Subspace Descriptors** 4324
Lei Wang (*Australian National University & Data61/CSIRO*),
Piotr Koniusz (*Data61/CSIRO & Australian National University*)
- **Learning Multi-Granular Spatio-Temporal Graph Network for Skeleton-based Action Recognition** 4334
Tailin Chen (*Newcastle University*), Desen Zhou, Jian Wang (*Baidu, Inc.*),
Shidong Wang, Yu Guan (*Newcastle University*), Xuming He (*ShanghaiTech University*),
Errui Ding (*Baidu, Inc.*)
- **ION: Instance-level Object Navigation** 4343
Wei jie Li, Xinhang Song, Yubing Bai, Sixian Zhang, Shuqiang Jiang (*Institute of Computing Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences*)
- **Skeleton-Aware Neural Sign Language Translation** 4353
Shiwei Gan, Yafeng Yin, Zhiwei Jiang, Lei Xie, Sanglu Lu (*Nanjing University*)
- **Fingerspelling Recognition in the Wild with Fixed-Query based Visual Attention**..... 4362
Srinivas Kruthiventi S S, George Jose, Nitya Tandon, Rajesh Biswal, Aashish Kumar
(*Harman International India Pvt. Ltd.*)
- **Deep Human Dynamics Prior** 4371
Qiongjie Cui, Huaijiang Sun, Yue Kong, Xiaoning Sun (*Nanjing University of Science and Technology*)
- **Exploiting Invariance of Mining Facial Landmarks** 4380
Jiangming Shi, Zixian Gao, Hao Liu (*Ningxia University*), Zekuan Yu (*Fudan University*),
Fengjun Li (*Ningxia University*)
- **Joint Implicit Image Function for Guided Depth Super-Resolution**..... 4390
Jiaxiang Tang, Xiaokang Chen, Gang Zeng (*Peking University*)
- **Transformer-based Feature Reconstruction Network for Robust Multimodal Sentiment Analysis** 4400
Ziqi Yuan, Wei Li, Hua Xu, Wenmeng Yu (*Tsinghua University & Beijing National Research Center for Information Science and Technology(BNRist)*)

- **Self-feature Learning: An Efficient Deep Lightweight Network for Image Super-resolution** 4408
Jun Xiao (*The Hong Kong Polytechnic University*), Qian Ye (*Tohoku University*),
Rui Zhao, Kin-Man Lam (*The Hong Kong Polytechnic University*), Kao Wan (*Peng Cheng Laboratory*)
- **DAWN: Dynamic Adversarial Watermarking of Neural Networks** 4417
Sebastian Szyller, Buse Gul Atli (*Aalto University*), Samuel Marchal (*Aalto University & F-Secure Corporation*),
N. Asokan (*University of Waterloo & Aalto University*)
- **Visible Watermark Removal via Self-calibrated Localization and Background Refinement** 4426
Jing Liang, Li Niu, Fengjun Guo, Teng Long (*INTSIG*), Liqing Zhang (*Shanghai Jiao Tong University*)
- **Learning to Decode Contextual Information for Efficient Contour Detection**..... 4435
Ruoxi Deng (*Wenzhou University*), Shengjun Liu (*Central South University*), Jinxin Wang (*Wenzhou University*),
Huibing Wang (*Dalian Maritime University*), Hanli Zhao, Xiaoqin Zhang (*Wenzhou University*)
- **Fast, High-Quality Hierarchical Depth-Map Super-Resolution**..... 4444
Yiguo Qiao (*University of Bath*), Licheng Jiao (*Xidian University*),
Wenbin Li, Christian Richardt, Darren Cosker (*University of Bath*)
- **TsFPS: An Accurate and Flexible 6DoF Tracking System with Fiducial Platonic Solids**..... 4454
Nan Xiang, Xiaosong Yang, Jian J Zhang (*Bournemouth University*)
- **Consistency-Constancy Bi-Knowledge Learning for Pedestrian Detection in Night Surveillance** 4463
Xiao Wang (*Wuhan University of Science and Technology*), Zheng Wang (*Wuhan University*),
Wu Liu (*AI Research of JD.com*), Xin Xu (*Wuhan University of Science and Technology*),
Jing Chen (*Wuhan University*), Chia-Wen Lin (*National Tsing Hua University*)
- **SSconv: Explicit Spectral-to-Spatial Convolution for Pansharpening**..... 4472
Yudong Wang, Liang-Jian Deng, Tian-Jing Zhang, Xiao Wu (*University of Electronic Science and Technology of China*)
- **TriTransNet: RGB-D Salient Object Detection with a Triplet Transformer Embedding Network**..... 4481
Zhengyi Liu, Yuan Wang, Zhengzheng Tu, Yun Xiao (*Anhui University*), Bin Tang (*Hefei University*)
- **Learning Sample-Specific Policies for Sequential Image Augmentation** 4491
Pu Li , Xiaobai Liu (*San Diego State University*), Xiaohui Xie (*University of California, Irvine*)
- **Image Quality Caption with Attentive and Recurrent Semantic Attractor Network** 4501
Wen Yang, Jinjian Wu, Leida Li, Weisheng Dong, Guangming Shi (*Xidian University*)
- **Triangle-Reward Reinforcement Learning: A Visual-Linguistic Semantic Alignment for Image Captioning** 4510
Weizhi Nie (*Tianjin University & People's Daily Online*),
Jiesi Li, Ning Xu, An-An Liu (*Tianjin University*), Xuanya Li (*Baidu Inc.*),
Yongdong Zhang (*University of Science and Technology of China*)
- **Stacked Semantically-Guided Learning for Image De-distortion** 4519
Huiyuan Fu, Changhao Tian (*Beijing University of Posts and Telecommunications*),
Xin Wang (*Stony Brook University*), Huadong Ma (*Beijing University of Posts and Telecommunications*)
- **Focal and Composed Vision-semantic Modeling for Visual Question Answering** 4528
Yudong Han, Yangyang Guo, Jianhua Yin (*Shandong University*), Meng Liu (*Shandong Jianzhu University*),
Yupeng Hu, Liqiang Nie (*Shandong University*)
- **Pose-Guided Feature Learning with Knowledge Distillation for Occluded Person Re-Identification** 4537
Kecheng Zheng (*University of Science and Technology of China*),
Cuiling Lan, Wenjun Zeng (*Microsoft Research Asia*), Jiawei Liu (*University of Science and Technology of China*),
Zhizheng Zhang (*Microsoft Research Asia*), Zheng-Jun Zha (*University of Science and Technology of China*)
- **Multiple Objects-Aware Visual Question Generation** 4546
Jiayuan Xie, Yi Cai (*South China University of Technology & Ministry of Education*),
Qingbao Huang (*South China University of Technology, Ministry of Education, & Guangxi University*),
Tao Wang (*King's College London*)
- **VASTile: Viewport Adaptive Scalable 360-Degree Video Frame Tiling**..... 4555
Chamara Madarasingha, Kanchana Thilakarathna (*The University of Sydney*)

- **Delving into Deep Image Prior for Adversarial Defense: A Novel Reconstruction-based Defense Framework** 4564
 Li Ding (*Xi'an Jiaotong University & University of British Columbia*),
 Yongwei Wang, Xin Ding, Kaiwen Yuan (*University of British Columbia*),
 Ping Wang (*Xi'an Jiaotong University*), Hua Huang (*Beijing Normal University*),
 Z. Jane Wang (*University of British Columbia*)
- **Fine-Grained Language Identification in Scene Text Images**..... 4573
 Yongrui Li, Shilian Wu, Jun Yu (*University of Science and Technology of China*),
 Zengfu Wang (*Institute of Intelligent Machines, Chinese Academy of Sciences*)
- **CARE: Cloudified Android OSEs on the Cloud Rendering** 4582
 Dongjie Tang (*Shanghai Jiao Tong University*),
 Cathy Bao, Yong Yao, Chao Xie, Qiming Shi, Marc Mao, Randy Xu (*Intel Corporation*),
 Linsheng Li (*Shanghai Jiao Tong University*), Mohammad R. Haghghat (*Intel Corporation*),
 Zhengwei Qi, Haibing Guan (*Shanghai Jiao Tong University*)
- **Context-Aware Selective Label Smoothing for Calibrating Sequence Recognition Model** ... 4591
 Shuangping Huang (*South China University of Technology & Pazhou Laboratory*),
 Yu Luo, Zhenzhou Zhuang, Jin-Gang Yu (*South China University of Technology & Pazhou Laboratory*),
 Mengchao He (*DAMO Academy, Alibaba Group*), Yongpan Wang (*Alibaba Group*)
- **Image Search with Text Feedback by Deep Hierarchical Attention Mutual Information Maximization** 4600
 Chunbin Gu, Jiajun Bu, Zhen Zhang (*Zhejiang University & Alibaba-Zhejiang University Joint Institute of Frontier Technologies*), Zhi Yu, Dongfang Ma, Wei Wang (*Zhejiang University*)

Panel 2

- **Social Signals and Multimedia: Past, Present, Future** 4610
 Hayley Hung (*Delft University of Technology*), Cathal Gurrin (*Dublin City University*),
 Martha Larson (*Radboud University*), Hatice Gunes (*University of Cambridge*),
 Fabien Ringeval (*Université Grenoble Alpes*), Elisabeth Andre (*University of Augsburg*),
 Louis-Philippe Morency (*Carnegie Mellon University*)

Session 31: Multimedia Telepresence and Virtual/Augmented Reality

- **Learning Spatial-angular Fusion for Compressive Light Field Imaging in a Cycle-consistent Framework** 4613
 Xianqiang Lyu, Zhiyu Zhu, Mantang Guo, Jing Jin, Junhui Hou (*City University of Hong Kong*),
 Huanqiang Zeng (*Huaqiao University*)
- **From Voxel to Point: IoU-guided 3D Object Detection for Point Cloud with Voxel-to-Point Decoder** 4622
 Jiale Li (*Zhejiang University*), Hang Dai (*Mohamed bin Zayed University of Artificial Intelligence*), Ling Shao (*Inception Institute of Artificial Intelligence*), Yong Ding (*Zhejiang University*)
- **Extending 6-DoF VR Experience Via Multi-Sphere Images Interpolation**..... 4632
 Jisheng Li, Yuze He, Jinghui Jiao, Yubin Hu (*Tsinghua University*), Yuxing Han (*Research Institute of Tsinghua University in Shenzhen*), Jiangtao Wen (*Tsinghua University & Research Institute of Tsinghua University in Shenzhen*)
- **iButter: Neural Interactive Bullet Time Generator for Human Free-viewpoint Rendering** ... 4641
 Liao Wang, Ziyu Wang, Pei Lin, Yuheng Jiang, Xin Suo, Minye Wu, Lan Xu, Jingyi Yu (*ShanghaiTech University*)
- **Neural Free-Viewpoint Performance Rendering under Complex Human-object Interactions** 4651
 Guoxing Sun (*ShanghaiTech University*), Xin Chen (*ShanghaiTech University; Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences; & University of Chinese Academy of Sciences*),
 Yizhang Chen (*ShanghaiTech University*),
 Anqi Pang (*ShanghaiTech University; Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences; & University of Chinese Academy of Sciences*),
 Pei Lin, Yuheng Jiang, Lan Xu, Jingyi Yu, Jingya Wang (*ShanghaiTech University*)

