

	#	Paper title	Authors	affiliation
<b>GMT+8</b>		<b>ICMR 2021 Workshop Program</b>		
<b>Nov. 16</b>				
<b>9:00–17:00</b>		<b>Workshop on Intelligent Cross-Data Analysis and Retrieval</b>		
<b>9:00–12:00</b>		<b>Workshop on Multimedia for Cooking and Eating Activities</b>		
<b>9:00–12:00</b>		<b>Workshop on Multi-Modal Pre-Training for Multimedia Understanding</b>		
<b>13:00–17:00</b>		<b>Joint Workshop on Multimedia Artworks Analysis and Attractiveness Computing in Multimedia</b>		
<b>GMT+8</b>		<b>ICMR 2021 Main Program</b>		
<b>Nov. 17</b>				
<b>8:30–9:00</b>		<b>Opening</b>		
<b>9:00–10:00</b>		<b>Keynote 1: Allen Lu, MediaTek Inc.</b>		
<b>11:00–12:00</b>	<b>SS1</b>	<b>SS on Content-Based Medical Image Retrieval</b>	<b>Session Chair: Jian Wang (jwang@sdnu.edu.cn)</b>	
	icmrsp200	M-DFNet: Multi-phase Discriminative Feature Network for Retrieval of Focal Liver Lesions	Yingying Xu, Jing Liu, Lanfen Lin, Hongjie Hu, Ruofeng Tong, Jingsong Li and Yen-Wei Chen	Research Center for Healthcare Data Science, Zhejiang Lab, College of Computer Science and Technology, Zhejiang University, Department of Radiology, Sir Run Run Shaw Hospital, Zhejiang University, College of Information Science and Engineering, Ritsumeikan University
	icmrsp053	Contextualized Keyword Representations for Multi-modal Retinal Image Captioning	Jia-Hong Huang, Ting-Wei Wu and Marcel Worring	University of Amsterdam, Georgia Institute of Technology
	icmrsp204	A Tensor Sparse Representation-Based CBMR System for Computer-Aided Diagnosis of Focal Liver Lesions and Its Pilot Trial	Jian Wang, Xian-Hua Han, Lanfen Lin, Hongjie Hu and Yen-Wei Chen	Shandong Normal University, Yamaguchi University, Zhejiang University, Ritsumeikan University, Peking University
<b>11:00–12:30</b>	<b>SS2</b>	<b>SS on Delving into Vision and Language Intelligence</b>	<b>Session Chair: Huibing Wang (huibing.wang@dlnu.edu.cn)</b>	
	icmrsp101	Human Pose Estimation based on Attention Multi-resolution Network	Congcong Zhang, Ning He, Qixiang Sun, Xiaojie Yin and Ke Lu	Beijing Union University, Beijing Union University, Beijing Union University, Beijing Union University, University of Chinese Academy of Sciences
	icmrsp112	MSAV: An Unified Framework for Multi-view Subspace Analysis with View Consistency	Huibing Wang, Guangqi Jiang, Jinjia Peng and Xianping Fu	Dalian Maritime University, Dalian Maritime University, Hebei University, Dalian Maritime University & Pengcheng Laboratory
	icmrsp141	M <sup>2</sup> GUDA: Multi-Metrics Graph-Based Unsupervised Domain Adaptation for Cross-Modal Hashing	Chengyuan Zhang, Zhi Zhong, Lei Zhu, Shichao Zhang, Da Cao and Jianfeng Zhang	Hunan University, Hunan University, Hunan Agricultural University, Central South University, Hunan University, National University of Defense Technology
	icmrsp159	Visible-infrared Person Re-identification with Human Body Parts Assistance	Huangpeng Dai, Qing Xie, Jiachen Li, Yanchun Ma, Lin Li and Yonglian Liu	Wuhan University of Technology, Wuhan University of Technology, Wuhan University of Technology, Wuhan University of Technology, Wuhan University of Technology
