

# AGENDA OVERVIEW

V

WED. | OCT. 20, 2021

Time	Arrangement	Room
8:00 AM-9:00 PM	Sign-in & Conference Materials Collection	Hall of The Langbo Chengdu

Time	Arrangement	Room	Zoom Meeting ID
9:00 AM-12:00 PM	<b>Workshop-MMSports'21</b>	Langbo Room A	817 9260 5901
	<b>Workshop -Trustworthy AI 2021</b>	Langbo Room B	864 1291 4191
	<b>Workshop -MuCAI 2021</b>	Tianfu Room A	830 6390 3376
	<b>Workshop-ADVM 2021</b>	Tianfu Room B	842 2630 3581
	<b>Tutorial-Multimedia classifiers: Behind the Scenes</b>	Jincheng Room	852 4844 0073

12:00 PM-2:00PM	Lunch Break @ Hall Way
-----------------	------------------------

2:00 PM-5:00 PM	<b>Workshop-SUMAC 2021</b>	Langbo Room A	817 9260 5901
	<b>Workshop-HUMA 2021</b>	Langbo Room B	864 1291 4191
	<b>Workshop-WAB 2021</b>	Tianfu Room A	830 6390 3376
	<b>Workshop-UrbanMM 2021</b>	Rongcheng Room	879 4652 5187
	<b>Tutorial: Trustworthy Multimedia Analysis</b>	Tianfu Room B	842 2630 3581
	<b>Tutorial: Out-of-distribution Generalization and Its Applications for Multimedia</b>	Jincheng Room	852 4844 0073

7:00 PM~9:00 PM	Welcome Reception @ Langbo (Whole Room)
-----------------	---

THU. | OCT. 21, 2021 | MORNING

Time	Arrangement	Room	Zoom Meeting ID
8:15 AM-8:30 AM	<b>Opening Ceremony</b>	Langbo (Whole Room)	817 9260 5901

**Keynote Session Chair:** Heng Tao Shen, University of Electronic Science and Technology of China

8:30 AM-9:20 AM	<b>Keynote 1: Wen Gao</b> Video Coding for Machine	Langbo (Whole Room)	817 9260 5901
9:20 AM-10:10 AM	<b>Keynote 2: H. V. Jagadish</b> Semantic Media Conversion: Possibilities and Limits		

10:10 AM-10:30 AM	Coffee Break
-------------------	--------------

10:30 AM-12:00 PM	<b>Session 1: Deep Learning for Multimedia-I</b> <b>Session Chair:</b> Cong Bai, Zhejiang University of Technology	Langbo Room A	817 9260 5901
	<b>Session 2: Deep Learning for Multimedia-II</b> <b>Session Chair:</b> Jianxin Sun, Institute of Automation, Chinese Academy of Sciences	Langbo Room B	864 1291 4191
	<b>Session 3: Deep Learning for Multimedia-III</b> <b>Session Chair:</b> Jin Wang, Beijing University of Technology	Tianfu Room A	830 6390 3376
	<b>Session 4: Brave New Idea</b> <b>Session Chair:</b> Qin Jin, Renmin University of China	Tianfu Room B	842 2630 3581
	<b>Session 5: Emerging Multimedia Applications-I</b> <b>Session Chair:</b> Weiqing Min, Chinese Academy of Sciences	Jiaozi Room	870 2448 7906
	<b>Session 6: Emerging Multimedia Applications-II</b> <b>Session Chair:</b> Le Zhang, University of Electronic Science and Technology of China	Jincheng Room	852 4844 0073
	<b>Session 7: Multimedia HCI and Quality of Experience</b> <b>Session Chair:</b> Qingbo Wu, University of Electronic Science and Technology of China	Rongcheng Room	879 4652 5187

12:00 PM-2:00 PM	Lunch Break @ Hall Way
------------------	------------------------



**THU. | OCT. 21, 2021 | AFTERNOON**

Time	Arrangement	Room	Zoom Meeting ID
2:00 PM-3:00 PM	<b>Panel 1: Social Signals and Multimedia: Past, Present, Future</b> <b>Session Chairs:</b> Jie Shao, University of Electronic Science and Technology of China Hayley Hung, Delft University of Technology	Langbo (Whole Room)	817 9260 5901
2:00 PM-4:00 PM	<b>Session 8: Industrial Track</b> <b>Session Chairs:</b> Rongrong Ji, Xiamen University Jianlong Fu, Microsoft Research Asia	Tianfu Room A	830 6390 3376
3:00 PM-3:30 PM	Coffee Break		
3:30 PM-5:00 PM	<b>Session 9: Emerging Multimedia Applications-III</b> <b>Session Chair:</b> Shuai Li, Shandong University	Langbo Room A	817 9260 5901
	<b>Session 10: Emerging Multimedia Applications-IV</b> <b>Session Chair:</b> Jing Dong, Institute of Automation	Langbo Room B	864 1291 4191
	<b>Session 11: Emotional and Social Signals in Multimedia</b> <b>Session Chair:</b> Jia Jia, Tsinghua University	Tianfu Ballroom B	842 2630 3581
	<b>Session 12: Multimodal Analysis and Description-I</b> <b>Session Chair:</b> Haonan Cheng, Communication University of China	Jiaozi Room	870 2448 7906
	<b>Session 13: Multimodal Analysis and Description-II</b> <b>Session Chair:</b> Zhigang Tu, Wuhan University	Jincheng Room	852 4844 0073
	<b>Session 14: Multimedia Cloud, Edge and Device Computing</b> <b>Session Chair:</b> Yandong Guo, OPPO Research Institute	Rongcheng Room	879 4652 5187



FRI. | OCT. 22, 2021 | MORNING

Time	Arrangement	Room	Zoom Meeting ID
------	-------------	------	-----------------

**Session Chairs:** Tat-Seng Chua, National University of Singapore

Yang Yang, University of Electronic Science and Technology of China

08:30 AM-10:00 AM	<b>Session 15:</b> Best Paper Session	Langbo (Whole Room)	817 9260 5901
-------------------	---------------------------------------	------------------------	---------------

10:10 AM-10:30 AM	Coffee Break		
-------------------	--------------	--	--

**Session Chairs:** Phoebe Chen (Vice-Chair), The ACM SIGMM Executive Committee

Yueting Zhuang (Onsite Chair), Zhejiang University

10:30 AM-12:00 PM	<b>Session 16:</b> ACM SIGMM 2021 Awards	Langbo (Whole Room)	817 9260 5901
	10:30-11:00   Outstanding PhD Thesis Presentation		
	10:30-11:00   ACM SIGMM 2021 Test of Time Paper Award		
	11:00-11:30   Rising Star Talk		
	11:30-12:00   Technical Achievement Awards		

12:00 PM-2:00 PM	Lunch Break		
------------------	-------------	--	--

**FRI. | OCT. 22, 2021 | AFTERNOON**

Time	Arrangement	Room	Zoom Meeting ID
------	-------------	------	-----------------

**Keynote Session Chair:** Yang Yang, University of Electronic Science and Technology of China

2:00 PM-2:50 PM	<b>Keynote 3:</b> Cordelia Schmid Do you see what I see? Large-scale Learning from Multimodal Videos	Langbo (Whole Room)	817 9260 5901
2:50 PM-3:40 PM	<b>Keynote 4:</b> Jingren Zhou Large-scale Multi-Modality Pretrained Models: Applications and Experiences		

3:40 PM-4:00 PM	Coffee Break		
-----------------	--------------	--	--

4:00 PM-5:00 PM	<b>Panel 2:</b> The Next Generation of Multimedia Search and Recommendation <b>Session Chairs:</b> Kun Kuang, Zhejiang University Alan Smeaton, Dublin City University	Langbo (Whole Room)	817 9260 5901
4:00 PM-6:00 PM	<b>Session 17:</b> Video Program and Demo Session <b>Session Chair:</b> Jingkuan Song, University of Electronic Science and Technology of China	Tianfu Room A	830 6390 3376
	<b>Session 18:</b> Doctoral Symposium <b>Session Chair:</b> Shuaicheng Liu, University of Electronic Science and Technology of China	Tianfu Room B	842 2630 3581
3:00 PM-4:00 PM	<b>TOMM Executive Meeting</b>	Online	
4:00 PM-5:00 PM	<b>ACM SIGMM Executive Meeting</b>	Online	
5:30 PM-6:30 PM	<b>ACM SIGMM Business Meeting</b> <b>Session Chairs:</b> The ACM SIGMM Executive Committee Alberto del Bimbo (Chair) Phoebe Chen (Vice-Chair) Miriam Redi (Director of Conferences)	Jincheng Room	879 4652 5187

7:00 PM-9:00 PM	Banquet Dinner @ Langbo Ballroom		
-----------------	----------------------------------	--	--



SAT. | OCT. 23, 2021 | MORNING

Time	Arrangement	Room	Zoom Meeting ID
------	-------------	------	-----------------

**Keynote Session Chair:** Jingkuan Song, University of Electronic Science and Technology of China

08:30 AM-09:20 AM	<b>Keynote 5:</b> James Lester AI and the Future of Education	Langbo (Whole Room)	817 9260 5901
09:20 AM-10:10 AM	<b>Keynote 6:</b> Zhengyou Zhang Digital Human in an Integrated Physical-Digital World (IPhD)		

10:10 AM-10:30 AM	Coffee Break		
-------------------	--------------	--	--

10:30 AM-12:00 PM	<b>Session 19:</b> Media Interpretation-I <b>Session Chair:</b> Xinchun Liu, JD AI Research	Langbo Room A	817 9260 5901
	<b>Session 20:</b> Media Interpretation-II <b>Session Chair:</b> Xiaoshuang Shi, University of Electronic Science and Technology of China	Langbo Room B	864 1291 4191
	<b>Session 21:</b> Open Source Competition <b>Session Chair:</b> Qi Dai, Microsoft Research Asia	Tianfu Room A	830 6390 3376
	<b>Session 22:</b> Multimedia Search and Recommendation-I <b>Session Chair:</b> Shuhui Wang, Institute of Computing Technology, CAS	Tianfu Room B	842 2630 3581
	<b>Session 23:</b> Multimedia Search and Recommendation-II <b>Session Chair:</b> Hao Liu, Ningxia University	Jiaozi Room	870 2448 7906
	<b>Session 24:</b> Music, Speech and Audio Processing in Multimedia <b>Session Chair:</b> Qin Jin, Renmin University of China	Jincheng Room	852 4844 0073
	<b>Session 25:</b> Multimedia Transport and Delivery <b>Session Chair:</b> Jie Shao, University of Electronic Science and Technology of China	Rongcheng Room	879 4652 5187

12:00 PM-2:00 PM	Lunch Break		
------------------	-------------	--	--

12:00 PM-1:30 PM	Rising Star Lunch Meeting (Supported by Kuaishou)	Jiaozi Room	
------------------	---	-------------	--



**SAT. | OCT. 23, 2021 | AFTERNOON**

Time	Arrangement	Room	Zoom Meeting ID
2:00 PM-4:40 PM	<b>Session 26:</b> Multimedia Grand Challenge <b>Session Chair:</b> Xirong Li, Renmin University of China	Langbo Room A	817 9260 5901
2:00 PM-3:30 PM	<b>Session 27:</b> Media Interpretation-III <b>Session Chair:</b> Yi Jin, Beijing Jiaotong University	Langbo Room B	864 1291 4191
	<b>Session 28:</b> Multimodal Fusion and Embedding-I <b>Session Chair:</b> Junying Chen, South China University of Technology	Tianfu Room A	830 6390 3376
	<b>Session 29:</b> Multimodal Fusion and Embedding-II <b>Session Chair:</b> Wen Li, University of Electronic Science and Technology of China	Tianfu Room B	842 2630 3581
	<b>Session 30:</b> Multimodal Fusion and Embedding-III <b>Session Chair:</b> Lin Zhang, Tongji University	Jiaozi Room	870 2448 7906
	<b>Session 31:</b> Multimedia Art, Entertainment and Culture <b>Session Chair:</b> Xin Jin, Beijing Electronic Science and Technology Institute	Rongcheng Room	879 4652 5187
3:30 PM-4:00 PM	Coffee Break		
4:00 PM-5:30 PM	<b>Session 32:</b> Social Multimedia <b>Session Chair:</b> Zhao Kang, University of Electronic Science and Technology of China	Langbo Room B	864 1291 4191
	<b>Session 33:</b> Summarization, Analytics, and Storytelling <b>Session Chair:</b> Lifang Wu, Beijing University of Technology	Tianfu Room A	830 6390 3376
	<b>Session 34:</b> Vision and Language-I <b>Session Chair:</b> Zhou Zhao, Zhejiang University	Tianfu Room B	842 2630 3581
	<b>Session 35:</b> Vision and Language-II <b>Session Chair:</b> Lianli Gao, University of Electronic Science and Technology of China	Jiaozi Room	870 2448 7906
	<b>Session 36:</b> Vision and Language-III <b>Session Chair:</b> Xinhang Song, Institute of Computing Technology, CAS	Jincheng Room	852 4844 0073
	<b>Session 37:</b> Multimedia Telepresence and Virtual/Augmented Reality <b>Session Chair:</b> Xuejin Chen, University of Science and Technology of China	Rongcheng Room	879 4652 5187



SUN. | OCT. 24, 2021

Time	Arrangement	Room	Zoom Meeting ID
9:00 AM-12:00 PM	<b>Workshop-ADGD 2021</b>	Tianfu Room A	817 9260 5901
	<b>Workshop-AI &amp; Food 2021</b>	Tianfu Room B	864 1291 4191
	<b>Tutorial: Deep Learning for Visual Data Compression</b>	Jiaozi Room	870 2448 7906
	<b>Tutorial: Few-shot Learning for Multi-Modality Tasks</b>	Jincheng Room	852 4844 0073
	<b>Tutorial: A Tutorial on AI Music Composition</b>	Rongcheng Room	879 4652 5187





12:00 PM-2:00 PM	Lunch Break @ Hall Way
------------------	------------------------

2:00 PM-5:00 PM	<b>Workshop-MuSe 2021</b>	Tianfu Room A	817 9260 5901
	<b>Workshop-MULL 2021</b>	Tianfu Room B	864 1291 4191
	<b>Workshop-FME 2021</b>	Jiaozi Room	870 2448 7906
	<b>Tutorial: Image Quality Assessment in the Modern Age</b>	Jincheng Room	852 4844 0073
	<b>Tutorial: Plenoptic Quality Assessment: The JPEG Pleno Experience</b>	Rongcheng Room	879 4652 5187

## TECHNICAL SESSIONS

October 20, 2021



OCT.  
20

9:00 AM-12:00 PM	Room	Zoom Meeting ID	Website QR
<b>Workshop-MMSports'21</b>	Langbo Room A	817 9260 5901	
<ul style="list-style-type: none"> <li>✦ <i>A Unified Taxonomy and Multimodal Dataset for Events in Invasion Games</i></li> <li>✦ <i>Multi-Task Learning for Jersey Number Recognition in Ice Hockey</i></li> <li>✦ <i>Automated Offside Detection by Spatio-Temporal Analysis of Football Videos</i></li> <li>✦ <i>Star: Noisy Semi-Supervised Transfer Learning for Visual Classification</i></li> <li>✦ <i>Three-stream 3d/1d CNN for Fine-Grained Action Classification and Segmentation in Table Tennis</i></li> <li>✦ <i>Speed21: Speed Climbing Motion Dataset</i></li> </ul>			
<b>Workshop -Trustworthy AI 2021</b>	Langbo Room B	864 1291 4191	
<ul style="list-style-type: none"> <li>✦ <i>An Empirical Study of Uncertainty Gap for Disentangling Factors</i></li> <li>✦ <i>Patch Replacement: A Transformation-Based Method to Improve Robustness Against Adversarial Attack</i></li> <li>✦ <i>Dataset Diversity: Measuring and Mitigating Geographical Bias in Image Search and Retrieval</i></li> <li>✦ <i>Hierarchical Semantic Enhanced Directional Graph Network for Visual Commonsense Reasoning</i></li> </ul>			
<b>Workshop -MuCAI 2021</b>	Tianfu Room A	830 6390 3376	
<ul style="list-style-type: none"> <li>✦ <i>Enriching Responses with Crowd-sourced Knowledge for Task-oriented Dialogues</i></li> <li>✦ <i>The Design of A Trust-Based Game As A Conversational Component of Interactive Environment for A Human-Agent Negotiation</i></li> <li>✦ <i>Towards A Real-Time Measure of The Perception of Anthropomorphism in Human-Robot Interaction</i></li> </ul>			
<b>Workshop-ADVM 2021</b>	Tianfu Room B	842 2630 3581	

- ✧ *Comparative Study of Adversarial Training Methods for Long-tailed Classification*
- ✧ *Imperceptible Adversarial Examples by Spatial Chroma-Shift*
- ✧ *Generating Adversarial Remote Sensing Images via Pan-Sharpening Technique*
- ✧ *Improving Generalization of Deepfake Detection with Domain Adaptive Batch Normalization*
- ✧ *Comparative Study of Adversarial Training Methods for Cold-Start Recommendation*
- ✧ *Detecting Adversarial Patch Attacks through Global-local Consistency*
- ✧ *Real World Robustness from Systematic Noise*
- ✧ *Enhancing Adversarial Examples Transferability via Ensemble Feature Manifolds*
- ✧ *An Investigation on Sparsity of CapsNets for Adversarial Robustness*
- ✧ *Frequency Centric Defense Mechanisms against Adversarial Examples*

<b>Tutorial-Multimedia classifiers: Behind the Scenes</b>	Jincheng Room	852 4844 0073	
---	---------------	---------------	--

OCT.  
20

2:00 PM-5:00 PM	Room	Zoom Meeting ID	Website QR
<b>Workshop-SUMAC 2021</b>	Langbo Room A	817 9260 5901	
<ul style="list-style-type: none"> <li>✧ <i>Deep Learning for Historical Data Analysis</i></li> <li>✧ <i>Analyzing CHANGE in cultural heritage objects through images</i></li> <li>✧ <i>Built Year Prediction from Buddha Face with Heterogeneous Labels</i></li> <li>✧ <i>Software and Content Design of a Browser-based Mobile 4D VR Application to Explore Historical City Architecture</i></li> <li>✧ <i>Evaluation of Deep Learning Techniques for Content Extraction in Spanish Colonial Notary Records</i></li> <li>✧ <i>How to Spatialize Geographical Iconographic Heritage</i></li> <li>✧ <i>Searching Silk Fabrics by Images Leveraging on Knowledge Graph and Domain Expert Rules</i></li> </ul>			
<b>Workshop-HUMA 2021</b>	Langbo Room B	864 1291 4191	
<ul style="list-style-type: none"> <li>✧ <i>A Closer Look at Temporal Sentence Grounding in Videos: Dataset and Metric</i></li> <li>✧ <i>Modeling 3D Objects: Implications for Neuroscience, Behavioral and Medical Studies: A Case Demo</i></li> <li>✧ <i>Using Feature Interaction among GPS Data for Road Intersection Detection</i></li> <li>✧ <i>NLOS Imaging Assisted Navigation for BVI</i></li> <li>✧ <i>Fine-grained Vehicle Recognition by Convolutional Neural Network</i></li> <li>✧ <i>Learning Positional Priors for Pretraining 2D Pose Estimators</i></li> </ul>			



<b>Workshop-WAB 2021</b>	Tianfu Room A	830 6390 3376	
<ul style="list-style-type: none"> <li>✧ <i>Domain Balanced Sampling and Iterative Search for Product Identification</i></li> <li>✧ <i>Multimodal Product Identification: Submission to Watch and Buy 2021 Challenge</i></li> <li>✧ <i>A Two Stage Model Using only Pixel-level Features for Video-Image Retrieval</i></li> <li>✧ <i>Multimodal Region-level Clothing Re-identification in E-commerce Livestreaming</i></li> <li>✧ <i>Watch and buy: A practical solution for real-time fashion product identification in live stream</i></li> </ul>			
<b>Workshop-UrbanMM 2021</b>	Rongcheng Room	879 4652 5187	
<ul style="list-style-type: none"> <li>✧ <i>Zurich Like New: Analyzing Open Urban Multimodal Data</i></li> <li>✧ <i>MBDF-Net: Multi branch deep fusion network for 3D Object Detection</i></li> <li>✧ <i>UrbanAccess: Query Driven Urban Analytics Platform for Detecting Complex Accessibility Event Patterns using Tactile Surfaces</i></li> <li>✧ <i>Urban Footpath Image Dataset to Assess Pedestrian Mobility</i></li> </ul>			
<b>Tutorial: Trustworthy Multimedia Analysis</b>	Tianfu Room B	842 2630 3581	
<b>Tutorial: Out-of-distribution Generalization and Its Applications for Multimedia</b>	Jincheng Room	852 4844 0073	

OCT.  
20

October 21, 2021

10:30 AM-12:00 PM	Session 1: Deep Learning for Multimedia-I	▼ Langbo Room A
<b>Session Chair: Cong Bai, Zhejiang University of Technology</b>		
10:30 AM-10:45 AM	<b>Image Re-composition via Regional Content-Style Decoupling</b> Author(s): Rong Zhang (Zhejiang University); Wei Li (Peking University); Yiqun Zhang (Zhejiang University); Hong Zhang (SenseTime Group Ltd); Jinhui Yu (Zhejiang University); Ruigang <b>Yang</b> (University of Kentucky, USA); Weiwei Xu (Zhejiang University) *	
10:45 AM-11:00 AM	<b>Deep Clustering based on Bi-Space Association Learning</b> Author(s): Hao Huang (GE Research) *; Shinjae Yoo (Brookhaven National Lab); Chenxiao Xu (Stony Brook University)	
11:00 AM-11:15 AM	<b>Feature Stylization and Domain-aware Contrastive Learning for Domain Generalization</b> Author(s): Seogkyu Jeon (Yonsei University) *; Kibeom Hong (Yonsei University); Pilhyeon Lee (Yonsei University); Jewook Lee (Yonsei University); Hyeran Byun (Yonsei University)	
11:15 AM-11:30 AM	<b>HDA-Net: Horizontal Deformable Attention Network for Stereo Matching</b> Author(s): Qi Zhang (Beijing University of Posts and Telecommunications) *; Xuesong Zhang (Beijing University of Posts and Telecommunications); Baoping Li (Beijing University of Posts and Telecommunications); Yuzhong Chen (Beijing University of Posts and Telecommunications); Anlong Ming (Beijing University of Posts and Telecommunications)	
11:30 AM-11:45 AM	<b>MBRS: Enhancing Robustness of DNN-based Watermarking by Mini-Batch of Real and Simulated JPEG Compression</b> Author(s): Zhaoyang Jia (University of Science and Technology of China) *; Han Fang (University of Science and Technology of China); Weiming Zhang (University of Science and Technology of China)	
11:45 AM-12:00 PM	<b>From Synthetic to Real: Image Dehazing Collaborating with Unlabeled Real Data</b> Author(s): Ye Liu (College of Intelligence and Computing, Tianjin University); Lei Zhu (The Chinese University of Hong Kong); Shunda Pei (College of Intelligence and Computing, Tianjin University); Huazhu Fu (IHPC, ASTAR); Jing Qin (The Hong Kong Polytechnic University); Qing Zhang (Sun Yat-sen University); Liang Wan (College of Intelligence and Computing, Tianjin University)*; Wei Feng (College of Intelligence and Computing, Tianjin University, China)	

OCT.  
21

10:30 AM-12:00 PM

Session 2: Deep Learning for Multimedia-II

▼ Langbo Room B

Session Chair: Jianxin Sun, Institute of Automation, Chinese Academy of Sciences

10:30 AM-10:45 AM

**Video Semantic Segmentation with Sparse Temporal Transformer**

Author(s): Jiangtong Li (Shanghai Jiao Tong University) \*; Wentao Wang (Shanghai Jiao Tong University); Junjie Chen (Shanghai Jiao Tong University); Li Niu (Shanghai Jiao Tong University); Jianlou Si (SenseTime); Chen Qian (SenseTime); Liqing Zhang (Shanghai Jiao Tong University)

10:45 AM-11:00 AM

**Diverse Image Inpainting with Bidirectional and Autoregressive Transformers**

Author(s): Yingchen Yu (Nanyang Technological University); Fangneng Zhan (Nanyang Technological University); Rongliang WU (Nanyang Technological University); Jianxiong Pan (Alibaba Group); Kaiwen Cui (Nanyang Technological University); Shijian Lu (Nanyang Technological University) \*; Feiyang Ma (alibaba); Xuansong Xie (Alibaba); Chunyan Miao (NTU)

11:00 AM-11:15 AM

**SSFlow: Style-guided neural Spline Flows for Face Image Manipulation**

Author(s): Hanbang Liang (Shenzhen University); Xianxu Hou (Shenzhen University); Linlin Shen (Shenzhen University) \*

11:15 AM-11:30 AM

**Constrained Graphic Layout Generation via Latent Optimization**

Author(s): Kotaro Kikuchi (Waseda University) \*; Edgar Simo-Serra (Waseda University); Mayu Otani (CyberAgent, Inc.); Kota Yamaguchi (Cyberagent)

11:30 AM-11:45 AM

**Transfer Vision Patterns for Multi-Task Pixel Learning**

Author(s): Xiaoya Zhang (Nanjing University of Science and Technology); Ling Zhou (Nanjing University of Science and Technology); Yong Li (Nanjing University of Science and Technology); Zhen Cui (Nanjing University of Science and Technology) \*; Jin Xie (Nanjing University of Science and Technology); Jian Yang (Nanjing University of Science and Technology)

11:45 AM-12:00 PM

**Object-aware Long-short-range Spatial Alignment for Few-Shot Fine-Grained Image Classification**

Author(s): Yike Wu (Fudan University); Bo Zhang (Fudan University); Gang Yu (Tencent); Weixi Zhang (Fudan University); Bin Wang (Fudan University); Tao Chen (Fudan University) \*; Jiayuan Fan (Fudan University)

OCT.  
21

10:30 AM-12:00 PM	Session 3: Deep Learning for Multimedia-III	▼ Tianfu Room A
<b>Session Chair: Jin Wang, Beijing University of Technology</b>		
10:30 AM-10:45 AM	<b>Enhanced Invertible Encoding for Learned Image Compression</b> Author(s): Yueqi Xie (The Hong Kong University of Science and Technology); Ka Leong Cheng (The Hong Kong University of Science and Technology); Qifeng Chen (HKUST)*	
10:45 AM-11:00 AM	<b>DC-GNet: Deep Mesh Relation Capturing Graph Convolution Network for 3D Human Shape Reconstruction</b> Author(s): Shihao Zhou (Xiamen University) *; Mengxi Jiang (Xiamen university); shanshan cai (Xiamen university); Yunqi Lei (Xiamen university)	
11:00 AM-11:15 AM	<b>Deep Marginal Fisher Analysis based CNN for Image Representation and Classification</b> Author(s): Xun Cai (Shandong university); Jiajing Chai (Shandong University); Yanbo Gao (Shandong University) *; Shuai Li (Shandong University); Bo Zhu (Shandong University)	
11:15 AM-11:30 AM	<b>Learning Structure Affinity for Video Depth Estimation</b> Author(s): Yuanzhouhan Cao (Beijing Jiaotong University); Yidong Li (Beijing Jiaotong Univeristy, China)*; Haokui Zhang (Intellifusion); Chao Ren (Sichuan University); Yifan Liu (University of Adelaide)	
11:30 AM-11:45 AM	<b>X-GGM: Graph Generative Modeling for Out-of-distribution Generalization in Visual Question Answering</b> Author(s): Jingjing Jiang (Xi'an Jiaotong University) *; Ziyi Liu (Xi'an Jiaotong University); Yifan Liu (Xi'an Jiaotong University); Zhixiong Nan (Xi'an Jiaotong University); Nanning Zheng (Xi'an Jiaotong University)	
11:45 AM-12:00 PM	<b>DSSL: Deep Surroundings-person Separation Learning for Text-based Person Retrieval</b> Author(s): Aichun Zhu (Nanjing Tech University) *; Zijie Wang (Nanjing Tech University); Yifeng Li (Nanjing Tech University); Xili Wan (Nanjing Tech University); Jing Jin (Nanjing Tech University); Tian Wang (Beihang University); Fangqiang Hu (Nanjing Tech University); Gang Hua (China University of Mining and Technology)	



10:30 AM-12:00 PM

Session 4: Brave New Idea

▼ Tianfu Room B

Session Chair: Qin Jin, Renmin University of China

**10:30 AM-10:48 AM Recycling Discriminator: Towards Opinion-Unaware Image Quality Assessment Using Wasserstein GAN**

Author(s): Zhu Yunan; Ma Haichuan; Peng Jialun; Liu Dong\*; Xiong Zhiwei

**10:48 AM-11:06 AM Learning Kinematic Formulas from Multiple View Videos**

Author(s): Song, Liangchen\*; Liu, Sheng; Liu, Celong; Li, Zhong; Ding, Yuqi; Xu, Yi; Yuan, Junsong

**11:06 AM-11:24 AM Depa: Self-Supervised Audio Embedding for Depression Detection**

Author(s): Zhang, Pingyue; Wu, Mengyue\*; Dinkel, Heinrich; Yu, Kai

**11:24 AM-11:42 AM Retinomorphic Sensing: A Novel Paradigm for Future Multimedia Computing**

Author(s): Kang, Zhaodong\*; Li, Jianing; Zhu, Lin; Tian, Yonghong

**11:42 AM-12:00 PM Metaverse for Social Good: A University Campus Prototype**

Author(s): Duan, Haihan; Li, Jiaye; Fan, Sizheng; Lin, Zhonghao; Wu, Xiao; Cai, Wei\*

OCT.  
21

10:30 AM-12:00 PM	Session 5: Emerging Multimedia Applications-I	▼ Jiaozi Room
<b>Session Chair: Weiqing Min, Chinese Academy of Sciences</b>		
10:30 AM-10:45 AM	<b>Diverse Multimedia Layout Generation with Multi Choice Learning</b> Author(s): David D Nguyen (UNSW)*; Surya Nepal (CSIRO); salil kanhere (UNSW Sydney)	
10:45 AM-11:00 AM	<b>Viewing From Frequency Domain: A DCT-based Information Enhancement Network for Video Person Re-Identification</b> Author(s): Liangchen Liu (Xidian University); Xi Yang (Xidian University)*; Nannan Wang (Xidian University); Xinbo Gao (Chongqing University of Posts and Telecommunications)	
11:00 AM-11:15 AM	<b>Unsupervised Portrait Shadow Removal via Generative Prior</b> Author(s): Yingqing HE (HKUST); Yazhou Xing (HKUST); Tianjia Zhang (HKUST); Qifeng Chen (HKUST)*	
11:15 AM-11:30 AM	<b>Multimodal Global Relation Knowledge Distillation for Egocentric Action Anticipation</b> Author(s): Yi Huang (Institute of Automation, Chinese Academy of Sciences); Xiaoshan Yang (CASIA); Changsheng Xu (CASIA)*	
11:30 AM-11:45 AM	<b>Exploring Pathologist Knowledge for Automatic Assessment of Breast Cancer Metastases in Whole-slide Image</b> Author(s): Liuan Wang (FUJITSU R&D CENTER CO.,LTD.)*; Li Sun (FUJITSU R&D CENTER CO.,LTD.); Mingjie Zhang (FUJITSU R&D CENTER CO.,LTD.); Huigang Zhang (FUJITSU R&D CENTER CO.,LTD.); Wang Ping (Fujitsu R&D Center); Rong Zhou (Fujitsu research and development center); Jun Sun (Fujitsu R&D Center CO., Ltd)	
11:45 AM-12:00 PM	<b>Towards Multiple Black-boxes Attack via Adversarial Example Generation Network</b> Author(s): Duan Mingxing (Hong Kong Polytechnic University); Kenli Li (Hunan University); Lingxi Xie (Huawei Inc.); Qi Tian (Huawei Cloud & AI) *; Bin Xiao (The Hong Kong Polytechnic University)	



10:30 AM-12:00 PM

Session 6: Emerging Multimedia Applications-II

▼ Jincheng Room

Session Chair: Le Zhang, University of Electronic Science and Technology of China

10:30 AM-10:45 AM

**DocTr: Document Image Transformer for Geometric Unwarping and Illumination Correction**

Author(s): Hao Feng (University of Science and Technology of China)\*; Yuechen Wang (University of Science and Technology of China); Wengang Zhou (University of Science and Technology of China); Jiajun Deng (University of Science and Technology of China); Houqiang Li (University of Science and Technology of China)