- **Semi-supervised Learning via Improved Teacher-Student Network for Robust 3D Reconstruction of Stereo Endoscopic Image**..... 4661
Hongkuan Shi, Zhiwei Wang, Jinxin Lv, Yilang Wang, Peng Zhang, Fei Zhu, Qiang Li
(Huazhong University of Science and Technology)

Session 32: Social Multimedia

- **FoodLogoDet-1500: A Dataset for Large-Scale Food Logo Detection via Multi-Scale Feature Decoupling Network**..... 4670
Qiang Hou (Shandong Normal University), Weiqing Min (Institute of Computing Technology, Chinese Academy of Sciences), Jing Wang, Sujuan Hou, Yuanjie Zheng (Shandong Normal University), Shuqiang Jiang (Institute of Computing Technology, Chinese Academy of Sciences)
- **Cross-View Representation Learning for Multi-View Logo Classification with Information Bottleneck**..... 4680
Jing Wang, Yuanjie Zheng, Jingqi Song, Sujuan Hou (Shandong Normal University)
- **Parametric Reshaping of Portraits in Videos**..... 4689
Xiangjun Tang, WenXin Sun (Zhejiang University), Yong-Liang Yang (University of Bath), Xiaogang Jin (Zhejiang University)
- **Human Attributes Prediction under Privacy-preserving Conditions**..... 4698
Anshu Singh, Shaojing Fan, Mohan Kankanhalli (National University of Singapore)
- **Multi-Modal Sarcasm Detection with Interactive In-Modal and Cross-Modal Graphs**..... 4707
Bin Liang, Chenwei Lou, Xiang Li (Harbin Institute of Technology, Shenzhen & Joint lab of CMS-HITSZ), Lin Gui (University of Warwick), Min Yang (Chinese Academy of Sciences), Ruifeng Xu (Harbin Institute of Technology & Peng Cheng Lab)
- **Linking the Characters: Video-oriented Social Graph Generation via Hierarchical-cumulative GCN** 4716
Shiwei Wu, Joya Chen, Tong Xu, Liyi Chen (University of Science and Technology of China), Lingfei Wu (JD.COM Silicon Valley Research Center), Yao Hu (Alibaba Youku Cognitive and Intelligent Lab), Enhong Chen (University of Science and Technology of China)

Session 33: Multimedia Grand Challenge

- **Overview of Tencent Multi-modal Ads Video Understanding**..... 4725
Zhenzhi Wang (Nanjing University), Zhimin Li (Huazhong University of Science and Technology), Liyu Wu (Peking University), Jiangfeng Xiong (Tencent Data Platform), Qinglin Lu (Tencent Data Platform)
- **Better Learning Shot Boundary Detection via Multi-task**..... 4730
Haoxin Zhang (Tencent Data Platform), Zhimin Li (Huazhong University of Science and Technology), Qinglin Lu (Tencent Data Platform)
- **Facial Micro-Expression Generation based on Deep Motion Retargeting and Transfer Learning**..... 4735
Xinqi Fan (City University of Hong Kong), Ali Raza Shahid (City University of Hong Kong & COMSATS University Islamabad), Hong Yan (City University of Hong Kong)
- **Deadline and Priority-aware Congestion Control for Delay-sensitive Multimedia Streaming**..... 4740
Chao Zhou, Wenjun Wu, Dan Yang (Kuaishou), Tianchi Huang (Tsinghua University & Kuaishou), Liang Guo, Bing Yu (Kuaishou)
- **LSSNet: A Two-stream Convolutional Neural Network for Spotting Macro- and Micro-expression in Long Videos** 4745
Wang-Wang Yu (University of Electronic Science And Technology of China), Jingwen Jiang (West China Biomedical Big Data Center & Sichuan University), Yong-Jie Li (University of Electronic Science and Technology of China)

- **Multi-Level Visual Representation with Semantic-Reinforced Learning for Video Captioning** 4750
Chengbo Dong, Xinru Chen, Aozhu Chen, Fan Hu, Zihan Wang, Xirong Li (*Renmin University of China*)
- **Facial Prior Based First Order Motion Model for Micro-expression Generation** 4755
Yi Zhang, Youjun Zhao, Yuhang Wen, Zixuan Tang, Xinhua Xu (*Sun Yat-Sen University*),
Mengyuan Liu (*Sun Yat-Sen University & Guangdong Provincial Key Laboratory of Fire Science and Technology*)
- **Rethinking the Impacts of Overfitting and Feature Quality on Small-scale Video Classification** 4760
Xuansheng Wu (*University of Georgia*), Feichi Yang (*University of South California*),
Tong Zhou (*Central University of Finance and Economics*), Xinyue Lin (*University of Virginia*)
- **A Gradient Balancing Approach for Robust Logo Detection** 4765
Fuxing Leng (*Huazhong University of Science and Technology & ByteDance*)
- **Multi-modal Representation Learning for Video Advertisement Content Structuring** 4770
Daya Guo, Zhaoyang Zeng (*Sun Yat-sen University*)
- **Phoenix: Combining Highest-Profit First Scheduling and Responsive Congestion Control for Delay-sensitive Multimedia Transmission** 4775
Haozhe Li (*Beijing University of Posts and Telecommunications*)
- **VidVRD 2021: The Third Grand Challenge on Video Relation Detection**..... 4779
Wei Ji, Yicong Li, Meng Wei, Xindi Shang, Junbin Xiao (*National University of Singapore*),
Tongwei Ren (*Nanjing University*), Tat-Seng Chua (*National University of Singapore*)
- **A Simple and Effective Baseline for Robust Logo Detection** 4784
Weipeng Xu (*Shanghai Jiao Tong University*), Ye Liu (*University of Electronic Science and Technology of China*),
Daquan Lin (*ShanghaiTech University*)
- **Robust Logo Detection in E-Commerce Images by Data Augmentation** 4789
Hang Chen, Xiao Li, Zefan Wang, Xiaolin Hu (*Tsinghua University*)
- **Facial Action Unit-based Deep Learning Framework for Spotting Macro- and Micro-expressions in Long Video Sequences** 4794
Bo Yang, Jianming Wu (*KDDI Research, Inc.*), Zhiguang Zhou, Megumi Komiya (*Systemsoft, Inc.*),
Koki Kishimoto, Jianfeng Xu, Keisuke Nonaka, Toshiharu Horiuchi, Satoshi Komorita, Gen Hattori, Sei Naito,
Yasuhiro Takishima (*KDDI Research, Inc.*)
- **NJU MCG – SenseTime Team Submission to Pre-training for Video Understanding Challenge Track II** 4799
Liwei Jin, Haoyue Cheng (*Nanjing University*), Su Xu, Wayne Wu (*SenseTime Research*),
Limin Wang (*Nanjing University*)
- **Research on Micro-Expression Spotting Method Based on Optical Flow Features** 4803
He Yuhong (*Harbin Institute of Technology*)
- **A Solution to Multi-modal Ads Video Tagging Challenge** 4808
Hao Wu, Jiajie Wang, Yuanzhe Gu (*Alibaba Group*), Peisen Zhao (*Shanghai Jiao Tong University*),
Zhonglin Zu (*Alibaba Group*)
- **FAMGAN: Fine-grained AUs Modulation based Generative Adversarial Network for Micro-Expression Generation**..... 4813
Yifan Xu, Sirui Zhao, Huaying Tang, Xinglong Mao, Tong Xu, Enhong Chen (*University of Science and Technology of China*)
- **Semantic Tag Augmented XlanV Model for Video Captioning** 4818
Yiqing Huang (*Tsinghua University*), Hongwei Xue (*University of Science and Technology of China*),
Jiansheng Chen (*University of Science and Technology Beijing, Beijing National Research Center for Information Science and Technology*), Huimin Ma (*University of Science and Technology Beijing*),
Hongbing Ma (*Tsinghua University*)
- **Automated Multi-Modal Video Editing for Ads Video** 4823
Qin Lin, Nuo Pang, Zhiying Hong (*Tencent Data Platform*)
- **Rate Adaptation and Block Scheduling for Delay-sensitive Multimedia Applications** 4828
Dongyuan Su, Laizhong Cui, Lei Zhang, Yanyan Suo, Yan Qiu (*Shenzhen University*)

- **Video Relation Detection via Tracklet based Visual Transformer** 4833
Kaifeng Gao (*Zhejiang University*), Long Chen (*Columbia University*),
Yifeng Huang, Jun Xiao (*Zhejiang University*)
- **Group-Level Focus of Visual Attention for Improved Next Speaker Prediction** 4838
Chris Birmingham (*University of Southern California*), Kalin Stefanov (*Monash University*),
Maja J. Matarić (*University of Southern California*)
- **A Multimodal Framework for Video Ads Understanding** 4843
ZeJia Weng, Lingchen Meng, Rui Wang, Zuxuan Wu, Yu-Gang Jiang (*Fudan University*)
- **Joint Learning for Relationship and Interaction Analysis in Video with Multimodal Feature Fusion** 4848
Beibei Zhang (*Nanjing University*),
Fan Yu (*Nanjing University & Shenzhen Research Institute of Nanjing University*),
Yanxin Gao (*Nanjing University*),
Tongwei Ren (*Nanjing University & Shenzhen Research Institute of Nanjing University*),
Gangshan Wu (*Nanjing University*)
- **MM21 Pre-training for Video Understanding Challenge: Video Captioning with Pretraining Techniques** 4853
Sihan Chen (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Xinxin Zhu (*Institute of Automation, Chinese Academy of Sciences*),
Dongze Hao, Wei Liu, Jiawei Liu, Zijia Zhao, Longteng Guo, Jing Liu (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*)
- **CLIP4Caption: CLIP for Video Caption** 4858
Mingkang Tang (*Tencent & Tsinghua University*),
Zhanyu Wang, Zhenhua Liu, Fengyun Rao, Dian Li (*Tencent*), Xiu Li (*Tsinghua University*)
- **The ACM Multimedia 2021 Meet Deadline Requirements Grand Challenge** 4863
Jie Zhang (*Tsinghua University*), Junjie Deng (*Beijing University of Posts and Telecommunications*),
Mowei Wang, Yong Cui (*Tsinghua University*), Wei Tsang Ooi (*National University of Singapore*),
Jiangchuan Liu (*Simon Fraser University*), Xinyu Zhang (*University of California, San Diego*),
Kai Zheng (*Huawei Technologies*), Yi Li (*PowerInfo*)
- **MultiModal Language Modelling on Knowledge Graphs for Deep Video Understanding** 4868
Vishal Anand (*Columbia University*),
Raksha Ramesh, Boshen Jin (*Columbia University & Graphen Inc.*), Ziyin Wang (*Columbia University*),
Xiaoxiao Lei (*Graphen Inc.*), Ching-Yung Lin (*Columbia University & Graphen Inc.*)
- **Using Motion Histories for Eye Contact Detection in Multiperson Group Conversations** 4873
Eugene Yujun Fu (*The Hong Kong Polytechnic University*), Michael W. Ngai (*Phillips Exeter Academy*)
- **MULTIMEDIATE: Multi-modal Group Behaviour Analysis for Artificial Mediation** 4878
Philipp Müller (*DFKI GmbH*), Michael Dietz, Dominik Schiller (*Augsburg University*),
Dominike Thomas, Guanhua Zhang (*University of Stuttgart*), Patrick Gebhard (*DFKI GmbH*),
Elisabeth André (*University of Augsburg*), Andreas Bulling (*University of Stuttgart*)