	icmrsp217	Look Back Again: Dual Parallel Attention Network for Accurate and Robust Scene Text Recognition	Zilong Fu, Hongtao Xie, Guoqing Jin and Junbo Guo	University of Science and Technology of China, State Key Laboratory of Communication Content Cognition, People's Daily Online, Beijing, China 100733
<b>13:30–15:00</b>	<b>Oral 1</b>	<b>Cross Modal Retrieval</b>	<b>Session Chair: Chih-Chung Hsu (cchs@gs.ncku.edu.tw)</b>	
	icmr065	Global Relation-Aware Attention Network for Image-Text Retrieval	Jie Cao, Shengsheng Qian, Hualwen Zhang, Quan Fang and Changsheng Xu	School of Artificial Intelligence, Institute of Chinese Academy of Sciences, National Lab of Pattern Recognition Institute of Automation, Chinese Academy of Sciences
	icmr047	Multi-Feature Graph Attention Network for Cross-Modal Video-Text Retrieval	Xiaoshuai Hao, Yucan Zhou, Dayan Wu, Wanqian Zhang, Bo Li and Weiping Wang	Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences, Institute of Information Engineering, Chinese Academy of Sciences, Institute of Information Engineering, Chinese Academy of Sciences, Institute of Information Engineering, Chinese Academy of Sciences, Institute of Information Engineering, Chinese Academy of Sciences
	icmr106	TEACH: Attention-Aware Deep Cross-Modal Hashing	Hong-Lei Yao, Yu-Wei Zhan, Zhen-Duo Chen, Xin Luo and Xin-Shun Xu	Shandong University, Shandong University, Shandong University, Shandong University
	icmr213	HSGMP: Heterogeneous Scene Graph Message Passing for Cross-modal Retrieval	Yu Duan, Yun Xiong, Yao Zhang, Yuwei Fu and Yangyong Zhu	Fudan University, Fudan University Fudan University, Fudan University, Fudan University & Shanghai Institute for Advanced Communication and Data Science
	icmr109	Unsupervised Deep Cross-Modal Hashing by Knowledge Distillation for Large-scale Cross-modal Retrieval	Mingyong Li and Hongya Wang	Donghua University & Chongqing Normal University, Donghua University & Shanghai Key Laboratory of Computer Software Evaluating and Testing
<b>13:30–15:00</b>	<b>Challenge 1</b>	<b>Grand Challenges on ROD</b>	<b>Session Chair: Yizhou Wang (ywang26@uw.edu)</b>	
	icmrch237	Radar Object Detection Using Data Merging, Enhancement and Fusion	Jun Yu, Xinlong Hao, Xinjian Gao, Qiang Sun, Yuyu Liu, Peng Chang, Zhong Zhang, Fang Gao and Feng Shuang	University of Science and Technology of China, Ping An Technology Co., Ltd, PAII labs, Palo Alto, Hefei ZhanDa Intelligence Technology Co., Ltd and Guangxi University
	icmrch238	Squeeze-and-Excitation network-Based Radar Object Detection With Weighted Location Fusion	Pengliang Sun, Xuelong Niu, Pengfei Sun and Kele Xu	The Chinese University of Hong Kong, King's College London, Institute of Neuroinformatics, University of Zurich and ETH Zurich, National Key Laboratory of Parallel and Distributed Processing
	icmrch239	Scene-aware Learning Network for Radar Object Detection	Zangwei Zheng, Xiangyu Yue, Kurt Keutzer and Alberto Sangiovanni Vincentelli	Nanjing University, University of California, Berkeley, University of California, Berkeley, University of California, Berkeley
	icmrch240	DANet: Dimension Apart Network for Radar Object Detection	Bo Ju, Wei Yang, Jinrang Jia, Xiaojing Ye, Qu Chen, Xiao Tan, Hao Sun, Yifeng Shi and Erui Ding	Baidu Inc., Baidu Inc., Baidu Inc., Baidu Inc., Baidu Inc., Baidu Inc.
