ACM Multimedia 2022
Lisbon, Portugal | 10-14 October

ACM Multimedia 2022 in Lisbon: Detailed Program

Session	Local Time	Paper ID	Regis/Pres	Paper Title & Authors
Monday, October 10th (Workshops & Tutorials)				
Workshops				
Auditorio II	**9h-18h**			
MADiMa 2022: 7th International Workshop on Multimedia Assisted Dietary Management	9h05	madima25k	On-site	The Quest towards Automated Dietary Monitoring : Intervention in Free-living -- Oliver Amft (University of Freiburg)
	9h35	madima67	On-site	Real Scale Hungry Networks: Real Scale 3D Reconstruction of a Dish and a Plate using Implicit Function and a Single RGB-D Image -- Shu Naritomi (NTT Communications Corporation), Keiji Yanai (The University of Electro-Communications)
	9h50	madima43	On-site	Chewing Detection from Commercial Smart-glasses -- Vasileios Papapanagiotou (Aristotle University of Thessaloniki), Anastasia Liapi (Aristotle University of Thessaloniki), Anastasios Delopoulos (Aristotle University of Thessaloniki)
	10h05	madima42	On-site	Learning Multi-Subset of Classes for Fine-Grained Food Recognition -- Javier Ródenas (Universtiat de Barcelona), Bhalaji Nagarajan (Universtiat de Barcelona), Marc Bolaños (AIGecko Technologies SL), Petia Radeva (Universtiat de Barcelona : Computer Vision Center)
	10h20	**Short Break**		
	10h30	madima22k	On-site	Using AI to Assess and Improve Mediterranean Diet Adherence -- Arindam Ghosh (Oviva AG)
	11h	madima72	On-site	Text-based Image Editing for Food Images with CLIP -- Kohei Yamamoto (The University of Electro-Communications), Keiji Yanai (The University of Electro-Communications)
	11h15	madima59	On-site	World Food Atlas for Food Navigation -- Ali Rostami (University of California, Irvine), Nitish Nagesh (University of California, Irvine), Amir M. Rahmani (University of California, Irvine), Ramesh Jain (University of California, Irvine)
	11h30	madima28k	Virtual	MADiMA'22 Keynote Talk by George Hadjigeorgiouh / Sheridan Will Update Title on Submission Page -- George Hadjigeorgiou (ZOE)
	12h	**Virtual Poster Presentations**		
	12h30	cea003s	Virtual	Comparable Recipes: A Construction and Analysis of a Dataset of Recipes Described by Different People for the Same Dish -- Rina Kagawa (University of Tsukuba, Japan), Rei Miyata (Nagoya University, Japan), Yoko Yamakata (University of Tokyo, Japan)
	12h40	cea009s	Virtual	Few-shot Food Recognition with Pre-trained Model -- Yanqi Wu (Fudan University), Xue Song (Fudan University), Jingjing Chen (Fudan University)
	12h50	cea002s	Virtual	ABLE: Aesthetic Box Lunch Editing -- Yutong Zhou (Ritsumeikan University), Nobutaka Shimada (Ritsumeikan University)
&	**13h**	**Lunch Break**		
CEA++ 2022: 1st International Workshop on Multimedia for Cooking, Eating, and related APPlications	14h	cea011	Virtual	Recipe Recommendation for Balancing Ingredient Preference and Daily Nutrients -- Sara Ozeki (Tokyo University of Agriculture and Technology), Masaaki Kotera (Preferred Networks), Katushiko Ishiguro (Preferred Networks), Taichi Nishimura (Kyoto University), Keita Higuchi (Preferred Networks)
	14h20	cea010	Virtual	Learning Sequential Transformation Information of Ingredients for Fine-Grained Cooking Activity Recognition -- Atsushi Okamoto (Osaka Metropolitan University), Katsufumi Inoue (Osaka Metropolitan University), Michifumi Yoshioka (Osaka Metropolitan University)
	14h40	cea001	On-site	Multimodal Dish Pairing: Predicting Side Dishes to Serve with a Main Dish -- Taichi Nishimura (Kyoto University), Katsuhiko Ishiguro (Preferred Networks, Inc.), Keita Higuchi (Preferred Networks, Inc.), Masaaki Kotera (Preferred Networks, Inc.)