Session 34: Summarization, Analytics, and Storytelling

- **MeshNet++: A Network with a Face** 4883
Vinit Veerendraveer Singh, Shivanand Venkanna Sheshappanavar, Chandra Kambhamettu
(*University of Delaware*)
- **Latent Memory-augmented Graph Transformer for Visual Storytelling** 4892
Mengshi Qi (*Beijing University of Posts and Telecommunications*),
Jie Qin (*Nanjing University of Aeronautics and Astronautics*), Di Huang (*Beihang University*),
Zhiqiang Shen (*Carnegie Mellon University*), Yi Yang (*University of Technology Sydney*),
Jiebo Luo (*University of Rochester*)
- **TSA-Net: Tube Self-Attention Network for Action Quality Assessment** 4902
Shunli Wang (*Fudan University & Engineering Research Center of AI and Robotics*),
Dingkang Yang (*Fudan University & Ji Hua Laboratory*),
Peng Zhai (*Fudan University & Jilin Provincial Key Laboratory of Intelligence Science and Engineering*),
Chixiao Chen (*Fudan University*), Lihua Zhang (*Ji Hua Laboratory & Fudan University*)

- **Exploring Contextual-Aware Representation and Linguistic-Diverse Expression for Visual Dialog** 4911
Xiangpeng Li, Lianli Gao, Lei Zhao, Jingkuan Song (*University of Electronic Science and Technology of China*)
- **Automated Playtesting with a Cognitive Model of Sensorimotor Coordination** 4920
Injung Lee, Hyunchul Kim (*Korea Advanced Institute of Science and Technology*),
Byungjoo Lee (*Yonsei University*)
- **CAA: Candidate-Aware Aggregation for Temporal Action Detection** 4930
Yifan Ren, Xing Xu, Fumin Shen (*University of Electronic Science and Technology of China*),
Yazhou Yao (*Nanjing University of Science and Technology*), Huimin Lu (*Kyushu Institute of Technology*)

Session 35: Vision and Language-I

- **Disentangle Your Dense Object Detector** 4939
Zehui Chen (*University of Science and Technology of China*), Chenhongyi Yang (*University of Edinburgh*),
Qiaofei Li (*SenseTime*), Feng Zhao, Zheng-Jun Zha, Feng Wu (*University of Science and Technology of China*)
- **Do We Really Need Frame-by-Frame Annotation Datasets for Object Tracking?** 4949
Lei Hu (*Northwest A&F University*), Shaoli Huang (*University of Sydney*),
Shilei Wang (*Northwest A&F University*), Wei Liu (*University of Technology Sydney*),
Jifeng Ning (*Northwest A&F University*)
- **Video-to-Image Casting: A Flattening Method for Video Analysis** 4958
Xu Chen, Chenqiang Gao, Feng Yang (*Chongqing University of Posts and Telecommunications
& Chongqing Key Laboratory of Signal and Information Processing*),
Xiaohan Wang (*Zhejiang University & Baidu Research*), Yi Yang (*Zhejiang University*),
Yahong Han (*Tianjin University & Peng Cheng Laboratory*)
- **Complementary Trilateral Decoder for Fast and Accurate Salient Object Detection** 4967
Zhirui Zhao (*Beihang University*), Changqun Xia (*Peng Cheng Laboratory*),
Chenxi Xie (*Beihang University*), Jia Li (*Beihang University & Peng Cheng Laboratory*)
- **Learning Human Motion Prediction via Stochastic Differential Equations** 4976
Kedi Lyu (*Jilin University*), Zhenguang Liu (*Zhejiang University*),
Shuang Wu (*Nanyang Technological University*), Haipeng Chen (*Jilin University*),
Xuhong Zhang (*Zhejiang University*), Yuyu Yin (*Hangzhou Dianzi University*)
- **Spatio-Temporal Interaction Graph Parsing Networks for Human-Object Interaction Recognition** 4985
Ning Wang, Guangming Zhu (*Xidian University*), Liang Zhang (*Xidian University & Shanghai BNC*),
Peiyi Shen, Hongsheng Li, Cong Hua (*Xidian University*)

Session 36: Vision and Language-II

- **Learning Hierarchical Channel Attention for Fine-grained Visual Classification** 5011
Xiang Guan, Guoqing Wang, Xing Xu, Yi Bin (*University of Electronic Science and Technology of China*)
- **Group-based Distinctive Image Captioning with Memory Attention** 5020
Jiuniu Wang (*City University of Hong Kong, Aerospace Information Research Institute, Chinese Academy of
Sciences, & University of Chinese Academy of Sciences*), Wenjia Xu (*Aerospace Information Research Institute,
Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Qingzhong Wang (*City University of Hong Kong & Baidu Research*),
Antoni B. Chan (*City University of Hong Kong*)
- **VQMG: Hierarchical Vector Quantised and Multi-hops Graph Reasoning for Explicit Representation Learning** 5029
Lei Li, Chun Yuan (*Tsinghua University*)
- **Structure-aware Mathematical Expression Recognition with Sequence-Level Modeling** 5038
Minli Li (*South China University of Technology*), Peilin Zhao (*Tencent AI Lab*),
Yifan Zhang (*National University of Singapore*),
Shuaicheng Niu, Qingyao Wu, Mingkui Tan (*South China University of Technology*)

- **Exploring Logical Reasoning for Referring Expression Comprehension**..... 5047
Ying Cheng, Ruize Wang, Jiashuo Yu, Rui-Wei Zhao, Yuejie Zhang, Rui Feng (*Fudan University*)
- **Direction Relation Transformer for Image Captioning** 5056
Zeliang Song, Xiaofei Zhou, Linhua Dong, Jianlong Tan, Li Guo (*Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences*)

Session 37: Vision and Language-III

- **Contrastive Disentangled Meta-Learning for Signer-Independent Sign Language Translation** 5065
Tao Jin, Zhou Zhao (*Zhejiang University*)
- **Scene Graph with 3D Information for Change Captioning**..... 5074
Zeming Liao (*Guangxi University*), Qingbao Huang (*Guangxi University; South China University of Technology; & MOE of China*), Yu Liang, Mingyi Fu (*Guangxi University*), Yi Cai (*South China University of Technology & MOE of China*), Qing Li (*Hong Kong Polytechnic University*)
- **Progressive Semantic Matching for Video-Text Retrieval** 5083
Hongying Liu (*Xidian University & Peng Cheng Laboratory*), Ruyi Luo (*Xidian University*), Fanhua Shang (*Xidian University & Peng Cheng Laboratory*), Mantang Niu, Yuanyuan Liu (*Xidian University*)
- **Multimodal Asymmetric Dual Learning for Unsupervised Eyeglasses Removal** 5092
Qing Lin, Bo Yan, Weimin Tan (*Fudan University*),
- **Neighbor-view Enhanced Model for Vision and Language Navigation** 5101
Dong An (*Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Yuankai Qi (*University of Adelaide*), Yan Huang (*Institute of Automation, Chinese Academy of Sciences*), Qi Wu (*University of Adelaide*), Liang Wang (*Institute of Automation, Chinese Academy of Sciences, CEBSIT, & CAS-AIR*), Tieniu Tan (*Institute of Automation, Chinese Academy of Sciences & CEBSIT*),
- **Multi-Perspective Video Captioning**..... 5110
Yi Bin (*University of Electronic Science and Technology of China*), Xindi Shang (*National University of Singapore*), Bo Peng (*Tianjin University*), Yujuan Ding (*Hong Kong Polytechnic University*), Tat-Seng Chua (*National University of Singapore*)

Poster Session 6

- **Pairwise VLAD Interaction Network for Video Question Answering**..... 5119
Hui Wang, Dan Guo (*Hefei University of Technology*), Xian-Sheng Hua (*Alibaba Group*), Meng Wang (*Hefei University of Technology*)
- **Attention-guided Temporally Coherent Video Object Matting** 5128
Yunke Zhang, Chi Wang (*Zhejiang University*), Miaomiao Cui, Peiran Ren, Xuansong Xie (*Alibaba Group*), Xian-Sheng Hua (*Damo Academy, Alibaba Group*), Hujun Bao (*Zhejiang University*), Qixing Huang (*The University of Texas at Austin*), Weiwei Xu (*Zhejiang University*)
- **Disentangling Hate in Online Memes** 5138
Roy Ka-Wei Lee, Rui Cao (*Singapore Management University*), Ziqing Fan (*University of Electronic Science and Technology of China*), Jing Jiang, Wen-Haw Chong (*Singapore Management University*)
- **Robust Real-World Image Super-Resolution against Adversarial Attacks** 5148
Jiutao Yue (*Sun Yat-sen University*), Haofeng Li (*The Chinese University of Hong Kong*), Pengxu Wei, Guanbin Li, Liang Lin (*Sun Yat-sen University*)
- **Towards Robust Deep Hiding Under Non-Differentiable Distortions for Practical Blind Watermarking**..... 5158
Chaoning Zhang, Adil Karjauv, Philipp Benz, In So Kweon (*Korea Advanced Institute of Science and Technology*)
- **Bottom-Up and Bidirectional Alignment for Referring Expression Comprehension** 5167
Liuwu Li, Yuqi Bu, Yi Cai (*South China University of Technology & Ministry of Education of China*)
- **SalS-GAN: Spatially-Adaptive Latent Space in StyleGAN for Real Image Embedding** 5176
Lingyun Zhang, Xiuxiu Bai, Yao Gao (*Xi'an Jiaotong University*)