	icmrch242	Efficient-ROD: Efficient Radar Object Detection based on Densely Connected Residual Network	Chih-Chung Hsu, Chieh Lee, Lin Chen, Min-Kai Hung, Yu-Lun Chen and Xian-Yu Wang	National Cheng Kung University, National Cheng Kung University, National Pingtung University of science and technology, National Pingtung University of Science and Technology, National Cheng Kung University
	icmrch245	ROD2021 Challenge: A Summary for Radar Object Detection Challenge for Autonomous Driving Applications	Yizhou Wang, Jeng-Neng Hwang, Gaoyang Wang, Hui Liu, Kwang-Ju Kim, Hung-Min Hsu, Jiarui Cai, Haotian Zhang, Zhongyu Jiang and Renshu Gu.	University of Washington, University of Washington, Zhejiang University, University of Washington & Silkwave Holdings Limited, Electronics and Telecommunications Research Institute (ETRI), University of Washington, University of Washington, University of Washington, University of Washington, Hangzhou Dianzi University
<b>15:00–16:30</b>	<b>Challenge 3</b>	<b>Real DSR Challenge</b>	<b>Session Chair: Yao Zhao (yzhao@bjtu.edu.cn), Runmin Cong (rmcong@bjtu.edu.cn)</b>	
<b>Nov. 18</b>				
<b>9:00–10:30</b>	<b>Oral 2</b>	<b>Image Synthesis and Analysis</b>	<b>Session Chair: Keiji Yanai (yanai@cs.uec.ac.jp)</b>	
	icmr054	Image-to-Image Transfer Makes Chaos to Order	Sanbi Luo, Tao Guo	Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences, Institute of Information Engineering, Chinese Academy of Sciences & University of Chinese Academy of Sciences
	icmr223	Generative Adversarial Networks with Bi-directional Normalization for Semantic Image Synthesis	Jia Long and Hongtao Lu	Shanghai Jiao Tong University, Shanghai Jiao Tong University
	icmr212	Nested Dense Attention Network for Single Image Super-Resolution	Cheng Qiu, Yirong Yao and Yuntao Du	Nanjing University, Nanjing University, Nanjing University
	icmr184	Leveraging EfficientNet and Contrastive Learning for Accurate Global-scale Location Estimation	Giorgos Kordopatis-Zilos, Panagiotis Galopoulos, Symeon Papadopoulos and Ioannis Kompatsiaris	Information Technologies Institute, CERTH, Information Technologies Institute, CERTH, Information Technologies Institute, CERTH
	icmr143	GCNBoost: Artwork Classification by Label Propagation through a Knowledge Graph	Cheikh Ibrahim El Vaigh, Noa Garcia, Benjamin Renoust, Chenhui Chu, Yuta Nakashima and Hajime Nagahara	Univ Rennes 1, IRISA, National University of Defense Technology, Osaka University, Institute for Datability Science, Median Technologies and Osaka University, Institute for Datability Science, Kyoto University, and Osaka University, Institute for Datability Science, Osaka University, Institute for Datability Science
<b>9:00–10:30</b>	<b>Oral 3</b>	<b>Image Recognition and Retrieval</b>	<b>Session Chair: Ichiro Ide (ide@i.nagoya-u.ac.jp)</b>	
	icmr077	G-CAM: Graph Convolution Network Based Class Activation Mapping for Multi-label Image Recognition	Yangtao Wang, Yanzhao Xie and Yu Liu	Guangzhou University, Huazhong University of Science and Technology, Huazhong University of Science and Technology
	icmr132	Know Yourself and Know Others: Efficient Common Representation Learning for Few-shot Cross-modal Retrieval	Shaoying Wang, Hanjiang Lai and Zhenyu Shi	Sun Yat-sen University, Sun Yat-Sen University, Sun Yat-Sen University
	icmr189	A Denoising Convolutional Neural Network for Self-Supervised Rank Effectiveness Estimation on Image Retrieval	Lucas Valem and Daniel C. G. Pedronette	São Paulo State University, São Paulo State University, São Paulo State University
	icmr069	Cross-Modal Image-Recipe Retrieval via Intra- and Inter-Modality Hybrid Fusion	Jiao Li, Xing Xu, Wei Yu, Jialiang Sun and Fumin She	University of Electronic Science and Technology of China, University of Electronic Science and Technology of China, University of Electronic Science and Technology of China, University of Electron