10:45 AM-11:00 AM

**Self-supervised Multi-view Multi-Human Association and Tracking**

Author(s): Yiyang Gan (College of Intelligence and Computing, Tianjin University, China); Ruize Han (College of Intelligence and Computing, Tianjin University)\*; Liqiang Yin (Tianjin University); Wei Feng (College of Intelligence and Computing, Tianjin University, China); Song Wang (University of South Carolina)

11:00 AM-11:15 AM

**Learning Fine-Grained Motion Embedding for Landscape Animation**

Author(s): Hongwei Xue (University of Science and Technology of China); Bei Liu (Microsoft Research)\*; Huan Yang (Microsoft Research); Jianlong Fu (Microsoft Research); Houqiang Li (University of Science and Technology of China); Jiebo Luo (U. Rochester)

11:15 AM-11:30 AM

**Multi-label Pattern Image Retrieval via Attention Mechanism Driven Graph Convolutional Network**

Author(s): Ying Li (Nanjing Normal University)\*; Hongwei Zhou (Nanjing Normal University); Yeyu Yin (Nanjing Normal University); Jiaquan Gao (Nanjing Normal University)

11:30 AM-11:45 AM

**Collocation and Try-on Network: Whether an outfit is Compatible**

Author(s): Na Zheng (Shandong University)\*; Xueming Song (Shandong University); Qingying Niu (Shandong University); Xue Dong (Shandong University); Yibing Zhan (Jingdong ); Liqiang Nie (Shandong University )

11:45 AM-12:00 PM

**MeronymNet: A Hierarchical Model for Unified and Controllable Multi-Category Object Generation**

Author(s): Rishabh Baghel (IIIT Hyderabad); Abhishek Trivedi (IIIT Hyderabad); Tejas Ravichandran (IIIT Hyderabad); Ravi Kiran Sarvadevabhatla (IIIT Hyderabad)\*

OCT.  
21

10:30 AM-12:00 PM	Session 7: Multimedia HCI and Quality of Experience	▼ Rongcheng Room
<b>Session Chair: Qingbo Wu, University of Electronic Science and Technology of China</b>		
10:30 AM-10:45 AM	<b>Fast video visual quality and resolution improvement using SR-UNet</b> Author(s): Federico Vaccaro (University of Florence); Marco Bertini (University of Florence)*; Tiberio Uricchio (University of Florence); Alberto Del Bimbo (University of Florence)	
10:45 AM-11:00 AM	<b>MS-GraphSIM: Inferring Point Cloud Quality via Multiscale Graph Similarity</b> Author(s): Yujie Zhang (Shanghai Jiao Tong University)*; Qi Yang (Shanghai Jiao Tong University); Yiling Xu (Shanghai Jiao Tong University)	
11:00 AM-11:15 AM	<b>I Know Your Keyboard Input: A Robust Keystroke Eavesdropper Based-on Acoustic Signals</b> Author(s): Jia-Xuan Bai (University of Science and Technology of China)*; Bin Liu (University of Science and Technology of China); Luchuan Song (University of Science and Technology of China)	
11:15 AM-11:30 AM	<b>Perceptual Quality Assessment of Internet Videos</b> Author(s): Jiahua Xu (University of Science and Technology of China); Jing Li (Alibaba Group)*; 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)	
11:30 AM-11:45 AM	<b>Using Interaction Data to Predict Engagement with Interactive Media</b> Author(s): Jonathan Carlton (University of Manchester)*; Andy Brown (BBC); Dr.Caroline Jay (The University Of Manchester); John Keane (University of Manchester)	
11:45 AM-12:00 PM	<b>Air-Text: Air-Writing and Recognition System</b> Author(s): Sun-Kyung Lee (KAIST); Jong-Hwan Kim (KAIST)*	



2:00 PM-4:00 PM

Session 8: Industrial Track

▼ Tianfu Room A

Session Chair: Rongrong Ji, Xiamen University

Jianlong Fu, Microsoft Research Asia

2:00 PM-2:30 PM

**Invited Talk 1**  
Yandong Guo, OPPO

2:30 PM-3:00 PM

**Invited Talk 2**  
Xiansheng Hua, Alibaba Group

3:00 PM-3:15 PM

**Once and for All: Self-supervised Multi-modal Co-training on One-billion Videos at Alibaba**

Author(s): Lianghua Huang (Alibaba Group)\*; Yu Liu (Alibaba Group); Xiangzeng Zhou (Alibaba Group); Ansheng You (Alibaba Group); Ming Li (Alibaba); bin wang (alibaba group); Yingya Zhang (Alibaba Group); Pan Pan (Alibaba Group); Xu Yinghui (Alibaba Group)

3:15 PM-3:30 PM

**L2RS: A Learning-to-Rescore Mechanism for Hybrid Speech Recognition**

Author(s): Yuanfeng Song (The Hong Kong University of Science and Technology)\*; Di Jiang (WeBank); Xuefang Zhao (WeBank Co., Ltd); Qian Xu (WeBank); Raymond Chi-Wing Wong (Hong Kong University of Science and Technology); Lixin Fan (WeBank); Qiang Yang (Webank)

3:30 PM-3:45 PM

**Boosting End-to-end Multi-Object Tracking and Person Search via Knowledge Distillation**

Author(s): Wei Zhang (CRIPAC)\*; Lingxiao He (AI Research of JD.com); Peng Chen (AI Research of JD.com); Xingyu Liao (AI Research of JD.com); Wu Liu (AI Research of JD.com); Qi Li (CASIA); Zhenan Sun (Chinese of Academy of Sciences)

3:45 PM-4:00 PM

**RecycleNet: An Overlapped Text Instance Recovery Approach**

Author(s): Yiqing Hu (Tencent)\*; Yan Zheng (Tencent); Xinghua Jiang (Tencent); Hao Liu (Tencent); Deqiang Jiang (Tencent); Yinsong Liu (Tencent); Bo Ren (Tencent); Rongrong Ji (Xiamen University, China)

OCT.  
21

3:30 PM-5:00 PM	Session 9: Emerging Multimedia Applications-III	▼ Langbo Room A
<b>Session Chairs: Shuai Li, Shandong University</b>		
3:30 PM-3:45 PM	<b>Ada-VSR: Adaptive Video Super-Resolution with Meta-Learning</b> Author(s): Akash Gupta (University of California, Riverside)*; Padmaja Jonnalagedda (UC Riverside); Bir Bhanu (University of California Riverside); Amit K. Roy-Chowdhury (University of California, Riverside)	
3:45 PM-4:00 PM	<b>CoReD: Generalizing Fake Media Detection with Continual Representation using Distillation</b> Author(s): Minha Kim (Sungkyunkwan University); Shahroz Tariq (Sungkyunkwan University)*; Simon Woo (Sungkyunkwan University (SKKU))	
4:00 PM-4:15 PM	<b>SRNet: Spatial Relation Network for Efficient Single-stage Instance Segmentation in Videos</b> Author(s): Xiaowen Ying (Lehigh University)*; Xin Li (Lehigh University); Mooi Choo Chuah (Lehigh University)	
4:15 PM-4:30 PM	<b>Personality Recognition by Modelling Person-specific Cognitive Processes using Graph Representation</b> Author(s): zilong shao (Shenzhen University); Siyang Song (University of Cambridge)*; Shashank Jaiswal (University of Nottingham); Linlin Shen (Shenzhen University); Michel Valstar (University of Nottingham); Hatice Gunes (University of Cambridge)	
4:30 PM-4:45 PM	<b>Enhancing Knowledge tracing via adversarial training</b> Author(s): Xiaopeng Guo (Peking University)*; Zhijie Huang (Peking University); Jie Gao (Peking University); Mingyu Shang (Peking University); Maojing shu (Peking University); Jun Sun (Peking University)	
4:45 PM-5:00 PM	<b>Beyond OCR + VQA: Involving OCR into the Flow for Robust and Accurate TextVQA</b> Author(s): Gangyan Zeng (School of Information and Communication Engineering, Communication University of China); Yuan Zhang (Communication University of China); Yu Zhou (Institute of Information Engineering, CAS; Also with University of Chinese Academy of Sciences)*; Xiaomeng Yang (Institute of Information Engineering, CAS)	



3:30 PM-5:00 PM	Session 10: Emerging Multimedia Applications-IV	▼ Langbo Room B
Session Chair: Jing Dong, Institute of Automation		
3:30 PM-3:45 PM	<b>VoteHMR: Occlusion-Aware Voting Network for Robust 3D Human Mesh Recovery from Partial Point Clouds</b> Author(s): Guanze Liu (Beihang University); Yu Rong (The Chinese University of Hong Kong); Lu Sheng (Beihang University)*	
3:45 PM-4:00 PM	<b>MageAdd: Real-Time Interaction Simulation for Scene Synthesis</b> Author(s): Shao-Kui Zhang (Tsinghua University); Yi-Xiao Li (Tsinghua University); Yu He (Tsinghua University); Yongliang Yang (University of Bath); Song-Hai Zhang (Tsinghua University)*	
4:00 PM-4:15 PM	<b>Cross-View Exocentric to Egocentric Video Synthesis</b> Author(s): Gaowen Liu (Cisco Research); Hao Tang (ETH Zurich)*; Hugo M Latapie (Cisco); Jason J Corso (University of Michigan); Yan Yan (Texas State University)	
4:15 PM-4:30 PM	<b>EVRNet: Efficient Video Restoration on Edge Devices</b> Author(s): Sachin Mehta (University of Washington)*; Amit Kumar (Facebook ); Fitsum Reda (Google); Varun Nasery (Facebook); Vikram Mulukutla (Facebook); Rakesh Ranjan (Facebook); Vikas Chandra (Facebook)	
4:30 PM-4:45 PM	<b>Multimodal Entity Linking: A New Dataset and A Baseline</b> Author(s): Jingru Gan ( Key Lab of Intell. Info. Process., Inst. of Comput. Tech., CAS)*; Jinchang Luo (Baidu Inc., China); Haiwei Wang (Baidu Inc., China); Shuhui Wang (VIPL,ICT,Chinese academic of science); Wei He (Baidu Inc., China); Qingming Huang (University of Chinese Academy of Sciences)	
4:45 PM-5:00 PM	<b>AI-Lyricist: Generating Music and Vocabulary Constrained Lyrics</b> Author(s): Xichu Ma (National University of Singapore); Ye Wang (National University of Singapore)*; Min-Yen Kan (National University of Singapore); Wee Sun Lee (National University of Singapore)	

OCT.  
21

3:30 PM-5:00 PM	Session 11: Emotional and Social Signals in Multimedia	▼ Tianfu Room B
Session Chair: Jia Jia, Tsinghua University		
3:30 PM-3:45 PM	<b>CaFGraph: Context-aware Facial Multi-graph Representation for Facial Action Unit Recognition</b>	
	Author(s): Yingjie Chen (Peking University); Diqi Chen (AIIT); Yizhou Wang (PKU); Tao Wang (Peking University)*; Yun Liang (Peking University)	
3:45 PM-4:00 PM	<b>Self-Supervised Regional and Temporal Auxiliary Tasks for Facial Action Unit Recognition</b>	
	Author(s): Jingwei Yan (Hikvision Research Institute)*; Jingjing Wang (Hikvision Research Institute); Qiang Li (Hikvision Research Institute); Chunmao Wang (Hikvision Research Institute); Shiliang Pu (Hikvision Research Institute)	
4:00 PM-4:15 PM	<b>HetEmotionNet: Two-Stream Heterogeneous Graph Recurrent Neural Network for Multi-modal Emotion Recognition</b>	
	Author(s): Ziyu Jia (Beijing Jiaotong University); Youfang Lin (Beijing Jiaotong University); Jing Wang (Beijing Jiaotong University)*; Zhiyang Feng (Beijing Jiaotong University); Xiangheng Xie (Beijing Jiaotong University); Caijie Chen (Beijing Jiaotong University)	
4:15 PM-4:30 PM	<b>Simplifying Multimodal Emotion Recognition with Single Eye Movement Modality</b>	
	Author(s): Xu Yan (University of Washington)*; Liming Zhao (Shanghai Jiao Tong University); Bao-Liang Lu (Shanghai Jiao Tong University)	
4:30 PM-4:45 PM	<b>Learning What and When to Drop: Adaptive Multimodal and Contextual Dynamics for Emotion Recognition in Conversation</b>	
	Author(s): Feiyu Chen (University of Electronic Science and Technology of China); 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)*	
4:45 PM-5:00 PM	<b>Zero-shot Video Emotion Recognition via Multimodal Protagonist-aware Transformer Network</b>	
	Author(s): Fan Qi (HeFei University of Technology, CASIA)*; Xiaoshan Yang (CASIA); Changsheng Xu (CASIA)	



3:30 PM-5:00 PM

Session 12: Multimodal Analysis and Description-I

▼ Jiaozi Room

Session Chair: Haonan Cheng, Communication University of China

3:30 PM-3:45 PM

**How to Learn a Domain-Adaptive Event Simulator?**

Author(s): Daxin Gu (Beihang University); Jia Li (Beihang University)\*; Yu Zhang (Beihang University); Yonghong Tian (Peking University)

3:45 PM-4:00 PM

**A Stepwise Matching Method for Multimodal Image based on Cascaded Network**

Author(s): Jinming Mu (Xidian University); Shuiping Gou (Xidian University); Shasha Mao (Xidian University)\*; Shankui Zheng (Beijing Interstellar Glory Space Technology Co., Ltd.)

4:00 PM-4:15 PM

**DL-Easy: An Easy-to-Use Framework for MultiModal Analysis**

Author(s): Naili Xing (National University of Singapore); Chris Yeung (National University of Singapore); Cheng-Hao Cai (National University of Singapore); Teck Khim Ng (National University of Singapore); Wei Wang (ByteDance); Kaiyuan Yang (National University of Singapore); Nan Yang (National University of Singapore); Meihui Zhang (Beijing Institute of Technology); Gang Chen (Zhejiang University); Beng Chin Ooi (NUS)\*

4:15 PM-4:30 PM

**Informative Class-Conditioned Feature Alignment for Unsupervised Domain Adaptation**

Author(s): Wanxia Deng (National University of Defense Technology); Yawen Cui (University of Oulu); Zhen Liu (National University of Defense Technology); Gangyao Kuang (National University of Defense Technology); Dewen Hu (National University of Defense Technology); Matti Pietikäinen (University of Oulu); Li Liu (National University of Defense Technology)\*

4:30 PM-4:45 PM

**Hierarchical Multi-Task Learning for Diagram Question Answering with Multi-Modal Transformer**

Author(s): Zhaoquan Yuan (School of Computing and Artificial Intelligence, Southwest Jiaotong University); Xiao Peng (Southwest Jiaotong University); Xiao Wu (Southwest Jiaotong University)\*; Changsheng Xu (CASIA)

4:45 PM-5:00 PM

**Differentiated Learning for Multi-Modal Domain Adaptation**

Author(s): Jianming Lv (South China University of Technology)\*; Kaijie Liu (South China University of Technology); Shengfeng He (South China University of Technology)

OCT.  
21

3:30 PM-5:00 PM	Session 13: Multimodal Analysis and Description-II	▼ Jincheng Room
Session Chair: Zhigang Tu, Wuhan University		
3:30 PM-3:45 PM	<b>Two-stage Visual Cues Enhancement Network for Referring Image Segmentation</b> Author(s): Yang Jiao (Fudan University); Zequn Jie (Meituan inc.); Weixin Luo (Meituan); Jingjing Chen (Fudan University)*; Yu-Gang Jiang (Fudan University); Xiaolin Wei (Meituan); Lin Ma (Meituan)	
3:45 PM-4:00 PM	<b>Partial Tubal Nuclear Norm Regularized Multi-view Learning</b> Author(s): 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)	
4:00 PM-4:15 PM	<b>Deep Unsupervised 3D SfM Face Reconstruction Based on Massive Landmark Bundle Adjustment</b> Author(s): Yuxing Wang (Rochester Institute of Technology); Yawen Lu (Rochester Institute of Technology); Zhihua Xie (Jiangxi Sciences and Technology Normal University); Guoyu Lu (Rochester Institute of Technology)*	
4:15 PM-4:30 PM	<b>SimuLR: Simultaneous Lip Reading Transducer with Attention-Guided Adaptive Memory</b> Author(s): Zhijie Lin (Zhejiang University)*; Zhou Zhao (Zhejiang University); Haoyuan Li (Zhejiang University); Jinglin Liu (Zhejiang University); Meng Zhang (Huawei Noah's Ark Lab); Xingshan Zeng (Huawei Noah's Ark Lab); Xiaofei He (Zhejiang University)	
4:30 PM-4:45 PM	<b>Dense Semantic Contrast for Self-Supervised Visual Representation Learning</b> Author(s): Xiaoni Li (Institute of Information Engineering, Chinese Academy of Sciences; School of Cyber Security, University of Chinese Academy of Sciences); Yu Zhou (Institute of Information Engineering, CAS; Also with University of Chinese Academy of Sciences)*; Yifei Zhang (Institute of Information Engineering, Chinese Academy of Sciences; School of Cyber Security, University of Chinese Academy of Sciences); aoting zhang (Institute of Information Engineering); Wei Wang (Institute of Information Engineering, Chinese Academy of Sciences; School of Cyber Security, University of Chinese Academy of Sciences); Ning Jiang (Mashang Consumer Finance Co., Ltd.); Haiying Wu (Mashang Consumer Finance Co., Ltd.); Weiping Wang (Institute of Information Engineering, CAS, China)	
4:45 PM-5:00 PM	<b>Multiple Object Tracking by Trajectory Map Regression with Temporal Priors Embedding</b> Author(s): Xingyu Wan (Xi'an Jiaotong University)*; Sanping Zhou (Xi'an Jiaotong University); Jinjun Wang (Xi'an Jiaotong University); Rongye Meng (IAIR, Xi'an Jiaotong University)	



3:30 PM-5:00 PM	<b>Session 14: Multimedia Cloud, Edge and Device Computing</b>	▼ Rongcheng Room
<b>Session Chair: Yandong Guo, OPPO Research Institute</b>		
3:30 PM-3:45 PM	<b>DeepGame: Efficient Video Encoding for Cloud Gaming</b> Author(s): Omar Mossad (Simon Fraser University)*; Khaled Diab (Simon Fraser University); Ihab Amer (AMD); Mohamed Hefeeda (Simon Fraser University)	
3:45 PM-4:00 PM	<b>ChartPointFlow for Topology-Aware 3D Point Cloud Generation</b> Author(s): Takumi Kimura (Kobe University); Takashi Matsubara (Osaka University)*; Kuniaki Uehara (Osaka Gakuin University)	
4:00 PM-4:15 PM	<b>Co-learning: Learning from Noisy Labels with Self-Supervision</b> Author(s): Cheng Tan (Westlake University)*; Jun Xia (Westlake University and Zhejiang University); Lirong Wu (Westlake University); Stan Z. Li (Westlake University)	
4:15 PM-4:30 PM	<b>Graph Convolutional Multi-modal Hashing for Flexible Multimedia Retrieval</b> Author(s): Xu Lu (Shandong Normal University); Lei Zhu (Shandong Normal University); Li Liu (Shandong Normal University); Liqiang Nie (Shandong University); Huaxiang Zhang (Shandong Normal University)*	
4:30 PM-4:45 PM	<b>Hybrid Network Compression via Meta-Learning</b> Author(s): Jianming Ye (Peking University); Shiliang Zhang (Peking University)*; Jingdong Wang (Microsoft)	
4:45 PM-5:00 PM	<b>Two-pronged Strategy: Lightweight Augmented Graph Network Hashing for Scalable Image Retrieval</b> Author(s): Hui Cui (Shandong Normal University); Lei Zhu (Shandong Normal University)*; Jingjing Li (University of Electronic Science and Technology of China); Zhiyong Cheng (Shandong Academy of Sciences); Zheng Zhang (Harbin Institute of Technology, Shenzhen)	

OCT.  
21

9:00 AM-12:00 PM	Poster Session 1	▼ Hall Way
P001	<b>Reconstruction: A Motion Driven Interactive Artwork Inspired by Chinese Shadow Puppet</b> Author(s): Jiang Wenli (Beihang University); Chong CAO (Beihang University)*	
P002	<b>JPGNet: Joint Predictive Filtering and Generative Network for Image Inpainting</b> Author(s): Guo, Qing*; Li, Xiaoguang; Juefei-Xu, Felix; Yu, Hongkai; Liu, Yang; Wang, Song	
P003	<b>AdvFilter: Predictive Perturbation-aware Filtering against Adversarial Attack via Multi-domain Learning</b> Author(s): Huang, Yihao; Guo, Qing*; Juefei-Xu, Felix; Ma, Lei; Miao, Weikai; Liu, Yang; Pu, Geguang	
P004	<b>Pixel-level Intra-domain Adaptation for Semantic Segmentation</b> Author(s): Yan, Zizheng*; Yu, Xianggang; Qin, Yipeng; Wu, Yushuang; Han, Xiaoguang; Cui, Shuguang	
P005	<b>Mask is All You Need: Rethinking Mask R-CNN for Dense and Arbitrary-Shaped Scene Text Detection</b> Author(s): Qin, Xugong; Zhou, Yu*; Guo, Youhui; Wu, Dayan; Tian, Zhihong; Jiang, Ning; Wang, Hongbin; Wang, Weiping	
P006	<b>Windowing Decomposition Convolutional Neural Network for Image Enhancement</b> Author(s): Zheng, Chuanjun*; Shi, Daming; Liu, Yukun	
P007	<b>Joint Optimization in Edge-Cloud Continuum for Federated Unsupervised Person Re-identification</b> Author(s): Zhuang, Weiming*; Wen, Yonggang; zhang, shuai	
P008	<b>Multi-view 3D Smooth Human Pose Estimation based on Heatmap Filtering and Spatio-temporal Information</b> Author(s): Niu, Zehai; Lu, Ke; Xue, Jian*; Ma, Haifeng; Wei, Runchen	
P009	<b>Imitative Learning for Multi-Person Action Forecasting</b> Author(s): Li, Yuke*; Wang, Pin; YE, Mang; Chan, Ching-Yao	
P010	<b>Stereo Video Super-Resolution via Exploiting View-Temporal Correlations</b> Author(s): Xu, Ruikang; Xiao, Zeyu; Yao, Mingde; Zhang, Yueyi*; Xiong, Zhiwei	
P011	<b>M3TR: Multi-modal Multi-label Recognition with Transformer</b> Author(s): Zhao, Jiawei; Zhao, Yifan; Li, Jia*	
P012	<b>TACR-Net: Editing on Deep Video and Voice Portraits</b> Author(s): Song, Luchuan*; Liu, Bin; Yin, Guojun; Dong, Xiaoyi; Zhang, Yufei; Bai, Jia-Xuan	



P013	<b>Annotation-Efficient Untrimmed Video Action Recognition</b> Author(s): Zou, Yixiong*; Zhang, Shanghang; Chen, Guangyao; Tian, Yonghong; Keutzer, Kurt; Moura, José M. F.
P014	<b>Face-based Voice Conversion: Learning the Voice behind a Face</b> Author(s): Lu, Hsiao-Han*; Weng, Shao-En ; Yen, Ya-Fan; Shuai, Hong-Han; Cheng, Wen-Huang
P015	<b>A Large-Scale Benchmark for Food Image Segmentation</b> Author(s): Wu Xiongwei*; Fu Xin; Liu Ying; Lim Ee-peng; Hoi Steven; Sun Qianru
P016	<b>HAT: Hierarchical Aggregation Transformers for Person Re-identification</b> Author(s): Zhang, Guowen; Zhang, Pingping; Qi, Jinqing; Lu, Huchuan*
P017	<b>Long-Range Feature Propagating for Natural Image Matting</b> Author(s): Liu, Qinglin*; Xie, Haozhe; Zhang, Shengping; Zhong, Bineng; Ji, Rongrong
P018	<b>Towards Controllable and Photorealistic Region-wise Image Manipulation</b> Author(s): You, Ansheng*; Zhou, Chenglin; Zhang, Qixuan; Xu, Lan
P019	<b>Information-Growth Attention Network for Image Super-Resolution</b> Author(s): Li Zhuangzi*; Li Ge; Li, Thomas H; Liu Shan; Gao Wei
P020	<b>Anchor-free 3D Single Stage Detector with Mask-Guided Attention for Point Cloud</b> Author(s): Li, Jiale; Dai, Hang*; Shao, Ling; Ding, Yong
P021	<b>Shape Controllable Virtual Try-on for Underwear Models</b> Author(s): Gao, Xin*; Liu, Zhenjiang; Feng, Zunlei; Shen, Chengji; Ou, Kairi; Tang, Haihong; Song, Mingli
P022	<b>E<sup>2</sup>Net: Excitative-Expansile Learning for Weakly Supervised Object Localization</b> Author(s): Chen, Zhiwei*; Cao, Liujuan; Shen, Yunhang; Lian, Feihong; Wu, Yongjian; Ji, Rongrong
P023	<b>Few-shot Fine-Grained Action Recognition via Bidirectional Attention and Contrastive Meta-Learning</b> Author(s): Wang, Jiahao*; Wang, Yunhong; liu, sheng; Li, Annan
P024	<b>Selective Dependency Aggregation for Action Classification</b> Author(s): Tan, Yi; Hao, Yanbin*; He, Xiangnan; Wei, Yinwei; Yang, Xun
P025	<b>Conditional Directed Graph Convolution for 3D Human Pose Estimation</b> Author(s): Hu, Wenbo*; Zhang, Changgong; Zhan, Fangneng; Zhang, Lei; Wong, Tien-Tsin
P026	<b>Cross Chest Graph for Disease Diagnosis with Structural Relational Reasoning</b> Author(s): Zhao, Gangming*

P027	<b>ZiGAN: Fine-grained Chinese Calligraphy Font Generation via a Few-shot Style Transfer Approach</b> Author(s): Wen, Qi; Li, Shuang; han, bingfeng; Yuan, Yi*
P028	<b>Cycle-Consistent Inverse GAN for Text-to-Image Synthesis</b> Author(s): Wang, Hao*; Lin, Guosheng; Hoi, Steven; Miao, Chunyan
P029	<b>Fully Quantized Image Super-Resolution Networks</b> Author(s): Wang, Hu*; Chen, Peng; Zhuang, Bohan; Shen, Chunhua
P030	<b>AKECP: Adaptive Knowledge Extraction from Feature Maps for Fast and Efficient Channel Pruning</b> Author(s): Zhang, Haonan; Liu, Longjun*; Zhou, Hengyi; Hou, Wenxuan; Sun, Hongbin; Zheng, Nanning
P031	<b>Dynamic Momentum Adaptation for Zero-Shot Cross-Domain Crowd Counting</b> Author(s): Wu, Qiangqiang*; Wan, Jia; Chan, Antoni
P032	<b>Auto-MSFNet: Search Multi-scale Fusion Network for Salient Object Detection</b> Author(s): Zhang, Miao; Liu, Tingwei; Piao, Yongri*; Yao, Shunyu; Lu, Huchuan
P033	<b>Few-shot Unsupervised Domain Adaptation with Image-to-Class Sparse Similarity Encoding</b> Author(s): Huang, Shengqi; Yang, Wanqi*; Wang, Lei; Zhou, Luping; Yang, Ming
P034	<b>Semantic-aware Transfer with Instance-adaptive Parsing for Crowded Scenes Pose Estimation</b> Author(s): wang, xuanhan*; Gao, Lianli; Dai, Yan; Zhou, Yixuan; Song, Jingkuan
P035	<b>Multimodal Dialog System: Relational Graph-based Context-aware Question Understanding</b> Author(s): Zhang, Haoyu; Liu, Meng; Gao, Zan; Lei, Xiaoqiang; Wang, Yinglong; Nie, Liqiang*
P036	<b>Shadow Detection via Predicting the Confidence Maps of Shadow Detection Methods</b> Author(s): Liao, JingWei*; Liu, Yanli; Xing, GuanYu; wei, housheng; Chen, JueYu; Xu, Songhua
P037	<b>TEID: A New State-of-the-art Approach for Future Motion Prediction</b> Author(s): Su, Pengxiang*; Liu, Zhenguang; Wu, Shuang; Zhu, Lei; Yin, Yifang; Shen, Xuanjing
P038	<b>Q-Art Code: Generating Scanning-robust Art-style QR Codes by Deformable Convolution</b> Author(s): Su, Hao*; Niu, Jianwei; Liu, Xuefeng; Li, Qingfeng; Wan, Ji; Xu, Mingliang



P039	<b>Depth Quality-Inspired Feature Manipulation for Efficient RGB-D Salient Object Detection</b> Author(s): Zhang, Wenbo; Ji, Ge-Peng; Wang, Zhuo; Fu, Keren*; Zhao, Qijun
P040	<b>Revisiting Mid-Level Patterns for Cross-Domain Few-Shot Recognition</b> Author(s): Zou, Yixiong*; Zhang, Shanghang; Yu, Jianpeng; Tian, Yonghong; Moura, José M. F.
P041	<b>Space-Angle Super-Resolution for Multi-View Images</b> Author(s): Sun, Yuqi; Cheng, Ri; Yan, Bo*; Zhou, Shili
P042	<b>Weakly-Supervised Video Object Grounding via Stable Context Learning</b> Author(s): Wang, Wei*; Gao, Junyu; Xu, Changsheng
P043	<b>Modeling the Uncertainty for Self-supervised 3D Skeleton Action Representation Learning</b> Author(s): Su, Yukun; Lin, Guosheng; Sun, RuiZhou; Hao, Yun; Wu, Qingyao*
P044	<b>D3Net: Dual-Branch Disturbance Disentangling Network for Facial Expression Recognition</b> Author(s): Mo, Rongyun; Yan, Yan*; Xue, Jing-Hao; Chen, Si; Wang, Hanzi
P045	<b>Towards a Unified Middle Modality Learning for Visible-Infrared Person Re-Identification</b> Author(s): yukang, zhang; Yan, Yan; Lu, Yang; Wang, Hanzi*
P046	<b>ROSITA: Enhancing Vision-and-Language Semantic Alignments via Cross-and Intra-modal Knowledge Integration</b> Author(s): Cui, Yuhao; Yu, Zhou*; Wang, Chunqi; Zhao, Zhongzhou; Zhang, Ji; Wang, Meng; Yu, Jun
P047	<b>Object Point Cloud Classification via Poly-ConvolutionalArchitecture Search</b> Author(s): Lin, Xuanxiang*; Chen, Ke; Jia, Kui
P048	<b>Semantic-Guided Relation Propagation Network for Few-shot Action Recognition</b> Author(s): Wang, Xiao; Ye, Weirong; Qi, Zhongang; Zhao, Xun; Wang, Guangge; Shan, Ying; Wang, Hanzi*
P049	<b>Anti-Distillation Backdoor Attacks: Backdoors Really Can Survive in The Knowledge Distillation</b> Author(s): Ge, Yunjie*; Wang, Qian; Zheng, Baolin; Zhuang, Xinlu; Li, Qi; Shen, Chao; Wang, Cong
P050	<b>One-stage Context and Identity Hallucination Network</b> Author(s): Liu, Yinglu*; Xiang, Mingcan; Shi, Hailin; Mei, Tao
P051	<b>Enhancing Generalized Zero-Shot Learning by Visual Perturbation</b> Author(s): Chen Zhi*; Luo Yadan; Wang Sen; Qiu Ruihong; Li Jingjing; Huang Zi