- **Structured Multi-modal Feature Embedding and Alignment for Image-Sentence Retrieval** 5185
Xuri Ge (*University of Glasgow*), Fuhai Chen (*National University of Singapore*),
Joemon M. Jose (*University of Glasgow*), Zhilong Ji, Zhongqin Wu, Xiao Liu (*Tomorrow Advancing Life*)
- **Keyframe Extraction from Motion Capture Sequences with Graph based Deep Reinforcement Learning**..... 5194
Clinton Mo, Kun Hu (*The University of Sydney*), Shaohui Mei (*Northwestern Polytechnical University*),
Zebin Chen (*Electronic Arts*), Zhiyong Wang (*The University of Sydney*)
- **Dense Contrastive Visual-Linguistic Pretraining**..... 5203
Lei Shi, Kai Shuang (*Beijing University of Posts and Telecommunications*),
Shijie Geng (*Rutgers University*),
Peng Gao (*Shanghai AI Laboratory*), Zuohui Fu (*Rutgers University*),
Gerard de Melo (*Hasso Plattner Institute, University of Potsdam*), Yunpeng Chen (*YITU Technology*),
Sen Su (*Beijing University of Posts and Telecommunications*)
- **Hybrid Reasoning Network for Video-based Commonsense Captioning**..... 5213
Weijiang Yu (*Sun Yat-sen University*), Jian Liang (*Peking University*), Lei Ji (*Microsoft Research Asia*),
Lu Li (*Zhejiang University*), Yuejian Fang (*Peking University*), Nong Xiao (*Sun Yat-sen University*),
Nan Duan (*Microsoft Research Asia*)
- **Learning Regularizer for Monocular Depth Estimation with Adversarial Guidance** 5222
Guibao Shen (*Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Yingkui Zhang (*Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences*),
Jialu Li (*Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences*), Mingqiang Wei (*Nanjing University of Aeronautics and Astronautics*),
Qiong Wang, Guangyong Chen (*Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences*),
Pheng-Ann Heng (*The Chinese University of Hong Kong*)
- **Pixel-wise Graph Attention Networks for Person Re-identification** 5231
Wenyu Zhang, Qing Ding, Jian Hu, Yi Ma, Mingzhe Lu (*University of Science and Technology of China*)
- **Neighbor-Vote: Improving Monocular 3D Object Detection through Neighbor Distance Voting**..... 5239
Xiaomeng Chu, Jiajun Deng, Yao Li (*University of Science and Technology of China*),
Zhenxun Yuan (*The University of Sydney*),
Yanyong Zhang, Jianmin Ji, Yu Zhang (*University of Science and Technology of China*)
- **Remember and Reuse: Cross-Task Blind Image Quality Assessment via Relevance-aware Incremental Learning**..... 5248
Rui Ma, Hanxiao Luo, Qingbo Wu, King Ngi Ngan, Hongliang Li, Fanman Meng, Linfeng Xu
(*University of Electronic Science and Technology of China*),
- **MSO: Multi-Feature Space Joint Optimization Network for RGB-Infrared Person Re-Identification** 5257
Yajun Gao, Tengfei Liang, Yi Jin (*Beijing Jiaotong University*),
Xiaoyan Gu (*Institute of Information Engineering, Chinese Academy of Sciences*),
Wu Liu (*JD AI Research*), Yidong Li, Congyan Lang (*Beijing Jiaotong University*)
- **Point Cloud Projection and Multi-Scale Feature Fusion Network Based Blind Quality Assessment for Colored Point Clouds** 5266
Wen-xu Tao, Gang-yi Jiang, Zhi-di Jiang, Mei Yu (*Ningbo University*)
- **Multi-branch Channel-wise Enhancement Network for Fine-grained Visual Recognition** 5273
Guangjun Li, Yongxiong Wang (*University of Shanghai for Science and Technology*),
Fengting Zhu (*Shanghai University*)
- **General Approximate Cross Validation for Model Selection: Supervised, Semi-supervised and Pairwise Learning** 5281
Bowe Zhu, Yong Liu (*Renmin University of China*)
- **Progressive and Selective Fusion Network for High Dynamic Range Imaging**..... 5290
Qian Ye (*Tohoku University*), Jun Xiao, Kin-man Lam (*Hong Kong Polytechnic University*),
Takayuki Okatani (*Tohoku University & RIKEN Center for AIP*)

• Multimodal Relation Extraction with Efficient Graph Alignment	5298
Changmeng Zheng (<i>The Hong Kong Polytechnic University & Ministry of Education</i>), Junhao Feng, Ze Fu, Yi Cai (<i>Ministry of Education & South China University of Technology</i>), Qing Li (<i>The Hong Kong Polytechnic University</i>), Tao Wang (<i>King’s College London</i>)	
• Legitimate Adversarial Patches: Evading Human Eyes and Detection Models in the Physical World	5307
Jia Tan (<i>Xiangtan University & China Academy of Space Technology</i>), Nan Ji, Haidong Xie (<i>China Academy of Space Technology</i>), Xueshuang Xiang (<i>China Academy of Space Technology</i>)	
• Unsupervised Vehicle Search in the Wild: A New Benchmark	5316
Xian Zhong, Shilei Zhao (<i>Wuhan University of Technology</i>), Xiao Wang (<i>Wuhan University of Science and Technology</i>), Kui Jiang (<i>Wuhan University</i>), Wenxuan Liu (<i>Wuhan University of Technology</i>), Wenxin Huang (<i>Hubei University</i>), Zheng Wang (<i>Wuhan University</i>)	
• Meta-FDMixup: Cross-Domain Few-Shot Learning Guided by Labeled Target Data	5326
Yuqian Fu, Yanwei Fu, Yu-Gang Jiang (<i>Fudan University</i>)	
• Target-guided Adaptive Base Class Reweighting for Few-Shot Learning	5335
Jiliang Yan, Deming Zhai, Junjun Jiang, Xianming Liu (<i>Harbin Institute of Technology</i>)	
• Deep Reasoning Network for Few-shot Semantic Segmentation	5344
Yunzhi Zhuge, Chunhua Shen (<i>The University of Adelaide</i>)	
• Heterogeneous Feature Fusion and Cross-modal Alignment for Composed Image Retrieval	5353
Gangjian Zhang, Shikui Wei, Huaxin Pang, Yao Zhao (<i>Beijing Jiaotong University & Beijing Key Laboratory of Advanced Information Science and Network Technology</i>)	
• Similar Scenes Arouse Similar Emotions: Parallel Data Augmentation for Stylized Image Captioning	5363
Guodun Li, Yuchen Zhai, Zehao Lin, Yin Zhang (<i>Zhejiang University</i>)	
• Trajectory is not Enough: Hidden Following Detection	5373
Danni Xu, Ruimin Hu (<i>Wuhan University</i>), Zixiang Xiong (<i>Texas A&M University</i>), Zheng Wang (<i>Wuhan University</i>), Linbo Luo (<i>Xidian University</i>), Dengshi Li (<i>Jiangnan University</i>)	
• Contrastive Learning for Cold-Start Recommendation	5382
Yinwei Wei, Xiang Wang (<i>National University of Singapore</i>), Qi Li (<i>Shandong University</i>), Liqiang Nie (<i>Shandong University</i>), Yan Li, Xuanping Li (<i>Beijing Kuaishou Technology Co., Ltd.</i>), Tat-Seng Chua (<i>National University of Singapore</i>)	
• CG-GAN: Class-Attribute Guided Generative Adversarial Network for Old Photo Restoration	5391
Jixin Liu, Rui Chen, Shipeng An, Heng Zhang (<i>Tianjin University</i>)	
• Get The Best of the Three Worlds: Real-Time Neural Image Compression in a Non-GPU Environment	5400
Zekun Zheng, Xiaodong Wang (<i>National University of Defense Technology</i>), Xinye Lin (<i>National Key Laboratory of Science and Technology on Blind Signal Processing</i>), Shaohe Lv (<i>Seeingdo tech. corp</i>)	
• Visual Language Based Succinct Zero-Shot Object Detection	5410
Ye Zheng (<i>Institute of Computing Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences</i>), Xi Huang, Li Cui (<i>Institute of Computing Technology, Chinese Academy of Sciences</i>)	
• GAMnet: Robust Feature Matching via Graph Adversarial-Matching Network	5419
Bo Jiang (<i>Anhui University & Hefei Comprehensive National Science Center</i>), Pengfei Sun, Ziyang Zhang, Jin Tang, Bin Luo (<i>Anhui University</i>)	
• MCCN: Multimodal Coordinated Clustering Network for Large-Scale Cross-modal Retrieval	5427
Zhixiong Zeng, Ying Sun, Wenji Mao (<i>Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences</i>)	
• AFEC: Adaptive Feature Extraction Modules for Learned Image Compression	5436
Yi Ma, Yongqi Zhai, Jiayu Yang, Chunhui Yang, Ronggang Wang (<i>Peking University</i>)	