	icmr161	Scene Text Recognition with Cascade Attention Network	Min Zhang, Meng Ma and Ping Wang	Peking University, Peking University, Ministry of Education
<b>9:00–10:30</b>	<b>Doc.</b>	<b>Doctoral Symposium</b>	<b>Session Chair: Chih-Yang Lin (andrewlin@saturn.yzu.edu.tw)</b>	
	icmrdoc229	Discrete Techebief Transform for Versatile Video Coding	Ka-Hou Chan	Macao Polytechnic Institute, Macao Polytechnic Institute
	icmrdoc236	Fire Detection using Transformer Network	Mohammad Shahid and Kai-Lung Hua	National Taiwan University of Science and Technology, National Taiwan University of Science and Technology
	icmrdoc244	A Beneficial Dual Transformation Approach for Deep Learning Networks Used in Steel Surface Defect Detection	Fityanul Akhyar, Chih-Yang Lin and Gagan S Kathiresan	Yuan Ze University & Telkom University, Yuan Ze University, Vellore Institute of Technology

11:00-12:00	Oral 4	Query, Reasoning, and Recommendation	Session Chair: Hong-Han Shuai (hhshuai@g2.nctu.edu.tw)	
	icmr136	Social Relation Analysis from Videos via Multi-entity Reasoning	Chenghao Yan, Zihe Liu, Fangtao Li, Chenyu Cao, Zheng Wang and Bin Wu	Beijing University of Posts and Telecommunications, Beijing University of Posts and Telecommunications, Beijing University of Posts and Telecommunications, Beijing University of Posts and Telecommunications, Beijing University of Posts and Telecommunications, Beijing University of Posts and Telecommunications
	icmr137	Relation-aware Hierarchical Attention Framework for Video Question Answering	Fangtao Li, Ting Bai, Chenyu Cao, Zihe Liu, Chenghao Yan and Bin Wu	Beijing University of Posts and Telecommunications, Beijing University of Posts and Telecommunications, Beijing University of Posts and Telecommunications, Beijing University of Posts and Telecommunications
	icmr195	Question-Guided Semantic Dual-Graph Visual Reasoning with Novel Answers	Xinzhe Zhou and Yadong Mu	Wangxuan Institute of Computer Technology, Peking University, Wangxuan Institute of Computer Technology, Peking University
	icmr129	Heterogeneous Side Information-based Iterative Guidance Model for Recommendation	Feifei Dai, Xiaoyu Gu, Zhuo Wang, Mingda Qian, Bo Li and Weiping Wang	University of Chinese Academy of Sciences & Institute of Information Engineering, Chinese Academy of Sciences, Institute of Information Engineering, Chinese Academy of Sciences, Sangfor Inc., University of Chinese Academy of Sciences & Institute of Information Engineering, Chinese Academy of Sciences, Institute of Information Engineering, Chinese Academy of Sciences, Institute of Information Engineering, Chinese Academy of Sciences
11:00-12:00	Challenge 2	Grand Challenges on Embedded Object Detection	Session Chair: Jiun-In Guo (jigu@nctu.edu.tw)	
	icmr313	Summary of the 2021 Embedded Deep Learning Object Detection Model Compression Competition for Traffic in Asian Countries	Yu-Shu Ni, Chia-Chi Tsai, Jiun-In Guo, Jenq-Neng Hwang, Bo-Xun Wu, Po-Chi Hu, Ted T. Kuo, Po-Yu Chen and Hsien-Kai Kuo	National Yang Ming Chiao Tung University, MediaTek Inc., National Yang Ming Chiao Tung University, University of Washington, National Yang Ming Chiao Tung University, Pervasive Artificial Intelligence Research Labs (PAIR Labs), National Yang Ming Chiao Tung University, MediaTek Inc., MediaTek Inc
	icmrch250	Bag of Tricks for Building an Accurate and Slim Object Detector for Embedded Applications	Yongkun Du, Zhiheng Chen, Caiyan Jia, Xuanya Li and Yu-Gang Jiang	Beijing Jiaotong University, Fudan University, Beijing Jiaotong University, Baidu Inc., Fudan University
	icmrch251	Embedded YOLO: Faster and Lighter Object Detection	Wu Wen-Kai, Chen Chien-Yu and Lee Jiann-Shu	National University of Tainan, National University of Tainan, National University of Tainan
	icmrch253	Object Detection on Embedded Systems for Traffic in Asian Countries	Bao-Hong Lai and Hsun-Ping Hsieh	National Cheng Kung University, National Cheng Kung University
13:00-14:00	Best	Best Paper Candidate Session	Session Chair: Marcel Worring (m.worring@uva.nl)	