	15h	cea005	On-site	Prediction of Mental State from Food Images -- Kei Nakamoto (The University of Tokyo), Sosuke Amano (foo.log Inc.), Hiroaki Karasawa (Hongo Software Development), Yoko Yamakata (The University of Tokyo), Kiyoharu Aizawa (The University of Tokyo)
	15h20	cea012s	On-site	Recipe Recording by Duplicating and Editing Standard Recipe -- Akihisa Ishino (The University of Tokyo), Yoko Yamakata (The University of Tokyo), Kiyoharu Aizawa (The University of Tokyo)
	15h30	cea006s	On-site	A Multimedia Recipe Dataset with Ingredient Annotation at Each Instructional Step -- Yixin Zhang (Kyoto University), Yoko Yamakata (University of Tokyo), Keishi Tajima (Kyoto University)
	15h50	cea022k	On-site	Creating a World Food Atlas -- Ramesh Jain (University of California, Irvine)
	16h20	cea025p	On-site	CEA++2022 Panel — Toward Building a Global Food Network -- Yoko Yamakata (The University of Tokyo), Ramesh Jain (University of Bern), Ramesh Jain (University of California, Irvine)
	17h	**Joint MADiMA & CEA++ Poster Session**		
Auditorio III	**9h-18h**			
ACM MMSports 2022: 5th International ACM Workshop on Multimedia Content Analysis in Sports	9h	mmspor32	On-site	DeepSportradar-v1 -- All authors
	9h30	mmspor21	On-site	Towards Automated Key-Point Detection in Images with Partial Pool View -- Tim Woinoski (Simon Fraser University), Ivan V. Bajić (Simon Fraser University)
	9h50	mmspor10	On-site	Improving Exertion and Wellbeing Prediction in Outdoor Running Conditions using Audio-based Surface Recognition -- Alexander Gebhard (University of Augsburg), Andreas Triantafyllopoulos (University of Augsburg), Shahin Amiriparian (University of Augsburg), Sandra Ottl (University of Augsburg), Valerie Dieter (University Hospital Tuebingen), Maurice Gerczuk (University of Augsburg), Mirko Jaumann (HB Technologies AG), David Hildner (HB Technologies AG), Patrick Schneeweiß (University Hospital Tuebingen), Inka Rösel (University Hospital Tuebingen), Inga Krauß (University Hospital Tuebingen), Björn W. Schuller (University of Augsburg)
	10h10	mmspor12	On-site	Video- and Location-based Analysis of Cycling Routes for Safety Measures and Fan Engagement -- Pirlouit Dumez (ISEN Lille), Guillaume Prevost (ISEN Lille), Maarten Slembrouck (Ghent University - imec, IDLab), Jelle De Bock (Ghent University - imec, IDLab), Julien Marbaix (Ghent University - imec, IDLab), Steven Verstockt (Ghent University - imec, IDLab)
	11h	mmspor16	On-site	Action Recognition using Time-series Heat Maps of Joint Positions from Volleyball Match Videos -- Akimasa Kondo (Keio University), Hideo Saito (Keio University), Shoji Yachida (NEC), Ryo Fujiwara (NEC)
	11h20	mmspor22	On-site	BadmintonDB -- Kar-Weng Ban (Multimedia University), John See (Heriot-Watt University Malaysia), Junaidi Abdullah (Multimedia University), Yuen Peng Loh (Multimedia University)
	11h40	mmspor23	On-site	Video-Based Detection of Combat Positions and Automatic Scoring in Jiu-jitsu -- Valter Hudovernik (University of Ljubljana), Danijel Skočaj (University of Ljubljana)
	12h	mmspor09	On-site	Pass Evaluation in Women's Olympic Ice Hockey -- Robyn Ritchie (Simon Fraser University), Alon Harell (Simon Fraser University), Phillip Shreeves (Simon Fraser University)
	12h20	**Lunch Break**		
	14h	mmspor05	On-site	SoccerNet 2022 Challenges Results -- Many Authors
	14h20	mmspor11	On-site	Spatio-Temporal Encoder for Action Spotting in Soccer Videos -- Abdulrahman Darwish (German University in Cairo), Tallal El-Shabrway (German University in Cairo)
	14h40	mmspor17		A Graph-Based Method for Soccer Action Spotting Using Unsupervised Player Classification -- Alejandro Cartas (Universitat Pompeu Fabra), Coloma Ballester (Universitat Pompeu Fabra), Gloria Haro (Universitat Pompeu Fabra)
	15h	mmspor07		A Transformer-based System for Action Spotting in Soccer Videos -- He Zhu (Tshinghua University), Junwei Liang (Tencent Youtu Lab), Chengzhi Lin (Sun Yat-sen University), Jun Zhang (Tencent Youtu Lab), Jianming Hu (Tshinghua University)