P052	<b>Weakly-Supervised Temporal Action Localization via Cross-Stream Collaborative Learning</b> Author(s): Ji, Yuan*; Jia, Xu; Lu, Huchuan; Ruan, Xiang
P053	<b>Deep Interactive Video Inpainting: an Invisibility Cloak for Harry Potter</b> Author(s): Chen, Cheng*; Cai, Jiayin; Hu, Yao; Tang, Xu; Wang, Xinggang; Yuan, Chun; Bai, Xiang; Bai, Song
P054	<b>Searching Motion Graphs for Human Motion Synthesis</b> Author(s): Liu, Chenchen; Mu, Yadong*
P055	<b>When Video Classification Meets Incremental Classes</b> Author(s): Zhao, Hanbin; Qin, Xin; Su, Shihao; Fu, Yongjian; Lin, Zibo; Li, Xi*
P056	<b>Fast and Accurate Lane Detection via Frequency Domain Learning</b> Author(s): he, yulin*; Chen, Wei; Liang, Zhengfa; Chen, Dan; Tan, Yusong; Luo, Xin; Li, Chen; Guo, Yulan
P057	<b>Learning Multi-context Aware Location Representations from Large-scale Geotagged Images</b> Author(s): Yin, Yifang*; Zhang, Ying; Liu, Zhenguang; Liang, Yuxuan; Wang, Sheng; Shah, Rajiv Ratn; Zimmermann, Roger
P058	<b>MV-TON: Memory-based Video Virtual Try-on Network</b> Author(s): Zhong, Xiaojing; Wu, Zhonghua; Tan, Taizhe; Lin, Guosheng; Wu, Qingyao*
P059	<b>Token Shift Transformer for Video Classification</b> Author(s): Zhang, Hao; Hao, Yanbin*; Ngo, Chong-Wah
P060	<b>Attribute-specific Control Units in StyleGAN for Fine-grained Image Manipulation</b> Author(s): Wang, Rui; Chen, Jian; Yu, Gang; Sun, Li; Yu, Changqian; Gao, Changxin*; Sang, Nong
P061	<b>Attention-driven Graph Clustering Network</b> Author(s): Peng, Zhihao; Liu, Hui; Jia, Yuheng; Hou, Junhui*
P062	<b>Lifting the Veil of Frequency in Joint Segmentation and Depth Estimation</b> Author(s): Fu, Tianhao*; Li, Yingying; Ye, Xiaoqing; Tan, Xiao; Sun, Hao; Shen, Fumin; Ding, Errui
P063	<b>Show, Read and Reason: Table Structure Recognition with Flexible Context Aggregator</b> Author(s): Hao Liu (Tencent)*; Xin Li (Tencent); Bing Liu (Tencent); Deqiang Jiang (Tencent); Yinsong Liu (Tencent); Bo Ren (Tencent); Rongrong Ji (Xiamen University, China)
P064	<b>TransFusion: Multi-Modal Fusion for Video Tag Inference via Translation-based Knowledge Embedding</b> Author(s): Di Jin (University of Michigan)*; Zhongang Qi (Tencent); Yingmin Luo (Tencent); Ying Shan (Tencent)



- P065 **ARShoe: Real-Time Augmented Reality Shoe Try-on System on Smartphones**  
Author(s): Shan An (Beihang University); Guangfu Che (J.D.com); Jinghao Guo (JD.COM Inc); Haogang Zhu (Beihang University)\*; Junjie Ye (Tongji University); Fangru Zhou (JD.COM); Zhaoqi Zhu (J.D.com); Dong Wei (JD.COM Inc); Aishan Liu (Beihang University); Wei Zhang (JD)
- 
- P066 **Inferring the Importance of Product Appearance with Semi-supervised Multi-modal Enhancement: A Step Towards the Screenless Retailing**  
Author(s): Yongshun Gong (Shandong University)\*; Jinfeng Yi (JD AI Research); Dong-Dong Chen (Nanjing Univ); Jian Zhang (UTS); Jiayu Zhou (Michigan State University); Zhihua Zhou (Nanjing University)
- 
- P067 **AsyNCE: Disentangling False-Positives for Weakly-Supervised Video Grounding**  
Author(s): Cheng Da (Alibaba.inc)\*; Yanhao Zhang (Alibaba Group); Yun Zheng (Alibaba group); Pan Pan (Alibaba Group); Yinghui Xu (Alibaba DAMO Academy); Chunhong Pan (Institute of Automation, Chinese Academy of Sciences)
- 
- P068 **Unifying Multimodal Transformer for Bi-directional Image and Text Generation**  
Author(s): Yupan Huang (Sun Yat-sen University)\*; Hongwei Xue (University of Science and Technology of China); Bei Liu (Microsoft Research); Yutong Lu (Sun Yat-sen University)
- 
- P069 **Distantly Supervised Semantic Text Detection and Recognition for Broadcast Sports Videos Understanding**  
Author(s): Avijit Shah (Verizon Media)\*; Topojoy Biswas (Verizon Media); Sathish Ramadoss (Verizon Media); Deven Santosh Shah (Stony Brook University)
- 
- P070 **Focusing on Persons: Colorizing Old Images Learning from Modern Historical Movies**  
Author(s): Xin Jin (Beijing Electronic Science and Technology Institute); Zhonglan Li (Beijing Electronic Science and Technology Institute); Ke Liu (Beijing Electronic Science and Technology Institute); Dongqing Zou (SenseTime Research); Qing Yuan Research Institute, Shanghai Jiao Tong University)\*; Xiaodong Li (Beijing Electronic Science and Technology Institute); Xingfan Zhu (Beijing Electronic Science and Technology Institute); Ziyin Zhou (Beijing Electronic Science and Technology Institute); Qilong Sun (Beijing Electronic Science and Technology Institute); Qingyu Liu (Beijing Electronic Science and Technology Institute)
- 
- P071 **Personalized Multi-modal Video Retrieval on Mobile Devices**  
Author(s): Haotian Zhang (Samsung AI Center Toronto)\*; Allan D Jepson (Samsung Toronto AIC); Iqbal Mohamed (Samsung AI Centre Toronto); Konstantinos Derpanis (SAIC Toronto); Ran Zhang (Samsung AI Center Toronto); Afsaneh Fazly (SAIC Toronto)
- 
- P072 **A Virtual Character Generation and Animation System for E-Commerce Live Streaming**  
Author(s): Li Hu (Alibaba)\*; Bang Zhang (Damo Academy Alibaba); Peng Zhang (Alibaba); Jinwei Qi (Alibaba Group); jian cao (Alibaba); Daiheng Gao (Alibaba); haoming zhao (Alibaba); xiaoduan feng (Alibaba); Qi Wang (Alibaba group); Li'an Zhuo (Alibaba Group); Pan Pan (Alibaba Group); Yinghui Xu (Alibaba DAMO Academy)

P073

**Improving Fake News Detection by Using an Entity-enhanced Framework to Fuse Diverse Multimodal Clues**

Author(s): Peng Qi (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences ); 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); Qin He (Hangzhou ZhongkeRuijian Technology Co., Ltd.); Yongbiao Lv (Hangzhou ZhongkeRuijian Technology Co., Ltd.); Chenyang Guo (Hangzhou ZhongkeRuijian Technology Co., Ltd.); Yingchao Yu (Hangzhou ZhongkeRuijian Technology Co., Ltd.)

---

OCT.  
21

2:00 PM-5:00 PM	Poster Session 2	▼ Hall Way
P074	<b>Visual Co-Occurrence Alignment Learning for Weakly-Supervised Video Moment Retrieval</b> Author(s): Zheng Wang (Fudan University)*; Jingjing Chen (Fudan University); Yu-Gang Jiang (Fudan University)	
P075	<b>Adaptive Normalized Representation Learning for Generalizable FaceAnti-Spoofing</b> Author(s): ShuBao Liu (East China Normal University)*; Ke-Yue Zhang (East China Normal University); Taiping Yao (Tencent YouTu); mingwei Bi (Tencent); Shouhong Ding (Tencent); Jilin Li (Tencent); Feiyue Huang (Tencent); Lizhuang Ma (East China Normal University)	
P076	<b>Imitating Arbitrary Talking Style for Realistic Audio-Driven Talking Face Synthesis</b> Author(s): Haozhe Wu (Tsinghua University)*; Jia Jia (Tsinghua University); Haoyu Wang (Tsinghua University); Yishun Dou (HiSilicon Company); Chao Duan (HiSilicon Company); Qingshan Deng (HiSilicon Company)	
P077	<b>Pose-guided Inter-and Intra-part Relational Transformer for Occluded Person Re-Identification</b> Author(s): Zhongxing Ma (Beihang University); Yifan Zhao (Beihang University); Jia Li (Beihang University)*	
P078	<b>VLAD-VSA: Cross-Domain Face Presentation Attack Detection with Vocabulary Separation and Adaptation</b> Author(s): Jiong Wang (ZJU); Zhou Zhao (Zhejiang University)*; Weike Jin (Zhejiang University ); Xinyu Duan (Huawei Cloud); Zhen Lei (NLPR, CASIA, China); Baoxing Huai (Huawei Cloud); Yiling Wu (Huawei Technology Co., Ltd); Xiaofei He (Zhejiang University)	
P079	<b>End-to-End Video Object Detection with Spatial-Temporal Transformers</b> Author(s): Lu He (Shanghai Jiao Tong University )*; Qianyu Zhou (Shanghai Jiao Tong University); Xiangtai Li (Peking University); Li Niu (Shanghai Jiao Tong University); Guangliang Cheng (SenseTime Group Limited); xiao li (SenseTime); Wenxuan Liu (UCLA); Yunhai Tong (Peking University); Lizhuang Ma (Shanghai Jiao Tong University); Liqing Zhang (Shanghai Jiao Tong University)	
P080	<b>Joint-teaching: Learning to Refine Knowledge for Resource-constrained Unsupervised Cross-modal Retrieval</b> Author(s): Peng-Fei Zhang (University of Queensland)*; Jiasheng Duan (The University of Queensland); Zi Huang (University of Queensland); Hongzhi Yin (The University of Queensland)	

- P081 **AggNet for Self-supervised Monocular Depth Estimation: Go An Aggressive Step Furthe**  
 Author(s): Zhi Chen (University of Science and Technology of China); Xiaoqing Ye (baidu); Liang Du (Fudan University); Wei Yang (University of Science and Technology of China)\*; 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.)
- 
- P082 **Boosting Lightweight Single Image Super-resolution via Joint-distillation**  
 Author(s): Xiaotong Luo (Xiamen University)\*; Qiuyuan Liang (XiaMen University); Ding Liu (Bytedance); Yanyun Qu (XMU)
- 
- P083 **Discriminator-free Generative Adversarial Attack**  
 Author(s): Shaohao Lu (Sun Yat-Sen University); Yuqiao Xian (Sun Yat-sen University); Ke Yan (Tencent); Yi Hu (Tencent); Xing Sun (Tencent); Xiaowei Guo (Tencent Youtu Lab); Feiyue Huang (Tencent); WEI-SHI ZHENG (Sun Yat-sen University, China)\*
- 
- P084 **Former-DFER: Dynamic Facial Expression Recognition Transformer**  
 Author(s): Zengqun Zhao (Nanjing University of Information Science & Technology); Qingshan Liu (Nanjing University of Information Science & Technology)\*
- 
- P085 **Discovering Density-Preserving Latent Space Walks in GANs for Semantic Image Transformations**  
 Author(s): Guanyue Li (SCUT); Yi Liu (South China University of Technology); Xiwen Wei (South China University of Technology); Yang Zhang (City University of Hong Kong); Si Wu (South China University of Technology)\*; Yong Xu (South China University of Technology); Hau San Wong (City University of Hong Kong)
- 
- P086 **MGH: Metadata Guided Hypergraph Modeling for Unsupervised Person Re-identification**  
 Author(s): Yiming Wu (Zhejiang University); Xintian Wu (Zhejiang University); Xi Li (Zhejiang University)\*; Jian Tian (Zhejiang University)
- 
- P087 **Recovering the Unbiased Scene Graphs from the Biased Ones**  
 Author(s): Meng-Jiun Chiou (National University of Singapore)\*; Henghui Ding (Nanyang Technological University); Hanshu YAN (NUS); Changhu Wang (ByteDance.Inc); Roger Zimmermann (NUS); Jiashi Feng (NUS)
- 
- P088 **Cross-modal Consensus Network for Weakly Supervised Temporal Action Localization**  
 Author(s): Fa-Ting Hong (Sun Yat-Sen University); Jia-Chang Feng (Sun Yat-Sen University); Dan Xu (The Hong Kong University of Science and Technology); Ying Shan (Tencent); WEI-SHI ZHENG (Sun Yat-sen University, China)\*
- 
- P089 **Searching a Hierarchically Aggregated Fusion Architecture for Fast Multi-Modality Image Fusion**  
 Author(s): Risheng Liu (Dalian University of Technology)\*; Zhu Liu (Dalian University of Technology); Jinyuan Liu (Dalian University of Technology); Xin Fan (Dalian University of Technology)



- P090 **SuperFront: From Low-resolution to High-resolution Frontal Face Synthesis**  
Author(s): Yu Yin (Northeastern University)\*; Joseph Robinson (Northeastern University); Songyao Jiang (Northeastern University); Yue Bai (Northeastern University); Can Qin (Northeastern University); YUN FU (Northeastern University)
- 
- P091 **Learning Segment Similarity and Alignment in Large-Scale Content Based Video Retrieval**  
Author(s): Chen Jiang (Antgroup)\*; kaiming huang (ant group); Sifeng He (Ant Group); Xudong Yang (AntGroup); Wei Zhang (Ant Financial Services Group); Xiaobo Zhang (Ant Financial); Yuan Cheng (Alibaba Group); Lei Yang (Ant Group); QING WANG (Ant Group); Furong Xu (Ant Group); Tan Pan (Ant Group); Wei Chu (Ant Group)
- 
- P092 **Cut-Thumbnail: A Novel Data Augmentation for Convolutional Neural Network**  
Author(s): Tianshu Xie (University of Electronic Science and Technology of China); Xuan Cheng (University of Electronic Science and Technology of China); Xiaomin Wang (University of Electronic Science and Technology of China); Minghui Liu (University of Electronic Science and Technology of China); Jiali Deng (University of Electronic Science and Technology of China); Tao Zhou (University of Electronic Science and Technology of China); Ming Liu (University of Electronic Science and Technology of China)\*
- 
- P093 **Diffusing the Liveness Cues for Face Anti-spoofing**  
Author(s): Sheng Li (Fudan University); Xun Zhu (Fudan University); Guorui Feng (Shanghai University); Xinpeng Zhang (School of Computer Science, Fudan University)\*; Zhenxing Qian (School of Computer Science, Fudan University)
- 
- P094 **Co-Transport for Class-Incremental Learning**  
Author(s): Da-Wei Zhou (Nanjing University)\*; Han-Jia Ye (Nanjing University); De-Chuan Zhan (Nanjing University)
- 
- P095 **Skeleton-Contrastive 3D Action Representation Learning**  
Author(s): Fida Mohammad Thoker (University of Amsterdam)\*; Hazel Doughty (University of Amsterdam); Cees Snoek (University of Amsterdam)
- 
- P096 **Fast-forwarding, Rewinding, and Path Exploration in Interactive Branched Video Streaming**  
Author(s): Albin Vogel (Linkoping University); Erik Kronberg (Linkoping University); Niklas Carlsson (Linkoping University)\*
- 
- P097 **Multiview Detection with Shadow Transformer (and View-Coherent Data Augmentation)**  
Author(s): Yunzhong Hou (Australian National University)\*; Liang Zheng (Australian National University)
- 
- P098 **Domain Generalization via Feature Variation Decorrelation**  
Author(s): Chang Liu (Northeastern University)\*; Lichen Wang (Northeastern University); Kai Li (Northeastern University); YUN FU (Northeastern University)

- P099 **Occlusion-aware Bi-directional Guided Network for Light Field Salient Object Detection**  
 Author(s): Dong Jing (Beijing Jiaotong University)\*; Shuo Zhang (Beijing Jiaotong University); Runmin Cong (Beijing Jiaotong University); Youfang Lin (Beijing Jiaotong University)
- 
- P100 **One-Stage Visual Grounding via Semantic-Aware Feature Filter**  
 Author(s): Jiabo Ye (East China Normal University); Xin Lin (ECNU)\*; Liang He (ECNU); Dingbang Li (ECNU); Qin Chen (East China Normal University)
- 
- P101 **Few-Shot Multi-Agent Perception**  
 Author(s): Chenyou Fan (Shenzhen Institute of Artificial Intelligence and Robotics for Society)\*; Junjie Hu (The Chinese University of Hong Kong, Shenzhen); Jianwei Huang (The Chinese University of Hong Kong, Shenzhen)
- 
- P102 **SI3DP: Source Identification Challenges and Benchmark for Consumer-Level 3D Printer Forensics**  
 Author(s): Bo Seok Shim (Hallym University); You Seung Shin (Hallym University); Seong wook Park (Hallym University); Jong-Uk Hou (Hallym University)\*
- 
- P103 **Exploring Sequence Feature Alignment for Domain Adaptive Detection Transformers**  
 Author(s): Wen Wang (University of Science and Technology of China); Yang Cao (University of Science and Technology of China)\*; 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.com)
- 
- P104 **Towards Realistic Visual Dubbing with Heterogeneous Sources**  
 Author(s): Tianyi Xie (Bytedance)\*; Liucheng Liao (University of Electronic Science and Technology of China); Cheng Bi (Bytedance Ltd.); Benlai Tang (ByteDance AI LAB); Xiang Yin (ByteDance AI LAB); Jianfei Yang (Nanyang Technological University); Mingjie Wang (Memorial University of Newfoundland); Jiali Yao (Bytedance); Yang Zhang (ByteDance AI LAB); Zejun Ma (Bytedance)
- 
- P105 **Deep Self-Supervised t-SNE for Multi-modal Subspace Clustering**  
 Author(s): Qianqian Wang (Xidian University)\*; Wei Xia (Xidian University); Zhiqiang Tao (Santa Clara University); Quanxue Gao (State Key Laboratory of Integrated Services Networks, Xidian University); Xiaochun Cao (Chinese Academy of Sciences)
- 
- P106 **Multimodal Video Summarization via Time-Aware Transformers**  
 Author(s): Xindi Shang (National University of Singapore)\*; Zehuan Yuan (Bytedance.Inc); Anran Wang (Bytedance.Inc); Changhu Wang (ByteDance.Inc)
- 
- P107 **State-aware Video Procedural Captioning**  
 Author(s): Taichi Nishimura (Kyoto University)\*; Atsushi Hashimoto (OMRON SINIC X Corp.); Yoshitaka Ushiku (OMRON SINIC X ); Hirotaka Kameko (Kyoto University); Shinsuke Mori (Academic Center for Computing and Media Studies, Kyoto University)



- P108      **AMSS-Net: Audio Manipulation on User-Specified Sources with Textual Queries**  
Author(s): Woosung Choi (Korea University)\*; Minseok Kim (Korea University); Marco Martínez Ramírez (Sony Corporation); Jaehwa Chung (Korea National Open University); Soonyoung Jung (Korea University)
- 
- P109      **Fully Functional Image Manipulation Using Scene Graphs in A Bounding-Box Free Way**  
Author(s): sitong su (University of Electronic Science and Technology of China)\*; Lianli Gao (The University of Electronic Science and Technology of China); junchen zhu (University of Electronic Science and Technology of China); Jie Shao (University of Electronic Science and Technology of China); Jingkuan Song (UESTC)
- 
- P110      **Multi-Level Counterfactual Contrast for Visual Commonsense Reasoning**  
Author(s): Xi Zhang (CASIA)\*; Feifei Zhang (Institute of Automation, Chinese Academy of Sciences); Changsheng Xu (CASIA)
- 
- P111      **Data-Free Ensemble Knowledge Distillation for Privacy-conscious Multimedia Model Compression**  
Author(s): Zhiwei Hao (Beijing Institute of Technology)\*; Yong Luo (Wuhan University); Han Hu (Beijing Institute of Technology, China); Jianping An (Beijing Institute of Technology); Yonggang Wen (Nanyang Technological University)
- 
- P112      **SM-SGE: A Self-Supervised Multi-Scale Skeleton Graph Encoding Framework for Person Re-Identification**  
Author(s): Haocong Rao (Shenzhen Institutes of Advanced Technology)\*; Xiping Hu (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences); Jun Cheng (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences); Bin Hu (Lanzhou University)
- 
- P113      **Video Transformer for Deepfake Detection with Incremental Learning**  
Author(s): Sohail Ahmed Khan (Mohamed bin Zayed University of Artificial Intelligence); Hang Dai (Mohamed bin Zayed University of Artificial Intelligence)\*
- 
- P114      **Chinese Character Inpainting with Contextual Semantic Constraints**  
Author(s): Jiahao Wang (Tianjin University); Gang Pan (Tianjin University)\*; Di Sun (Tianjin University of Science and Technology); Jiawan Zhang (Tianjin University)
- 
- P115      **Curriculum-Based Meta-learning**  
Author(s): Ji Zhang (University of Electronic Science and Technology of China); Jingkuan Song (UESTC)\*; Yazhou Yao (Nanjing University of Science and Technology); Lianli Gao (The University of Electronic Science and Technology of China)
- 
- P116      **Ego-Deliver: A Large-Scale Dataset for Egocentric Video Analysis**  
Author(s): Haonan Qiu (East China Normal University); Pan He (University of Florida); Shuchun Liu (Alibaba Group); weiyuan shao (eleme); Feiyun Zhang (eleme AILab ); Jiajun Wang (Eleme Inc.); Liang He (ECNU); Feng Wang (ECNU)\*

- P117      **Adversarial Pixel Masking: A Defense against Physical Attacks for Pre-trained Object Detectors**  
Author(s): Bing-Han Chiang (NTHU); Chi-Shen Chan (NTHU); Shan-Hung Wu (NTHU)\*
- 
- P118      **Knowledge-Supervised Learning: Knowledge Consensus Constraints for Person Re-Identification**  
Author(s): Li Wang (inspur); Baoyu Fan (Inspur Electronic Information Industry Co.,Ltd.)\*; Zhenhua Guo (Inspur Electronic Information Industry Co.,Ltd.); Yaqian Zhao (Inspur); Runze Zhang (Inspur Electronic Information Industry Co.,Ltd.); Rengang Li (Inspur Electronic Information Industry Co.,Ltd.); Weifeng Gong (Inspur Electronic Information Industry Co.,Ltd.); Endong Wang (Inspur Electronic Information Industry Co.,Ltd.)
- 
- P119      **View-normalized Skeleton Generation for Action Recognition**  
Author(s): Qingzhe Pan (Xidian University); Zhifu Zhao (Xidian University); Xuemei Xie (Xidian University)\*; Jianan Li (Xidian University); Yuhan Cao (Xidian University); Guangming Shi (Xidian University)
- 
- P120      **Learning Hierarchical Embedding for Video Instance Segmentation**  
Author(s): Zheyun Qin (Shandong university); Xiankai Lu (Shandong University)\*; Xiushan Nie (Shandong Jianzhu University); Xiantong Zhen (University of Amsterdam); Yilong Yin (Shandong University)
- 
- P121      **Text as Neural Operator:Image Manipulation by Text Instruction**  
Author(s): Tianhao Zhang (Google Research)\*; Hung-Yu Tseng (University of California, Merced); Lu Jiang (Google Research); Weilong Yang (Google Inc. ); Honglak Lee (Google / U. Michigan); Irfan Essa (Google)
- 
- P122      **DSANet: Dynamic Segment Aggregation Network for Video-Level Representation Learning**  
Author(s): Wenhao Wu (Baidu)\*; Yuxiang Zhao (SIAT); Yanwu Xu (University of Pittsburgh); Xiao Tan (Baidu Inc.); Dongliang He (Baidu); Zhikang Zou (Huazhong university of science and technology); Jin Ye (Baidu Inc.); Yingying Li (Baidu); Mingde Yao (University of Science and Technology of China); Dong Zichao (Baidu); yifeng shi (baidu)
- 
- P123      **Structured Text Understanding with Multi-Modal Transformers**  
Author(s): yulin li (Baidu.Inc)\*; yuxi qian (Beijing University of Posts and Telecommunications); yuechen yu (Baidu.Inc); xiameng Qin (Baidu); Chengquan Zhang (Baidu Inc); Yan Liu (Taikang Insurance Group); Kun Yao (Baidu Inc.); Junyu Han (Baidu Inc.); jingtuo liu (baidu); Errui Ding (Baidu Inc.)
- 
- P124      **Local Graph Convolutional Networks for Cross-Modal Hashing**  
Author(s): Yudong Chen (The University of Queensland)\*; Sen Wang (The University of Queensland); Jianglin Lu (Shenzhen University); Zhi Chen (The University of Queensland); Zheng Zhang (Harbin Institute of Technology, Shenzhen); Zi Huang (University of Queensland)
- 
- P125      **Metric Learning for Anti-Compression Facial Forgery Detection**  
Author(s): Cao Shenhao (Wuhan University); Qin Zou (Wuhan University)\*; Xiuqing Mao (Information Engineering University); Dengpan Ye (School of Computer, Wuhan University); Zhongyuan Wang (Wuhan University)



- P126 **ASFM-Net: Asymmetrical Siamese Feature Matching Network for Point Completion**  
Author(s): Yaqi Xia (Xidian University)\*; Yan Xia (Technical University of Munich); Wei Li (Peking University); Song Rui (Xidian University); Kailang Cao (Xidian University); Uwe M Stilla (Technical University of Munich)
- 
- P127 **Capsule-based Object Tracking with Natural Language Specification**  
Author(s): Ding Ma (Harbin Institute of Technology)\*; XIANGQIAN WU (Harbin Institute of Technology, China)
- 
- P128 **Faster-PPN: Towards Real-Time Semantic Segmentation with Dual Mutual Learning for Ultra-High Resolution Images**  
Author(s): Bicheng Dai (Xiamen University); Kaisheng Wu (Xiamen University); Tong Wu (Xiamen University); Kai Li (Northeastern University); Yanyun Qu (XMU)\*; Yuan Xie (East China Normal University); YUN FU (Northeastern University)
- 
- P129 **Distributed Attention for Grounded Image Captioning**  
Author(s): Nenglun Chen (The University of Hong Kong)\*; Xingjia Pan (Tencent); Runnan Chen (The University of Hong Kong); LEI YANG (University of Hong Kong); Zhiwen Lin (Zhejiang University); Yuqiang Ren (Tencent Youtu Lab); Haolei Yuan (Tencent Youtu Lab); Xiaowei Guo (Tencent Youtu Lab); Feiyue Huang (Tencent); Wenping Wang (The University of Hong Kong)
- 
- P130 **Multi-initialization Optimization Network for Accurate 3D Human Pose and Shape Estimation**  
Author(s): Zhiwei Liu (Institute of Automation Chinese Academy of Sciences)\*; Xiangyu Zhu (Chinese Academy of Science); Lu Yang (Beijing University of Posts and Telecommunications); Xiang Yan (Dilusense Technology Corporation); Ming Tang (Institute of Automation, Chinese Academy of Sciences); Zhen Lei (NLPR, CASIA, China); Guibo Zhu (Institute of Automation, Chinese Academy of Sciences); Xuetao Feng (Alibaba Group); Yan Wang (Alibaba Group); Jinqiao Wang (Institute of Automation, Chinese Academy of Sciences)
- 
- P131 **Feedback Network for Mutually Boosted Stereo Image Super-Resolution and Disparity Estimation**  
Author(s): Qinyan Dai (East China Normal University); Juncheng Li (East China Normal University); Qiaosi Yi (East China Normal University); Faming Fang (East China Normal University)\*; Guixu Zhang (East China Normal University)
- 
- P132 **Merging Multiple Template Matching Predictions in Intra Coding with Attentive Convolutional Neural Network**  
Author(s): Qijun Wang (Anhui University)\*; GuoDong Zheng (Anhui University)
- 
- P133 **Camera-Agnostic Person Re-Identification via Adversarial Disentangling Learning**  
Author(s): Hao Ni (The University of Electronic Science and Technology of China)\*; Jingkuan Song (UESTC); Xiaosu Zhu (University of Electronic Science and Technology of China); Feng Zheng (SUSTech); Lianli Gao (The University of Electronic Science and Technology of China)

2:00 PM-5:00 PM	Art Session Poster	▼ Hall Way
P134	<p><b>Syntropic Counterpoints: Metaphysics of The Machines</b></p> <p>Author(s): Predrag K. Nikolic (ShanghaiTech University); Predrag K. Nikolic (School of Creativity and Art, ShanghaiTech University, Shanghai, China) *; RUIYANG LIU (ShanghaiTech University)</p>	
P135	<p><b>Kandinsky Mobile: Abstract Art-Inspired Interactive Visualization of Social Discussions on Mobile Devices</b></p> <p>Author(s): Castillo Clarence Fitzgerald Gumtang (Nanyang Technological University); Sourav S Bhowmick (Nanyang Technological University) *</p>	
P136	<p><b>Sand Scope: An Interactive Installation for Revealing the Connection Between Mental Space and Life Space in a Microcosm of the World</b></p> <p>Author(s): Lyn Chao-ling Chen (Artist; Researcher(Fu Jen Catholic University))*</p>	
P137	<p><b>Heraclitus's Forest: An Interactive Artwork for Oral History</b></p> <p>Author(s): Lin Wang (Guangzhou Academy of Fine Arts); Zhonghao Lin (The Chinese University of Hong Kong, Shenzhen ); Wei Cai (The Chinese University of Hong Kong, Shenzhen)*</p>	
P138	<p><b>Affective Color Fields: Reimagining Rothkoesque Artwork As An Interactive Companion For Artistic Self-Expression</b></p> <p>Author(s): Aiden Kang (Tufts University)*; Liang Wang (Tufts University); Ziyu Zhou (Tufts University); zhe huang (Tufts University); Robert Jacob (Tufts University)</p>	
P139	<p><b>Apercevoir: Bio Internet of Things Interactive System</b></p> <p>Author(s): You-Yang Hu (NTU)*; Chiao-Chi Chou (NTHU)</p>	



October 22, 2021

8:30 AM-10:00 AM	Session 15: Best Paper Session	▼ Langbo (Whole Room)
<b>Session Chairs: Tat-Seng Chua, National University of Singapore Yang Yang, University of Electronic Science and Technology of China</b>		
8:30 AM-8:48 AM	<b>Speech2AffectiveGestures: Synthesizing Co-Speech Gestures with Generative Adversarial Affective Expression Learning</b> Author(s): Uttaran Bhattacharya (University of Maryland, College Park) *; Elizabeth Childs (University of Maryland, College Park); Nicholas S Rewkowski (UMD College Park); Dinesh Manocha (University of Maryland at College Park)	
8:48 AM-9:06 AM	<b>Video Background Music Generation with Controllable Music Transformer</b> Author(s): Shangzhe Di (Beihang University); Zeren Jiang (Megvii); Si Liu (Beihang University) *; Zhaokai Wang (beihang university); Leyan Zhu (Beihang University); Zexin He (Beihang University); Hongming Liu (Sea AI Labs); Shuicheng Yan (Sea AI Labs)	
9:06 AM-9:24 AM	<b>PIMNet: A Parallel, Iterative and Mimicking Network for Scene Text Recognition</b> Author(s): Zhi Qiao (Institute of Information Engineering, Chinese Academy of Sciences); Yu Zhou (Institute of Information Engineering, CAS; Also with University of Chinese Academy of Sciences)*; Jin Wei (School of Information and Communication Engineering, 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 (Mashang Consumer Finance Co., Ltd.); Hongbin Wang (Mashang Consumer Finance Co, Ltd); Weiping Wang (Institute of Information Engineering, CAS, China)	
9:24 AM-9:42 AM	<b>Theophany: Multimodal Speech Augmentation in Instantaneous Privacy Channels</b> Author(s): Abhishek Kumar (University of Helsinki) *; Tristan Braud (HKUST); Lik Hang Lee (KAIST); Pan Hui (Hong Kong University of Science and Technology)	
9:42 AM-10:00 AM	<b>aBio: Active Bi-Olfactory Display Using Subwoofers for Virtual Reality</b> Author(s): You-Yang Hu (NTU)*; Yao-Fu Jan (NTU); Kuan-Wei Tseng (NTU); You-Shin Tsai (NTU); Hung-Ming Sung (NTU); Jin-Yao Lin (Tainan National University of the Arts); Yi-Ping Hung (NTU)	