- **How Video Super-Resolution and Frame Interpolation Mutually Benefit**..... 5445
Chengcheng Zhou, Zongqing Lu (*Tsinghua University*),
Linge Li, Qiangyu Yan (*Huawei Technologies Co., Ltd*), Jing-Hao Xue (*University College London*)
- **FOCAS: Practical Video Super Resolution using Foveated Rendering**..... 5454
Lingdong Wang, Mohammad Hajiesmaili, Ramesh K. Sitaraman (*University of Massachusetts, Amherst*)
- **Adaptive Affinity Loss and Erroneous Pseudo-Label Refinement for Weakly Supervised Semantic Segmentation** 5463
Xiangrong Zhang, Zelin Peng, Peng Zhu, Tianyang Zhang (*Xidian University*),
Chen Li (*Xi'an Jiaotong University*), Huiyu Zhou (*University of Leicester*), Licheng Jiao (*Xidian University*)
- **Relationship-Preserving Knowledge Distillation for Zero-Shot Sketch Based Image Retrieval**..... 5473
Jialin Tian, Xing Xu (*University of Electronic Science and Technology of China*),
Zheng Wang (*University of Electronic Science and Technology of China & Institute of Electronic and Information Engineering of UESTC in Guangdong*),
Fumin Shen (*University of Electronic Science and Technology of China*), Xin Liu (*Huaqiao University*)
- **Partially Fake it Till you Make It: Mixing Real and Fake Thermal Images for Improved Object Detection** 5482
Francesco Bongini, Lorenzo Berlincioni, Marco Bertini, Alberto Del Bimbo (*Università degli Studi di Firenze*)
- **CDP: Towards Optimal Filter Pruning via Class-wise Discriminative Power** 5491
Tianshuo Xu, Yuhang Wu (*Xiamen University*), Xiawu Zheng (*Xiamen University & Peng Cheng Laboratory*),
Teng Xi, Gang Zhang, Errui Ding (*Baidu Inc.*), Fei Chao (*Xiamen University*),
Rongrong Ji (*Xiamen University & Peng Cheng Laboratory*)
- **Face Hallucination via Split-Attention in Split-Attention Network**..... 5501
Tao Lu, Yuanzhi Wang, Yanduo Zhang, Yu Wang, Liu Wei (*Wuhan Institute of Technology*),
Zhongyuan Wang (*Wuhan University*), Junjun Jiang (*Harbin Institute of Technology*)
- **Fake Gradient: A Security and Privacy Protection Framework for DNN-based Image Classification** 5510
Xianglong Feng, Yi Xie, Mengmei Ye, Zhongze Tang, Bo Yuan, Sheng Wei (*Rutgers University*)
- **Integrating Semantic and Temporal Relationships in Facial Action Unit Detection** 5519
Zhihua Li, Xiang Deng, Xiaotian Li, Lijun Yin (*Binghamton University*)
- **Sparse to Dense Depth Completion using a Generative Adversarial Network with Intelligent Sampling Strategies** 5528
Md Fahim Faysal Khan, Nelson Daniel Troncoso Aldas, Abhishek Kumar (*The Pennsylvania State University*),
Siddharth Advani (*Samsung Electronics America*),
Vijaykrishnan Narayanan (*The Pennsylvania State University*)
- **How does Color Constancy Affect Target Recognition and Instance Segmentation?** 5537
Siyao Xue, Shaobing Gao, Minjie Tan (*Sichuan University*),
Zhen He (*Beijing Institute of Basic Medical Sciences*), Liangtian He (*Anhui University*)
- **Convolutional Transformer based Dual Discriminator Generative Adversarial Networks for Video Anomaly Detection** 5546
Xinyang Feng (*Columbia University*), Dongjin Song (*University of Connecticut*),
Yuncong Chen, Zhengzhang Chen, Jingchao Ni, Haifeng Chen (*NEC Laboratories America, Inc.*)
- **Salient Error Detection based Refinement for Wide-baseline Image Interpolation**..... 5555
Yuan Chang, Yisong Chen, Guoping Wang (*Peking University*)
- **A Multi-Domain Adaptive Graph Convolutional Network for EEG-based Emotion Recognition** 5565
Rui Li, Yiting Wang, Bao-Liang Lu (*Shanghai Jiao Tong University*)
- **Interpolation Variable Rate Image Compression** 5574
Zhenhong Sun, Zhiyu Tan, Xiuyu Sun, Fangyi Zhang, Yichen Qian, Dongyang Li, Hao Li (*Alibaba Group*)
- **Armor: A Benchmark for Meta-evaluation of Artificial Music**..... 5583
Songhe Wang, Zheng Bao, Jingtong E (*University of North Carolina at Chapel Hill*)

- **DRDF: Determining the Importance of Different Multimodal Information with Dual-Router Dynamic Framework** 5591
Haiwen Hong (*Zhejiang University*), Xuan Jin (*Alibaba Group*),
Yin Zhang, Yunqing Hu, Jingfeng Zhang (*Zhejiang University*), Yuan He, Hui Xue (*Alibaba Group*)
- **CoCo-BERT: Improving Video-Language Pre-training with Contrastive Cross-modal Matching and Denoising** 5600
Jianjie Luo (*Sun Yat-set University & Ministry of Education*),
Yehao Li, Yingwei Pan, Ting Yao (*JD AI Research*),
Hongyang Chao (*Sun Yat-set University & Ministry of Education*), Tao Mei (*JD AI Research*)
- **DLA-Net for FG-SBIR: Dynamic Local Aligned Network for Fine-Grained Sketch-Based Image Retrieval**..... 5609
Jiaqing Xu, Haifeng Sun, Qi Qi, Jingyu Wang, Ce Ge, Lejian Zhang, Jianxin Liao
(*Beijing University of Posts and Telecommunications*)
- **Pareto Optimality for Fairness-constrained Collaborative Filtering**..... 5619
Qianxiu Hao (*Institute of Computing Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences*),
Qianqian Xu (*Institute of Computing Technology, Chinese Academy of Sciences*),
Zhiyong Yang (*University of Chinese Academy of Sciences*),
Qingming Huang (*Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences; & Peng Cheng Laboratory*)
- **Decoupled IoU Regression for Object Detection** 5628
Yan Gao (*Alibaba Group*), Qimeng Wang (*Huazhong University of Science and Technology*),
Xu Tang, Haochen Wang, Fei Ding, Jing Li, Yao Hu (*Alibaba Group*)
- **RCNet: Reverse Feature Pyramid and Cross-scale Shift Network for Object Detection** 5637
Zhuofan Zong, Qianggang Cao, Biao Leng (*Beihang University*)
- **Recursive Fusion and Deformable Spatiotemporal Attention for Video Compression Artifact Reduction** 5646
Minyi Zhao, Yi Xu, Shuigeng Zhou (*Fudan University*)
- **JokerGAN: Memory-Efficient Model for Handwritten Text Generation with Text Line Awareness** 5655
Jan Zdenek, Hideki Nakayama (*The University of Tokyo*)

Tutorials

- **Image Quality Assessment in the Modern Age** 5664
Kede Ma (*City University of Hong Kong*), Yuming Fang (*Jiangxi University of Finance and Economics*)
- **Trustworthy Multimedia Analysis**..... 5667
Xiaowen Huang, Jiaming Zhang, Yi Zhang, Xian Zhao, Jitao Sang (*Beijing Jiaotong University*)
- **Multimedia Classifiers: Behind the Scenes**..... 5670
Manjunath Iyer (*Wipro*)
- **Few-shot Learning for Multi-Modality Tasks** 5673
Jie Chen (*Peking University & Peng Cheng Laboratory*),
Qixiang Ye, Xiaoshan Yang, S. Kevin Zhou (*University of Science and Technology of China & Peng Cheng Laboratory*),
Xiaopeng Hong (*Xi'an Jiaotong University & Peng Cheng Laboratory*), Li Zhang (*Fudan University*)
- **Plenoptic Quality Assessment: The JPEG Pleno Experience**..... 5675
Antonio M. G. Pinheiro (*Universidade da Beira Interior and Instituto de Telecomunicacoes*)
- **A Tutorial on AI Music Composition** 5678
Xu Tan (*Microsoft Research Asia*), Xiaobing Li (*Central Conservatory of Music*)
- **Out-of-distribution Generalization and Its Applications for Multimedia**..... 5681
Xin Wang, Peng Cui, Wenwu Zhu (*Tsinghua University*),
- **Deep Learning for Visual Data Compression** 5683
Guo Lu (*Beijing Institute of Technology*), Ren Yang (*ETH Zürich*),
Shenlong Wang (*University of Illinois at Urbana-Champaign*), Shan Liu (*Tencent*),
Radu Timofte (*ETH Zürich & University of Würzburg*)

Workshop Summaries

- **ADVM’21: 1st International Workshop on Adversarial Learning for Multimedia**..... 5686
 Aishan Liu (*Beihang University*), Xinyun Chen (*University of California, Berkeley*),
 Yingwei Li (*Johns Hopkins University*), Chaowei Xiao (*NVIDIA Research*),
 Xun Yang (*National University of Singapore*), Xianglong Liu (*Beihang University*),
 Dawn Song (*University of California, Berkeley*), Dacheng Tao (*JD Explore Academy*),
 Alan Yuille (*Johns Hopkins University*), Anima Anandkumar (*California Institute of Technology*)
- **AIxFood’21: 3rd Workshop on AIxFood**..... 5688
 Ricardo Guerrero (*Samsung AI Center*), Michael Spranger (*Sony Corporation*),
 Shuqiang Jiang (*Chinese Academy of Sciences*), Chong-Wah Ngo (*Singapore Management University*)
- **HUMA’21: 2nd International Workshop on Human-centric Multimedia Analysis**..... 5690
 Wu Liu, Xinchun Liu (*AI Research of JD.com*),
 Jingkuan Song (*University of Electronic Science and Technology of China*),
 Dingwen Zhang (*Northwestern Polytechnical University*),
 Wenbing Huang (*Tsinghua University*), Junbo Guo (*People’s Daily Online*), John Smith (*IBM*)
- **MMSports’21: 4th International Workshop on Multimedia Content Analysis in Sports**..... 5692
 Rainer Lienhart (*University of Augsburg*), Thomas B. Moeslund (*Aalborg University*),
 Hideo Saito (*Keio University*)
- **SUMAC’21: 3rd Workshop on Structuring and Understanding of
 Multimedia heritAge Contents**..... 5694
 Valérie Gouet-Brunet (*University Gustave Eiffel, IGN-ENSG/LASTIG*), Margarita Khokhlova (*Fujitsu*),
 Ronak Kosti (*Friedrich-Alexander-Universität*), Li Weng (*Hangzhou Dianzi University*)
- **UrbanMM’21: 1st International Workshop on Multimedia Computing for Urban Data**..... 5696
 Stevan Rudinac (*University of Amsterdam*), Alessandro Bozzon (*Delft University of Technology*),
 Tat-Seng Chua (*National University of Singapore*), Suzanne Little (*Dublin City University*),
 Daniel Gatica-Perez (*Idiap Research Institute & EPFL*), Kiyoharu Aizawa (*University of Tokyo*)
- **ADGD’21: 1st Workshop on Synthetic Multimedia – Audiovisual Deepfake Generation
 and Detection**..... 5698
 Stefan Winkler, Weiling Chen (*National University of Singapore*),
 Abhinav Dhall (*Monash University & Indian Institute of Technology Ropar*),
 Pavel Korshunov (*Idiap Research Institute*)
- **FME’21: 1st Workshop on Facial Micro-Expression: Advanced Techniques for Facial
 Expressions Generation and Spotting**..... 5700
 Jingting Li (*CAS Key Laboratory of Behavioral Science, Institute of Psychology*),
 Moi Hoon Yap (*Manchester Metropolitan University*),
 Wen-Huang Cheng (*National Yang Ming Chiao Tung University*),
 John See (*Heriot-Watt University*), Xiaopeng Hong (*Xi’an Jiaotong University*),
 Xiaobai Li (*University of Oulu*),
 Su-Jing Wang (*Chinese Academy of Sciences & University of the Chinese Academy of Sciences*),
- **MuCAI’21: 2nd ACM Multimedia Workshop on Multimodal Conversational AI**..... 5702
 Joao Magalhaes (*Universidade Nova de Lisboa*), Alex Hauptmann (*Carnegie Mellon University*),
 Ricardo G. Sousa (*Farfetch*), Carlos Santiago (*Instituto Superior Tecnico, Universidade Lisboa*)
- **MULL’21: First International Workshop on Multimedia Understanding
 with Less Labeling**..... 5704
 Xiu-Shen Wei (*Nanjing University of Science and Technology*), Jufeng Yang (*Nankai University*),
 Han-Jia Ye (*Nanjing University*), Jian Yang (*Nanjing University of Science and Technology*)
- **MuSe 2021 Challenge: Multimodal Emotion, Sentiment, Physiological-Emotion,
 and Stress Detection**..... 5706
 Lukas Stappen (*University of Augsburg*), Eva-Maria Meßner (*University of Ulm*),
 Erik Cambria (*Nanyang Technological University*),
 Guoying Zhao (*University of Oulu*), Björn W. Schuller (*Imperial College London*)