	icmr007	RGB-D Scene Recognition based on Object-Scene Relation and Semantics-Preserving Attention	Yuhui Guo and Xun Liang	Renmin University of China, Renmin University of China
	icmr056	Reading Scene Text by Fusing Visual Attention with Semantic Representations	Zhiguang Liu, Liangwei Wang and Jian Qiao	Noah's Ark Lab & Huawei Technologies, Noah's Ark Lab & Huawei Technologies, Huawei Technologies
	icmr157	Leveraging Two Types of Global Graph for Sequential Fashion Recommendation	Yujuan Ding, Yunshan Ma, Wai Keung Wong and Tat-Seng Chua	Shenzhen University, National University of Singapore, The Hong Kong Polytechnic University, National University of Singapore
	icmr162	A Smart Adversarial Attack on Deep Hashing Based Image Retrieval	Junda Lu, Mingyang Chen, Yifang Sun, Wei Wang, Yi Wang and Xiaochun Yang	University of New South Wales, University of New South Wales, Northeastern University, Dongguan University of Technology, Northeastern University
13:00-15:30	Poster	Poster Session	Session Chair: Shintami Hidayati (shintami.c.hidayati@gmail.com)	
	icmr011s	2.5D Pose Guided Human Image Generation	Kang Yuan and Sheng Li	University of Georgia, Renmin University of China
	icmr023s	Weakly Supervised Sketch Based Person Search	Lan Yan, Wenbo Zheng, Fei-Yue Wang and Chao Gou	Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences, Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences, Institute of Automation, Chinese Academy of Sciences & University of Chinese Academy of Sciences, Sun Yat-sen University
	icmr025s	Personal Knowledge Base Construction from Multimodal Data	An-Zi Yen, Chia-Chun Chang, Hen-Hsen Huang and Hsin-Hsi Chen	National Taiwan University, National Taiwan University, Academia Sinica & MOST Joint Research Center for AI Technology and All Vista Healthcare, National Taiwan University & MOST Joint Research Center for AI Technology and All Vista Healthcare
	icmr028s	NMS-Loss: Learning with Non-Maximum Suppression for Crowded Pedestrian Detection	Zekun Luo, Zheng Fang, Sixiao Zheng, Yabiao Wang and Yanwei Fu	Youtu Lab, Tencent, Behang University, Academy for Engineering Technology, Fudan University, Youtu Lab, Fudan University
	icmr035s	Image Retrieval by Hierarchy-aware Deep Hashing Based on Multi-task Learning	Bowen Wang, Liangzhi Li, Yuta Nakashima, Takehiro Yamamoto, Hiroaki Ohshima, Yoshiyuki Shoji, Kenro Aihara and Noriko Kando	Osaka University, Osaka University, University of Hyogo, National Institute of Informatics, Aoyama Gakuin University, National Institute of Informatics, Aoyama Gakuin University, National Institute of Informatics, The Graduate University for Advanced Studies
	icmr050s	Spatio-Temporal Activity Detection and Recognition in Untrimmed Surveillance Videos	Konstantinos Gkountakos, Despoina Touska, Konstantinos Ioannidis, Theodora Tsirikia, Stefanos Vrochidis and Ioannis Kompatsiaris	Information Technologies Institute - Centre for Research and Technology Hellas, Information Technologies Institute - Centre for Research and Technology Hellas, Information Technologies Institute - Centre for Research and Technology Hellas, Information Technologies Institute - Centre for Research and Technology Hellas, Information Technologies Institute - Centre for Research and Technology Hellas, Information Technologies Institute - Centre for Research and Technology Hellas
	icmr064s	Cross-Modal Self-Attention with Multi-Task Pre-Training for Medical Visual Question Answering	Haifan Gong, Guanqi Chen, Sishuo Liu, Yizhou Yu and Guanbin Li	Sun Yat-sen University, Sun Yat-sen University, The University of Hong Kong, The University of Hong Kong
	icmr108s	Evaluating Contrastive Models for Instance-based Image Retrieval	Tarun Krishna, Kevin McGuinness and Noel O'Connor	Dublin City University, Dublin City University, Dublin City University