OCT.  
22

4:00 PM-6:00 PM	Session 17: Video Program and Demo Session	▼ Tianfu Room A
<b>Session Chair: Jingkuan Song, University of Electronic Science and Technology of China</b>		
4:00 PM-4:05 PM	<b>Text2Video: Automatic Video Generation Based on Text Scripts</b> Author(s): Yipeng Yu (Tencent)*; Zirui Tu (Tencent); Longyu Lu (Tencent); Xiao Chen (Tencent); Hui Zhan (Tencent); Zixun Sun (tencent)	
4:05 PM-4:10 PM	<b>A System for Interactive and Intelligent AD Auxiliary Screening</b> Author(s): sen yang (UESTC)*; Qike Zhao (Center for Future Media, University of Electronic Science and Technology of China); Lanxin Miao (UESTC); Min Chen (University of Electronic Science and Technology of China); Lianli Gao (The University of Electronic Science and Technology of China); Jingkuan Song (UESTC); Weidong Le (Institute of Neurology, Sichuan Academy of Medical Sciences & Provincial People's Hospital)	
4:10 PM-4:15 PM	<b>Move As You Like: Image Animation in E-Commerce Scenario</b> Author(s): 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 (Alibaba Group); Yuning Jiang (Alibaba Group); Wen Li (University of Electronic Science and Technology of China); Lixin Duan (University of Electronic Science and Technology of China)	
4:15 PM-4:20 PM	<b>MDMS: Music Data Matching System for Query Variant Retrieval</b> Author(s): Rinita Roy (Technical University of Munich)*; Ruben Mayer (Technical University of Munich); Hans-Arno Jacobsen (University of Toronto)	
4:20 PM-4:25 PM	<b>Community Generated VR Painting using Eye Gaze</b> Author(s): Mu Mu (The University of Northampton)*; Murtada Dohan (The University of Northampton)	
4:25 PM-4:30 PM	<b>Sync Glass: Virtual Pouring and Toasting Experience with Multimodal Presentation</b> Author(s): Yuki TAJIMA (KDDI Research.)*; Toshiharu Horiuchi (KDDI Research, Inc.); Gen Hattori (KDDI Research)	
4:30 PM-4:35 PM	<b>VideoDiscovery: An Automatic Short-Video Generation System for E-commerce Live-streaming</b> Author(s): Yanhao Zhang (Alibaba Group)*; Qiang Wang (Alibaba Group); Yun Zheng (Alibaba group); Pan Pan (Alibaba Group); Xu Yinghui (Alibaba Group)	
4:35 PM-4:40 PM	<b>SmartSales: An AI-Powered Telemarketing Coaching System in FinTech</b> Author(s): Yuanfeng Song (The Hong Kong University of Science and Technology)*; Xuefang Zhao (WeBank Co., Ltd); Di Jiang (WeBank); Xiaoling Huang (WeBank Co., Ltd); Weiwei Zhao (WeBank); Qian Xu (WeBank); Raymond Chi-Wing Wong (Hong Kong University of Science and Technology); Qiang Yang (Hong Kong UST)	



- 4:40 PM-4:45 PM **SmartMeeting: Automatic Meeting Transcription and Summarization for In-Person Conversations**  
 Author(s): Yuanfeng Song (The Hong Kong University of Science and Technology)\*; Di Jiang (WeBank); Xuefang Zhao (WeBank Co., Ltd); Xiaoling Huang (WeBank Co., Ltd); Qian Xu (WeBank); Raymond Chi-Wing Wong (Hong Kong University of Science and Technology); Qiang Yang (Hong Kong UST)
- 
- 4:45 PM-4:50 PM **Aesthetic Evaluation and Guidance for Mobile Photography**  
 Author(s): Hao Lou (Beijing Electronic Science and Technology Institute); Heng Huang (Beijing Electronic Science and Technology Institute); Chaoen Xiao (Beijing Electronic Science and Technology Institute); Xin Jin (Beijing Electronic Science and Technology Institute)\*
- 
- 4:50 PM-4:55 PM **A Question Answering System for Unstructured Table Images**  
 Author(s): Wenyuan Xue (Beijing Jiaotong University)\*; Siqi Cai (Beijing Jiaotong University); Wen Wang (); Qingyong Li (Beijing Jiaotong University); Baosheng Yu (The University of Sydney); Yibing Zhan (Jingdong ); Dacheng Tao (Jingdong)
- 
- 4:55 PM-5:00 PM **Post2Story: Automatically Generating Storylines from Microblogging Platforms**  
 Author(s): Xujian Zhao (Southwest University of Science and Technology)\*; Chongwei Wang (Southwest University of Science and Technology); Peiquan Jin (University of Science and Technology of China); hui zhang (Southwest University of Science and Technology); Chunming Yang (Southwest University of Science and Technology); bo li (Southwest University of Science and Technology)
- 
- 5:00 PM-5:05 PM **ViDA-MAN: Visual Dialog with Digital Humans**  
 Author(s): Tong Shen (JD AI Research)\*; Jiawei Zuo (JD.com); Fan Shi (JD.com); Jin Zhang (Migu); Liqin Jiang (Migu); Meng Chen (JD AI); Zhengchen Zhang (JD.com); Wei Zhang (JD AI Research); Xiaodong He (JD AI Research); Tao Mei (AI Research of JD.com)
- 
- 5:05 PM-5:10 PM **A Picture is Worth a Thousand Words: A Unified System for Diverse Captions and Rich Images Generation**  
 Author(s): Yupan Huang (Sun Yat-sen University)\*; Bei Liu (Microsoft Research); Jianlong Fu (Microsoft Research); Yutong Lu (Sun Yat-sen University)
- 
- 5:10 PM-5:15 PM **Softly: Simulated Empathic Touch between an Agent and a Human**  
 Author(s): Maxime Grandidier (Université de Technologie de Compiègne)\*
- 
- 5:15 PM-5:20 PM **RecipeLog: Recipe Authoring App for Accurate Food Recording**  
 Author(s): Akihisa Ishino (The University of Tokyo)\*; Yoko Yamakata (University of Tokyo, Japan); Hiroaki KARASAWA (Hongo Software Development); Kiyoharu Aizawa (The University of Tokyo)
-

5:20 PM-5:25 PM	<p><b>iART: A Search Engine for Art-Historical Images to Support Research in the Humanities</b></p> <p>Author(s): Matthias Springstein (TIB – Leibniz Information Centre for Science and Technology)*; Stefanie Schneider (Ludwig Maximilian University of Munich); Javad Rahnema (University Paderborn); Eyke Hüllermeier (Ludwig Maximilian University of Munich); Hubertus Kohle (Ludwig Maximilian University of Munich); Ralph Ewerth (TIB - Leibniz Information Center for Science and Technology)</p>
5:25 PM-5:30 PM	<p><b>ArtiVisual: A Platform to Generate and Compare Art</b></p> <p>Author(s): Jardenna Mohazzab (University of Amsterdam); Abe Vos (University of Amsterdam); Jonathan van Westendorp (University of Amsterdam); Lucas Lageweg (University of Amsterdam); Dylan Prins (University of Amsterdam); Aritra Bhowmik (University of Amsterdam)*</p>
5:30 PM-5:35 PM	<p><b>GCNillustrator: Illustrating the effect of hyperparameters on Graph Convolutional Networks</b></p> <p>Author(s): Ivona Najdenkoska (University of Amsterdam)*; Jeroen den Boef (University of Amsterdam); Thomas Schneider (University of Amsterdam); Justo van der Werf (University of Amsterdam); Reinier de Ridder (University of Amsterdam); Fajar Fathurrahman (University of Amsterdam); Marcel Worrying (University of Amsterdam)</p>
5:35 PM-5:40 PM	<p><b>On-demand Action Detection System using Pose Information</b></p> <p>Author(s): Noboru Yoshida (NEC Corporation)*; Jianquan Liu (NEC Corporation)</p>
5:40 PM-5:45 PM	<p><b>APF: An Adversarial Privacy-preserving Filter to Protect Portrait Information</b></p> <p>Author(s): Xian Zhao (Beijing Jiaotong University, China)*; Jiaming Zhang (Beijing Jiaotong University); Xiaowen Huang (Beijing Jiaotong University)</p>
5:45 PM-5:50 PM	<p><b>Text-driven 3D Avatar Animation with Emotional and Expressive Behaviors</b></p> <p>Author(s): li hu (Alibaba)*; Jinwei Qi (Alibaba Group); Bang Zhang (Damo Academy Alibaba); Pan Pan (Alibaba Group); Xu Yinghui (Alibaba Group)</p>
5:50 PM-5:55 PM	<p><b>Text to Scene: A System of Configurable 3D Indoor Scene Synthesis</b></p> <p>Author(s): Xinyan Yang (Key Laboratory of Media Audio &amp; Video (Communication University of China), Ministry of Education); Fei Hu (State Key Laboratory of Media Convergence and Communication, Communication University of China)*; Long Ye (Communication University of China)</p>
5:55 PM-6:00 PM	<p><b>MovieREP: A New Movie Reproduction Framework for Film Soundtrack</b></p> <p>Author(s): Ruiqi Wang (Communication University of China); Long Ye (Communication University of China)*; Qin Zhang (Communication University of China)</p>



4:00 PM-4:15 PM

Session 18: Doctoral Symposium

▼ Tianfu Room B

Session Chair: Shuaicheng Liu, University of Electronic Science and Technology of China

4:00 PM-4:15 PM

**End-to-end Quality of Experience Evaluation for HTTP Adaptive Streaming**

Author(s): Babak Taraghi (Alpen-Adria-Universität Klagenfurt) \*

4:15 PM-4:30 PM

**Generative Adversarial Network for Text-to-Face Synthesis and Manipulation**

Author(s): ZHOU Yutong (Ritsumeikan University) \*

4:30 PM-4:45 PM

**Actor-Object Interaction and Scene Attentive Network for Action Recognition**

Author(s): Nitika Nigam (IIT BHU) \*

4:45 PM-5:00 PM

**GAN-aided Serial Dependence Study in Medical Image Perception**

Author(s): Zhihang Ren (UC Berkeley) \*

5:00 PM-5:15 PM

**Image Style Transfer with Generative Adversarial Networks**

Author(s): Ru Li (UESTC;Megvii)\*

5:15 PM-5:30 PM

**Annotation-Efficient Semantic Segmentation with Shape Prior Knowledge**

Author(s): Yuhang Lu (University of South Carolina) \*

5:30 PM-5:45 PM

**Neural-based Rendering and Application**

Author(s): Peng Dai (The University of Hong Kong) \*

5:45 PM-6:00 PM

**Towards Bridging Video and Language by Caption Generation and Sentence Localization**

Author(s): Shaoxiang Chen (Fudan University) \*

6:00 PM-6:15 PM

**Situational Anomaly Detection in Multimedia Data under Concept Drift**

Author(s): Pratibha Kumari (IIT Ropar) \*

6:15 PM-6:30 PM

**Dynamic Knowledge Distillation with Cross-Modality Knowledge Transfer**

Author(s): Guangzhi Wang (National University of Singapore) \*

OCT.  
22

9:00 AM-12:00 PM	Poster Session 3	▼ Hall Way
P140	<p><b>Learning to Understand Traffic Signs</b></p> <p>Author(s): Yunfei Guo (Institute of Automation, Chinese Academy of Sciences)*; Wei Feng (CAS); Fei yin (Institute of Automation of Chinese Academy of Sciences); TAO XUE (tencent); shuqi mei (Tencent); Cheng-Lin Liu (Institute of Automation of Chinese Academy of Sciences)</p>	
P141	<p><b>R-GAN: Exploring Human-like Way for Reasonable Text-to-Image Synthesis via Generative Adversarial Networks</b></p> <p>Author(s): Yanyuan Qiao (University of Adelaide)*; Qi Chen (University of Adelaide); Chaorui Deng (University of Adelaide); Ning Ding (South China University of Technology); Yuankai Qi (The University of Adelaide); Mingkui Tan (South China University of Technology); Xincheng Ren (Yanan University); Qi Wu (University of Adelaide)</p>	
P142	<p><b>Cross-modality Discrepant Interaction Network for RGB-D Salient Object Detection</b></p> <p>Author(s): Chen Zhang (Beijing Jiaotong University); Runmin Cong (Beijing Jiaotong University)*; Qinwei Lin (Beijing Jiaotong University); Lin Ma (Meituan); Feng Li (Beijing Jiaotong University); Yao Zhao (Beijing Jiaotong University); Sam Kwong (City Univeristy of Hong Kong)</p>	
P143	<p><b>Deconfounded and Explainable Interactive Vision-Language Retrieval of Complex Scenes</b></p> <p>Author(s): Junda Wu (New York University); Tong Yu (Carnegie Mellon University); Shuai Li (Shanghai Jiao Tong University)*</p>	
P144	<p><b>Long Short-term Convolutional Transformer for No-Reference Video Quality Assessment</b></p> <p>Author(s): Junyong You (Norwegian Research Centre)*</p>	
P145	<p><b>Automatic Channel Pruning with Hyper-parameter Search and Dynamic Masking</b></p> <p>Author(s): Baopu Li (BAIDU USA LLC)*; Yanwen Fan (Baidu); Zhihong Pan (Baidu USA LLC); Yuchen Bian (Baidu Research); gang zhang (Baidu Inc.)</p>	
P146	<p><b>SVHAN: Sequential View Based Hierarchical Attention Network for 3D Shape Recognition</b></p> <p>Author(s): Yue Zhao (Tianjin University); Weizhi Nie (Tianjin University) *; An-An Liu (Tianjin University); Zan Gao (Qilu University of Technology (Shandong Academy of Sciences), Shandong Computer Science Center (National Supercomputer Center in Jinan), Shandong Artificial Intelligence Institute, China); Yu-ting Su (Tianjin University)</p>	
P147	<p><b>ASFD: Automatic and Scalable Face Detector</b></p> <p>Author(s): Jian Li (Tencent Youtu)*; Bin Zhang (Southeast University); Yabiao Wang (Tencent); Ying Tai (Tencent YouTu); Zhenyu Zhang (Tencent); Chengjie Wang (Tencent); Jilin Li (Tencent); Xiaoming Huang (Tencent Youtu); Yili Xia (Southeast University)</p>	



- P148      **BridgeNet: A Joint Learning Network of Depth Map Super-Resolution and Monocular Depth Estimation**  
Author(s): Qi Tang (Beijing Jiaotong University); Runmin Cong (Beijing Jiaotong University)\*; Ronghui Sheng (Beijing Jiaotong University); Lingzhi He (Beijing Jiaotong University); Dan Zhang (UISEE Technology (Beijing) Co., Ltd.); Yao Zhao (Beijing Jiaotong University); Sam Kwong (City Univeristy of Hong Kong)
- 
- P149      **LSTC: Boosting Atomic Action Detection with Long-Short-Term Context**  
Author(s): Yuxi Li (Tencent)\*; Boshen Zhang (Tencent); Jian Li (Tencent Youtu); Yabiao Wang (Tencent); Weiyao Lin (Shanghai Jiao Tong university); Chengjie Wang (Tencent); Jilin Li (Tencent); Feiyue Huang (Tencent)
- 
- P150      **UACANet: Uncertainty Augmented Context Attention for Polyp Semgnetaiion**  
Author(s): Taehun Kim (Pohang University of Science and Technology)\*; Hyemin Lee (Pohang University of Science and Technology); Daijin Kim (Pohang University of Science and Technology)
- 
- P151      **Weight Evolution: Improving Deep Neural Networks Training through Evolving Inferior Weight Values**  
Author(s): Zhenquan Lin (South China University of Technology); Kailing Guo (South China University of Technology)\*; Xiaofen Xing ( South China University of Technology); Xiangmin Xu (South China University of Technology)
- 
- P152      **Coarse to Fine: Domain Adaptive Crowd Counting via Adversarial Scoring Network**  
Author(s): Zhikang Zou (Huazhong University of Science and Technology) \*; Xiaoye Qu (Huawei Cloud); Pan Zhou (Huazhong University of Science and Technology); Shuangjie Xu (Deeproute.ai); Xiaoqing Ye (baidu); Wenhao Wu (Baidu); Jin Ye (Baidu Inc.)
- 
- P153      **Towards Adversarial Patch Analysis and Certified Defense against Crowd Counting**  
Author(s): Qiming Wu (Huazhong University of Science and Technology) \*; Zhikang Zou (Huazhong university of science and technology); Pan Zhou (Huazhong University of Science and Technology); Xiaoqing Ye (baidu); binghui wang (duke); Ang Li (Duke University)
- 
- P154      **Conceptual and Syntactical Cross-modal Alignment with Cross-level Consistency for Image-Text Matching**  
Author(s): Pengpeng Zeng (University of Electronic Science and Technology of China); Lianli Gao (The University of Electronic Science and Technology of China); Xinyu Lyu (University of Electronic Science and Technology of China); Shuaiqi Jing (University of Electronic Science and Technology of China); Jingkuan Song (UESTC)\*
- 
- P155      **SSPU-Net: Self-Supervised Point Cloud Upsampling via Differentiable Rendering**  
Author(s): YiFan Zhao (Nanjing University Of Science And Technology); Le Hui (Nanjing University of Science and Technology); Jin Xie (Nanjing University of Science and Technology) \*
-

- P156      **VmAP: A Fair Metric for Video Object Detection**  
 Author(s): Anupam Sobti (IIT Delhi) \*; Vaibhav Mavi (Indian Institute of Technology Delhi); M Balakrishnan (IIT Delhi); Chetan Arora (Indian Institute of Technology Delhi)
- 
- P157      **Source Data-free Unsupervised Domain Adaptation for Semantic Segmentation**  
 Author(s): Mucong Ye (Beihang University); Jing Zhang (Beihang University)\*; Jinpeng Ouyang (Beihang University); Yuan Ding (Beihang university)
- 
- P158      **Yes, “Attention Is All You Need”, for Exemplar based Colorization**  
 Author(s): Wang Yin (Beijing University of Posts and Telecommunications)\*; Peng Lu (Beijing University of Posts and Telecommunications); Zhaoran Zhao (Beijing University of Posts and Telecommunications); Xujun Peng (Information Sciences Institute, University of Southern California)
- 
- P159      **Heuristic Depth Estimation via Recurrent Prediction with Confidence-Aware Loss**  
 Author(s): jiehua zhang (Hangzhou Dianzi University)\*; Liang Li (Institute of Computing Technology, Chinese Academy of Sciences); Chenggang Yan (Hangzhou Dianzi University); Yaoqi Sun (Hangzhou Dianzi University); tao shen (Kunming University of Science and Technology); Jiyong Zhang (Hangzhou Dianzi University); Zhan Wang (Moreal Pte. Ltd)
- 
- P160      **Unsupervised Cross-Modal Distillation for Thermal Infrared Tracking**  
 Author(s): Jingxian Sun (Northwestern Polytechnical University); Lichao Zhang (Air Force Engineering University); Yufei Zha (Northwestern Polytechnical University)\*; Abel Gonzalez-Garcia (Computer Vision Center); Peng Zhang (Northwestern Polytechnical University); Wei Huang (School of Information Engineering, Nanchang University); Yanning Zhang (Northwestern Polytechnical University)
- 
- P161      **ABPNet: Adaptive Background Modeling for Generalized Few Shot Segmentation**  
 Author(s): Kaiqi Dong (University of Science and Technology of China)\*; Wei Yang (University of Science and Technology of China); Zhenbo Xu (University of Science and Technology of China); Liusheng Huang (University of Science and Technology of China); Zhidong Yu (University of Science and Technology of China)
- 
- P162      **Towards Reasoning Ability in Scene Text Visual Question Answering**  
 Author(s): Qingqing Wang (Shanghai Jiao Tong University)\*; Liqiang Xiao (Artificial Intelligence Institute, SJTU); Yue Lu (East China Normal University); Yaohui Jin (Shanghai Jiao Tong University); Hao He (Shanghai Jiao Tong University)
- 
- P163      **Multi-caption Text-to-Face Synthesis: Dataset and Algorithm**  
 Author(s): Jianxin Sun (Center for Research on Intelligent Perception and Computing, NLPR, CASIA)\*; Qi Li (CASIA); Weining Wang (Institute of Automation, Chinese Academy of Sciences); Jian Zhao (Institute of North Electronic Equipment); Zhenan Sun (Chinese of Academy of Sciences)



- P164 **Multimodal Compatibility Modeling via Exploring the Consistent and Complementary Correlations**  
Author(s): Weili Guan (Monash University)\*; Haokun Wen (Shandong university); Xuemeng Song (Shandong University); Chung-Hsing Yeh (Monash University); Xiaojun Chang (Monash University); Liqiang Nie (Shandong University)
- P165 **CDD: Multi-view Subspace Clustering via Cross-view Diversity Detection**  
Author(s): Shudong Huang (Sichuan University); Ivor Tsang (University of Technology Sydney); Zenglin Xu (Harbin Institute of Technology); Jiancheng Lv (Sichuan University); Quan-Hui Liu (Sichuan University)\*
- P166 **Learning Spatio-temporal Representation by Channel Aliasing Video Perception**  
Author(s): YiQi Lin (School of Data and Computer Science, Sun Yat-sen University); Jinpeng Wang (Sun Yat-sen University); Manlin Zhang (Sun Yat-sen University); Andy J Ma (Sun Yat-sen University)\*
- P167 **Efficient Sparse Attacks on Videos using Reinforcement Learning**  
Author(s): Huanqian Yan (Beihang University); Xingxing Wei (Beihang University)\*
- P168 **AdvHash: Set-to-set Targeted Attack on Deep Hashing with One Single Adversarial Patch**  
Author(s): Shengshan Hu (Huazhong University of Science and Technology)\*; Yechao Zhang (Huazhong University of Science and Technology); 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)
- P169 **TransRefer3D: Entity-and-Relation Aware Transformer for Fine-Grained 3D Visual Grounding**  
Author(s): Dailan He (Beihang University)\*; Yusheng Zhao (Beihang University); Junyu Luo (Beihang University); Tianrui Hui (Institute of Information Engineering, Chinese Academy of Sciences); Shaofei Huang (Institute of Information Engineering, Chinese Academy of Sciences); aixi zhang (Alibaba Group); Si Liu (Beihang University)
- P170 **Single Image 3D Object Estimation with Primitive Graph Networks**  
Author(s): Qian He (ShanghaiTech University)\*; Desen Zhou (ShanghaiTech University); Bo Wan (ShanghaiTech University); Xuming He (ShanghaiTech University)
- P171 **Boosting Mobile CNN Inference through Semantic Memory**  
Author(s): Yun Li (University of Science and Technology of China)\*; Chen Zhang (Microsoft Research Asia); Shihao Han (Rose-Hulman Institute of Technology); Li Lina Zhang (Microsoft Research); Baoqun YIN (USTC); Yunxin Liu (MSRA); Mengwei Xu (Beijing University of Posts and Telecommunications)
- P172 **Knowing When to Quit: Selective Cascaded Regression with Patch Attention for Real-Time Face Alignment**  
Author(s): Gil Shapira (Samsung)\*; Noga Levy (Samsung); Ishay Goldin (Samsung); Roy J Jevnisek (Samsung)

- 
- P173      **End-to-end Boundary Exploration for Weakly-supervised Semantic Segmentation**
- Author(s): jianjun chen (National Engineering Laboratory for Information Security Technologies, Institute of Information Engineering, Chinese Academy of Sciences)\*; Shancheng Fang (University of Science and Technology of China); Hongtao Xie (University of Science and Technology of China); Zheng-Jun Zha (University of Science and Technology of China); Yue Hu (Institute of Information Engineering, Chinese Academy of Sciences); Jianlong Tan (Institute of Information Engineering, CAS, China)
- 
- P174      **SFE-Net: EEG-based emotion recognition with symmetrical spatial feature extraction**
- Author(s): Xiangwen Deng (University of Electronic Science and Technology of China); Junlin Zhu (University of Electronic Science and Technology of China); Shangming Yang (University of Electronic Science and Technology of China)\*
- 
- P175      **Bridging the Gap between Low-Light Scenes: Bilevel Learning for Fast Adaptation**
- Author(s): Dian Jin (Dalian University of Technology); Long Ma (Dalian University of Technology); Risheng Liu (Dalian University of Technology)\*; Xin Fan (Dalian University of Technology)
- 
- P176      **Handling Difficult Labels for Multi-label Image Classification via Uncertainty Distillation**
- Author(s): Liangchen Song (University at Buffalo) \*; Jialian Wu (State University of New York at Buffalo); Ming Yang (Horizon Robotics); Qian Zhang (Horizon Robotics); Yuan Li (Google); Junsong Yuan ("State University of New York at Buffalo, USA")
- 
- P177      **Perception-Oriented Stereo Image Super-Resolution**
- Author(s): Chenxi Ma (Fudan University); Bo Yan (Fudan University) \*; Weimin Tan (Fudan University); Xuhao Jiang (Fudan University)
- 
- P178      **ReLLIE: Deep Reinforcement Learning for Customized Low-Light Image Enhancement**
- Author(s): Rongkai Zhang (Nanyang Technological University) \*; Lanqing Guo (Nanyang Technological University); Siyu Huang (Nanyang Technological University); Bihan Wen (Nanyang Technological University)
- 
- P179      **Intrinsic Temporal Regularization for High-resolution Human Video Synthesis**
- Author(s): Lingbo Yang (Peking University) \*; Zhanning Gao (Alibaba Group); Siwei Ma (Peking University, China); Wen Gao (PKU)
- 
- P180      **A2W: Context-Aware Recommendation System for Mobile Augmented Reality Web Browser**
- Author(s): Kit Yung Lam (Hong Kong University of Science and Technology) \*; Lik Hang Lee (The University of Oulu); Pan Hui (Hong Kong University of Science and Technology)
- 



- P181 **Cross-modal Self-Supervised Learning for Lip Reading: When Contrastive Learning meets Adversarial Training**  
Author(s): changchong sheng (University of Oulu); Matti Pietikäinen (University of Oulu); Qi Tian (Xidian University); Li Liu (National University of Defense Technology)\*
- P182 **OsGG-Net: One-step Graph Generation Network for Unbiased Head Pose Estimation**  
Author(s): Shentong Mo (Carnegie Mellon University); Miao Xin (Institute of Automation, Chinese Academy of Sciences (CASIA))\*
- P183 **Multi-Modal Multi-Instance Learning for Retinal Disease Recognition**  
Author(s): Xirong Li (Renmin University of China)\*; Yang Zhou (Visionary Intelligence Ltd.); Jie Wang (Renmin University of China); Hailan Lin (Renmin University of China); Jianchun Zhao (Vistel Inc.); Dayong Ding (Vistel Inc.); Weihong Yu (Peking Union Medical College Hospital); Youxin Chen (Peking Union Medical College Hospital)
- P184 **Locally Adaptive Structure and Texture Similarity for Image Quality Assessment**  
Author(s): Keyan Ding (City University of Hongkong)\*; Yi Liu (Huawei Noah's Ark Lab); Xueyi Zou (Huawei Noah's Ark Lab); Shiqi Wang (City University of Hong Kong); Kede Ma (City University of Hong Kong)
- P185 **CALLip: Lipreading using Contrastive and Attribute Learning**  
Author(s): Yiyang Huang (Xidian University); Xuefeng Liang (Xidian University)\*; Chaowei Fang (Xidian University)
- P186 **Cross-Modal Recipe Embeddings by Disentangling Recipe Contents and Dish Styles**  
Author(s): Yu Sugiyama (Univ. Electro-Comm., Tokyo); Keiji Yanai (Univ. Electro-Comm., Tokyo)\*
- P187 **TDI TextSpotter: Taking Data Imbalance into Account in Scene Text Spotting**  
Author(s): Yu Zhou (University of Science and Technology of China)\*; Hongtao Xie (University of Science and Technology of China); Shancheng Fang (University of Science and Technology of China); Jing Wang (Huawei Cloud & AI); Zheng-Jun Zha (University of Science and Technology of China); Yongdong Zhang (University of Science and Technology of China)
- P188 **Position-Augmented Transformers with Entity-Aligned Mesh for TextVQA**  
Author(s): Xuanyu Zhang (Du Xiaoman Financial)\*; Ching Yang (DXM)
- P189 **Learning Contextual Transformer Network for Image Inpainting**  
Author(s): ye deng (Xi'an Jiaotong University) \*; Siqi Hui (Xi'an Jiaotong University); Sanping Zhou (Xi'an Jiaotong University); Deyu Meng (Xi'an Jiaotong University); Jinjun Wang (Xi'an Jiaotong University)
- P190 **Milliseconds Color Stippling**  
Author(s): Lei Ma (Peking University); Jian Shi (Institute of Automation, Chinese Academy of Sciences) \*; Yanyun Chen (Institute of Software, Chinese Academy of Sciences)

- P191 **AFD-Net: Adaptive Fully-Dual Network for Few-Shot Object Detection**  
Author(s): Longyao Liu (BIT); Bo Ma (Beijing Institute of Technology) \*; Yulin Zhang (Beijing Institute of Technology); Xin Yi (Beijing Institute of Technology); Haozhi Li (BIT)
- P192 **Missing Data Imputation for Solar Yield Prediction using Temporal Multi-Modal Variational Auto-Encoder**  
Author(s): Meng Shen (Nanyang Technological University) \*; Huaizheng Zhang (Nanyang Technological University); yixin cao (Nanyang Technological University); Fan Yang (Nanyang Technological University); Yonggang Wen (Nanyang Technological University)
- P193 **Understanding Chinese Video and Language via Contrastive Multimodal Pre-Training**  
Author(s): Chenyi Lei (University of Science and Technology of China, Alibaba Group) \*; Shixian Luo (Alibaba Group); Yong Liu (Nanyang Technological University); wanggui He (Alibaba Group); Jiamang Wang (Alibaba); Guoxin Wang (Alibaba Group); Haihong Tang (Alibaba Group); Chunyan Miao (NTU); Houqiang Li (University of Science and Technology of China)
- P194 **DehazeFlow: Multi-scale Conditional Flow Network for Single Image Dehazing**  
Author(s): HongYu Li (Beihang University)\*; Jia Li (Beihang University); Dong Zhao (Beihang University); Long Xu (National Astronomical Observatories)
- P195 **GCM-Net: Towards Effective Global Context Modeling for Image Inpainting**  
Author(s): Zheng Huan (Hefei University of Technology); Zhao Zhang (Hefei University of Technology)\*; Yang Wang (Hefei University of Technology); Zheng Zhang (Harbin Institute of Technology, Shenzhen); Mingliang Xu (Zhengzhou University); Yi Yang (UTS); Meng Wang (Hefei University of Technology)
- P196 **Embracing the Dark Knowledge: Domain Generalization Using Regularized Knowledge Distillation**  
Author(s): Yufei Wang (Nanyang Technological University)\*; Haoliang Li (CityU); Lap-Pui Chau (Nanyang Technological University); Alex Kot (Nanyang Technological University)
- P197 **Cluster and Scatter: A Multi-grained Active Semi-supervised Learning Framework for Scalable Person Re-identification**  
Author(s): Bingyu Hu (University of Science and Technology of China)\*; Zheng-Jun Zha (University of Science and Technology of China); Jiawei Liu (University of Science and Technology of China); Xierong Zhu (University of Science and Technology of China); Hongtao Xie (University of Science and Technology of China)
- P198 **Dual Graph Convolutional Networks with Transformer and Curriculum Learning for Image Captioning**  
Author(s): Xinzhi Dong (Wuhan University); Chengjiang Long (JD Finance America Corporation); Wenju Xu (University of Kansas); Chunxia Xiao (Wuhan University)\*



P199

**Build Your Own Bundle - A Neural Combinatorial Optimization Method**

Author(s): Qilin Deng (NetEase)\*; Kai Wang (Netease); Minghao Zhao (NetEase); Runze Wu (NetEase Fuxi AI Lab); Yu Ding (Netease Fuxi AI Lab); Zhene Zou (Netease); Yue Shang (Netease Fuxi AI Lab); Jianrong Tao (Netease); Changjie Fan (NetEase Fuxi AI Lab)

P200

**Unsupervised Image Deraining: Optimization Model Driven Deep CNN**

Author(s): Changfeng Yu (Huazhong University of Science and Technology); Yi Chang (PengCheng Laboratory); Yi Li (Huazhong University of Science and Technology); Xi-Le Zhao (School of Mathematical Sciences, University of Electronic Science and Technology of China); Luxin Yan (Huazhong University of Science and Technology)\*