- **Trustworthy AI'21: 1st International Workshop on Trustworthy AI for Multimedia Computing** 5708
 Teddy Furon (*INRIA*), Jingen Liu (*JD AI Research*), Yogesh Rawat (*University of Central Florida*),
 Wei Zhang (*JD AI Research*), Qi Zhao (*University of Minnesota*)
- **WAB'21: 1st Workshop on Multimodal Product Identification in Livestreaming and WAB Challenge** 5710
 Yueting Zhuang (*Zhejiang University*), Xing Tang, Guilin Wu (*Alibaba Group*),
 Yahong Han (*Tianjin University*), Haihong Tang, Xiaobo Li (*Alibaba Group*),
 Xiaohan Wang (*Zhejiang University*), Baoming Yan, Bo Gao (*Alibaba Group*),
 Yi Yang (*University of Technology Sydney*)
- Author Index** 5712

ACM Multimedia 2021 Organization

Honorary Chairs: Yunhe Pan (Zhejiang University, China)
Qionghai Dai (Tsinghua University, China)

General Chairs: Heng Tao Shen (University of Electronic Science & Technology of China, China)
Yueting Zhuang (Zhejiang University, China)
John R. Smith (IBM, USA)

Program Chairs: Yang Yang (University of Electronic Science and Technology of China, China)
Pablo Cesar (CWI & TU Delft, The Netherlands)
Florian Metzger (FACEBOOK, Inc., USA)
Balakrishnan Prabhakaran (University of Texas at Dallas, USA)

Panels Chairs: Xian-Sheng Hua (Alibaba Cloud, China)
Fei Wu (Zhejiang University, China)
Stevan Rudinac (University of Amsterdam, The Netherlands)
Xiaochun Cao (Chinese Academy of Sciences, China)

Tutorial Chairs: Changsheng Xu (Chinese Academy of Sciences, China)
Elisa Ricci (University of Trento & Fondazione Bruno Kessler, Italy)
Dong Xu (The University of Sydney, Australia)
Cathal Gurrin (Dublin City University, Ireland)

Workshop Chairs: Meng Wang (Hefei University of Technology, China)
Lianli Gao (University of Electronic Science and Technology of China, China)
Peng Wang (The University of Wollongong, Australia)
Piotr Koniusz (Data61/CSIRO & Australian National University, Australia)

Brave New Idea Track Chairs: Chang Wen Chen (University at Buffalo, USA)
Zi Huang (The University of Queensland, Australia)
Qin Jin (Renmin University of China, China)

Multimedia Grand Challenge Xirong Li (Renmin University of China, China)
Chairs: Lu Fang (Tsinghua University, China)
Lixin Duan (University of Electronic Science and Technology of China, China)
David Brady (University of Arizona, USA)

Open Source Software Chairs: Jingdong Wang (Microsoft Research Asia, China)
Roger Zimmermann (National University of Singapore, Singapore)
Toshihiko Yamasaki (The University of Tokyo, Japan)
Qi Dai (Microsoft Research Asia, China)

Technical Demo Chairs: Jingkuan Song (University of Electronic Science & Technology of China, China)
Richang Hong (Hefei University of Technology, China)

Video Program Chairs: Jitao Sang (Beijing Jiaotong University, China)
Rainer Lienhart (Universität Augsburg, Germany)

Interactive Arts Chairs: Klaus Schöffmann (Klagenfurt University, Austria)
Minjeong Kim (The University of North Carolina at Greensboro, USA)
Ning Xie (University of Electronic Science and Technology of China, China)

Local Organization Chairs: Ling Tian (University of Electronic Science and Technology of China, China)
Xing Xu (University of Electronic Science and Technology of China, China)

Sponsorship Chairs: Tao Mei (JD.COM, China)
Qi Tian (Huawei, China)

Web and Social Media Chairs: Benoit Huet (MEDIAN Technologies, France)
Xiao Wu (Southwest Jiaotong University, China)

Publicity Chairs: Jinhui Tang (Nanjing University of Science and Technology, China)
Islem Rekik (Istanbul Technical University & BASIRA Lab, Turkey)
Phoebe Chen (La Trobe University, Australia)
Xiaofeng Zhu (University of Electronic Science and Technology of China, China)

Finance Chairs: Yanli Ji (University of Electronic Science and Technology of China, China)
Tatsuya Harada (The University of Tokyo, Japan)

Reproducibility Companion Naoko Nitta (Osaka University, Japan)
Chairs: Michael Alexander Riegler (SimulaMet, Norway)

Doctoral Symposium Chairs: Mohan Kankanhalli (National University of Singapore, Singapore)
Bing Zeng (University of Electronic Science and Technology of China, China)

Travel Grant Chairs: Fumin Shen (University of Electronic Science and Technology of China, China)
Huimin Lu (Kyushu Institute of Technology, Japan)

Industry Track Chairs: Changhu Wang (ByteDance AI Lab, China)
Dong Yu (Tencent AI Lab, USA)
Rongrong Ji (Xiamen University, China)
Jianlong Fu (Microsoft Research Asia, China)

Proceedings Chairs: Liqiang Nie (Shandong University, China)
Qianru Sun (Singapore Management University, Singapore)
Peng Cui (Tsinghua University, China)

Best Paper Committee Chair: Tat-Seng Chua (National University of Singapore, Singapore)

Registration Chairs: Jingjing Chen (Fudan University, China)
Wen Li (University of Electronic Science and Technology of China, China)

ACM Multimedia 2021 Area Chairs

Theme: Engaging Users with Multimedia

- Emotional and Social Signals:** Alan Hanjalic (TU Delft, Netherlands)
Albert Ali Salah (Utrecht University)
Ali Karime (The Royal Military College of Canada)
Bruno Lepri (FBK, Trento, Italy)
Karan Sikka (SRI International)
Laura Cabrera-Quiros (Instituto Tecnológico de Costa Rica)
Mohamed Daoudi (IMT Lille Douai)
Oya Celiktutan (King's College London)
Ramanathan Subramanian (University of Canberra)
Roland Goecke (University of Canberra)
Symeon Papadopoulos (Information Technologies Institute, Centre for Research and Technology Hellas)
Weihong Deng (Beijing University of Posts & Telecommunications)
- Social Multimedia:** Akisato Kimura (NTT Communication Science Laboratories)
Jia Wu (Macquarie University)
Pradeep Atrey (University at Albany, SUNY)
- Search and Recommendation:** An-An Liu (Tianjin University)
Changsheng Xu (CASIA)
Fumin Shen (University of Electronic Science and Technology of China)
Hongtao Xie (University of Science and Technology of China)
Jiajun Liu (CSIRO)
Jingkuan Song (University of Electronic Science and Technology of China)
Lei Zhu (Shandong Normal University)
Markus Schedl (Johannes Kepler University Linz)
Minh-Son Dao (National Institute of Information & Communications Technology)
Qi Tian (Huawei Cloud & AI)
Qianqian Xu (Chinese Academy of Sciences)
Qingming Huang (University of Chinese Academy of Sciences)
Rongrong Ji (Xiamen University)
Shin'ichi Satoh (National Institute of Informatics)
Stevan Rudinac (University of Amsterdam)
Tong Yu (Carnegie Mellon University)
Wei Liu (Tencent)
Werner Bailer (JOANNEUM RESEARCH)
Wu Liu (AI Research of JD.com)
Xin-Shun Xu (Shandong University)
Xirong Li (Renmin University of China)
Zheng Zhang (Harbin Institute of Technology, Shenzhen)
Zhineng Chen (Fudan University)
Zhiyong Cheng (Shandong Academy of Sciences)

Summarization, Analytics, and Storytelling: Benjamin Renoust (Osaka University)
Frank Nack (University of Amsterdam)
Guillaume Gravier (CNRS, IRISA)
Ichiro Ide (Nagoya University)
Joao Magalhaes (Universidade NOVA Lisboa)
Marcel Worring (University of Amsterdam)
Mike Zheng Shou (Facebook Research)
Mukesh Saini (IIT Ropar)

Theme: Multimedia Experience

Emerging Multimedia Applications: Cathal Gurrin (Dublin City University)
Abdulmoteleb EI Saddik (University of Ottawa)
Forouzan Golshani (California State University, Long Beach)
Min Wu (University of Maryland)
Mohammad Akbari (University College London)
Ron Vetter (University of North Carolina in Wilmington)
Wen-Huang Cheng (National Chiao Tung University)

Art, Entertainment and Culture: Asreen Rostami (Research Institute of Sweden)
Jia Jia (Tsinghua University)
John Kender (Columbia University)
Lorenzo Seidenari (University of Florence)
Miriam Redi (Wikimedia Foundation)

HCI and Quality of Experience: Axel Carlier (IRIT, University of Toulouse)
Emin Zerman (Trinity College Dublin)
Estêvão Bissoli Saleme (Federal Institute of Espirito Santo)
Irene Viola (CWI Amsterdam)
Jesús Gutiérrez (Universidad Politécnica de Madrid)
Maria Torres Vega (Universiteit Gent)
Muhammad Shahid Farid (University of the Punjab)
Mylene Farias (UnB)
Rodrigo Laiola Guimaraes (Federal University of Espirito Santo)
Tingting Jiang (Peking University)

Music, Speech, Audio Processing: Mustafa SERT (Baskent University)