	icmr144s	Unsupervised Video Summarization via Multi-source Features	Hussain Kanafani, Junaid Ahmed Ghauri, Sherzod Hakimov and Ralph Ewerth	Leibniz University Hannover, TIB - Leibniz Information Centre for Science and Technology
	icmr163s	Text-Enhanced Attribute-Based Attention for Generalized Zero-Shot Fine-Grained Image Classification	Yan-He Chen and Mei-Chen Yeh	National Taiwan Normal University, National Taiwan Normal University
	icmr166s	Learning to Select: A Fully Attentive Approach for Novel Object Captioning	Learning to Select: A Fully Attentive Approach for Novel Object Captioning	University of Modena and Reggio Emilia, University of Modena and Reggio Emilia, University of Modena and Reggio Emilia, University of Modena and Reggio Emilia, University of Modena and Reggio Emilia
	icmr179s	Semi-supervised Many-to-many Music Timbre Transfer	Yu-Chen Chang, Wen-Cheng Chen and Min-Chun Hu	National Cheng Kung University, National Cheng Kung University, National Tsing Hua University
	icmr196s	Collaborative Representation for Deep Meta Metric Learning	Min Zhu, Wefeng Liu, Kai Zhang, Ye Li and Baodi Liu	China University of Petroleum (East China), China University of Petroleum (East China), China University of Petroleum (East China), Qilu University of Technology (Shandong Academy of Sciences), China University of Petroleum (East China)
	icmr199s	AWFA-LPD: Adaptive Weight Feature Aggregation for Multi-frame License Plate Detection	Xiaocheng Lu, Yuan Yuan and Qi Wang	School of Computer Science and School of Artificial Intelligence, Optics and Electronics (IOPEN), School of Computer Science and School of Artificial Intelligence, Optics and Electronics (IOPEN), School of Computer Science and School of Artificial Intelligence, Optics and Electronics (IOPEN)
	icmr216s	Body Shape Calculator: Understanding the Type of Body Shapes from Anthropometric Measurements	Shintami Chusnul Hidayati and Yeni Anistasyari	Insttitut Teknologi Sepuluh Nopember, Universitas Negeri Surabaya
	icmrp005	Reproducibility Companion Paper: Knowledge Enhanced Neural Fashion Trend Forecasting	Yunshan Ma, Yujuan Ding, Xun Yang, Lizi Liao, Wai Keung Wong and Tat-Seng Chua	National University of Singapore, The Hong Kong Polytechnic University
13:00-15:30	Demo	Demo Session	Session Chair: Yi-Ling Chen (yiling@mail.ntust.edu.tw)	
	icmrde231	Automatic Baseball Pitch Overlay	Ting-Hsuan Chou and Wei-Ta Chu	National Cheng Kung University, National Cheng Kung University
	icmrde232	Video Action Retrieval Using Action Recognition Model	Yuko Iinuma and Shirohichi Satoh	The University of Tokyo, National Institute of Informatics
	icmrde233	IR Questioner: QA-based Interactive Retrieval System	Rintaro Yanagi, Ren Togo, Takahiro Ogawa and Miki Haseyama	Hokkaido University, Hokkaido University, Hokkaido University, Hokkaido University
	icmrde241	MeTILDA: Platform for Melodic Transcription in Language Documentation and Application	Mitchell Lee, Praveena Avula and Min Chen	University of Washington Bothell, University of Washington Bothell, University of Washington Bothell
14:00-15:30	Oral 5	Learning Strategies	Session Chair: Vasileios Mezaris (bmezaris@iti.gr)	
	icmr029	Multi-Initialization Graph Meta-Learning for Node Classification	Feng Zhao, Donglin Wang and Xintao Xiang	Zhejiang University & Westlake University, Westlake University & Westlake Institute for Advanced Study, Australian National University
	icmr036	MLFont: Few-Shot Chinese Font Generation via Deep Meta-Learning	Xu Chen, Lei Wu, Minggang He, Lei Meng and Xiangxu Meng	Shandong University, Shandong University, Shandong Survey and Design Institute of Water Conservancy, Shandong University & Key Laboratory of Shandong Province for Software Engineering, Shandong University
	icmr116	Combining Adversarial and Reinforcement Learning for Video Thumbnail Selection	Evangelos Apostolakis, Eleni Adamantidou, Vasileios Mezaris and Ioannis Patras	Information Technologies Institute - Centre for Research and Technology, Hellas