OCT.  
22

9:00 AM-12:00 PM	Reproducibility Session Poster	▼ Hall Way
9:00 AM-12:00 PM	<b>Reproducibility Companion Paper: Campus3D: A Photogrammetry Point Cloud Benchmark for Outdoor Scene Hierarchical Understanding</b> Author(s): Yuqing Liao (Wuhan University); Xinke Li (National University of Singapore)*; Zekun Tong (National University of Singapore); Andrew Lim (National University of Singapore); Yabang Zhao (National University of Singapore); Zhenzhong Kuang (Hangzhou Dianzi University); Cise Midoglu (Simula Research Laboratory (SRL))	
9:00 AM-12:00 PM	<b>Reproducibility Companion Paper: Norm-in-Norm Loss with Faster Convergence and Better Performance for Image Quality Assessment</b> Author(s): Dingquan Li (Peng Cheng Laboratory)*; Tingting Jiang (Peking University); Ming Jiang (Peking University); Vajira Lasantha Thambawita (SimulaMet); Haoliang Wang (Adobe Research)	
9:00 AM-12:00 PM	<b>Kalman Filter-Based Head Motion Prediction for Cloud-Based Mixed Reality</b> Author(s): Serhan Gül (Fraunhofer HHI)*; Sebastian Bosse (Fraunhofer HHI); Dimitri Podborski (Fraunhofer HHI); Thomas Schierl (Fraunhofer HHI); Cornelius Hellge (Fraunhofer HHI); Marc A. Kastner (National Institute of Informatics); Jan Zahálka (Czech Technical University in Prague)	
9:00 AM-12:00 PM	<b>Blind Natural Video Quality Prediction via Statistical Temporal Features and Deep Spatial Features</b> Author(s): Jari Korhonen (Shenzhen University)*; Yicheng Su (Shenzhen University); Junyong You (Norwegian Research Centre); Steven Hicks (SimulaMet); Cise Midoglu (Simula Research Laboratory (SRL))	
9:00 AM-12:00 PM	<b>Describing Subjective Experiment Consistency by p-Value P-P Plot</b> Author(s): Jakub Nawala (AGH University of Science and Technology)*; Lucjan Janowski (AGH University of Science and Technology); Bogdan Ćmiel (); Krzysztof Rusek (AGH University of Science and Technology); Marc A. Kastner (National Institute of Informatics); Jan Zahálka (Czech Technical University in Prague)	
9:00 AM-12:00 PM	<b>Self-supervised Video Representation Learning Using Inter-intra Contrastive Framework</b> Author(s): Li Tao (The University of Tokyo)*; Xueting Wang (CyberAgent, Inc.); Toshihiko Yamasaki (The University of Tokyo); Jingjing Chen (Fudan University); Steven Hicks (SimulaMet)	
9:00 AM-12:00 PM	<b>Visual Relation of Interest Detection</b> Author(s): Fan Yu (Nanjing University); Haonan Wang (Nanjing University); Tongwei Ren (Nanjing University)*; Jinhui Tang (Nanjing University of Science and Technology); Gangshan Wu (Nanjing University); Jingjing Chen (Fudan University); Zhenzhong Kuang (Hangzhou Dianzi University)	



9:00 AM-12:00 PM

**On Learning Disentangled Representation for Acoustic Event Detection**

Author(s): Lijian Gao (Jiangsu University); Qirong Mao (Jiangsu University)\*; Ming Dong (Wayne State University); Jing Yu (Wayne State University); Jingjing Chen (Jiangsu University); Ratna Babu Chinnam (Wayne State University); Miguel Fabian Romero Rondon (Université Côte d'azur); Lucile Sassatelli (Universite Cote d'Azur); Ujjwal Sharma (University of Amsterdam)

---

OCT.  
22

2:00 AM-5:00 AM	Poster Session 4	▼ Hall Way
P201	<p><b>An EM Framework for Online Class Incremental Semantic Segmentation with Dynamic Sampling</b></p> <p>Author(s): Shipeng Yan (ShanghaiTech University) *; Jiale Zhou (Shanghaitech University); Jiangwei Xie (ShanghaiTech University); Songyang Zhang (ShanghaiTech University); Xuming He (ShanghaiTech University)</p>	
P202	<p><b>I2V-GAN: Unpaired Infrared-to-Visible Video Translation</b></p> <p>Author(s): Shuang Li (Beijing Institute of Technology) *; bingfeng han (Beijing Institute of Technology); zhenjie yu (Beijing Institute of Technology); Chi Harold Liu (Beijing Institute of Technology); kai chen (IRay); Shuigen Wang (Yantai IRay Technologies Lt. Co.)</p>	
P203	<p><b>Implicit feedbacks are not always favorable: Iterative Relabeled One-Class Collaborative Filtering against Noisy Interactions</b></p> <p>Author(s): Zitai Wang (SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences; SCS, University of CAS)*; Qianqian Xu (Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences); Zhiyong Yang (SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences; SCS, University of Chinese Academy of Sciences); Xiaochun Cao (Chinese Academy of Sciences); Qingming Huang (University of Chinese Academy of Sciences)</p>	
P204	<p><b>InsPose: Instance-Aware Networks for Single-Stage Multi-Person Pose Estimation</b></p> <p>Author(s): Dahu Shi (Hikvision Research Institute)*; Xing Wei (Xi'an Jiaotong University); xiaodong yu (Hikvision Research Institute); Wenming Tan (Hikvision Research Institute); Ye Ren (Hikvision Research Institute); Shiliang Pu (Hikvision Research Institute)</p>	
P205	<p><b>Implicit Feature Refinement for Instance Segmentation</b></p> <p>Author(s): Lufan Ma (Tsinghua University) *; Tiancai Wang (Megvii Technology); Bin Dong (Megvii Technology); Jiangpeng Yan (Tsinghua University); Xiu Li (Tsinghua University); Xiangyu Zhang (Megvii Technology)</p>	
P206	<p><b>Question-controlled Text-aware Image Captioning</b></p> <p>Author(s): Anwen Hu (Renming University of China) *; Shizhe Chen (INRIA); Qin Jin (Renmin University of China)</p>	
P207	<p><b>Style-Aware Image Recommendation for Social Media Marketing</b></p> <p>Author(s): Yiwei Zhang (The University of Tokyo) *; Toshihiko Yamasaki (The University of Tokyo)</p>	
P208	<p><b>WePerson: Learning a Generalized Re-identification Model from All-weather Virtual Data</b></p> <p>Author(s): He Li (Wuhan University); Mang YE (Wuhan University) *; Bo Du (Wuhan University)</p>	



- P209      **Polar Ray: A Single-stage Angle-free Detector for Oriented Object Detection in Aerial Images**  
Author(s): Shuai Liu (Dalian University of Technology); Lu Zhang (Dalian University of Technology); Shuai Hao (Dalian University of Technology); Huchuan Lu (Dalian University of Technology)\*; You He (Naval Aviation University)
- 
- P210      **Self-Contrastive Learning with Hard Negative Sampling for Self-supervised Point Cloud Learning**  
Author(s): Bi'an Du (Peking University); Xiang Gao (Peking University); Wei Hu (Peking University)\*; Xin Li (West Virginia University)
- 
- P211      **Generally Boosting Few-Shot Learning with HandCrafted Features**  
Author(s): Yi Zhang (Chongqing University); Sheng Huang (Chongqing University)\*; Fengtao Zhou (Chongqing University)
- 
- P212      **ROECS: A Robust Semi-direct Pipeline Towards Online Extrinsic Correction of the Surround-view System**  
Author(s): Tianjun Zhang (Tongji University); Nlong Zhao (University of Southern California); Ying Shen (Tongji University); Xuan Shao (Tongji University); Lin Zhang (Tongji University, China)\*; Yicong Zhou (University of Macau)
- 
- P213      **Pseudo Graph Convolutional Network for Vehicle ReID**  
Author(s): Wen Qian (Institute of Automation, Chinese Academy of Sciences)\*; zhiqun He (SenseTime Group Limited); Silong Peng (The Chinese academy of science); Chen Chen (The Chinese academy of science); Wei Wu (SenseTime Group Limited)
- 
- P214      **Towards Fast and High-Quality Sign Language Production**  
Author(s): Wencan Huang (Zhejiang University)\*; Wenwen Pan (Zhejiang University); Zhou Zhao (Zhejiang University); Qi Tian (Huawei Cloud & AI)
- 
- P215      **Effective De-identification Generative Adversarial Network for Face Anonymization**  
Author(s): Zhenzhong Kuang (Hangzhou Dianzi University & Osaka University)\*; Huigui Liu (Hangzhou Dianzi University); Jun Yu (HDU); Aikui Tian (Shandong University of Technology); Lei Wang (Shandong University of Technology); Jianping Fan (UNCC); Noboru Babaguchi (Osaka University, Japan)
- 
- P216      **Cross-modal Retrieval and Synthesis (X-MRS): Closing the modality gap in shared subspace learning**  
Author(s): Ricardo Guerrero (Samsung AI Center Cambridge)\*; Hai X Pham (Samsung AI Center); Vladimir Pavlovic (Rutgers University)
- 
- P217      **When Face Completion Meets Irregular Holes: An Attributes Guided Deep Inpainting Network**  
Author(s): Jie Xiao (Sun Yat-sen University); Dandan Zhan (Sun Yat-sen University); Haoran Qi (Sun Yat-sen University); Zhi Jin (Sun Yat-sen University)\*
-

- P218      **Non-Linear Fusion for Self-Paced Multi-View Clustering**  
 Author(s): Zongmo Huang (University of Electronic Science and Technology of China); Yazhou Ren (University of Electronic Science and Technology of China)\*; Xiaorong Pu (University of Electronic Science and Technology of China); Lifang He (Lehigh University)
- 
- P219      **Counterfactual Debiasing Inference for Compositional Action Recognition**  
 Author(s): Pengzhan Sun (University of Electronic Science and Technology of China)\*; Bo Wu (MIT-IBM Watson AI Lab); Xunsong Li (University of Electronic Science and Technology of China); Wen Li (University of Electronic Science and Technology of China); Lixin Duan (University of Electronic Science and Technology of China); Chuang Gan (MIT-IBM Watson AI Lab)
- 
- P220      **STST: Spatial-Temporal Specialized Transformer for Skeleton-based Action Recognition**  
 Author(s): Yuhan Zhang (University of Electronic Science and Technology of China)\*; Bo Wu (MIT-IBM Watson AI Lab); Wen Li (University of Electronic Science and Technology of China); Lixin Duan (University of Electronic Science and Technology of China); Chuang Gan (MIT-IBM Watson AI Lab)
- 
- P221      **Exploring Gradient Flow Based Saliency for DNN Model Compression**  
 Author(s): Xinyu Liu (City University of Hong Kong); Baopu Li (BAIDU USA LLC)\*; Zhen Chen (City University of Hong Kong); Yixuan YUAN (City University of Hong Kong)
- 
- P222      **An adaptive iterative inpainting method with more information exploration**  
 Author(s): Shengjie Chen (Tsinghua University)\*; Zhenhua Guo (Graduate School at Shenzhen, Tsinghua University, China); Bo Yuan (Tsinghua University)
- 
- P223      **Assisting News Media Editors with Cohesive Visual Storylines**  
 Author(s): Gonçalo Marcelino (Univer); David Semedo (Universidade NOVA de Lisboa); Andre Mourao (Universidade NOVA Lisboa); Saverio Blasi (BBC); Joao Magalhaes (Universidade NOVA Lisboa)\*; Marta Mrak (BBC)
- 
- P224      **MM-Flow: Multi-modal Flow Network for Point Cloud Completion**  
 Author(s): Yiqiang Zhao (Tianjin University); Yiyao Zhou (Tianjin University)\*; Rui Chen (Tianjin University); Bin Hu (Tianjin university); Arxidn Akbar (Tianjin University)
- 
- P225      **Long-tailed Distribution Adaptation**  
 Author(s): Zhiliang Peng (University of Chinese Academy of Sciences); Wei Huang (University of Chinese Academy of Sciences); Zonghao Guo (University of Chinese Academy of Sciences); Xiaosong Zhang (University of Chinese Academy of Sciences); Jianbin Jiao (University of Chinese Academy of Sciences); Qixiang Ye (University of Chinese Academy of Sciences, China)\*
- 
- P226      **Lesion-Inspired Denoising Network: Connecting Medical Image Denoising and Lesion Detection**  
 Author(s): Kecheng Chen (University of Electronic Science And Technology Of China)\*; Kun Long (University of Electronic Science and Technology of China); 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)



- P227      **Domain Adaptive Semantic Segmentation without Source Data**  
Author(s): Fuming You (University of Electronic Science and Technology of China); Jingjing Li (University of Electronic Science and Technology of China)\*; Lei Zhu (Shandong Normal University); Zhi Chen (The University of Queensland); Zi Huang (University of Queensland)
- 
- P228      **Cross-modal Joint Prediction and Alignment for Composed Query Image Retrieval**  
Author(s): Yuchen Yang (University of Science and Technology of China)\*; Min Wang (Hefei Comprehensive National Science Center Institute of Artificial Intelligence); Wengang Zhou (University of Science and Technology of China); Houqiang Li (University of Science and Technology of China)
- 
- P229      **JDMAN: Joint Discriminative and Mutual Adaptation Networks for Cross-Domain Facial Expression Recognition**  
Author(s): Yingjian Li (Harbin Institute of Technology (Shenzhen), Shenzhen); Yingnan Gao (Harbin Institute of Technology, Shenzhen); Bingzhi Chen (Harbin Institute of Technology, Shenzhen); Zheng Zhang (Harbin Institute of Technology, Shenzhen); Lei Zhu (Shandong Normal University); Guangming Lu (Harbin Institute of Technology, Shenzhen)\*
- 
- P230      **Improving Weakly Supervised Object Localization via Causal Intervention**  
Author(s): Feifei Shao (Zhejiang University); Yawei Luo (Zhejiang University)\*; Li Zhang (Zhejiang Insigma digital technology CO.LTD); Lu Ye (Zhejiang University of Science and Technology); Siliang Tang (Zhejiang University); Yi Yang (University of Technology, Sydney); Jun Xiao (Zhejiang University)
- 
- P231      **Imbalanced Source-free Domain Adaptation**  
Author(s): Li Xinhao (UESTC); 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)
- 
- P232      **Learning Transferrable and Interpretable Representations for Domain Generalization**  
Author(s): Zhekai Du (University of Electronic Science and Technology of China); Jingjing Li (University of Electronic Science and Technology of China) \*; Ke Lu (University of Electronic Science and Technology of China); Lei Zhu (Shandong Normal University); Zi Huang (University of Queensland)
- 
- P233      **WAS-VTON: Warping Architecture Search for Virtual Try-on Network**  
Author(s): Zhenyu Xie (Sun Yat-sen University); Xujie Zhang (Sun Yat-sen University); Fuwei Zhao (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 (Sun Yat-sen University)\*
- 
- P234      **DFR-Net: A Novel Multi-Task Learning Network for Real-Time Multi-Instrument Segmentation**  
Author(s): Yan-Jie Zhou (Institute of Automation, Chinese Academy of Sciences)\*; Shiqi Liu (Institute of Automation, Chinese Academy of Sciences); Xiaoliang Xie (Institute of Automation, Chinese Academy of Sciences); Zengguang Hou (Chinese Academy of Sciences)

- P235 **From Superficial to Deep: Language Bias driven Curriculum Learning for Visual Question Answering**  
Author(s): Mingrui Lao (Leiden University); Yanming Guo (National University of Defense Technology) \*; Yu Liu (Dalian University of Technology); Wei Chen (Leiden University); Nan Pu (Leiden University); Michael S Lew (Leiden University)
- 
- P236 **Pairwise Emotional Relationship Recognition in Drama Videos: Dataset and Benchmark**  
Author(s): Xun Gao (Alibaba Group); Yin Zhao (Alibaba Group)\*; Jie Zhang (Alibaba Group); LongJun Cai (Alibaba)
- 
- P237 **Block Popularity Prediction for Multimedia Storage Systems Using Spatial-Temporal-Sequential Neural Networks**  
Author(s): Yingying Cheng (Huawei)\*; Fan Zhang (Huawei); Gang Hu (Huawei); Yiwen Wang (Hong Kong University of Science and Technology); Hanhui Yang (Huawei); Gong Zhang (Huawei); Zhuo Cheng (Huawei Storage Product Line)
- 
- P238 **Transferrable Contrastive Learning for Visual Domain Adaptation**  
Author(s): Yang Chen (University of Science and Technology of China); Yingwei Pan (JD AI Research) \*; Yu Wang (JD AI Research); Ting Yao (JD AI Research); Xinmei Tian (USTC); Tao Mei (AI Research of JD.com)
- 
- P239 **Weighted Gaussian Loss based Hamming Hashing**  
Author(s): Rong-Cheng Tu (Beijing Institute of Technology)\*; Xian-Ling Mao (Beijing Institute of Technology); Cihang Kong (Beijing Institute of Technology); Zihang Shao (Beijing Institute of Technology); Ze-Lin Li (Beijing Institute of technology); Wei Wei (Huazhong University of Science and Technology); Heyan Huang (Beijing Engineering Research Center of High Volume Language Information Processing and Cloud Computing Applications, Department of Computer Science and Technology, Beijing Institute of technology)
- 
- P240 **Domain-Aware SE Network for Sketch-based Image Retrieval with Multiplicative Euclidean Margin Softmax**  
Author(s): Peng Lu (Tsinghua University); Gao Huang (Tsinghua); Hangyu Lin (Fudan University); Wenming Yang (Tsinghua University)\*; Guodong Guo (IDL, Baidu Research); Yanwei Fu (Fudan University)
- 
- P241 **FTAface: Context-enhanced Face Detector with Fine-grained Task Attention**  
Author(s): Deyu Wang (Canon Innovative Solution (Beijing) Co., Ltd.)\*; Dongchao Wen (Canon Innovative Solution (Beijing) Co., Ltd.); Wei Tao (Canon Innovative Solution (Beijing) Co., LTD); Lingxiao Yin (Canon Innovative Solution (Beijing) Co., Ltd.); Tse-Wei Chen (Canon Inc.); Tadayuki Ito (Canon Inc.); Kinya Osa (Canon Inc.); Masami Kato (Canon Inc.)
- 
- P242 **Identity-aware Graph Memory Network for Action Detection**  
Author(s): JingCheng Ni (Behang University)\*; Jie Qin (Inception Institute of Artificial Intelligence); Di Huang (Beihang University, China)
- 



- P243 **Improving Robustness and Accuracy via Relative Information Encoding in 3D Human Pose Estimation**  
Author(s): Wenkang Shan (Peking University)\*; Haopeng Lu (Shanghai Jiao Tong University); Shanshe Wang (Peking University); xinfeng zhang (University of Chinese Academy of Sciences); Wen Gao (PKU)
- 
- P244 **Deep Neural Network Retrieval**  
Author(s): Nan Zhong (Fudan University); Zhenxing Qian (School of Computer Science, Fudan University); Xinpeng Zhang (School of Computer Science, Fudan University)\*
- 
- P245 **Adversarial Learning with Mask Reconstruction for Text-Guided Image Inpainting**  
Author(s): Xingcai Wu (Guangdong University of Technology)\*; yucheng xie (gdut); Jiaqi Zeng (Guangdong University of Technology); Zhenguo Yang (Guangdong University of Technology); Yi Yu (NII); Qing Li (The Hong Kong Polytechnic University); Wenyin Liu (Guangdong University of Technology)
- 
- P246 **Spatiotemporal Inconsistency Learning for DeepFake Video Detection**  
Author(s): Zhihao Gu (Shanghai Jiao Tong University)\*; Yang Chen (Tencent); Taiping Yao (Tencent YouTu); Shouhong Ding (Tencent); Jilin Li (Tencent); Feiyue Huang (Tencent); Lizhuang Ma (Shanghai Jiao Tong University)
- 
- P247 **VeloCity: Using Voice Assistants for Cyclists to Provide Traffic Reports**  
Author(s): Gian-Luca Savino (University of Bremen)\*; Jesse Moraes Braga (University of Bremen); Johannes Schoening (University of Bremen)
- 
- P248 **Edit Like A Designer: Modeling Design Workflows for Unaligned Fashion Editing**  
Author(s): Qiyu Dai (Peking University); Shuai Yang (Peking University); Wenjing Wang (Peking University); Wei Xiang (BIGO Ltd.); Jiaying Liu (Peking University)\*
- 
- P249 **Privacy-Preserving Portrait Matting**  
Author(s): Jizhizi Li (The University of Sydney); Sihan Ma (The University of Sydney); Jing Zhang (The University of Sydney)\*; Dacheng Tao (JD.com)
- 
- P250 **A Transformer based Approach for Image Manipulation Chain Detection**  
Author(s): Jiayang You (Shenzhen University); Yuanman Li (Shenzhen University)\*; Jiantao Zhou (University of Macau); Zhongyun Hua (Harbin Institute of Technology, Shenzhen); Weiwei Sun (Alibaba Group); Xia Li (Shenzhen University)
- 
- P251 **HANet: Hierarchical Alignment Networks for Video-Text Retrieval**  
Author(s): Peng Wu (Xidian University)\*; Xiangteng He (Alibaba Group); Mingqian Tang (Alibaba Group); Yiliang Lv (Alibaba Group); Jing Liu (Xidian University)
-

- P252      **Scalable Multi-view Subspace Clustering with Unified Anchors**  
 Author(s): Mengjing Sun (National University of Defense Technology); Pei Zhang (National University of Defense Technology)\*; Siwei Wang (National University of Defense Technology); Sihang Zhou (NUDT); Wenxuan Tu (NUDT); Xinwang Liu (National University of Defense Technology); En Zhu (National University of Defense Technology); Changjian Wang (National University Of Defense Technology)
- 
- P253      **PRNet: A Progressive Recovery Network for Revealing Perceptually Encrypted Images**  
 Author(s): Tao Xiang (Chongqing University); Ying Yang (Chongqing University); Shangwei Guo (Chongqing University)\*; Hangcheng Liu (Chongqing University); Hantao LIU (Cardiff University)
- 
- P254      **FakeTagger: Robust Safeguards against DeepFake Dissemination via Provenance Tracking**  
 Author(s): Run Wang (Wuhan University)\*; Felix Juefei-Xu (Alibaba Group, USA); Meng Luo (Northeastern University); Yang Liu (Nanyang Technology University, Singapore); Lina Wang (Wuhan University)
- 
- P255      **Discriminative Latent Semantic Graph for Video Captioning**  
 Author(s): Yang Bai (Newcastle University)\*; Junyan Wang (UNSW Sydney); Yang Long (Durham University); Bingzhang Hu (Newcastle University); Yang Song (University of New South Wales); Maurice Pagnucco (UNSW); Yu Guan (Newcastle University)
- 
- P256      **From Image to Imuge: Immunized Image Generation**  
 Author(s): Qichao Ying (School of Computer Science, Fudan University); Zhenxing Qian (School of Computer Science, Fudan University)\*; Hang Zhou (Simon Fraser University); Haisheng Xu (Nvidia); Xinpeng Zhang (School of Computer Science, Fudan University); Siyi Li (Nvidia)
- 
- P257      **Wisdom of (Binned) Crowds: A Bayesian Stratification Paradigm for Crowd Counting**  
 Author(s): Sravya V Shivapuja (IIIT Hyderabad); Mansi P Khamkar (IIIT Hyderabad); Divij Bajaj (IIIT Hyderabad); Ganesh Ramakrishnan (IIT Bombay); Ravi Kiran Sarvadevabhatla (IIIT Hyderabad)\*
- 
- P258      **Demystifying Commercial Video Conferencing Applications**  
 Author(s): Insoo Lee (University of Colorado Boulder)\*; Jinsung Lee (University of Colorado Boulder); Kyunghan Lee (Seoul National University); Dirk Grunwald (university of colorado at boulder); Sangtae Ha (University of Colorado at Boulder)
- 
- P259      **LightFEC: Network Adaptive FEC with A Lightweight Deep-Learning Approach**  
 Author(s): Han Hu (Peking University)\*; Sheng Cheng (Peking University); Xinggong Zhang (PKU); Zongming Guo (Peking University)
- 



P260

**SOGAN: 3D-Aware Shadow and Occlusion Robust GAN for Makeup Transfer**

Author(s): Yueming Lyu (CASIA); Jing Dong (Chinese Academy of Sciences)\*; Bo Peng (Institute of Automation, Chinese Academy of Sciences); Wei Wang (Center for Research on Intelligent Perception and Computing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences); Tieniu Tan (NLPR, China)

---

OCT.  
22

October 23, 2021

10:30 AM-12:00 PM	Session 19: Media Interpretation-I	▼ Langbo Room A
<b>Session Chair: Xinchun Liu, JD AI Research</b>		
10:30 AM-10:45 AM	<b>Mix-order Attention Networks for Image Restoration</b> Author(s): Tao Dai (Tsinghua University); Yalei Lv (Tsinghua University); Bin Chen (Tsinghua University)*; Zhi Wang (Tsinghua University); Zexuan Zhu (Shenzhen University); Shu-Tao Xia (Tsinghua University)	
10:45 AM-11:00 AM	<b>Vehicle Counting Network with Attention-based Mask Refinement and Spatial-awareness Block Loss</b> Author(s): Ji Zhang (Southwest Jiaotong University); Jian-Jun Qiao (Southwest Jiaotong University); Xiao Wu (Southwest Jiaotong University)*; Wei Li (Southwest Jiaotong University)	
11:00 AM-11:15 AM	<b>DPT: Deformable Patch-based Transformer for Visual Recognition</b> Author(s): Zhiyang Chen (Institute of Automation, Chinese Academy of Sciences)*; Yousong Zhu (Institute of Automation, Chinese Academy of Sciences); Chaoyang Zhao (National Laboratory of Pattern Recognition, CASIA); Guosheng Hu (AnyVision); Wei Zeng (Peking University, China); Jinqiao Wang (Institute of Automation, Chinese Academy of Sciences); Ming Tang (Institute of Automation, Chinese Academy of Sciences)	
11:15 AM-11:30 AM	<b>Scene Text Image Super-Resolution via Parallely Contextual Attention Network</b> Author(s): cairong zhao (Tongji University) *; shuyang feng (tongji university); Brain Nlong Zhao (Viterbi School of Engineering University of Southern California); Zhijun Ding (Tongji University); Jun Wu (Fudan University); Fumin Shen (UESTC); Heng Tao Shen (University of Electronic Science and Technology of China (UESTC))	
11:30 AM-11:45 AM	<b>Improving Pedestrian Detection from a Long-tailed Domain Perspective</b> Author(s): Mengyuan Ding (Nanjing University of Science and Technology) *; Shanshan Zhang (Nanjing University of Science and Technology); Jian Yang (Nanjing University of Science and Technology)	
11:45 AM-12:00 PM	<b>Robust Shadow Detection by Exploring Effective Shadow Contexts</b> Author(s): Xianyong Fang (Anhui University)*; Xiaohao He (Anhui University); Linbo Wang (Anhui University); Jianbing Shen (Inception Institute of Artificial Intelligence)	

 OCT.  
 23


10:30 AM-12:00 PM

Session 20: Media Interpretation-II

▼ Langbo Room B

**Session Chair: Xiaoshuang Shi, University of Electronic Science and Technology of China**

10:30 AM-10:45 AM

**WeClick: Weakly-Supervised Video Semantic Segmentation with Click Annotations**

Author(s): Peidong Liu (Tsinghua University)\*; Zibin He (Tsinghua University); Xiyu Yan (Tsinghua University); Yong Jiang (Tsinghua University); Shu-Tao Xia (Tsinghua University); Feng Zheng (SUSTech); Maowei Hu (THU)

10:45 AM-11:00 AM

**Towards Cross-Granularity Few-Shot Learning: Coarse-to-Fine Pseudo-Labeling with Visual-Semantic Meta-Embedding**

Author(s): Jinhai Yang (Shanghai Jiao Tong University); Hua Yang (Shanghai Jiao Tong University)\*; Lin Chen (Shanghai Jiao Tong University)

11:00 AM-11:15 AM

**Disentangled Representation Learning and Enhancement Network for Single Image De-Raining**

Author(s): Guoqing Wang (University of Electronic Science and Technology of China); Changming Sun (CSIRO Data61); Xing Xu (University of Electronic Science and Technology of China); Jingjing Li (University of Electronic Science and Technology of China); Zheng Wang (UESTC)\*; Zeyu Ma (UESTC)

11:15 AM-11:30 AM

**Towards Robust Cross-Domain Image Understanding with Unsupervised Noise Removal**

Author(s): Lei Zhu (NUS Graduate School Integrative Sciences and Engineering Programme); Zhaojing Luo (National University of Singapore) \*; Wei Wang (National University of Singapore); Meihui Zhang (Beijing Institute of Technology); Gang Chen (Zhejiang University); Kaiping Zheng (National University of Singapore)

11:30 AM-11:45 AM

**Exploiting BERT For Multimodal Target Sentiment Classification Through Input Space Translation**

Author(s): Zaid Khan (Northeastern University) \*; YUN FU (Northeastern University)

11:45 AM-12:00 PM

**Video Representation Learning with Graph Contrastive Augmentation**

Author(s): Jingran Zhang (University of Electronic Science and Technology of China) \*; Xing Xu (University of Electronic Science and Technology of China); Fumin Shen (UESTC); Yazhou Yao (Nanjing University of Science and Technology); Jie Shao (University of Electronic Science and Technology of China); Xiaofeng Zhu (Guangxi Normal University)

OCT.  
23

10:30 AM-12:00 PM

Session 21: Open Source Competition

▼ Tianfu Room A

Session Chair: Qi Dai, Microsoft Research Asia

10:30 AM-10:37 AM

**MMFashion: An Open-Source Toolbox for Visual Fashion Analysis**

Author(s): Xin Liu (Duke University); Jiancheng Li (Tsinghua University); Jiaqi Wang (The Chinese University of Hong Kong); Ziwei Liu (Nanyang Technological University)\*

10:37 AM-10:44 AM

**Efficient Reinforcement Learning Development with RLzoo**

Author(s): Zihan Ding (Princeton University) \*; Tianyang Yu (Nanchang University); Hongming Zhang (PKU); Yanhua Huang (Xiaohongshu); Guo Li (Imperial College London); Quancheng Guo (The University of Edinburgh); Luo Mai (University of Edinburgh); Hao Dong (Peking University)

10:44 AM-10:51 AM

**Fast and Flexible Human Pose Estimation with HyperPose**

Author(s): Yixiao Guo (Peking University) \*; Jiawei Liu (Tongji University); Guo Li (Imperial College London); Luo Mai (The University of Edinburgh); Hao Dong (Peking University)

10:51 AM-10:58 AM

**SmartEye: An Open Source Framework for Real-Time Video Analytics with Edge-Cloud Collaboration**

Author(s): Xuezhi Wang (Nanjing University of Science and Technology); Guanyu Gao (Nanjing University of Science and Technology) \*

10:58 AM-11:05 AM

**ZoomSense: A Scalable Infrastructure for Augmenting Zoom**

Author(s): Tom Bartindale (Monash University) \*; Xiuzhi Chen (Monash University); Harrison Marshall (Monash University); Stanislav Pozdniakov (Monash University); Dan Richardson (Monash University)

11:05 AM-11:12 AM

**Efficient Graph Deep Learning in TensorFlow with tf\_geometric**

Author(s): Jun Hu (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences) \*; Shengsheng Qian (institute of automation chinese academy of sciences); Quan Fang (Institute of Automation Chinese Academy of Sciences); Youze Wang (Hefei University of Technology); Quan Zhao (Hefei University of Technology); Huaiwen Zhang (Inner Mongolia University); Changsheng Xu (CASIA)

11:12 AM-11:19 AM

**FaceX-Zoo: A PyTorch Toolbox for Face Recognition**

Author(s): Jun Wang (JD AI Research) \*; Yinglu Liu (JD AI Research); Yibo Hu (JD AI Research); Hailin Shi (JD AI Research); Tao Mei (JD.com)