Theme: Multimedia Systems

Mobile Multimedia: Cise Midoglu (Simula Research Laboratory (SRL))

Cloud, Edge and Device Computing: Anand Bhojan (National University of Singapore)
Juan Gutierrez-Aguado (Universitat de València)
Yonggang Wen (Nanyang Technological University)

Data Systems Management and Indexing: Carl Taswell (UC San Diego)
Laurent Amsaleg (CNRS-IRISA)
Vincent Oria (NJIT)

Systems and Middleware: Debora Muchaluat-Saade (Universidade Federal Fluminense)
Herman Engelbrecht (University of Stellenbosch)
Hongbin Sun (Xi'an Jiaotong University)
Wu-chi Feng (Portland State University)
Yao Liu (SUNY Binghamton)
Zhisheng Yan (George Mason University)

Telepresence and Virtual/Augmented Reality: Fons Kuijk (CWI)
Kun Li (Tianjin University)
Mohammad Hosseini (UIUC)
Roberto Azevedo (ETH Zurich)
Zhihan Lv (Qingdao University)

Transport and Delivery: Abdelhak Bentaleb (National University of Singapore)
Ali Begen (Ozyegin University)
Christian Timmerer (Alpen-Adria-Universität Klagenfurt)
Fernando Boronat (Universitat Politècnica de València)
Hermann Hellwagner (University of Klagenfurt)
Hongkai Xiong (Shanghai Jiao Tong University)
Lucile Sassatelli (Universite Cote d'Azur)
Manuel Perez Malumbres (Miguel Hernandez University)
Ming Li (California State University, USA)
Roger Zimmermann (National University of Singapore)
Viswanathan (Vishy) Swaminathan (Adobe)
Wei Tsang Ooi (National University of Singapore)
Yong Cui (Tsinghua University)
Yuansong Qiao (Athlone Institute of Technology)

Theme: Understanding Multimedia Content

Deep Learning for Multimedia: Chang Xu (University of Sydney)
Chen Chen (University of Central Florida)
Cheng Jin (Fudan University)
Chengcui Zhang (The University of Alabama at Birmingham)
Dan Xu (The Hong Kong University of Science and Technology)
Diana Tobon Vallejo (Universidad de Medellín)
Elisa Ricci (University of Trento)
Feng Zheng (Southern University of Science and Technology)
Gabriel-Miro Muntean (Dublin City University)
Georg Thallinger (JOANNEUM RESEARCH, Austria)
Haojie Li (Dalian University of Technology)
Houwen Peng (Microsoft Research)
Hua Zhang (Chinese Academy of Sciences)
Huadong Ma (Beijing University of Posts and Telecommunications)
Humphrey Shi (University of Oregon | UIUC)
Jelena Tesic (Texas State University)
Chen Chen (University of Central Florida)
Cheng Jin (Fudan University)

Deep Learning for Multimedia Chengcui Zhang (The University of Alabama at Birmingham)
(continued): Dan Xu (The Hong Kong University of Science and Technology)
Diana Tobon Vallejo (Universidad de Medellín)
Elisa Ricci (University of Trento)
Feng Zheng (Southern University of Science and Technology)
Gabriel-Miro Muntean (Dublin City University)
Georg Thallinger (JOANNEUM RESEARCH, Austria)
Haojie Li (Dalian University of Technology)
Houwen Peng (Microsoft Research)
Hua Zhang (Chinese Academy of Sciences)
Huadong Ma (Beijing University of Posts and Telecommunications)
Humphrey Shi (University of Oregon | UIUC)
Jelena Tesic (Texas State University)
Jialie Shen (Queen's University, Belfast)
Jianlong Fu (Microsoft Research)
Joaquim Jorge (INESC-ID, Instituto Superior Técnico, Universidade de Lisboa)
Jong-Seok Lee (Yonsei University, Korea)
Jun Ye (Facebook)
Junhui Hou (City University of Hong Kong)
Junliang Xing (Institute of Automation, Chinese Academy of Sciences)
Koichi SHINODA (Tokyo Institute of Technology)
Lamberto Ballan (University of Padova)
Lei Zhang (International Digital Economy Academy (IDEA))
Lei Zhu (The Chinese University of Hong Kong)
Liang Zheng (Australian National University)
Liansheng Zhuang (University of Science and Technology of China)
Lin Sun (Samsung, USA)
Lorenzo Baraldi (University of Modena and Reggio Emilia)
Marco Bertini (University of Florence)
Mario Montagud (i2CAT & Universitat de València)
Michael Riegler (Simula)
Michael Zink (UMASS, USA)
Mingbo Zhao (Donghua University)
Noel Connor (Dublin City University)
Peng Wang (University of Wollongong)
Petia Radeva (University of Barcelona)
Qiang Qiu (Purdue University)
Qingzhong Liu (Sam Houston State University)
Rainer Lienhart (Universitat Augsburg, Germany)
Stefano Berretti (University of Florence, Italy)
Toshihiko Yamasaki (The University of Tokyo)
Vasileios Mezaris (Information Technologies Institute, Centre for Research
and Technology Hellas)
Weiyao Lin (Shanghai Jiao Tong university)
Xi Li (Zhejiang University)
Xiang Bai (Huazhong University of Science and Technology)
Xianglong Liu (Beihang University)

Deep Learning for Multimedia Xianglong Liu (Beihang University)
(continued): Xiao Wu (Southwest Jiaotong University)
Xiaolong Wang (IBM)
Xiaopeng Hong (Xi'an Jiaotong University)
Xueliang Liu (Hefei University of Technology)
Yao Zhao (Beijing Jiaotong University)
Yazhou Yao (Nanjing University of Science and Technology)
Yilin Wang (Arizona State University)
Yogesh Rawat (University of Central Florida)
Zechao Li (Nanjing University of Science and Technology)
Zhaowen Wang (Adobe)
Zhaoxiang Zhang (Chinese Academy of Sciences)
Zheng-Jun Zha (University of Science and Technology of China)
Zi Huang (The University of Queensland)

Media Interpretation: Di Huang (Beihang University)
Hailin Shi (JD AI Research)
Jitao Sang (Beijing Jiaotong University)
SUBHABRATA BHATTACHARYA (Netflix)
Xiaobai Liu (San Diego State University)
Yuchao Dai (Northwestern Polytechnical University)
Zhao Zhang (Hefei University of Technology)

Multimodal Analysis and Benoit Huet (MEDIAN Technologies)
Description: Bingkun Bao (Nanjing University of Posts and Telecommunications)
Bogdan Ionescu (Laboratorul de Analiza si Prelucrarea Imaginilor, Romania)
Chong-Wah Ngo (Singapore Management University)
Jie Wen (Harbin Institute of Technology, Shenzhen)
Jingjing Li (University of Electronic Science and Technology of China)
Jungong Han (Aberystwyth University)
Ralph Ewerth (TIB - Leibniz Information Center for Science and Technology)
Shuhui Wang (Chinese Academic of Science)
Shuqiang Jiang (Chinese Academy of Science)
Wei-Ta Chu (National Cheng Kung University)
Yuxin Peng (Peking University)
Zechao Li (Nanjing University of Science and Technology)

Multimodal Fusion and Junchi Yan (Shanghai Jiao Tong University)
Embedding: Liqiang Nie (Shandong University)
Ming Shao (University of Massachusetts Dartmouth)
Tianzhu Zhang (University of Science and Technology of China)
Xavier Alameda-Pineda (INRIA)
Xiaofeng Zhu (Guangxi Normal University)
Xing Xu (University of Electronic Science and Technology of China)
Xu Zhang (Amazon)
Yan Chen (University of Science and Technology of China)
Yi Yu (NII)

Vision and Language: Chuang Gan (MIT-IBM Watson AI Lab)
Feng Lu (Beihang University)
Hanli Wang (Tongji University)
Hanwang Zhang (Nanyang Technological University)
Jianfei Cai (Monash University)
Jingliang Peng (University of Jinan)
Jinhui Tang (Nanjing University of Science and Technology)
Junyu Gao (CASIA)
Lei Zhang (Chongqing University)
Lu Jiang (Google Research)
Marc A. Kastner (National Institute of Informatics)
Massimo Tistarelli (University of Sassari)
Meng Liu (Shandong Jianzhu University)
Nadia Kanwal (Athlone Institute of Technology)
Niall Murray (Athlone Institute of Technology)
Piotr Bilinski (University of Warsaw)
Qin Jin (Renmin University of China)
Richang Hong (Hefei University of Technology)
Sathyanarayanan Aakur (Oklahoma State University)
Si Liu (Beihang University)
Tao Mei (AI Research of JD.com)
Umapada Pal (Indian Statistical Institute, Kolkata)
Waqas Sultani (Information Technology University)
Weiqing Min (Chinese Academy of Sciences)
Xavier Giro-i-Nieto (Universitat Politècnica de Catalunya)
Xiangyu Zhu (Chinese Academy of Science)
Xiaojun Chang (RMIT University)
Xinxiao Wu (Beijing Institute of Technology)
Yadong Mu (Peking University)
Yi Bin (University of Electronic Science and Technology of China)
Yingwei Pan (JD AI Research)
Yu Cheng (Microsoft Research)
Yu Kong (Rochester Institute of Technology)
Yulan Guo (National University of Defense Technology)