	icmr164	Aligning Visual Prototypes with BERT Embeddings for Few-Shot Learning	Kun Yan, Zied Bouraoui, Ping Wang, Shoab Jameel and Steven Schockaert	Peking University, CRIL CNRS & Univ Arts, University of Essex, Cardiff University
	icmr062	HINFShot: A Challenge Dataset for Few-Shot Node Classification in Heterogeneous Information Network	Zifeng Zhuang, Xintao Xiang, Siteng Huang and Donglin Wang	Westlake University, The Australian National University
14:00-15:30	Oral 6	Human Action and Face	Session Chair: Li-Wei Kang (lw kang@ntnu.edu.tw)	

	icmr113	Dense Scale Network for Crowd Counting	Feng Dai, Hao Liu, Yike Ma, Xi Zhang and Qiang Zhao	Beijing University of Posts and Telecommunications & Institute of Computing Technology, Chinese Academy of Sciences, Institute of Computing Technology, Chinese Academy of Sciences, Institute of Computing Technology, Beijing University of Posts and Telecommunications, Institute of Computing Technology, Chinese Academy of Sciences
	icmr116	Distractor-Aware Tracker with A Domain-Special Optimized Benchmark for Soccer Player Tracking	Zikai Song, Zhiwen Wan, Wei Yuan, Ying Tang, Junqing Yu and Yi-Ping Pheobe Chen	Huazhong University of Science and Technology, Huazhong University of Science and Technology, Huazhong University of Science and Technology, Huazhong University of Science and Technology, La Trobe University
	icmr190	Efficient Indexing of 3D Human Motions	Petra Budikova, Jan Sedmidubsky and Pavel Zezula	Masaryk University, Masaryk University, Masaryk University
	icmr175	Few-Shot Action Localization without Knowing Boundaries	Tingting Xie, Christos Tzelepis, Fan Fu and Ioannis Patras	Queen Mary University of London, Queen Mary University of London, City University of London, Queen Mary University of London
	icmr167	Facial Structure Guided GAN for Identity-preserved Face Image De-occlusion	Yiu-Ming Cheung, Mengke Li and Rong Zou	Hong Kong Baptist University, Hong Kong Baptist University
<b>16:00-17:00</b>		<b>Keynote 2: Maja Pantic, Imperial College London</b>		
<b>17:00-18:00</b>		<b>ICMR 2021 Award Ceremony</b>		
<b>Nov. 19</b>				
<b>8:30-10:00</b>	<b>Oral 7</b>	<b>Information/Data Retrieval</b>	<b>Session Chair: Yusuke Matsui (matsui@hal.t.u-tokyo.ac.jp)</b>	
	icmr082	Efficient Nearest Neighbor Search by Removing Anti-hub	Kimihiro Tanaka, Yusuke Matsui and Shin'ichi Satoh	The University of Tokyo, The University of Tokyo, National Institute of Informatics
	icmrc096	Impact of Interaction Strategies on User Relevance Feedback	Omar Shahbaz Khan, Björn Pórr Jónsson, Jan Zahálka, Stevan Rudinac and Marcel Worring	IT University of Copenhagen, IT University of Copenhagen, Czech Technical University in Prague, University of Amsterdam, University of Amsterdam
	icmr068	Text-Guided Visual Feature Refinement for Text-Based Person Search	Liyang Gao, Kai Niu, Zehong Ma, Bingliang Jiao, Tonghao Tan and Peng Wang	Northwestern Polytechnical University, Northwestern Polytechnical University, Northwestern Polytechnical University, Northwestern Polytechnical University, Northwestern Polytechnical University
	icmr044bn	Ten Questions in Lifelog Mining and Information Recall	An-Zi Yen, Hen-Hsen Huang and Hsin-Hsi Chen	National Taiwan University, Academia Sinica & MOST Joint Research Center for AI Technology and All Vista Healthcare, National Taiwan University & MOST Joint Research Center for AI Technology and All Vista Healthcare
	icmr091	NASTER: Non-local Attentional Scene Text Recognizer	Lei Wu, Xueliang Liu, Yanbin Hao, Richang Hong and Yunjie Ma	Hefei University of Technology, Hefei University of Technology, Hefei University of Technology, Hefei University of Technology, Hefei University of Technology
<b>8:30-10:00</b>	<b>Oral 8</b>	<b>Representation Learning</b>	<b>Session Chair: Yung-Yao Chen (yungyaochen@gapps.ntust.edu.tw)</b>	