OCT.  
23

- 11:19 AM-11:26 AM **PyTorchVideo: A Deep Learning Library for Video Understanding**  
Author(s): Haoqi Fan (Facebook AI Research)\*; Tullie Murrell (Facebook AI Research); Heng Wang (Facebook AI); Kalyan Vasudev Alwala (Facebook); Yanghao Li (Facebook AI Research); Yilei Li (Facebook Inc); Bo Xiong (Facebook AI Research); Nikhila Ravi (Facebook AI Research); Meng Li (Facebook Inc); Haichuan Yang (Facebook); Jitendra Malik (University of California at Berkeley); Ross Girshick (FAIR); Matt Feiszli (Facebook Research); Aaron Adcock (Facebook AI Research); Wan-Yen Lo (Facebook AI Research); Christoph Feichtenhofer (Facebook AI Research)
- 
- 11:26 AM-11:33 AM **AICoacher: A System Framework for Online Realtime Workout Coach**  
Author(s): Haocong Ying (Capital Normal University, China); Tie Liu (Capital Normal University)\*; Mingxin Ai (Capital Normal University, China); Jiali Ding (Capital Normal University, China); Yuanyuan Shang (Capital Normal University)
- 
- 11:33 AM-11:40 AM **MMOCR: A Comprehensive Toolbox for Text Detection, Recognition and Understanding**  
Author(s): Zhanghui Kuang (Sensetime Ltd.)\*; Hongbin Sun (Sensetime Ltd.); Zhizhong Li (Sensetime); Xiaoyu Yue (CPII); Tsui Hin Lin (Sensetime); Jianyong Chen (South China University of Technology); Huaqiang Wei (Sensetime); Yiqin Zhu (South China University of Technology); Tong Gao (Sensetime); Wenwei Zhang (Nanyang Technological University); Kai Chen (SenseTime Research); Wayne Zhang (SenseTime Research); Dahua Lin (The Chinese University of Hong Kong)
- 
- 11:40 AM-11:47 AM **A Complete End To End Open Source Toolchain for the Versatile Video Coding (VVC) Standard**  
Author(s): Adam Wieckowski (HHI)\*; Christian Lehmann (HHI); Benjamin Bross (HHI); Detlev Marpe (HHI); Thibaud Biatek (ATEME); mikael raulet (Ateme); Jean Le Feuvre (Telecom Paris)
- 
- 11:47 AM-11:54 AM **X-modaler: A Versatile and High-performance Codebase for Cross-modal Analytics**  
Author(s): Yehao Li (Sun Yat-Sen University); Yingwei Pan (JD AI Research)\*; Jingwen Chen (Sun Yat-set University); Ting Yao (JD AI Research); Tao Mei (AI Research of JD.com)
- 
- 11:54 AM-12:01 PM **Interpreting Super-Resolution CNNs for Sub-Pixel Motion Compensation in Video Coding**  
Author(s): Luka Murn (British Broadcasting Corporation)\*; Alan Smeaton (Insight Centre for Data Analytics, Dublin City University); Marta Mrak (BBC)
-

10:30 AM-12:00 PM	Session 22: Multimedia Search and Recommendation-I	▼ Tianfu Room B
<b>Session Chair: Shuhui Wang, Institute of Computing Technology, CAS</b>		
10:30 AM-10:45 AM	<b>Towards Accurate Localization by Instance Search</b> Author(s): Yi-Geng Hong (Xiamen University); Hui-Chu Xiao (Xiamen University); Wan-Lei Zhao (Xiamen University)*	
10:45 AM-11:00 AM	<b>Database-adaptive Re-ranking for Enhancing Cross-modal Image Retrieval</b> Author(s): Rintaro Yanagi (Hokkaido University)*; Ren Togo (Hokkaido University); Takahiro Ogawa (Hokkaido University); Miki Haseyama (Hokkaido University)	
11:00 AM-11:15 AM	<b>Fine-grained Cross-modal Alignment Network for Text-Video Retrieval</b> Author(s): Ning Han (Hunan University); Jingjing Chen (Fudan University); Guangyi Xiao (College of Computer Science and Electronic Engineering, Hunan University, China); Hao Zhang (City University of Hong Kong); Yawen Zeng (Tencent Inc.); Hao Chen (Hunan University)*	
11:15 AM-11:30 AM	<b>Meta Self-Paced Learning for Cross-Modal Matching</b> Author(s): Jiwei Wei (University of Electronic Science and Technology of China); Xing Xu (University of Electronic Science and Technology of China)*; Zheng Wang (UESTC); Guoqing Wang (University of Electronic Science and Technology of China)	
11:30 AM-11:45 AM	<b>CausalRec: Causal Inference for Visual Debiasing in Visually-Aware Recommendation</b> Author(s): Ruihong Qiu (The University of Queensland)*; Sen Wang (The University of Queensland); Zhi Chen (The University of Queensland); Hongzhi Yin (The University of Queensland); Zi Helen Huang (University of Queensland)	
11:45 AM-12:00 PM	<b>Semi-supervised Domain Adaptive Retrieval via Discriminative Hashing Learning</b> Author(s): Haifeng Xia (Tulane University); Taotao Jing (Tulane University); Chen Chen (University of Central Florida); Zhengming Ding (Tulane University)*	



10:30 AM-12:00 PM

Session 23: Multimedia Search and Recommendation-II

▼ Jiaozi Room

Session Chair: Hao Liu, Ningxia University

10:30 AM-10:45 AM

**Hierarchical View Predictor: Unsupervised 3D Global Feature Learning through Hierarchical Prediction among Unordered Views**

Author(s): Zhizhong Han (Wayne State University); Xiyang Wang (Tsinghua University); Yu-Shen Liu (Tsinghua University)\*; Matthias Zwicker (University of Maryland)

10:45 AM-11:00 AM

**Mining Latent Structures for Multimedia Recommendation**

Author(s): Jinghao Zhang (CRIPAC, NLPR, CASIA); Yanqiao ZHU (Institute of Automation, Chinese Academy of Sciences)\*; Qiang Liu (Institute of Automation, Chinese Academy of Sciences); Shu Wu (NLPR, China); Shuhui Wang (VIPL,ICT,Chinese academic of science); Liang Wang (NLPR, China)

11:00 AM-11:15 AM

**Why Do We Click: Visual Impression-aware News Recommendation**

Author(s): Jiahao Xun (Zhejiang University) \*; Shengyu Zhang (Zhejiang University); Zhou Zhao (Zhejiang University); Jieming Zhu (Huawei Noah's Ark Lab); Qi Zhang (Huawei Noah's Ark Lab); Jingjie Li (Huawei Noah's Ark Lab); Xiuqiang He (Huawei Noah's Ark Lab); Xiaofei He (Zhejiang University); Tat-Seng Chua (National University of Singapore); Fei Wu (Zhejiang University, China)

11:15 AM-11:30 AM

**Identity-Preserving Face Anonymization via Adaptively Facial Attributes Obfuscation**

Author(s): Jingzhi Li (Institute of information engineering, CAS)\*; lutong han (Institute of Information Engineering,cas); Ruoyu Chen (Institute of Information Engineering, Chinese Academy of Sciences); Hua Zhang (Institute of Information Engineering,Chinese Academy of Sciences); Bing Han (University of Chinese Academy of Sciences); Lili Wang (State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering,Beihang University); Xiaochun Cao (Chinese Academy of Sciences)

11:30 AM-11:45 AM

**CONQUER: Contextual Query-aware Ranking for Video Corpus Moment Retrieval**

Author(s): Zhijian Hou (City University of Hong Kong) \*; Chong-Wah Ngo (Singapore Management University); W. K. Chan (City University of Hong Kong)

11:45 AM-12:00 PM

**Learning Unified Embeddings for Recommendation via Meta-path Semantics**

Author(s): Qianxiu Hao (University of Chinese Academy of Science) \*; Qianqian Xu (Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences); Zhiyong Yang (SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences); SCS, University of Chinese Academy of Sciences); Qingming Huang (University of Chinese Academy of Sciences)

OCT.  
23

10:30 AM-12:00 PM	Session 24: Music, Speech and Audio Processing in Multimedia	▼ Jincheng Room
<b>Session Chair: Qin Jin, Renmin University of China</b>		
10:30 AM-10:45 AM	<b>ReconVAT: A Semi-Supervised Automatic Music Transcription Framework for Low-Resource Real-World Data</b> Author(s): Kin Wai Cheuk (Singapore University of Technology and Design) *; Dorien Herremans (Singapore University of Technology and Design); Li Su (Academia Sinica)	
10:45 AM-11:00 AM	<b>Is Someone Speaking? Exploring Long-term Temporal Features for Audio-visual Active Speaker Detection</b> Author(s): Ruijie Tao (National University of Singapore)*; Zexu Pan (National University of Singapore); Rohan Kumar Das (Fortemedia); xinyuan qian (National University of Singapore); Mike Zheng Shou (National University of Singapore); Haizhou Li (National University of Singapore)	
11:00 AM-11:15 AM	<b>Actions speak louder than listening: evaluating music style transfer based on editing experience</b> Author(s): Wei-Tsung Lu (New Your University); Meng-Hsuan Wu (NTU); Yuh-Ming Chiu (KKLab); Li Su (Academia Sinica)*	
11:15 AM-11:30 AM	<b>FMSing: Fast Multi-Singer Vocoder with A large-Scale Singing Voice Corpus</b> Author(s): Rongjie Huang (Zhejiang University)*; Feiyang Chen (Zhejiang University); Yi Ren (Zhejiang University); Jinglin Liu (Zhejiang University); Chenye Cui (Zhejiang University); Zhou Zhao (Zhejiang University)	
11:30 AM-11:45 AM	<b>MusicBERT: A Self-supervised Learning of Music Representation</b> Author(s): Hongyuan Zhu (Alibaba)*; Ye Niu (Alibaba); Di Fu (Alibaba); Hao Wang (Alibaba Group)	
11:45 AM-12:00 PM	<b>UniCon: Unified Context Network for Robust Active Speaker Detection</b> Author(s): Yuanhang Zhang (Institute of Computing Technology, Chinese Academy of Sciences)*; Susan Liang (University of Chinese Academy of Sciences); Shuang Yang (ICT, CAS); Xiao Liu (Tomorrow Advancing Life); Zhongqin Wu (Tomorrow Advancing Life); Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences); Xilin Chen (Institute of Computing Technology, Chinese Academy of Sciences)	



10:30 AM-12:00 PM

Session 25: Multimedia Transport and Delivery

▼ Rongcheng Room

Session Chair: Jie Shao, University of Electronic Science and Technology of China

10:30 AM-10:45 AM

**AITransfer: Progressive AI-powered Transmission for Real-Time Point Cloud Video Streaming**

Author(s): Yakun Huang (Beijing University of Posts and Telecommunications)\*; Yuanwei Zhu (Beijing University of Posts and Telecommunications); Xiuquan Qiao (Beijing University of Posts and Telecommunications); Zhijie Tan (Beijing University of Posts and Telecommunications); Boyuan Bai (University of Birmingham)

10:45 AM-11:00 AM

**Game Theory-driven Rate Control for 360-Degree Video Coding**

Author(s): Tiesong Zhao (Fuzhou University)\*; Jielian Lin (Fuzhou University); Yanjie Song (Fuzhou University); Xu Wang (Shenzhen University); Yuzhen Niu (Fuzhou University)

11:00 AM-11:15 AM

**TBRA: Tiling and Bitrate Adaptation for Mobile 360-Degree Video Streaming**

Author(s): Lei Zhang (Shenzhen University)\*; Yanyan Suo (Shenzhen University); Ximing Wu (Shenzhen University); Feng Wang (The University of Mississippi); Yuchi Chen (Simon Fraser University); Laizhong Cui (Shenzhen University); Jiangchuan Liu (Simon Fraser University); Zhong Ming (Shenzhen University)

11:15 AM-11:30 AM

**QoE Ready to Respond: A QoE-aware MEC Selection Scheme for DASH-based Adaptive Video Streaming to Mobile Users**

Author(s): Wanxin Shi (Tsinghua University); Qing Li (Southern University of Science and Technology)\*; Ruishan Zhang (Tsinghua University); Gengbiao Shen (Tsinghua University); Yong Jiang (Tsinghua University); Zhenhui Yuan (Northumbria University); Gabriel-Miro Muntean (Dublin City University)

11:30 AM-11:45 AM

**Hierarchical Fusion for Practical Ghost-free High Dynamic Range Imaging**

Author(s): Pengfei Xiong (Tencent)\*; YU CHEN (Tencent.com)

11:45 AM-12:00 PM

**Edge-oriented Convolution Block for Real-time Super Resolution on Mobile Devices**

Author(s): Xindong Zhang (The Hong Kong Polytechnic University); Hui Zeng (The Hong Kong Polytechnic University); Lei Zhang ("Hong Kong Polytechnic University, Hong Kong, China")\*

OCT.  
23

2:00 PM-4:40 PM	Session 26: Multimedia Grand Challenge	▼ Langbo Room A
<b>Session Chair: Xirong Li, Renmin University of China</b>		
2:00 PM-2:05 PM	<b>Overview of Tencent Multi-modal Ads Video Understanding</b> Author(s): Zhenzhi Wang (Nanjing University); zhimin li (Huazhong University of Science and Technology); liyu wu (peking university); jiangfeng xiong (Tencent Inc); qinglin lu (Tencent Inc)*	
2:05 PM-2:10 PM	<b>Better Learning Shot Boundary Detection via Multi-task</b> Author(s): Haoxin Zhang (Tencent Inc.)*; zhimin li (Huazhong University of Science and Technology); qinglin lu (Tencent Inc)	
2:10 PM-2:15 PM	<b>Facial Micro-Expression Generation based on Deep Motion Re-targeting and Transfer Learning</b> Author(s): Xinqi Fan (City University of Hong Kong)*; Ali Shahid (City University of Hong Kong); Hong Yan (Hong Kong SAR)	
2:15 PM-2:20 PM	<b>Deadline and Priority-aware Congestion Control for Delay-sensitive Multimedia Streaming</b> Author(s): Chao Zhou (Beijing Kuaishou Technology Co., Ltd)*; Wenjun Wu (Kuaishou); Yang Dan (Kuaishou); Tianchi Huang (Tsinghua University); Liang Guo (Kuaishou); Bing Yu ( Beijing Kuaishou Technology Co., Ltd)	
2:20 PM-2:25 PM	<b>LSSNet: A Two-stream Convolutional Neural Network for Spotting Macro- and Micro-expression in Long Videos</b> Author(s): Wang-wang Yu (University Of Electronic Science And Technology Of China) *; Jingwen Jiang (West China Hospital, Sichuan University); Yongjie Li (Univ of Electronic Science and Technology of China)	
2:25 PM-2:30 PM	<b>Multi-Level Visual Representation with Semantic-Reinforced Learning for Video Captioning</b> Author(s): Chengbo Dong (Renmin University of China) *; Xinru Chen (Renmin University); Aozhu Chen (Renmin University of China); Hu Fan (RUC); Zihan Wang (Renmin University); Xirong Li (Renmin University of China)	
2:30 PM-2:35 PM	<b>Facial Prior Based First Order Motion Model for Micro-expression Generation</b> Author(s): Yi Zhang (Sun Yat-sen University) *; Youjun Zhao (Sun Yat-sen University); Yuhang Wen (Sun Yat-sen University); Zixuan Tang (Sun Yat-sen University); Xinhua Xu (Sun Yat-sen University); Mengyuan Liu (Sun Yat-sen University)	
2:35 PM-2:40 PM	<b>Rethinking the Impacts of Overfitting and Feature Quality on Small-scale Video Classification</b> Author(s): Xuansheng WU (University of Georgia) *; Feichi Yang (University of Southern California); Tong Zhou (Center University of Finance and Economics); Xinyue Lin (UVA)	



- 2:40 PM-2:45 PM **A Gradient Balancing Approach for Robust Logo Detection**  
Author(s): Fuxing Leng (Huazhong University of Science and Technology) \*
- 
- 2:45 PM-2:50 PM **Multi-modal Representation Learning for Video Advertisement Content Structuring**  
Author(s): Daya Guo (Sun-Yat Sen University) \*; Zhaoyang Zeng (Sun Yat-sen University)
- 
- 2:50 PM-2:55 PM **Phoenix: Combining Highest-Profit First Scheduling and Responsive Congestion Control for Delay-sensitive Multimedia Transmission**  
Author(s): Haozhe Li (Beijing University of Posts and Telecommunications) \*
- 
- 2:55 PM-3:00 PM **VidVRD 2021: The Third Grand Challenge on Video Relation Detection**  
Author(s): Wei Ji (National University of Singapore) \*; Yicong Li (National University of Singapore); Meng Wei (National University of Singapore); Xindi Shang (ByteDance AI Lab); Junbin Xiao (National University of Singapore); Tongwei Ren (Nanjing University); Tat-Seng Chua (National Univ. of Singapore)
- 
- 3:00 PM-3:05 PM **A Simple and Effective Baseline for Robust Logo Detection**  
Author(s): Weipeng Xu (Shanghai Jiao Tong University) \*; Ye Liu (University of Electronic Science and Technology of China); Daquan Lin (ShanghaiTech)
- 
- 3:05 PM-3:10 PM **Robust Logo Detector for E-Commerce Images**  
Author(s): Hang Chen (Tsinghua University) \*; Xiao Li (Tsinghua University); Zefan Wang (Tsinghua University); Xiaolin Hu (Tsinghua University)
- 
- 3:10 PM-3:15 PM **Facial Action Unit-based Deep Learning Framework for Spotting Macro- and Micro-expressions in Long Video Sequences**  
Author(s): Bo Yang (KDDI Research, Inc.)\*; Jianming Wu (KDDI Research, Inc.); Zhiguang Zhou (Systemsoft, Inc.); Megumi Komiya (Systemsoft, Inc.); Koki Kishimoto (KDDI Research, Inc.); Jianfeng Xu (KDDI Research, Inc.); Keisuke Nonaka (KDDI Research Inc.); Toshiharu Horiuchi (KDDI Research, Inc.); Satoshi Komorita (KDDI Research, Inc.); Gen Hattori (KDDI Research); Sei Naito (KDDI Research Inc.); Yasuhiro Takishima (KDDI Research, Inc.)
- 
- 3:15 PM-3:20 PM **NJU MCG - Sensetime Team Silver Bullet's Submission to Pre-training for Video Understanding Challenge Track II**  
Author(s): Liwei Jin (Nanjing University) \*; Haoyue Cheng (Nanjing University); Su Xu (SenseTime Research); Wayne Wu (SenseTime Research); Limin Wang (Nanjing University)
- 
- 3:20 PM-3:25 PM **Research on Micro-Expression Spotting Method Based on Optical Flow Features**  
Author(s): He Yuhong (Harbin Institute of Technology) \*
- 
- 3:25 PM-3:30 PM **A Solution to Multi-modal Ads Video Tagging Challenge**  
Author(s): Hao Wu (Alibaba Group) \*; Jiajie Wang (Alibaba Group); Yuanzhe Gu (Alibaba Group); Peisen Zhao (Cooperative Medianet Innovation Center, Shanghai Jiao Tong University); Zhonglin Zu (alibaba)

OCT.  
23

- 3:30 PM-3:35 PM **FAMGAN: Fine-grained AUs Modulation based Generative Adversarial Network for Micro-Expression Generation**  
 Author(s): Yifan Xu (University of Science and Technology of China) \*; Sirui Zhao (University of Science and Technology of China); Huaying Tang (University of Science and Technology of China); Xinglong Mao (University of Science and Technology of China); Tong Xu (University of Science and Technology of China); Enhong Chen (University of Science and Technology of China)
- 
- 3:35 PM-3:40 PM **Semantic Tag Augmented XlanV Model for Video Captioning**  
 Author(s): Yiqing Huang (Tsinghua University); Hongwei Xue (University of Science and Technology of China); Jiansheng Chen (University of Science and Technology Beijing) \*; Huimin Ma (University of Science and Technology Beijing); Hongbing Ma (Tsinghua University)
- 
- 3:40 PM-3:45 PM **Automated Multi-Modal Video Editing for Ads Video**  
 Author(s): Qin Lin (Tencent)\*; Nuo Pang (Tencent); Zhiying Hong (Tencent)
- 
- 3:45 PM-3:50 PM **Rate Adaptation and Block Scheduling for Delay-sensitive Multimedia Applications**  
 Author(s): Dongyuan Su (Shenzhen University); Laizhong Cui (Shenzhen University) \*; Lei Zhang (Shenzhen University); Yanyan Suo (Shenzhen University); Yan Qiu (Shenzhen University)
- 
- 3:50 PM-3:55 PM **Video Relation Detection via Tracklet based Visual Transformer**  
 Author(s): Kaifeng Gao (Zhejiang University); Long Chen (Columbia University)\*; Yifeng Huang (Zhejiang University); Jun Xiao (Zhejiang University)
- 
- 3:55 PM-4:00 PM **Group-Level Focus of Visual Attention for Improved Next Speaker Prediction**  
 Author(s): Christopher Birmingham (University of Southern California) \*; Kalin Stefanov (Monash University); Maja Mataric ((University of Southern California, US))
- 
- 4:00 PM-4:05 PM **A Multimodal Framework for Video Ads Understanding**  
 Author(s): Zejia Weng (Fudan University) \*; Lingchen Meng (Fudan University); Rui Wang (Fudan University); Zuxuan Wu (UMD); Yu-Gang Jiang (Fudan University)
- 
- 4:05 PM-4:10 PM **Joint Learning for Relationship and Interaction Analysis in Video with Multimodal Feature Fusion**  
 Author(s): Beibei Zhang (Nanjing University); Fan Yu (Nanjing University); Yanxin Gao (Nanjing University); Tongwei Ren (Nanjing University) \*; Gangshan Wu (Nanjing University)
- 



- 4:10 PM-4:15 PM **MM21 Pre-training for Video Understanding Challenge: Video Captioning with Pretraining Techniques**  
Author(s): Sihan Chen (National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences and University of Chinese Academy of Sciences)\*; Xinxin Zhu (National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences); Dongze Hao (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences); Wei Liu (National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences and School of Artificial Intelligence, University of Chinese Academy of Sciences); Jiawei Liu (National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences and School of Artificial Intelligence, University of Chinese Academy of Sciences); Zijia Zhao (CAS); Longteng Guo ( Institute of Automation, Chinese Academy of Sciences); Jing Liu (National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences)
- 
- 4:15 PM-4:20 PM **CLIP4Caption: CLIP for Video Caption**  
Author(s): Mingkang Tang (Tsinghua University) \*; Zhanyu Wang (Tencent); Zhenhua LIU (encent Technology (Shenzhen) Co., Ltd.); Fengyun Rao (Tencent); Dian Li (Tencent PCG); Xiu Li (Tsinghua University)
- 
- 4:20 PM-4:25 PM **The ACM Multimedia 2021 Meet Deadline Requirements Grand Challenge**  
Author(s): Jie Zhang (Tsinghua University); Junjie Deng (Beijing University of Posts and Telecommunications); Mowei Wang (Tsinghua University); Yong Cui (Tsinghua University) \*; Wei Tsang Ooi (National University of Singapore); Jiangchuan Liu (Simon Fraser University); Xinyu Zhang (UCSD); Kai Zheng (Huawei); Yi Li (PowerInfo)
- 
- 4:25 PM-4:30 PM **MultiModal Language Modelling on Knowledge Graphs for Deep Video Understanding**  
Author(s): Vishal Anand (Columbia University) \*; Raksha Ramesh (Columbia University); Boshen Jin (Columbia University); Ziyin Wang (Columbia University); Xiaoxiao Lei (Graphen AI); Ching-Yung Lin (Columbia University)
- 
- 4:30 PM-4:35 PM **Using Motion Histories for Eye Contact Detection in Multiperson Group Conversations**  
Author(s): Eugene Yujun Fu (The Hong Kong Polytechnic University) \*; Michael Ngai (Phillips Exeter Academy)
- 
- 4:35 PM-4:40 PM **MultiMediate: Multi-modal Group Behaviour Analysis for Artificial Mediation**  
Author(s): Philipp Müller (DFKI GmbH)\*; Michael Dietz (University of Augsburg); Dominik Schiller (Augsburg University); Dominike Thomas (University of Stuttgart); Guanhua Zhang (University of Stuttgart); Patrick Gebhard ("DFKI, Germany"); Elisabeth André (University of Augsburg, Germany); Andreas Bulling (University of Stuttgart)
- 



2:00 PM-3:30 PM	Session 27: Media Interpretation-III	▼ Langbo Room B
<b>Session Chair: Yi Jin, Beijing Jiaotong University</b>		
2:00 PM-2:15 PM	<b>Cross-Camera Feature Prediction for Intra-Camera Supervised Person Re-identification across Distant Scenes</b> Author(s): Wenhang Ge (Sun Yat-sen University); Chunyan Pan (Sun Yat-sen University); Ancong Wu (Sun Yat-sen University) *; Hongwei Zheng (University of Chinese Academy of Sciences); WEI-SHI ZHENG (Sun Yat-sen University, China)	
2:15 PM-2:30 PM	<b>Video Visual Relation Detection via Iterative Inference</b> Author(s): Xindi Shang (ByteDance AI Lab) *; Yicong Li (National University of Singapore); Junbin Xiao (National University of Singapore); Wei Ji (National University of Singapore); Tat-Seng Chua (National university of Singapore)	
2:30 PM-2:45 PM	<b>Instance-wise or Class-wise? A Tale of Neighbor Shapley for Concept-based Explanation</b> Author(s): Jiahui Li (Zhejiang University) *; Kun Kuang (Zhejiang University); Lin Li (Zhejiang University); Long Chen (Columbia University); Songyang Zhang (University of Rochester); Jian Shao (Zhejiang University); Jun Xiao (Zhejiang University)	
2:45 PM-3:00 PM	<b>Multifocal Attention-Based Cross-Scale Network for Image De-raining</b> Author(s): Zheyu Zhang (University of Science and Technology of China); Yurui Zhu (University of Science and Technology of China); Xueyang Fu (University of Science and Technology of China)*; Zhiwei Xiong (University of Science and Technology of China); Zheng-Jun Zha (University of Science and Technology of China); Feng Wu (University of Science and Technology of China)	
3:00 PM-3:15 PM	<b>PFFN: Progressive Feature Fusion Network for Lightweight Image Super-Resolution</b> Author(s): Dongyang Zhang (University of Electronic Science and Technology of China); Changyu Li (University of Electronic Science and Technology of China); Ning Xie (Center for Future Media, the Department of Computer Science and Engineering, University of Electronic Science and Technology of China); Guoqing Wang (University of Electronic Science and Technology of China); Jie Shao (University of Electronic Science and Technology of China) *	
3:15 PM-3:30 PM	<b>InterBN: Channel Fusion for Adversarial Unsupervised Domain Adaptation</b> Author(s): Mengzhu Wang (NUDT); Wei Wang (Dalian University of Technology); Baopu Li (Baidu Research (USA)) *; Xiang Zhang (National University of Defense Technology); long lan (NUDT); Tan H Huibin (National Defense University of Science and Technology); Tianyi Liang (National University of Defense Technology); Wei Yu (National University of Defense Technology); Zhigang Luo (National University of Defense Technology)	



2:00 PM-3:30 PM

Session 28: Multimodal Fusion and Embedding-I

▼ Tianfu Room A

Session Chair: Junying Chen, South China University of Technology

2:00 PM-2:15 PM

**Multi-Source Fusion and Automatic Predictor Selection for Zero-Shot Video Object Segmentation**

Author(s): Xiaoqi Zhao (Dalian University of Technology)\*; Youwei Pang (Dalian University of Technology); Jiaying Yang (Dalian University of Technology); Lihe Zhang (Dalian University of Technology); Huchuan Lu (Dalian University of Technology)

2:15 PM-2:30 PM

**Self-supervised Consensus Representation Learning for Attributed Graph**

Author(s): Changshu Liu (University of Electronic Science and Technology of China)\*; liangjian Wen (the Noah's Ark Lab, Huawei Technologies Company Limited); Zhao Kang (University of Electronic Science and Technology of China); Guangchun Luo (University of Electronic Science and Technology of China); Ling Tian (University of Electronic Science and Technology of China)

2:30 PM-2:45 PM

**Efficient Multi-Modal Fusion with Diversity Analysis**

Author(s): Shuhui Qu (Samsung Display America Lab)\*; Yan Kang (Samsung Display America Lab); Janghwan Lee (Samsung)

2:45 PM-3:00 PM

**GCCN: Geometric Constraint Co-attention Network for 6D Object Pose Estimation**

Author(s): Yongming Wen (Sun Yat-Sen University); Yiquan Fang (Sun Yat-Sen University); Junhao Cai (Sun Yet-sen University); Kimwa Tung (Sun Yat-Sen University); Hui Cheng (Sun Yat-Sen University)\*

3:00 PM-3:15 PM

**Cross-Modal Generalization: Learning in Low Resource Modalities via Meta-Alignment**

Author(s): Paul Pu Liang (Carnegie Mellon University)\*; Peter Wu (Carnegie Mellon University); Liu Ziyin (University of Tokyo); Louis-Philippe Morency (Carnegie Mellon University); Ruslan Salakhutdinov (Carnegie Mellon University)

3:15 PM-3:30 PM

**Elastic Tactile Simulation Towards Tactile-Visual Perception**

Author(s): Yikai Wang (Tsinghua University)\*; Wenbing Huang (Tsinghua University); Bin Fang (Tsinghua University); Fuchun Sun (Tsinghua); Chang Li (JD Explore Academy)

OCT.  
23

2:00 PM-3:30 PM

Session 29: Multimodal Fusion and Embedding-II

▼ Tianfu Room B

**Session Chair: Wen Li, University of Electronic Science and Technology of China**

2:00 PM-2:15 PM

**A Novel Patch Convolutional Neural Network for View-based 3D Model Retrieval**

Author(s): Zan Gao (1. Shandong AI Institute, QiLU University of Technology, 2. Shandong Computer Science Center(National Supercomputer Center in Jinan),3. Tianjing University of Technology); Yuxiang Shao (Qilu University of Technology); Weili Guan (Monash University); Meng Liu (Shandong Jianzhu University); Zhiyong Cheng (Shandong Academy of Sciences)\*; Shengyong Chen (Tianjin University of Technology)

2:15 PM-2:30 PM

**Semi-Autoregressive Image Captioning**

Author(s): Xu Yan (VIPL, ICT, Chinese Academy of Sciences); Zhengcong Fei (Chinese Academy of Sciences, Institute of Computing Technology)\*; Zekang Li (Institute of Computing Technology, Chinese Academy of Sciences); Shuhui Wang (VIPL,ICT,Chinese academic of science); Qingming Huang (University of Chinese Academy of Sciences); Qi Tian (Huawei Cloud & AI)

2:30 PM-2:45 PM

**One-Stage Incomplete Multi-view Clustering via Late Fusion**

Author(s): Yi Zhang (National University of Defense Technology)\*; Xinwang Liu (National University of Defense Technology); Siwei Wang (National University of Defense Technology ); Jiyuan Liu (National University of Defense Technology); Sisi Dai (National University of Defense Technology); En Zhu (National University of Defense Technology)

2:45 PM-3:00 PM

**Self-Representation Subspace Clustering for Incomplete Multi-view Data**

Author(s): Jiyuan Liu (National University of Defense Technology)\*; Xinwang Liu (National University of Defense Technology); Yi Zhang (National University of Defense Technology); Pei Zhang (National University of Defense Technology); Wenxuan Tu (NUDT); Siwei Wang (National University of Defense Technology ); Sihang Zhou (NUDT); Weixuan Liang (National University of Defense Technology); Siqi Wang (College of Computer, National University of Defense Technology,China State Key Laboratory of High Performance Computing, National University of Defense Technology); Yuexiang Yang (National University of Defense Technology)

3:00 PM-3:15 PM

**Is Visual Context Really Helpful for Knowledge Graph? A Representation Learning Perspective**

Author(s): 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)

3:15 PM-3:30 PM

**Knowledge perceived multi-modal pretraining in E-commerce**

Author(s): Yushan Zhu (Zhejiang University)\*; Huaixiao Zhao (Alibaba Group); Wen Zhang (Zhejiang University); Ganqiang Ye (Zhejiang University); Hui Chen (Alibaba Group); Ningyu Zhang (Zhejiang University & AZFT Joint Lab for Knowledge Engine); Huajun Chen (Zhejiang University)

OCT.  
23

2:00 PM-3:30 PM

Session 30: Multimodal Fusion and Embedding-III

▼ Jiaozi Room

Session Chair: Lin Zhang, Tongji University

2:00 PM-2:15 PM

**DSP: Dual Soft-Paste for Unsupervised Domain Adaptive Semantic Segmentation**

Author(s): Li Gao (Wuhan University); Jing Zhang (The University of Sydney); Lefei Zhang (Wuhan University)\*; Dacheng Tao (JD.com)

2:15 PM-2:30 PM

**Generating Point Cloud from Single Image in The Few Shot Scenario**

Author(s): Yu Lin (University of Texas at Dallas)\*; Jinghui Guo (The University of Texas at Dallas); Yang Gao (University of Texas at Dallas); Yi-Fan Li (University of Texas at Dallas); Zhuoyi Wang (University of Texas at Dallas); Latifur Khan (The university of Texas at Dallas)

2:30 PM-2:45 PM

**Product-oriented Machine Translation with Cross-modal Cross-lingual Pre-training**

Author(s): Yuqing Song (Renmin University of China)\*; Shizhe Chen (INRIA); Qin Jin (Renmin University of China); Wei Luo (Alibaba); Jun Xie (Alibaba); Fei Huang (Alibaba)