	16h20	mmspor25		KaliCalib: A Framework for Basketball Court Registration -- Adrien Maglo (Université Paris-Saclay, CEA, List), Astrid Orcesi (Université Paris-Saclay, CEA, List), Quoc-Cuong Pham (Université Paris-Saclay, CEA, List)
		mmspor26		A Person Re-identification Approach Focusing on the Occlusion Problem and Ranking Optimization -- Wenkai Zheng (Beijing University of Posts and Telecommunications), Mei Yuan (Peking University)
		mmspor27		Dual Data Augmentation Method for Data-Deficient and Occluded Instance Segmentation -- Bo Yan (Ant Group), Yadong Li (Ant Group), Xingran Zhao (Ant Group), Hongbin Wang (Ant Group)
		mmspor28	On-site	CLIP-Reldent -- Konrad Habel (University of the Bundeswehr Munich), Fabian Deuser (University of the Bundeswehr Munich), Norbert Oswald (University of the Bundeswehr Munich)
		mmspor29	Virtual	Attention-Aware Multiple Granularities Network for Player Re-Identification -- Qi An (Beijing University of Posts and Telecommunications), Kuilong Cui (Beijing University of Posts and Telecommunications), Rongshuai Liu (Beijing University of Posts and Telecommunications), Chuanming Wang (Beijing University of Posts and Telecommunications), Mengshi Qi (Beijing University of Posts and Telecommunications), Huadong Ma (Beijing University of Posts and Telecommunications)
Auditorio IV	**9h-18h**			
MuSe 2022: 3rd International Multimodal Sentiment Analysis Workshop and Challenge	9h10	muse025k	On-site	The Dos and Don'ts of Affect Analysis -- Shahin Amiriparian (University of Augsburg)
	9h40	muse012	On-site	The MuSe 2022 Multimodal Sentiment Analysis Challenge: Humor, Emotional Reactions, and Stress -- Lukas Christ (University of Augsburg), Shahin Amiriparian (University of Augsburg), Alice Baird (Hume AI), Panagiotis Tzirakis (Hume AI), Alexander Kathan (University of Augsburg), Niklas Müller (University of Passau), Lukas Stappen (Recoro), Eva-Maria Meßner (University of Ulm), Andreas König (University of Passau), Alan Cowen (Hume AI), Erik Cambria (Nanyang Technological University), Björn W. Schuller (Imperial College London)
	9h55	muse004		Hybrid Multimodal Fusion for Humor Detection -- Haojie Xu (AHU-IAI AI Joint Laboratory, Anhui University : Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Weifeng Liu (AHU-IAI AI Joint Laboratory, Anhui University : Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), jiangwei liu (AHU-IAI AI Joint Laboratory, Anhui University : Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Mingzheng Li (University of Science and Technology of China : Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Yu Feng (School of Biomedical Engineering, Anhui Medical University : Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Yasi Peng (School of Biomedical Engineering, Anhui Medical University : Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Yunwei Shi (AHU-IAI AI Joint Laboratory, Anhui University : Institute of Artificial Intelligence, Hefei Comprehensive National Science Center), Xiao Sun (Hefei University of Technology : ZhongJuYuan Intelligent Technology Co., Ltd), Meng Wang (Hefei University of Technology : Institute of Artificial Intelligence, Hefei Comprehensive National Science Center)
	10h10	muse007		Integrating Cross-modal Interactions via Latent Representation Shift for Multi-modal Humor Detection -- Chengxin Chen (Institute of Acoustics, Chinese Academy of Sciences : University of Chinese Academy of Science), Pengyuan Zhang (Institute of Acoustics, Chinese Academy of Sciences : University of Chinese Academy of Science)
	10h25	muse002	On-site	A Personalised Approach to Audiovisual Humour Recognition and its Individual-level Fairness -- Alexander Kathan (University of Augsburg), Shahin Amiriparian (University of Augsburg), Lukas Christ (University of Augsburg), Andreas Triantafyllopoulos (University of Augsburg), Niklas Müller (University of Passau), Andreas König (University of Passau), Björn W. Schuller (University of Augsburg : Imperial College London)