	icmr078	Relation-Aware Keypoint Matching with Space-Spanned Feature	Jingwei Ou, Habin Ling, Tianle Wang, Xiaoqing Lv and Zhi Tang	Wangxuan Institute of Computer Technology, Peking University, Stony Brook University, Beijing University of Posts and Telecommunications, Wangxuan Institute of Computer Technology, Peking University
	icmr052	Learning Hierarchical Visual-Semantic Representation with Phrase Alignment	Baoming Yan, Qingheng Zhang, Liyu Chen, Lin Wang, Leihao Pei, Jiang Yang, Enjun Yu, Xiaobo Li and Binqiang Zhao	Alibaba Group, Alibaba Group, Alibaba Group, Alibaba Group, Alibaba Group, Alibaba Group, Alibaba Group, Alibaba Group, Alibaba Group
	icmr111	A Unified-Model via Block Coordinate Descent for Learning the Importance of Filter	Qinghua Li, Xue Zhang, Cuiqing Li and Hong Chen	Renmin University of China, Renmin University of China, Renmin University of China
	icmr152	Neural Symbolic Representation Learning for Image Captioning	Xiaomei Wang, Yanwei Fu, Lin Ma and Xiangyang Xue	Fudan University, Fudan University, Fudan University
<b>10:00-11:00</b>		<b>Keynote 3: Kristen Grauman, The University of Texas at Austin</b>		
<b>11:00-13:00</b>	<b>Oral 9</b>	<b>Video and Audio Analysis</b>	<b>Session Chair: Yong-Sheng Chen (yschen@nycu.edu.tw)</b>	
	icmrc018	GPT2MVS: Generative Pre-trained Transformer-2 for Multi-modal Video Summarization	Jia-Hong Huang, Luka Murn, Marta Mrak and Marcel Worring	University of Amsterdam, BBC Research and Development & Dublin City University, BBC Research and Development, University of Amsterdam
	icmr067	Local-enhanced Interaction for Temporal Moment Localization	Guoqiang Liang, Shiyu Ji and Yanning Zhang	Northwestern Polytechnical University, Northwestern Polytechnical University, Northwestern Polytechnical University
	icmr102	Multi-Attention Audio-Visual Fusion Network for Audio Spatialization	Wen Zhang and Jie Shao	University of Electronic Science and Technology of China, Sichuan Artificial Intelligence Research Institute
	icmr070	MS-SincResNet: Joint Learning of 1D and 2D Kernels Using Multi-scale SincNet and ResNet for Music Genre Classification	Pei-Chun Chang, Yong-Sheng Chen and Chang-Hsing Lee	National Yang Ming Chiao Tung University, National Yang Ming Chiao Tung University, Chung Hua University
<b>11:00-12:30</b>	<b>Oral 10</b>	<b>Human Action Analysis</b>	<b>Session Chair: Ching-Chun Huang (chingchun@cs.nctu.edu.tw)</b>	
	icmr048	Can Action be Imitated? Learn to Reconstruct and Transfer Human Dynamics from Videos	Yujian Fu, Yanwei Fu and Yu-Gang Jiang	Fudan University, Fudan University, Fudan University
	icmr033	HPOF: 3D Human Pose Recovery from Monocular Video with Optical Flow	Bin Ji, Ye Pan, Chen Yang and Yao Shunyu	Shanghai Jiao Tong University, Shanghai Jiao Tong University, Shanghai Jiao Tong University
	icmr058	Multi-scale Dynamic Network for Temporal Action Detection	Yifan Ren, Xing Xu, Fumin Shen, Zheng Wang, Yang Yang and Heng Tao Shen	University of Electronic Science and Technology of China, University of Electronic Science and Technology of China, University of Electronic Science and Technology of China, University of Electronic Science and Technology of China
	icmr133	SAGN: Semantic Adaptive Graph Network for Skeleton-Based Human Action Recognition	Ziwan Fu, Feng Liu, Jiahao Zhang, Hanyang Wang, Chengyi Yang, Qing Xu, Jiayin Qi, Xiangling Fu, Aimin Zhou and Zhibin Li	Beijing University of Posts and Telecommunications, East China Normal University, East China Normal University, East China Normal University, Shanghai University of International Business and Economics, Beijing University of Posts and Telecommunications, Shanghai University of International Business and Economics, Beijing University of Posts and Telecommunications
	icmr208	Joint Hand-Object Pose Estimation with Differentiably-Learned Physical Contact Point Analysis	Nan Zhuang and Yadong Mu	Peking University, Peking University
<b>9:00-12:00</b>		<b>Tutorial on Cross-modal Retrieval and Its Applications in Multimedia Advertising</b>		
<b>14:00-20:00</b>		<b>Annual Lifelog Search Challenge</b>		