2:45 PM-3:00 PM

**Pre-training Graph Transformer with Multimodal Side Information for Recommendation**

Author(s): Yong Liu (Nanyang Technological University)\*; Susen Yang (Alibaba Group); Chenyi Lei (University of Science and Technology of China, Alibaba Group); Guoxin Wang (Alibaba Group); Haihong Tang (Alibaba Group); Juyong Zhang (University of Science and Technology of China); Aixin Sun (Nanyang Technological University); Chunyan Miao (NTU)

3:00 PM-3:15 PM

**Learning Disentangled Factors from Paired Data in Cross-Modal Retrieval: An Implicit Identifiable VAE Approach**

Author(s): Minyoung Kim (Samsung AI Center Cambridge)\*; Ricardo Guerrero (Samsung AI Center Cambridge); Vladimir Pavlovic (Rutgers University)

3:15 PM-3:30 PM

**Progressive Graph Attention Network for Video Question Answering**

Author(s): Liang Peng (University of Electronic Science and Technology of China); Shuangji Yang (University of Electronic Science and Technology of China); Yi Bin (University of Electronic Science and Technology of China)\*; Guoqing Wang (University of Electronic Science and Technology of China)

OCT.  
23

2:00 PM-5:30 PM	<b>Session 31: Multimedia Art, Entertainment and Culture</b>	▼ Rongcheng Room
<b>Session Chair: Xin Jin, Beijing Electronic Science and Technology Institute</b>		
2:00 PM-2:15 PM	<b>Learning to Compose Stylistic Calligraphy Artwork with Emotions</b> Author(s): Shaozu Yuan (JD AI ); Ruixue Liu (JD AI ); Meng Chen (JD AI)*; Baoyang Chen (Faculty of School of Experimental Arts; Central Academy of Fine Arts); Zhijie Qiu (Faculty of School of Experimental Arts; Central Academy of Fine Arts); Xiaodong He (JD AI Research)	
2:15 PM-2:30 PM	<b>Graph Neural Networks for Knowledge Enhanced Visual Representation of Paintings</b> Author(s): Athanasios Efthymiou (University of Amsterdam)*; Stevan Rudinac (University of Amsterdam); Monika Kackovic (University of Amsterdam); Marcel Worring (University of Amsterdam); Nachoem Wijnberg (ABS, University of Amsterdam; CBE, University of Johannesburg)	
2:30 PM-2:45 PM	<b>ArtScience and the ICECUBE LED Display [ILDm^3]</b> Author(s): Mark-David Hosale (York University - Computational Arts)*; Robert Allison (York University); Jim Madsen (Wisconsin IceCube Particle Astrophysics Center); Marcus Gordon (York University)	
2:45 PM-3:00 PM	<b>PUGCQ: A Large Scale Dataset for Quality Assessment of Professional User-Generated Content</b> Author(s): Guo Li (Kingsoft Cloud); Baoliang Chen (City University of Hongkong); lingyu zhu (City University of Hongkong); Qinwen He (kingsoft cloud); hongfei fan (kingsoft.com); Shiqi Wang (City University of Hong Kong)*	
3:00 PM-3:15 PM	<b>Combining Attention with Flow for Person Image Synthesis</b> Author(s): Yurui Ren (Shenzhen Graduate School, Peking University); Yubo Wu (Dalian University of Technology); Thomas H Li (Advanced Institute of Information Technology, Peking University); Shan Liu (Tencent America); Ge Li (SECE, Shenzhen Graduate School, Peking University)*	
3:15 PM-3:30 PM	<b>Dual Learning Music Composition and Dance Choreography</b> Author(s): Shuang Wu (Nanyang Technological University) *; Zhenguang Liu (Zhejiang University); Shijian Lu (Nanyang Technological University); Li Cheng (ECE dept., University of Alberta)	



4:00 PM-5:30 PM

Session 32: Social Multimedia

▼ Langbo Room B

Session Chair: Zhao Kang, University of Electronic Science and Technology of China

4:00 PM-4:15 PM

**FoodLogoDet-1500: A Dataset for Large-Scale Food Logo Detection via Multi-Scale Feature Decoupling Network**

Author(s): Qiang Hou (Shandong Normal University); Weiqing Min (Institute of Computing Technology, Chinese Academy of Sciences); Jing Wang (Shandong Normal University); Sujuan Hou (Shandong Normal University) \*; Yuanjie Zheng (Shandong Normal University); Shuqiang Jiang (ICT, China Academy of Science)

4:15 PM-4:30 PM

**Cross-View Representation Learning for Multi-View Logo Classification with Information Bottleneck**

Author(s): Jing Wang (Shandong Normal University); Yuanjie Zheng (Shandong Normal University) \*; jingqi song (Shandong University of Finance and Economics); Sujuan Hou (Shandong Normal University)

4:30 PM-4:45 PM

**Parametric Reshaping of Portraits in Videos**

Author(s): Xiangjun Tang (Zhejiang University); WenXin Sun (State Key Lab of CAD&CG, Zhejiang University); Yongliang Yang (University of Bath); Xiaogang Jin (Zhejiang University) \*

4:45 PM-5:00 PM

**Human attributes prediction under privacy-preserving conditions**

Author(s): Anshu Singh (National University of Singapore) \*; Shaojing Fan (National University of Singapore); Mohan Kankanhalli (National University of Singapore,)

5:00 PM-5:15 PM

**Multi-Modal Sarcasm Detection with Interactive In-Modal and Cross-Modal Graphs**

Author(s): Bin Liang (Joint Lab of Harbin Institute of Technology and RICOH; Harbin Institute of Technology (Shenzhen))\*; Chenwei Lou (Harbin Institute of Technology(ShenZhen)); Xiang Li (Harbin Institute of Technology, Shenzhen); Lin Gui (Department of Computer Science, University of Warwick, Coventry CV4 7AL, U.K.); Min Yang ( Chinese Academy of Sciences); Ruifeng Xu (Harbin Institute of Technology (Shenzhen))

5:15 PM-5:30 PM

**Linking the Characters: Video-oriented Social Graph Generation via Hierarchical-cumulative GCN**

Author(s): Shiwei Wu (University of Science and Technology of China)\*; Joya Chen (National University of Singapore); Tong Xu (University of Science and Technology of China); Liyi Chen (University of Science and Technology of China); Lingfei Wu (IBM Research); Yao Hu (Alibaba Youku Cognitive and Intelligent Lab); Enhong Chen (University of Science and Technology of China)

OCT.  
23

4:00 PM-5:30 PM	<b>Session 33: Summarization, Analytics, and Storytelling</b>	▼ Tianfu Room A
<b>Session Chair: Lifang Wu, Beijing University of Technology</b>		
4:00 PM-4:15 PM	<b>MeshNet++: A Network with a Face</b> Author(s): Vinit Veerendraveer Singh (University Of Delaware) *; Shivanand Venkanna Sheshappanavar (University of Delaware); Chandra Kambhamettu (University of Delaware)	
4:15 PM-4:30 PM	<b>Latent Memory-augmented Graph Transformer for Visual Storytelling</b> Author(s): Mengshi Qi (Beijing University of Posts and Telecommunications) *; Jie Qin (Inception Institute of Artificial Intelligence); Di Huang (Beihang University, China); Zhiqiang Shen (Carnegie Mellon University); Yi Yang (UTS); Jiebo Luo (U. Rochester)	
4:30 PM-4:45 PM	<b>TSA-Net: Tube Self-Attention Network for Action Quality Assessment</b> Author(s): Shunli Wang (Fudan University); Ding kang Yang (Fudan University); Peng Zhai (Fudan university); Chixiao Chen (Fudan University); Lihua Zhang (Fudan University) *	
4:45 PM-5:00 PM	<b>Exploring Contextual-Aware Representation and Linguistic-Diverse Expression for Visual Dialog</b> Author(s): Xiangpeng Li (University of Electronic Science and Technology of China); Lianli Gao (The University of Electronic Science and Technology of China); Lei Zhao (The University of Electronic Science and Technology of China); Jingkuan Song (UESTC)*	
5:00 PM-5:15 PM	<b>Automated Playtesting with a Cognitive Model of Sensorimotor Coordination</b> Author(s): Injung Lee (KAIST); Hyunchul Kim (KAIST); Byungjoo Lee (Yonsei University) *	
5:15 PM-5:30 PM	<b>CAA: Candidate-Aware Aggregation for Temporal Action Detection</b> Author(s): Yifan Ren (University of Electronic Science and Technology of China); Xing Xu (University of Electronic Science and Technology of China) *; Fumin Shen (UESTC); Yazhou Yao (Nanjing University of Science and Technology); Huimin Lu (Kyushu Institute of Technology)	



4:00 PM-5:30 PM

Session 34: Vision and Language-I

▼ Tianfu Room B

Session Chair: Zhou Zhao, Zhejiang University

4:00 PM-4:15 PM

**Disentangle Your Dense Object Detector**

Author(s): Zehui Chen (University of Science and Technology of China); Chenhongyi Yang (University of Edinburgh); Qiaofei Li (Xidian University); Feng Zhao (University of Science and Technology of China) \*; Zheng-Jun Zha (University of Science and Technology of China); Feng Wu (University of Science and Technology of China)

4:15 PM-4:30 PM

**Do We Really Need Frame-by-Frame Annotation Datasets for Object Tracking?**

Author(s): 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) \*

4:30 PM-4:45 PM

**Video-to-Image Casting: A Flattening Method for Video Analysis**

Author(s): Xu Chen (Chongqing University of Posts and Telecommunications) \*; Chenqiang Gao (Chongqing University of Posts and Telecommunications); Feng Yang (Chongqing Key Laboratory of Signal and Information Processing, Chongqing University of Posts and Telecommunications); Xiaohan Wang (University of Technology, Sydney); Yi Yang (UTS); Yahong Han (Tianjin University)

4:45 PM-5:00 PM

**Complementary Trilateral Decoder for Fast and Accurate Salient Object Detection**

Author(s): Zhirui Zhao (Beihang University); Changqun Xia (Peng Cheng Laboratory) \*; Chenxi Xie (Peng Cheng Laboratory); Jia Li (Beihang University)

5:00 PM-5:15 PM

**Learning Human Motion Prediction via Stochastic Differential Equations**

Author(s): 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)

5:15 PM-5:30 PM

**Space-Time Interaction Graph Parsing Networks for Human-Object Interaction Recognition**

Author(s): Ning Wang (School of Computer Science and Technology, Xidian University, China); Guangming Zhu (School of Computer Science and Technology, Xidian University, China)\*; Liang Zhang (School of Computer Science and Technology, Xidian University, China); Peiyi Shen (School of Computer Science and Technology, Xidian University, China); Hongsheng Li (School of Computer Science and Technology, Xidian University, China); Cong Hua (School of Computer Science and Technology, Xidian University, China)

OCT.  
23

4:00 PM-5:30 PM

Session 35: Vision and Language-II

▼ Jiaozi Room

Session Chair: Lianli Gao, University of Electronic Science and Technology of China

4:00 PM-4:15 PM

**Learning Hierarchical Channel Attention for Fine-grained Visual Classification**

Author(s): Xiang Guan (University of Electronic Science and Technology of China); Guoqing Wang (University of Electronic Science and Technology of China) \*; Xing Xu (University of Electronic Science and Technology of China); Yi Bin (University of Electronic Science and Technology of China)

4:15 PM-4:30 PM

**Group-based Distinctive Image Captioning with Memory Attention**

Author(s): Jiuniu Wang (City University of Hong Kong) \*; Wenjia Xu (Institute of Electronics, Chinese Academy of Sciences); Qingzhong Wang (Department of Computer Science, City University of Hong Kong); Antoni Chan (City University of Hong Kong, Hong, Kong)

4:30 PM-4:45 PM

**VQMG: Hierarchical Vector Quantised and Multi-hops Graph Reasoning for Explicit Representation Learning**

Author(s): Lei Li (Tsinghua University) \*; Chun Yuan (Graduate school at ShenZhen, Tsinghua university)

4:45 PM-5:00 PM

**Structure-aware Mathematical Expression Recognition with Sequence-Level Modeling**

Author(s): Minli Li (School of Software Engineering, South China University of Technology); Peilin Zhao (Tencent AI Lab); Yifan Zhang (National University of Singapore); Shuaicheng Niu (South China University of Technology); Qingyao Wu (South China University of Technology); Mingkui Tan (South China University of Technology) \*

5:00 PM-5:15 PM

**Exploring Logical Reasoning for Referring Expression Comprehension**

Author(s): Ying Cheng (Fudan University) \*; Ruize Wang (Fudan University); Jiashuo Yu (Fudan University); Rui-Wei Zhao (Fudan University); Yuejie Zhang (Fudan University); Rui Feng (Fudan University)

5:15 PM-5:30 PM

**Direction Relation Transformer for Image Captioning**

Author(s): Zeliang Song (Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences, School of Cyber Security); Xiaofei Zhou (Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences, School of Cyber Security)\*; Linhua Dong (Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences, School of Cyber Security); Jianlong Tan (Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences, School of Cyber Security); Li Guo (Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences, School of Cyber Security)

OCT.  
23

4:00 PM-5:30 PM

Session 36: Vision and Language-III

▼ Jincheng Room

Session Chair: Xinhang Song, Institute of Computing Technology, CAS

4:00 PM-4:15 PM

**Contrastive Disentangled Meta-Learning for Signer-Independent Sign Language Translation**

Author(s): Tao Jin (Zhejiang University) \*; Zhou Zhao (Zhejiang University)

4:15 PM-4:30 PM

**Scene Graph with 3D Information for Change Captioning**

Author(s): Zeming Liao (School of Electrical Engineering, Guangxi University, Nanning, Guangxi, China); Qingbao Huang (School of Electrical Engineering, Guangxi University, Nanning, Guangxi, China) \*; Yu Liang (School of Electrical Engineering, Guangxi University); Mingyi Fu (School of Electrical Engineering, Guangxi University, Nanning, Guangxi); Yi Cai (School of Software Engineering, South China University of Technology); Qing Li (The Hong Kong Polytechnic University)

4:30 PM-4:45 PM

**Progressive Semantic Matching for Video-Text Retrieval**

Author(s): Hongying Liu (Key Lab. of Intelligent Perception and Image Understanding of Ministry of Education, School of Artificial Intelligence, Xidian University, China); ruyi luo (Xidian University); Fanhua Shang (Xidian University) \*; Mantang Niu (Xidian University); Yuanyuan Liu (Xidian University)

4:45 PM-5:00 PM

**Multimodal Asymmetric Dual Learning for Unsupervised Eyeglasses Removal**

Author(s): Qing Lin (Fudan University); Bo Yan (Fudan University) \*; Weimin Tan (Fudan University)

5:00 PM-5:15 PM

**Neighbor-view Enhanced Model for Vision and Language Navigation**

Author(s): Dong An (Institute of Automation, Chinese Academy of Sciences); Yuankai Qi (The University of Adelaide); Yan Huang (Institute of Automation, Chinese Academy of Sciences) \*; Qi Wu (University of Adelaide); Liang Wang (NLPR, China); Tieniu Tan (NLPR, China)

5:15 PM-5:30 PM

**Multi-Perspective Video Captioning**

Author(s): Yi Bin (University of Electronic Science and Technology of China) \*; Xindi Shang (ByteDance AI Lab); Bo Peng (Tianjin University); Yujuan Ding (The Hong Kong Polytechnic University); Tat-Seng Chua (National university of Singapore)

OCT.  
23

4:00 PM-5:30 PM	Session 37: Multimedia Telepresence and Virtual/Augmented Reality	▼ Rongcheng Room
<b>Session Chair: Xuejin Chen, University of Science and Technology of China</b>		
4:00 PM-4:15 PM	<b>Learning Spatial-angular Fusion for Compressive Light Field Imaging in a Cycle-consistent Framework</b> Author(s): Xianqiang LYU (City university); Zhiyu Zhu (City University of Hong Kong ); Mantang Guo (City University of Hong Kong); Jing Jin (City University of Hong Kong); Junhui Hou (City University of Hong Kong, Hong Kong)*; Huanqiang Zeng (Huaqiao University)	
4:15 PM-4:30 PM	<b>From Voxel to Point: IoU-guided 3D Object Detection for PointCloud with Voxel-to-Point Decoder</b> Author(s): Jiale Li (Zhejiang University); Hang Dai (Mohamed bin Zayed University of Artificial Intelligence)*; Ling Shao (Inception Institute of Artificial Intelligence); Yong Ding (Zhejiang University)	
4:30 PM-4:45 PM	<b>Extending 6-DoF VR Experience Via Multi-Sphere Images Interpolation</b> Author(s): Jisheng Li (Tsinghua University)*; Yuze He (Tsinghua University); Jinghui Jiao (Tsinghua University); Yubin Hu (Tsinghua University); Yuxing Han (Research Institute of Tsinghua University in Shenzhen); Jiangtao Wen (Tsinghua University)	
4:45 PM-5:00 PM	<b>iButter: Neural Interactive Bullet Time Generator for Human Free-viewpoint Rendering</b> Author(s): Liao Wang (ShanghaiTech University)*; Ziyu Wang (Shanghaitech University); Pei Lin (ShanghaiTech University); Yuheng Jiang (ShanghaiTech University); Xin Suo (Shanghaitech university); Minye Wu (Shanghaitech University); Lan Xu (ShanghaiTech); Jingyi Yu (Shanghai Tech University)	
5:00 PM-5:15 PM	<b>Neural Free-Viewpoint Performance Rendering under Complex Human-object Interactions</b> Author(s): Guoxing Sun (ShanghaiTech University)*; Xin Chen (ShanghaiTech University); Yizhang Chen (ShanghaiTech University); Anqi Pang (ShanghaiTech University); Pei Lin (ShanghaiTech University); Yuheng Jiang (ShanghaiTech University); Lan Xu (ShanghaiTech); Jingyi Yu (Shanghai Tech University); Jingya Wang ( ShanghaiTech University)	
5:15 PM-5:30 PM	<b>Semi-supervised Learning via Improved Teacher-Student Network For Robust 3D Reconstruction of Stereo Endoscopic Image</b> Author(s): Hongkuan Shi (Huazhong University of Science and Technology); Zhiwei Wang (Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology); Jinxin Lv (Huazhong University of Science and Technology); Yilang Wang (Huazhong University of Science and Technology); Peng Zhang (Huazhong University of Science and Technology); Fei Zhu (Huazhong University of Science and Technology); Qiang Li (Huazhong University of Science and Technology)*	



9:00 AM-12:00 PM	Poster Session 5	▼ Hall Way
P261	<b>Semantic Scalable Image Compression with Cross-Layer Priors</b> Author(s): Hanyue Tu (University of Science and Technology of China) *; Li Li (University of Science and Technology of China); Wengang Zhou (University of Science and Technology of China); Houqiang Li (University of Science and Technology of China)	
P262	<b>Cascade Cross-modal Attention Network for Video Actor and Action Segmentation from a Sentence</b> Author(s): Weidong Chen (University of Chinese Academy of Sciences) *; Guorong Li (University of Chinese Academy of Sciences); xinfeng zhang (University of Chinese Academy of Sciences); Hongyang Yu (Peng Cheng Laboratory); Shuhui Wang (VIPL,ICT,Chinese academic of science); Qingming Huang (University of Chinese Academy of Sciences)	
P263	<b>Extracting Useful Knowledge Form Noisy Web Images via Data Purification for Fine-Grained Recognition</b> Author(s): Chuanyi Zhang (Nanjing University of Science and Technology); Yazhou Yao (Nanjing University of Science and Technology) *; Xing Xu (University of Electronic Science and Technology of China); Jie Shao (University of Electronic Science and Technology of China); Jingkuan Song (UESTC); Zechao Li (Nanjing University of Science and Technology); Zhenmin Tang (Nanjing University of Science and Technology)	
P264	<b>Complementary Factorization towards Outfit Compatibility Modeling</b> Author(s): Tianyu Su (Shandong University) *; Xuemeng Song (Shandong University); Na Zheng (Shandong University); WEILI GUAN (Monash Univerisity); Yan Li (Kuaishou); Liqiang Nie (Shandong University)	
P265	<b>Open Set Face Anti-Spoofing in Unseen Attacks</b> Author(s): Xin Dong (Ningxia University); Hao Liu (Ningxia University) *; Weiwei Cai (Ningxia University); Pengyuan Lv (Ningxia University); Zekuan Yu (Fudan University)	
P266	<b>Interventional Video Relation Detection</b> Author(s): yicong Li (National University of Singapore) *; Xun Yang (National University of Singapore); Xindi Shang (National University of Singapore); Tat-Seng Chua (National University of Singapore)	
P267	<b>CanvasEmb: Learning Layout Representation with Large-scale Pre-training for Graphic Design</b> Author(s): Yuxi Xie (School of Computing, National University of Singapore) *; Danqing Huang (Microsoft Research); Jinpeng Wang (Microsoft Research Asia); Chin-Yew Lin (Microsoft Research Asia)	
P268	<b>Augmenting TV Shows via Uncalibrated Camera Small Motion Tracking in Dynamic Scene</b> Author(s): Yizhen Lao (Hunan University) *; Jie Yang (MGTV); xinying wang (MGTV); Jianxin Lin (Hunan University); Cao Yu (Hunan University); Song shien (MGTV)	

OCT.  
23

- P269 **SimuSLT: End-to-End Simultaneous Sign Language Translation**  
 Author(s): Aoxiong Yin (Zhejiang University); Zhou Zhao (Zhejiang University) \*; Jinglin Liu (Zhejiang University); Weike Jin (Zhejiang University); Meng Zhang (Huawei Noah's Ark Lab); Xingshan Zeng (Huawei Noah's Ark Lab); Xiaofei He (Zhejiang University)
- 
- P270 **Mask and Predict: Multi-step Reasoning for Scene Graph Generation**  
 Author(s): Hongshuo Tian (Tianjin University); Ning Xu (Tianjin University) \*; 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)
- 
- P271 **Heterogeneous Face Recognition with Attention-guided Feature Disentangling**  
 Author(s): Shanming Yang (Sichuan University); Xiao Yang (Sichuan University) \*; Yi Lin (Sichuan University); Peng Cheng (Sichuan University); Yi Zhang (Sichuan University); Jianwei Zhang (Sichuan University)
- 
- P272 **Exploring the Quality of GAN Generated Images for Person Re-Identification**  
 Author(s): Yiqi Jiang (Alibaba Group) \*; Weihua Chen (alibaba group); Xiuyu Sun (Alibaba Group); Xiaoyu Shi (Alibaba Group); Fan Wang (Alibaba Group); Hao Li (Alibaba Group)
- 
- P273 **Multi-view Clustering via Deep Matrix Factorization and Partition Alignment**  
 Author(s): Chen Zhang (NUDT)\*; Siwei Wang (National University of Defense Technology); Jiyuan Liu (National University of Defense Technology); Sihang Zhou (NUDT); Pei Zhang (National University of Defense Technology); Xinwang Liu (National University of Defense Technology); En Zhu (National University of Defense Technology); Changwang Zhang (CCF Theoretical Computer Science Technical Committee)
- 
- P274 **Video Similarity and Alignment Learning on Partial Video Copy Detection**  
 Author(s): Zhen Han (Alibaba Group) \*; Xiangteng He (Alibaba Group); Mingqian Tang (Alibaba Group); Yiliang Lv (Alibaba Group)
- 
- P275 **No-Reference Video Quality Assessment with Heterogeneous Knowledge Ensemble**  
 Author(s): Jinjian Wu (Xidian University) \*; Yongxu Liu (Xidian University); Leida Li (Xidian University); Weisheng Dong (Xidian University); Guangming Shi (Xidian University)
- 
- P276 **Seeing is Believing? Effects of Visualization on Smart Device Privacy Perceptions**  
 Author(s): Carlos Bermejo Fernandez (Hong Kong University of Science and Technology) \*; Petteri Nurmi (University of Helsinki); Pan Hui (Hong Kong University of Science and Technology)



- P277      **MHFC: Multi-Head Feature Collaboration for Few-Shot Learning**  
Author(s): Shuai Shao (China University of Petroleum (East China) College of Control Science and Engineering)\*; Lei Xing (China University of Petroleum(East China) College of Oceanography and Space Informatics); Yan WANG (Beihang University); Rui Xu (China University of Petroleum (East China) College of Control Science and Engineering); Chunyan Zhao (Suzhou Centennial College); yanjiang wang (China University of Petroleum (East China) College of Control Science and Engineering); baodi liu (China University of Petroleum (East China) College of Information and Control Engineering)
- 
- P278      **Vision-guided Music Source Separation via a Fine-grained Cycle-Separation Network**  
Author(s): Ma Shuo (UESTC); Yanli Ji (UESTC)\*; Xing Xu (University of Electronic Science and Technology of China); Xiaofeng Zhu (Guangxi Normal University)
- 
- P279      **GLM-Net: Global and Local Motion Estimation via Task-Oriented Encoder-Decoder Structure**  
Author(s): Yuchen Yang (Beijing University of Technology); ye xiang (BJUT); Shuaicheng Liu (UESTC; Megvii); Lifang Wu (Beijing University of Technology) \*; Boxuan Zhao (Beijing University of Technology); Bing Zeng (University of Electronic Science and Technology of China)
- 
- P280      **Sensor-Augmented Egocentric-Video Captioning with Dynamic Modal-Attention**  
Author(s): Katsuyuki Nakamura (Hitachi, Ltd.) \*; Hiroki Ohashi (Hitachi Ltd); Mitsuhiro Okada (Hitachi, Ltd.)
- 
- P281      **Cross Modal Compression: Towards Human-comprehensible Semantic Compression**  
Author(s): Jiguo Li (Institute of Computing Technology, Chinese Academy of Sciences) \*; Chuanmin Jia (Peking University); Xinfeng Zhang (University of Chinese Academy of Sciences); Siwei Ma (Peking University, China); Wen Gao (PKU)
- 
- P282      **RAMS-Trans: Recurrent Attention Multi-scale Transformer for Fine-grained Image Recognition**  
Author(s): Yunqing Hu (Zhejiang University); Xuan Jin (Alibaba Turing Lab, Alibaba Group); Yin Zhang (Zhejiang University) \*; Haiwen Hong (Zhejiang University); Jingfeng Zhang (Zhejiang University); Yuan He (Alibaba Group); Hui Xue (Alibaba)
- 
- P283      **Memory-Augmented Deep Unfolding Network for Compressive Sensing**  
Author(s): Jiechong Song (PKU); Bin Chen (School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School); Jian Zhang (Peking University Shenzhen Graduate School) \*
- 
- P284      **Underwater Species Detection using Channel Sharpening Attention**  
Author(s): Lihao Jiang (Dalian University of Technology); Yi Wang (Dalian University of Technology) \*; Qi Jia (Dalian University of Technology); Shengwei Xu (Dalian University of Technology); Yu Liu (Dalian University of Technology); Xin Fan (Dalian University of Technology); Haojie Li (Dalian University of Technology); Risheng Liu (Dalian University of Technology); Xinwei Xue (Dalian University of Technology); Ruili Wang (Massey University)

- P285 **Self-Supervised Pre-training on the Target Domain for Cross-Domain Person Re-identification**  
Author(s): Junyin Zhang (Chongqing University); Yongxin Ge (Chongqing University) \*; Xinqian Gu (University of Chinese Academy of Sciences); Boyu Hua (Chongqing University); Tao Xiang (Chongqing University)
- 
- P286 **Exploring Graph-Structured Semantics for Cross-Modal Retrieval**  
Author(s): Lei Zhang (School of Computer Science and Engineering, University of Electronic Science and Technology of China); Leiting Chen (School of Computer Science and Engineering, University of Electronic Science and Technology of China); Chuan Zhou (School of Computer Science and Engineering, University of Electronic Science and Technology of China) \*; Fan Yang (Group 42); Xin Li (Group 42)
- 
- P287 **Text is NOT Enough: Integrating Visual Impressions into Open-domain Dialogue Generation**  
Author(s): Lei Shen (Institute of Computing Technology, Chinese Academy of Sciences)\*; Haolan Zhan (University of Chinese Academy of Sciences); Xin Shen (Austrian National University); Yonghao Song (Institute of Computing Technology, Chinese Academy of Sciences); Xiaofang Zhao (Institute of Computing Technology, Chinese Academy of Sciences)
- 
- P288 **Quality Assessment of End-to-End Learned Image Compression: The Benchmark and Objective Measure**  
Author(s): Yang Li (Peking University) \*; Shiqi Wang (City University of Hong Kong); xinfeng zhang (University of Chinese Academy of Sciences); Shanshe Wang (Peking University); Siwei Ma (Peking University, China); Yue Wang (Beijing ByteDance Technology Co., Ltd.)
- 
- P289 **A Statistical Approach to Mining Semantic Similarity for Deep Unsupervised Hashing**  
Author(s): Xiao Luo (Peking University); Daqing Wu (Peking University); Zeyu Ma (Harbin Institute of Technology, Shenzhen); Chong Chen (Alibaba)\*; Minghua Deng (Peking University); Jianqiang Huang (Alibaba Group); Xian-Sheng Hua (Damo Academy, Alibaba Group)
- 
- P290 **BAM: Bilateral Activation Mechanism for Image Fusion**  
Author(s): Zi-Rong Jin (School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China); Liang-Jian Deng (School of Mathematical Sciences, University of Electronic Science and Technology of China)\*; Tian-Jing Zhang (Yingcai Honors College, University of Electronic Science and Technology of China); Xiaoxu Jin (Yingcai Honors College, University of Electronic Science and Technology of China)
- 
- P291 **Hallucinating Statistical Moment and Subspace Descriptors for Action Recognition**  
Author(s): Lei Wang (The Australian National University) \*; Piotr Koniusz (ANU College of Engineering and Computer Science)
- 



- P292      **Learning Multi-Granular Spatio-Temporal Graph Network for Skeleton-based Action Recognition**  
Author(s): Tailin Chen (Newcastle University) \*; Desen Zhou (Baidu, Inc.); Jian Wang (Baidu); Shidong Wang (Newcastle University); Yu Guan (Newcastle University); Xuming He (ShanghaiTech University); Errui Ding (Baidu Inc.)
- 
- P293      **ION: Instance-level Object Navigation**  
Author(s): Weijie Li (ICT, China Academy of Science) \*; Xinhang Song (ICT); Yubing Bai (ICT, China Academy of Science); Sixian Zhang (ICT, China Academy of Science); Shuqiang Jiang (ICT, China Academy of Science)
- 
- P294      **Skeleton-Aware Neural Sign Language Translation**  
Author(s): Shiwei Gan (Nanjing University); Yafeng Yin (Nanjing University) \*; Zhiwei Jiang (Nanjing University); Lei Xie (Nanjing University); Sanglu Lu (NJU)
- 
- P295      **Fingerspelling Recognition in the Wild with Fixed-Query based Visual Attention**  
Author(s): Srinivas Kruthiventi S S (Harman International India Pvt. Ltd.) \*; George Jose (Harman International India Pvt. Ltd.); Nitya Tandon (Harman International (India) Pvt. Ltd.); Rajesh Biswal (Harman International India Pvt. Ltd.); AASHISH KUMAR (HARMAN INDIA)
- 
- P296      **Deep Human Dynamics Prior**  
Author(s): Qiongjie Cui (Nanjing University of Science and Technology) \*; Huaijiang Sun (Nanjing University of Science and Technology); Yue Kong (Nanjing University of Science and Technology); Xiaoning Sun (Nanjing University of Science and Technology)
- 
- P297      **Exploiting Invariance of Mining Facial Landmarks**  
Author(s): Jiangming Shi (Ningxia University); Zixian Gao (Ningxia University); Hao Liu (Ningxia University) \*; Zekuan Yu (Fudan University); Fengjun Li (Ningxia University)
- 
- P298      **Joint Implicit Image Function for Guided Depth Super-Resolution**  
Author(s): Jiaxiang Tang (Peking University) \*; Xiaokang Chen (Peking University); Gang Zeng (Peking University)
- 
- P299      **Transformer-based Feature Reconstruction Network for Robust Multimodal Sentiment Analysis**  
Author(s): Ziqi Yuan (State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China); Wei Li (State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China); Hua Xu (State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)\*; Wenmeng Yu (State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