ACM Multimedia 2021 Program Committee

Bentaleb Abdelhak (National University of Singapore)
EI Saddik Abdulmotaleb (University of Ottawa)
Kimura Akisato (NTT Communication Science Laboratories)
Hanjalic Alan (TU Delft, Netherlands)
Salah Albert Ali (Utrecht University)
Begen C. Ali (Ozyegin University)
Karime Ali (The Royal Military College of Canada)
Liu An-An (Tianjin University)
BHOJAN ANAND (National University of Singapore)
Rostami Asreen (Research Institute of Sweden)
Carlier Axel (IRIT, University of Toulouse)
Renoust Benjamin (Osaka University)
Huet Benoit (MEDIAN Technologies)
Bao Bingkun (Nanjing University of Posts and Telecommunications)
Ionescu Bogdan (Laboratorul de Analiza si Prelucrarea Imaginilor, Romania)
Lepri Bruno (FBK, Trento, Italy)
Taswell Carl (University of California, San Diego)
Gurrin Cathal (Dublin City University)
Xu Chang (University of Sydney)
Xu Changsheng (CASIA)
Chen Chen (University of Central Florida)
Jin Cheng (Fudan University)
Zhang Chengcui (The University of Alabama at Birmingham)
Ngo Chong-Wah (Singapore Management University)
Timmerer Christian (Alpen-Adria-Universität Klagenfurt)
Gan Chuang (MIT-IBM Watson AI Lab)
Midoglu Cise (Simula Research Laboratory (SRL))
Xu Dan (The Hong Kong University of Science and Technology)
Muchaluat-Saade Debora (UFF, Brazil)
Huang Di (Beihang University, China)
Tobon Vallejo P Diana (Universidad de Medellín)
Ricci Elisa (University of Trento)
Zerman Emin (Trinity College Dublin)
Bissoli Saleme Estêvão (Federal Institute of Espirito Santo)
Lu Feng (Beihang University)
Zheng Feng (Southern University of Science and Technology)
Boronat Fernando (Universitat Politècnica de València)
Kuijk Fons (CWI)
Golshani Forouzan (California State University, Long Beach)
Nack Frank (University of Amsterdam, NL)
Shen Fumin (University of Electronic Science and Technology of China)
Muntean Gabriel-Miro (Dublin City University)
Thallinger Georg (JOANNEUM Research, Austria)
Gravier Guillaume (CNRS, IRISA)
Borgefors Gunilla (Uppsala University)
Shi Hailin (JD AI Research)
Wang Hanli (Tongji University)
Zhang Hanwang (Nanyang Technological University)
Li Haojie (Dalian University of Technology)
Engelbrecht Herman (University of Stellenbosch)
Hellwagner Hermann (University of Klagenfurt)
Sun Hongbin (Xi'an Jiaotong University)
Xiong Hongkai (Shanghai Jiao Tong University)
Xie Hongtao (University of Science and Technology of China)
Peng Houwen (Microsoft Research)
Zhang Hua (Institute of Information Engineering, Chinese Academy of Sciences)
Ma Huadong (Beijing University of Posts and Telecommunications)
Shi Humphrey (University of Oregon | UIUC)
Ide Ichiro (Nagoya University)
Viola Irene (CWI Amsterdam)
Testic Jelena (Texas State University)
Gutiérrez Jesús (Universidad Politécnica de Madrid)
Jia Jia (Tsinghua University)
Wu Jia (Macquarie University)
Liu Jiajun (CSIRO)
Shen Jialie (Queen's University, Belfast (QUB))
Cai Jianfei (Monash University)
Fu Jianlong (Microsoft Research)
Wen Jie (Harbin Institute of Technology, Shenzhen)
Li Jingjing (University of Electronic Science and Technology of China)
Song Jingkuan (UESTC)
Peng Jingliang (University of Jinan)
Tang Jinhui (Nanjing University of Science and Technology)

Sang Jitao (Beijing Jiaotong University, China)
Magalhaes Joao (Universidade NOVA Lisboa)
Jorge Joaquim (INESC-ID, Instituto Superior Técnico,
Universidade de Lisboa)
Kender John (Columbia University)
Lee Jong-Seok (Yonsei University, Korea)
Gutierrez-Aguado Juan (Universitat de València)
Ye Jun (Facebook)
Yan Junchi (Shanghai Jiao Tong University)
Han Jungong (Aberystwyth University)
Hou Junhui (City University of Hong Kong)
Xing Junliang (Institute of Automation, Chinese
Academy of Sciences)
Gao Junyu (CASIA)
Sikka Karan (SRI International)
Wac Katarzyna (University of Copenhagen)
Shinoda Koichi (Tokyo Institute of Technology)
Li Kun (Tianjin University)
Ballan Lamberto (University of Padova)
Cabrera-Quiros Laura (Instituto Tecnológico de Costa
Rica)
Amsaleg Laurent (CNRS-IRISA)
Zhang Lei (International Digital Economy Academy
(IDEA))
Zhang Lei (Chongqing University)
Zhu Lei (The Chinese University of Hong Kong)
Zhu Lei (Shandong Normal University)
Zheng Liang (Australian National University)
Zhuang Liansheng (University of Science and
Technology of China)
Sun Lin (Samsung, USA)
Nie Liqiang (Shandong University)
Baraldi Lorenzo (University of Modena and Reggio
Emilia)
Seidenari Lorenzo (University of Florence)
Jiang Lu (Google Research)
Sassatelli Lucile (Universite Cote d'Azur)
Perez Malumbres Manuel (Miguel Hernandez
University)
Kastner Marc A. (National Institute of Informatics)
Worring Marcel (University of Amsterdam)
Bertini Marco (University of Florence)
Torres Vega Maria (Universiteit Gent)
Montagud Mario (i2CAT & Universitat de València)
Schedl Markus (Johannes Kepler University Linz)
Tistarelli Massimo (University of Sassari)
Liu Meng (Shandong Jianzhu University)
Riegler Michael (Simula)
Zink Michael (UMASS)
Shou Mike Zheng (Facebook Research)
Wu Min (University of Maryland)
Li Ming (California State University)
Shao Ming (University of Massachusetts Dartmouth)
Zhao Mingbo (Donghua University)
Dao Minh-Son (National Institute of Information and
Communications Technology)
Redi Miriam (Wikimedia Foundation)
Daoudi Mohamed (IMT Lille Douai)
Akbari Mohammad (University College London)
Hosseini Mohammad (UIUC)
Farid Muhammad Shahid (University of the Punjab)
Saini Mukesh (IIT Ropar)
SERT Mustafa (Baskent University)
Farias Mylene (UnB)
Kanwal Nadia (Athlone Institute of Technology)
Murray Niall (Athlone Institute of Technology)
Connor O Noel (Home)
Celiktutan Oya (King's College London)
Wang Peng (University of Wollongong)
Radeva Petia (University of Barcelona)
Bilinski Piotr (University of Warsaw)
Bhattacharya Prabir (Thomas Edison State University)
Atrey K. Pradeep (University at Albany, SUNY)
Tian Qi (Huawei Cloud & AI)
Qiu Qiang (Purdue University)
Xu Qianqian (Key Laboratory of Intelligent
Information Processing, Institute of Computing
Technology, Chinese Academy of Sciences)
Jin Qin (Renmin University of China)
Huang Qingming (University of Chinese Academy of
Sciences)
Liu Qingzhong (Sam Houston State University)
Lienhart Rainer (Universitat Augsburg, Germany)
Ewerth Ralph (TIB - Leibniz Information Center for
Science and Technology)
Subramanian Ramanathan (University of Canberra)
Hong Richang (Hefei University of Technology)
Azevedo Roberto (ETH Zurich)
Laiola Guimaraes Rodrigo (Federal University of
Espírito Santo)
Zimmermann Roger (NUS)
Goecke Roland (University of Canberra)
Vetter Ron (University of North Carolina in
Wilmington)
Ji Rongrong (Xiamen University, China)
Aakur N Sathyanarayanan (OK State)

Satoh Shin'ichi (National Institute of Informatics)
Wang Shuhui (VIPL, ICT, Chinese Academic of Science)
Jiang Shuqiang (ICT, China Academy of Science)
Liu Si (Beihang University)
Berretti Stefano (University of Florence, Italy)
Rudinac Stevan (University of Amsterdam)
Bhattacharya Subhabrata (Netflix)
Papadopoulos Symeon (Information Technologies Institute / Centre for Research & Technology - Hellas, GR)
Mei Tao (AI Research of JD.com)
Zhang Tianzhu (University of Science and Technology of China)
Jiang Tingting (Peking University)
Yu Tong (Carnegie Mellon University)
Yamasaki Toshihiko (The University of Tokyo)
Pal Umapada (Indian Statistical Institute, Kolkata)
Mezaris Vasileios (Information Technologies Institute, Centre for Research and Technology Hellas, Greece)
Oria Vincent (NJIT)
Swaminathan Viswanathan (Vishy) (Adobe)
Sultani Waqas (Information Technology University)
Liu Wei (Tencent)
Ooi Wei Tsang (National University of Singapore)
Deng Weihong (Beijing University of Posts and Telecommunications)
Min Weiqing (Institute of Computing Technology, Chinese Academy of Sciences)
Chu Wei-Ta (National Cheng Kung University)
Lin Weiyao (Shanghai Jiao Tong university)
Cheng Wen-Huang (National Chiao Tung University)
Bailer Werner (JOANNEUM Research)
Liu Wu (AI Research of JD.com)
Feng Wu-chi (Portland State University)
Alameda-Pineda Xavier (INRIA)
Giro-i-Nieto Xavier (Universitat Politecnica de Catalunya)
Li Xi (Zhejiang University)
Bai Xiang (Huazhong University of Science and Technology)
Liu Xianglong (Beihang University)
Zhu Xiangyu (Chinese Academy of Science)
Wu Xiao (Southwest Jiaotong University)
Liu Xiaobai (San Diego State University)
Zhu Xiaofeng (Guangxi Normal University)
Chang Xiaojun (RMIT University)
Wang Xiaolong (IBM)
Hong Xiaopeng (Xi'an Jiaotong University)
Xu Xing (University of Electronic Science and Technology of China)
Xu Xin-Shun (Shandong University)
Wu Xinxiao (Beijing Institute of Technology)
Li Xirong (Renmin University of China)
Zhang Xu (Amazon)
Liu Xueliang (Hefei University of Technology)
Mu Yadong (Peking University)
Chen Yan (University of Science and Technology of China)
Liu Yao (SUNY Binghamton)
Zhao Yao (Beijing Jiaotong University)
Yao Yazhou (Nanjing University of Science and Technology)
Bin Yi (University of Electronic Science and Technology of China)
Yu Yi (NII)
Wang Yilin (Arizona State University)
Pan Yingwei (JD AI Research)
Rawat Yogesh (University of Central Florida)
Cui Yong (Tsinghua University)
Wen Yonggang (Nanyang Technological University)
Cheng Yu (Microsoft Research)
Kong Yu (Rochester Institute of Technology)
Qiao Yuansong (Athlone Institute of Technology)
Dai Yuchao (Northwestern Polytechnical University)
Guo Yulan (National University of Defense Technology)
Peng Yuxin (Peking University)
Li Zechao (NJUST)
Li Zechao (Nanjing University of Science and Technology)
Zhang Zhao (Hefei University of Technology)
Wang Zhaowen (Adobe)
Zhang Zhaoxiang (Chinese Academy of Sciences, China)
Zhang Zheng (Harbin Institute of Technology, Shenzhen)
Zha Zheng-Jun (University of Science and Technology of China)
Lv Zhihan (Qingdao University)
Chen Zhineng (CAS)
Yan Zhisheng (George Mason University)
Cheng Zhiyong (Shandong Academy of Sciences)
Huang Helen Zi (University of Queensland)

ACM Multimedia 2021 Sponsor & Supporters

Sponsor



Diamond
Supporters



Platinum
Supporters



Gold
Supporters