- P300      **Self-feature Learning: An Efficient Deep Lightweight Network for Image Super-Resolution**  
 Author(s): Jun Xiao (The Hong Kong Polytechnic University) \*; Qian Ye (Tohoku University); Rui Zhao (The Hong Kong Polytechnic University); Kin-Man Lam (The Hong Kong Polytechnic University); Kao Wan (the Peng Cheng Laboratory)
- 
- P301      **DAWN: Dynamic Adversarial Watermarking of Neural Networks**  
 Author(s): Sebastian Szyller (Aalto University) \*; Buse Gul Atli (Aalto University); Samuel Marchal (Aalto University); N. Asokan (University of Waterloo)
- 
- P302      **Visible Watermark Removal via Self-calibrated Localization and Background Refinement**  
 Author(s): Jing Liang (Shanghai Jiao Tong University); Li Niu (Shanghai Jiao Tong University) \*; Fengjun Guo (IntSig Information Co. Ltd); Teng Long (INTSIG); Liqing Zhang (Shanghai Jiao Tong University)
- 
- P303      **Learning to Decode Contextual Information for Efficient Contour Detection**  
 Author(s): Ruoxi Deng (WenZhou University) \*; Shengjun Liu (Central South University); Jinxin Wang (Wenzhou University); Huibing Wang (Dalian Maritime University); Hanli Zhao (Wenzhou University); Xiaoqin Zhang (-)
- 
- P304      **Fast, High-Quality Hierarchical Depth-Map Super-Resolution**  
 Author(s): Yiguo Qiao (University of Bath) \*; Licheng Jiao (Xidian University); Wenbin Li (University of Bath); Christian Richardt (University of Bath); Darren Cosker (University of Bath)
- 
- P305      **TsFPS: An Accurate and Flexible 6DoF Tracking System with Fiducial Platonic Solids**  
 Author(s): NAN XIANG (Bournemouth University) \*; Xiaosong Yang (Bournemouth University); Jian J Zhang (Bournemouth University)
- 
- P306      **Consistency-Constancy Bi-Knowledge Learning for Pedestrian Detection in Night Surveillance**  
 Author(s): Xiao Wang (Wuhan University); Zheng Wang (The University of Tokyo) \*; Wu Liu (AI Research of JD.com); Xin Xu (Wuhan University of Science and Technology); Jing Chen (Wuhan University); Chia-Wen Lin (NTHU)
- 
- P307      **SSconv: Explicit Spectral-to-Spatial Convolution for Pansharpening**  
 Author(s): Yudong Wang (Yingcai Honors college, University of Electronic Science and Technology of China); Liang-Jian Deng (School of Mathematical Sciences, University of Electronic Science and Technology of China) \*; Tian-Jing Zhang (Yingcai Honors College, University of Electronic Science and Technology of China); Xiao Wu (University of Electronic Science and Technology of China)
- 
- P308      **TriTransNet:RGB-D salient object detection with a triplet transformer embedding network**  
 Author(s): Zhengyi Liu (Anhui University) \*; Yuan Wang (Anhui University); Zhengzheng Tu (Anhui University); Yun Xiao (Anhui University); Bin Tang (Hefei University)
- 
- P309      **Learning Sample-Specific Policies for Sequential Image Augmentation**  
 Author(s): Pu Li (San Diego State University)\*; Xiaobai Liu (san diego state university); Xiaohui Xie (University of California, Irvine)



- P310 **Image Quality Caption with Attentive and Recurrent Semantic Attractor Network**  
Author(s): Wen Yang (Xidian University); Jinjian Wu (Xidian University)\*; Leida Li (Xidian University); Weisheng Dong (Xidian University); Guangming Shi (Xidian University)
- 
- P311 **Triangle-Reward Reinforcement Learning: A Visual-Linguistic Semantic Alignment for Image Captioning**  
Author(s): Weizhi Nie (Tianjin University); Jiesi Li (Tianjin University); Ning Xu (Tianjin University)\*; An-An Liu (Tianjin University); Xuanya Li (Baidu); Yongdong Zhang (University of Science and Technology of China)
- 
- P312 **Stacked Semantically-Guided Learning for Image De-distortion**  
Author(s): Huiyuan Fu (Beijing University of Posts and Telecommunications)\*; Changhao Tian (Beijing University of Posts and Telecommunications); Xin Wang (Stony Brook University); Huadong Ma (Beijing University of Posts and Telecommunications)
- 
- P313 **Focal and Composed Vision-semantic Modeling for Visual Question Answering**  
Author(s): Yudong Han (Shandong University)\*; Yangyang Guo (National University of Singapore); Jianhua Yin (Shandong University); Meng Liu (Shandong Jianzhu University); Yupeng Hu (Shandong University); Liqiang Nie (Shandong University)
- 
- P314 **Pose-Guided Feature Learning with Knowledge Distillation for Occluded Person Re-Identification**  
Author(s): Kecheng Zheng (University of Science and Technology of China)\*; Cuiling Lan (Microsoft Research); Wenjun Zeng (Microsoft Research); Jiawei Liu (University of Science and Technology of China); Zhizheng Zhang (Microsoft Research Asia); Zheng-Jun Zha (University of Science and Technology of China)
- 
- P315 **Multiple Objects-Aware Visual Question Generation**  
Author(s): Jiayuan Xie (School of Software Engineering, South China University of Technology); Yi Cai (School of Software Engineering, South China University of Technology)\*; Qingbao Huang (1.School of Electrical Engineering, Guangxi University, Nanning, Guangxi, China 2.School of Software Engineering, South China University of Technology, Guangzhou, Guangdong, China ); Tao Wang ( )
- 
- P316 **VASTile: Viewport Aware Scalable 360-Degree Video Frame Tiling.**  
Author(s): Chamara Manoj Madarasingha Kattadige (The University of Sydney)\*; Kanchana Thilakarathna (University of Sydney)
- 
- P317 **Delving into Deep Image Prior for Adversarial Defense: A Novel Reconstruction-based Defense Framework**  
Author(s): Li Ding (University of British Columbia)\*; Yongwei Wang (University of British Columbia); Xin Ding (University of British Columbia); Kaiwen Yuan (University of British Columbia); Ping Wang (Xi'an Jiaotong University); Hua Huang (Beijing Normal University); Z. Jane Wang (University of British Columbia)

- P318 **Fine-Grained Language Identification in Scene Text Images**  
Author(s): Yongrui Li (University of Science and Technology of China)\*; shilian wu (University of Science and Technology of China); Jun Yu (University of Science and Technology of China); Zengfu Wang ( Institute of Intelligent Machines, Chinese Academy of Sciences)
- 
- P319 **CARE: Cloudfied Android OSes on the Cloud Rendering**  
Author(s): Dongjie Tang (Shanghai Jiao Tong University); cathy bao (Intel); yong yao (Intel); Chao Xie (Intel); Qiming Shi (Intel); Marc Mao (Intel); randy xu (Intel); Linsheng Li (SJTU); Mohammad Reza R Haghighat (Intel); Zhengwei Qi (SJTU)\*; Haibing Guan (Shanghai Jiao Tong University)
- 
- P320 **Context-Aware Selective Label Smoothing for Calibrating Sequence Recognition Models**  
Author(s): Shuangping Huang (South China University of Technology); Yu Luo (South China University of Technology)\*; zhenzhou zhuang (South China University of Technology); Jin-Gang Yu (South China University of Technology); Mengchao He (DAMO Academy, Alibaba Group); Yongpan Wang (Alibaba Group)
- 
- P321 **Image Search with Text Feedback by Deep Hierarchical Attention Mutual Information Maximization**  
Author(s): Chunbin Gu (Zhejiang University)\*; Jiajun Bu (Zhejiang University); Zhen Zhang (Zhejiang University); Zhi Yu (Zhejiang University); Dongfang Ma (Zhejiang University); Wei Wang (Zhejiang University)
- 



2:00 AM-5:00 AM

Poster Session 6

▼ Hall Way

P322

**Pairwise VLAD Interaction Network for Video Question Answering**

Author(s): Hui Wang (Hefei University of Technology); Dan Guo (Hefei University of Technology) \*; Xian-Sheng Hua (Damo Academy, Alibaba Group); Meng Wang (Hefei University of Technology)

P323

**Attention-guided Temporally Coherent Video Object Matting**

Author(s): Yunke Zhang (Zhejiang University); Chi Wang (Zhejiang University); Miaomiao Cui (Alibaba-inc); Peiran Ren (Alibaba); Xuansong Xie (Alibaba); Xian-Sheng Hua (Damo Academy, Alibaba Group); Hujun Bao (Zhejiang University); Qixing Huang (The University of Texas at Austin); Weiwei Xu (Zhejiang unviersity)\*

P324

**Disentangling Hate in Online Memes**

Author(s): Roy Ka-Wei Lee (Singapore University of Technology and Design) \*; Rui Cao (Singapore Management University); Fan Ziqing (UESTC); Jing Jiang (Singapore Management University); Wen Haw Chong (SMU)

P325

**Robust Real-World Image Super-Resolution against Adversarial Attacks**

Author(s): Jiutao Yue (Sun Yat-sen University); Haofeng Li (Shenzhen Research Institute of Big Data, The Chinese University of Hong Kong (Shenzhen)); Pengxu Wei (Sun Yat-sen University); Guanbin Li (Sun Yat-sen University) \*; Liang Lin (Sun Yat-sen University)

P326

**Towards Robust Deep Hiding Under Non-Differentiable Distortions for Practical Blind Watermarking**

Author(s): Chaoning Zhang (KAIST)\*; Adil Karjauv (KAIST); Philipp Benz (KAIST); In So Kweon (KAIST)

P327

**Bottom-Up and Bidirectional Alignment for Referring Expression Comprehension**

Author(s): Liuwu Li (South China University of Technology); Yuqi Bu (South China University of Technology); Yi Cai (School of Software Engineering, South China University of Technology) \*

P328

**SaIS-GAN: Spatially-Adaptive Latent Space in StyleGAN for Real Image Embedding**

Author(s): Lingyun Zhang (Xi'an Jiaotong University); Xiuxiu Bai (Xi'an Jiaotong University)\*; Yao Gao (Xi'an Jiaotong University)

P329

**Structured Multi-modal Feature Embedding and Alignment for Image-Sentence Retrieval**

Author(s): Xuri Ge (University of Glasgow) \*; Fuhai Chen (National University of Singapore); Joemon Jose (University of Glasgow); zhilong ji (Tomorrow Advancing Life); Zhongqin Wu (Tomorrow Advancing Life); Xiao Liu (Tomorrow Advancing Life)

P330

**Keyframe Extraction from Motion Capture Sequences with Graph based Deep Reinforcement Learning**

Author(s): Clinton A Mo (The University of Sydney) \*; Kun Hu (The Univeristy of Sydney); Shaohui Mei (Northwestern Polytechnical University, China); Zebin Chen (Electronic Arts); Zhiyong Wang (The University of Sydney)

OCT.  
23

- P331 **Dense Contrastive Visual-Linguistic Pretraining**  
 Author(s): Lei Shi (Beijing University of Posts and Telecommunications (BUPT))\*; Kai Shuang (State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China.); Shijie Geng (Rutgers University); Peng Gao (Chinese university of hong kong); Zuohui Fu (Rutgers University); Gerard de Melo (Rutgers University); Yunpeng Chen (National University of Singapore); Sen Su (Beijing University of Posts and Telecommunications)
- P332 **Hybrid Reasoning Network for Video-based Commonsense Captioning**  
 Author(s): Weijiang Yu (Sun Yat-sen University) \*; Jian Liang (Peking University); Lei Ji (Microsoft); Lu Li (Zhejiang University); Yuejian Fang (Peking University); Nong Xiao (Sun Yat-sen University); Nan Duan (Microsoft Research)
- P333 **Learning Regularizer for Monocular Depth Estimation with Adversarial Guidance**  
 Author(s): Guibao Shen (Shenzhen Institutes of Advanced Technology, 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); Mingqiang Wei (Nanjing University of Aeronautics and Astronautics); Qiong Wang (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences)\*; Guangyong Chen (Shenzhen Institutes of Advanced Technology); Pheng-Ann Heng (The Chinese University of Hong Kong)
- P334 **Pixel-wise Graph Attention Networks for Person Re-identification**  
 Author(s): Wenyu Zhang (University of Science and Technology of China); Qing Ding (University of Science and Technology of China) \*; Jian Hu (University of Science and Technology of China); Yi Ma (University of Science and Technology of China); Mingzhe Lu (University of Science and Technology of China)
- P335 **Neighbor-Vote: Improving Monocular 3D Object Detection through Neighbor Distance Voting**  
 Author(s): Xiaomeng Chu (University of Science and Technology of China); Jiajun Deng (University of Science and Technology of China); Yao Li (University of Science and Technology of China); Zhenxun Yuan (The University of Sydney); Yanyong Zhang (University of Science and Technology of China)\*; Jianmin Ji (University of Science and Technology of China); Yu Zhang (University of Science and Technology of China)
- P336 **Remember and Reuse: Cross-Task Blind Image Quality Assessment via Relevance-aware Incremental Learning**  
 Author(s): Rui Ma (University of Electronic Science and Technology of China); Hanxiao Luo (UESTC (University of Electronic Science and Technology of China)); Qingbo Wu (University of Electronic Science and Technology of China)\*; King Ng (University of Electronic Science and Technology of China); Hongliang Li (University of Electronic Science and Technology of China); Fanman Meng (University of Electronic Science and Technology of China); LinFeng Xu (University of Electronic Science and Technology of China)



- P337 **MSO: Multi-Feature Space Joint Optimization Network for RGB-Infrared Person Re-Identification**  
Author(s): Yajun Gao (Beijing Jiaotong University); Tengfei Liang (Beijing Jiaotong University); Yi Jin (Beijing JiaoTong University) \*; Xiaoyan Gu (Institute of Information Engineering, Chinese Academy of Sciences); Wu Liu (AI Research of JD.com); Yidong Li (Beijing Jiaotong Univeristy, China); Congyan Lang (Beijing Jiaotong University)
- P338 **Point Cloud Projection and Multi-Scale Feature Fusion Network Based Blind Quality Assessment for Colored Point Clouds**  
Author(s): Wen-xu Tao (Ningbo University); Gang-yi Jiang (Ningbo University) \*; Zhi-di Jiang (Ningbo University); Mei Yu (Ningbo University)
- P339 **Multi-branch Channel-wise Enhancement Network for Fine-grained Visual Recognition**  
Author(s): Guangjun Li (University of Shanghai for Science and Technology) \*; Yongxiong Wang (University of Shanghai for Science and Technology); Fengting Zhu (Shanghai University)
- P340 **General Approximate Cross Validation for Model Selection: Supervised, Semi-supervised and Pairwise Learning**  
Author(s): Bowei Zhu (Renmin University of China); Yong Liu (Renmin University of China) \*
- P341 **Progressive and Selective Fusion Network for High Dynamic Range Imaging**  
Author(s): Qian Ye (Tohoku University) \*; Jun Xiao (The Hong Kong Polytechnic University); Kin-Man Lam (The Hong Kong Polytechnic University); Takayuki Okatani (Tohoku University/RIKEN AIP)
- P342 **Multimodal Relation Extraction with Efficient Graph Alignment**  
Author(s): Changmeng Zheng (South China University of Technology); Junhao Feng (South China University of Technology); Ze Fu (School of Software Engineering, South China University of Technology); Yi Cai (School of Software Engineering, South China University of Technology) \*; Qing Li (The Hong Kong Polytechnic University); Tao Wang ()
- P343 **Legitimate Adversarial Patches: Evading Human Eyes and Detection Models in the Physical World**  
Author(s): Jia Tan (Xiangtan University) \*; Nan Ji (Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology); Haidong Xie (Qian Xuesen Laboratory, China Academy of Space Technology); Xueshuang Xiang (Qian Xuesen Laboratory of Space Technology)
- P344 **Unsupervised Vehicle search in the Wild: A New Benchmark**  
Author(s): Xian Zhong (Wuhan University of Technology); shilei zhao (Wuhan University of Technology); Xiao Wang (Wuhan University) \*; kui jiang (Wuhan University); Wenxuan Liu (Wuhan University of Technology); Wenxin Huang (Hubei University); Zheng Wang (Wuhan University)
- P345 **Meta-FDMixup: Cross-Domain Few-Shot Learning Guided by Labeled Target Data**  
Author(s): Yuqian Fu (Fudan University) \*; Yanwei Fu (Fudan University); Yu-Gang Jiang (Fudan University)

- P346 **Target-guided Adaptive Base Class Reweighting for Few-Shot Learning**  
Author(s): Jiliang Yan (Harbin Institute of Technology); Deming Zhai (Harbin Institute of Technology); Junjun Jiang (Harbin Institute of Technology); Xianming Liu (Harbin Institute of Technology) \*
- P347 **Dynamic Reasoning Network for Few-shot Semantic Segmentation**  
Author(s): Yunzhi Zhuge (University of Adelaide); Chunhua Shen (University of Adelaide) \*
- P348 **Heterogeneous Feature Fusion and Cross-modal Alignment for Composed Image Retrieval**  
Author(s): Gangjian Zhang (Beijing Jiaotong University); Shikui Wei (Beijing Jiaotong University)\*; Huaxin Pang (BJTU); Yao Zhao (Beijing Jiaotong University)
- P349 **Similar Scenes arouse Similar Emotions: Parallel Data Augmentation for Stylized Image Captioning**  
Author(s): Guodun Li (Zhejiang University); Yuchen Zhai (Zhejiang University); Zehao Lin (Zhejiang University); Yin Zhang (Zhejiang University)\*
- P350 **Trajectory is not Enough: Hidden Following Detection**  
Author(s): Danni Xu (Wuhan University)\*; Ruimin Hu (Wuhan University); Zixiang Xiong (Texas A&M University); Zheng Wang (The University of Tokyo); Linbo Luo (Xidian University); Dengshi Li (Jiangnan University)
- P351 **Contrastive Learning for Cold-Start Recommendation**  
Author(s): Yinwei Wei (National University of Singapore)\*; Xiang Wang (National University of Singapore); Qi Li (Shandong University); Liqiang Nie (Shandong University ); Yan Li (Kuaishou); Xuanping Li (Kuaishou Technology); Tat-Seng Chua (National university of Singapore)
- P352 **CG-GAN: Class-Attribute Guided Generative Adversarial Network for Old Photo Restoration**  
Author(s): Jixin Liu (Tianjin University)\*; Rui Chen (Tianjin University); Shipeng An (Tianjin University); Heng Zhang (Tianjin University)
- P353 **Get The Best of the Three Worlds: Real-Time Neural Image Compression in a Non-GPU Environment**  
Author(s): Zekun Zheng (National University of Defense Technology)\*; Xiaodong Wang (National University of Defense Technology); Xinye Lin (National Key Laboratory of Science and Technology on Blind Signal Processing); Shaoh Lv (Seeingdo tech. corp)
- P354 **Visual Language Based Succinct Zero-Shot Object Detection**  
Author(s): Ye Zheng (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences)\*; Xi Huang (Institute of computing technology of the Chinese Academy of Sciences); Li Cui ( Institute of computing technology of the Chinese Academy of Sciences)
- P355 **GAMnet: Robust Feature Matching via Graph Adversarial-Matching Network**  
Author(s): Bo Jiang (Anhui University)\*; Pengfei Sun (Anhui University); Ziyang Zhang (Anhui University); Jin Tang (Anhui University); Bin Luo (Anhui University)



- P356 **MCCN: Multimodal Coordinated Clustering Network for Large-Scale Cross-modal Retrieval**  
Author(s): Zhixiong Zeng (Institute of Automation, Chinese Academy of Sciences)\*; Ying Sun (Institute of Automation, Chinese Academy of Sciences); Wenji Mao (Institute of Automation, Chinese Academy of Sciences)
- P357 **AFEC: Adaptive Feature Extraction Modules For Learned Image Compression**  
Author(s): Yi Ma (Peking university)\*; Yongqi Zhai (Peking University); Jiayu Yang (Peking University); Chunhui Yang (Peking University); Ronggang Wang (Peking University)
- P358 **How Video Super-Resolution and Frame Interpolation Mutually Benefit**  
Author(s): Chengcheng Zhou (Tsinghua University)\*; Zongqing Lu (Tsinghua University international Graduate School at Shenzhen); linge li (Huawei); qiangyu yan (Huawei); Jing-Hao Xue (University College London)
- P359 **FOCAS: Practical Video Super Resolution using Foveated Rendering**  
Author(s): Lingdong Wang (University of Massachusetts Amherst)\*; Mohammad Hajiesmaili (UMass Amherst); Ramesh K. Sitaraman (UMass Amherst & Akamai Technologies)
- P360 **Adaptive Affinity Loss and Erroneous Pseudo-Label Refinement for Weakly Supervised Semantic Segmentation**  
Author(s): Xiangrong Zhang (Xidian University)\*; zelin peng (xidian university); Peng Zhu (Xidian University); Tianyang Zhang (Xidian University); Chen Li (Xi'an Jiaotong University); Huiyu Zhou (University of Leicester); Licheng Jiao (Xidian University)
- P361 **Relationship-Preserving Knowledge Distillation for Zero-Shot Sketch Based Image Retrieval**  
Author(s): Jialin Tian (UESTC); Xing Xu (University of Electronic Science and Technology of China)\*; Zheng Wang (UESTC); Fumin Shen (UESTC); Xin Liu (Huaqiao University)
- P362 **Partially fake it till you make it: mixing real and fake thermal images for improved object detection**  
Author(s): Francesco Bongini (Università degli studi di Firenze); Lorenzo Berlincioni (University of Florence); Marco Bertini (University of Florence)\*; Alberto Del Bimbo (University of Florence)
- P363 **CDP: Towards Optimal Filter Pruning via Class-wise Discriminative Power**  
Author(s): Tianshuo Xu (Xiamen University); Yuhang Wu (Xiamen University); Xiwu Zheng (Xiamen University); Teng Xi (Baidu Inc.); gang zhang (Baidu Inc.); Errui Ding (Baidu Inc.); Fei Chao (School of Informatics)\*; Rongrong Ji (Xiamen University, China)
- P364 **Face Hallucination via Split-Attention in Split-Attention Network**  
Author(s): Tao Lu ( Wuhan Institute of Technology); Yuanzhi Wang (Wuhan Institute of Technology)\*; Yanduo Zhang (Wuhan Institute of Technology); Yu Wang (Wuhan Institute of technology); Liu Wei (SZU); Zhongyuan Wang (Wuhan University); Junjun Jiang (Harbin Institute of Technology)

P365	<p><b>Fake Gradient: A Security and Privacy Protection Framework for DNN-based Image Classification</b></p> <p>Author(s): Xianglong Feng (Rutgers University); Yi Xie (Rutgers University); Mengmei Ye (Rutgers University); Zhongze Tang (Rutgers University); bo yuan (rutgers university); Sheng Wei (Rutgers University)*</p>
P366	<p><b>Integrating Semantic and Temporal Relationships in Facial Action Unit Detection</b></p> <p>Author(s): Zhihua Li (Binghamton University)*; Xiang Deng (State University of New York at Binghamton); Xiaotian Li (Binghamton University); Lijun Yin (State University of New York at Binghamton)</p>
P367	<p><b>Sparse to Dense Depth Completion using a Generative Adversarial Network with Intelligent Sampling Strategies</b></p> <p>Author(s): Md Fahim Faysal Khan (Pennsylvania State University)*; Nelson D Troncoso Aldas (The Pennsylvania State University); Abhishek Kumar (Pennsylvania State University); Siddharth Advani (Samsung Research America); Vijaykrishnan Narayanan (Pennsylvania State University)</p>
P368	<p><b>How does Color Constancy Affect Target Recognition and Instance Segmentation?</b></p> <p>Author(s): Siyan Xue (Sichuan University); Shaobing Gao (Sichuan University)*; Minjie Tan (Sichuan University); Zhen He (Beijing Institute of Basic Medical Sciences); Liangtian He (Anhui University)</p>
P369	<p><b>Convolutional Transformer based Dual Discriminator Generative Adversarial Networks for Video Anomaly Detection</b></p> <p>Author(s): Xinyang Feng (Columbia University); Dongjin Song (University of Connecticut)*; Yuncong Chen (NEC Laboratories America, Inc.); Zhengzhang Chen (NEC Laboratories America, Inc.); Jingchao Ni (NEC Laboratories America); Haifeng Chen (NEC Labs)</p>
P370	<p><b>Salient Error Detection based Refinement for Wide-baseline Image Interpolation</b></p> <p>Author(s): Yuan Chang (Peking University)*; Yisong Chen (Peking University); Guoping Wang (Peking University)</p>
P371	<p><b>A Multi-Domain Adaptive Graph Convolutional Network for EEG-based Emotion Recognition</b></p> <p>Author(s): Rui Li (Shanghai Jiao Tong University)*; Yiting Wang (Center for Brain-like Computing and Machine Intelligence); Bao-Liang Lu (Shanghai Jiao Tong University)</p>
P372	<p><b>Interpolation Variable Rate Image Compression</b></p> <p>Author(s): Zhenhong Sun (Alibaba Group); Zhiyu Tan (Alibaba Group); Xiuyu Sun (Alibaba Group)*; Fangyi Zhang (Alibaba Group); Yichen Qian (Alibaba Group); Dongyang Li (Alibaba Group); Hao Li (Alibaba Group)</p>
P373	<p><b>Armor: A Benchmark for Testing the Effectiveness of Artificial Music Objective Evaluation Method</b></p> <p>Author(s): Songhe Wang (University of North Carolina at Chapel Hill)*; Zheng Bao (UNC Chapel Hill); Jingtong E (University of North Carolina at Chapel Hill)</p>



- P374 **DRDF: Determining the Importance of Different Multimodal Information with Dual-Router Dynamic Framework**  
Author(s): Haiwen Hong (Zhejiang University); Xuan Jin (Alibaba Turing Lab, Alibaba Group); Yin Zhang (Zhejiang University) \*; Yunqing Hu (Zhejiang University); Jingfeng Zhang (Zhejiang University); Yuan He (Alibaba Group); hui xue (Alibaba)
- P375 **CoCo-BERT: Improving Video-Language Pre-training with Contrastive Cross-modal Matching and Denoising**  
Author(s): Jianjie Luo (Sun Yat-set University); Yehao Li (Sun Yat-Sen University); Yingwei Pan (JD AI Research)\*; Ting Yao (JD AI Research); Hongyang Chao (Sun Yat-sen University); Tao Mei (AI Research of JD.com)
- P376 **DLA-Net for FG-SBIR: Dynamic Local Aligned Network for Fine-Grained Sketch-Based Image Retrieval**  
Author(s): Jiaqing Xu (Beijing University of Posts and Telecommunications)\*; Haifeng Sun (Beijing university of posts and telecommunications); Qi Qi (Beijing University of Posts and Telecommunications); Jingyu Wang (Beijing University of Posts and Telecommunications); Ce Ge (Beijing University of Posts and Telecommunications); Lejian Zhang (Beijing University of Posts and Telecommunications); Jianxin Liao (beijing university of posts and telecommunications)
- P377 **Pareto Optimality for Fairness-constrained Collaborative Filtering**  
Author(s): Qianxiu Hao (University of Chinese Academy of Science)\*; Qianqian Xu (Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences); Zhiyong Yang (SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences; SCS, University of Chinese Academy of Sciences); Qingming Huang (University of Chinese Academy of Sciences)
- P378 **Decoupled IoU Regression for Object Detection**  
Author(s): Yan Gao (Alibaba Youku Cognitive and Intelligent Lab)\*; Qimeng Wang (Huazhong University of Science and Technology); Xu Tang (Alibaba); Haochen Wang (Alibaba Youku Cognitive and Intelligent Lab); Fei Ding (Alibaba Youku Cognitive and Intelligent Lab); Jing Li (Alibaba); Yao Hu (Alibaba Youku Cognitive and Intelligent Lab)
- P379 **RCNet: Reverse Feature Pyramid and Cross-scale Shift Network for Object Detection**  
Author(s): Zhuofan Zong (Beihang University); Qianggang Cao (Beihang University); Biao Leng (Beihang University)\*
- P380 **Recursive Fusion and Deformable Spatiotemporal Attention for Video Compression Artifact Reduction**  
Author(s): Minyi Zhao (Fudan University); Yi Xu (Fudan University); Shuigeng Zhou (Fudan University)\*
- P381 **JokerGAN: Memory-Efficient Model for Handwritten Text Generation with Text Line Awareness**  
Author(s): Jan Zdenek (The University of Tokyo) \*; Hideki Nakayama (The University of Tokyo)

October 24, 2021

9:00 AM-12:00 PM	Room	Zoom Meeting ID	Website QR
<b>Workshop-ADGD 2021</b>	Tianfu Room A	817 9260 5901	
<ul style="list-style-type: none"> <li>✧ <i>DeepFake Detection with Inconsistent Head Poses: Reproducibility and Analysis</i></li> <li>✧ <i>Evaluation of an Audio-Video Multimodal Deepfake Dataset using Unimodal and Multimodal Detectors</i></li> <li>✧ <i>DmyT: Dummy Triplet Loss for Deepfake Detection</i></li> <li>✧ <i>Invertable Frowns: Modifying Affect in Existing Videos</i></li> </ul>			
<b>Workshop-AI &amp; Food 2021</b>	Tianfu Room B	864 1291 4191	
<ul style="list-style-type: none"> <li>✧ <i>A Generic Few-shot Solution for Food Shelf-Life Prediction using Meta-Learning</i></li> <li>✧ <i>Analyzing and Recognizing Food in Constrained and Unconstrained Environments</i></li> <li>✧ <i>An Integrated System for Mobile Image-Based Dietary Assessment</i></li> <li>✧ <i>3D Mesh Reconstruction of Foods from a Single Image</i></li> </ul>			
<b>Tutorial: Deep Learning for Visual Data Compression</b>	Jiaozi Room	870 2448 7906	
<b>Tutorial: Few-shot Learning for Multi-Modality Tasks</b>	Jincheng Room	852 4844 0073	
<b>Tutorial: A Tutorial on AI Music Composition</b>	Rongcheng Room	879 4652 5187	

OCT.  
24

2:00 PM-5:00 AM	Room	Zoom Meeting ID	Website QR
<b>Workshop-MuSe 2021</b>	Tianfu Room A	817 9260 5901	



- ✧ *A Physiologically-Adapted Gold Standard for Arousal during Stress*
- ✧ *Fusion of acoustic and linguistic information using supervised autoencoder for improved emotion recognition*
- ✧ *The MuSe 2021 Multimodal Sentiment Analysis Challenge: Sentiment, Emotion, Physiological-Emotion, and Stress*
- ✧ *MuSe-Toolbox: The Multimodal Sentiment Analysis Continuous Annotation Fusion and Discrete Class Transformation Toolbox*
- ✧ *Multimodal Fusion Strategies for Physiological-emotion Analysis*
- ✧ *Multi-modal Stress Recognition Using Temporal Convolution and Recurrent Network with Positional Embedding*
- ✧ *Multimodal Sentiment Analysis based on Recurrent Neural Network and Multi-modal Attention*
- ✧ *Multi-modal Fusion for Continuous Emotion Recognition by Using Auto-Encoders*
- ✧ *A Report for the MuSe-Wilder and MuSe-Sent sub-challenges in MuSe 2021*
- ✧ *Hybrid Mutimodal Fusion for Dimensional Emotion Recognition*

**sWorkshop-MULL 2021**

Tianfu Room B

864 1291 4191



- ✧ *Iterative Image Translation for Unsupervised Domain Adaptation*
- ✧ *Glocal Alignment for Unsupervised Domain Adaptation*
- ✧ *Occlusion Contrasts for Self-Supervised Facial Age Estimation*
- ✧ *Multi-Branch Convolution Network for Few-Shot Classification*
- ✧ *Incomplete label distribution learning by exploiting global sample correlation*
- ✧ *Improving Multimodal Data Labeling with Deep Active Learning for Post Classification in Social Networks*
- ✧ *Contextual Image Parsing via Panoptic Segment Sorting*

**Workshop-FME 2021**

Jiaozi Room

870 2448 7906



- ✧ *FaceTrack: Asymmetric Facial and Gesture Analysis Tool for Speech Language Pathologist Applications*
- ✧ *Transfer Spatio-Temporal Knowledge from Emotion-Related Tasks for Facial Expression Spotting*
- ✧ *Facial Action Unit Detection with Local Key Facial Sub-region based Multi-label Classification for Micro-expression Analysis*
- ✧ *Spatio-temporal Convolutional Attention Network for Spotting Macro- and Micro-expression Intervals*
- ✧ *A Brief Guide: Code for Spontaneous Expressions and Micro-Expressions in Videos*

**Tutorial: Image Quality Assessment in the Modern Age**

Jincheng Room

852 4844 0073

**Tutorial: Plenoptic Quality Assessment: The JPEG Pleno Experience**

Rongcheng Room

879 4652 5187