	10h50	muse016	Virtual	Comparing Biosignal and Acoustic feature Representation for Continuous Emotion Recognition -- Sarthak Yadav (Idiap Research Institute), Tilak Purohit (Idiap Research Institute : École polytechnique fédérale de Lausanne (EPFL)), Zohreh Mostaani (Idiap Research Institute : École polytechnique fédérale de Lausanne (EPFL)), Bogdan Vlasenko (Idiap Research Institute), Mathew Magimai.-Doss (Idiap Research Institute)
	11h05	muse009	On-site	Towards Multimodal Prediction of Time-continuous Emotion using Pose Feature Engineering and a Transformer Encoder -- Ho-min Park (Ghent University Global Campus), Ilho Yun (Ghent University Global Campus), Ajit Kumar (Soongsil University), Ankit Kumar Singh (Soongsil University), Bong Jun Choi (Soongsil University), Dhananjay Singh (Hankuk University of Foreign Studies), Wesley De Neve (Ghent University Global Campus)
	11h20	muse006	Virtual	Improving Dimensional Emotion Recognition via Feature-wise Fusion -- Yiping Liu (Hunan University), Wei Sun (Hunan University), Xing Zhang (Hunan University), Yebao Qin (Hunan University)
	11h35	muse015		Multimodal Temporal Attention in Sentiment Analysis -- Yu He (School of Artificial Intelligence, University of Chinese Academy of Sciences : NLPR, Institute of Automation, Chinese Academy of Sciences), Licai Sun (School of Artificial Intelligence, University of Chinese Academy of Sciences : NLPR, Institute of Automation, Chinese Academy of Sciences), Zheng Lian (NLPR, Institute of Automation, Chinese Academy of Sciences), Bin Liu (NLPR, Institute of Automation, Chinese Academy of Sciences), Jianhua Tao (NLPR, Institute of Automation, Chinese Academy of Sciences : School of Artificial Intelligence, University of Chinese Academy of Sciences : CAS Center for Excellence in Brain Science and Intelligence Technology), Meng Wang (Ant Financial Services Group), Yuan Cheng (Ant Financial Services Group)
	12h	muse022k	On-site	Uncovering the Nuanced Structure of Expressive Behavior Across Modalities -- Alan Cowen (Hume AI)
	12h30	muse008	On-site	ViPER: Video-based Perceiver for Emotion Recognition -- Lorenzo Vaiani (Politecnico di Torino), Moreno La Quatra (Politecnico di Torino), Luca Cagliero (Politecnico di Torino), Paolo Garza (Politecnico di Torino)

	12h45	muse014		Emotional Reaction Analysis based on Multi-Label Graph Convolutional Networks and Dynamic Facial Expression Recognition Transformer -- Kexin Wang (NLPR, Institute of Automation, Chinese Academy of Sciences : School of Artificial Intelligence, University of Chinese Academy of Sciences), Zheng Lian (NLPR, Institute of Automation, Chinese Academy of Sciences), Licai Sun (NLPR, Institute of Automation, Chinese Academy of Sciences : School of Artificial Intelligence, University of Chinese Academy of Sciences), Bin Liu (NLPR, Institute of Automation, Chinese Academy of Sciences), Jianhua Tao (Institute of Automation, Chinese Academy of Sciences; School of Artificial Intelligence, University of Chinese Academy of Sciences;), Yin Fan (iQIYI,Inc)
	13h	**Lunch Break**		
	14h	muse013		Hybrid Multimodal Feature Extraction, Mining and Fusion for Sentiment Analysis -- Jia Li (Hefei University of Technology), Ziyang Zhang (Hefei University of Technology), Junjie Lang (Hefei University of Technology), Yueqi Jiang (Hefei University of Technology), Liuwei An (Hefei University of Technology), Peng Zou (Hefei University of Technology), Yangyang Xu (Hefei University of Technology), Sheng Gao (Hefei University of Technology), Jie Lin (Hefei University of Technology), Chunxiao Fan (Hefei University of Technology), Xiao Sun (Hefei University of Technology : ZhongJuYuan Intelligent Technology Co., Ltd), Meng Wang (Hefei University of Technology : Hefei Comprehensive National Science Center)
	14h25	muse003	On-site	Transformer-based Non-Verbal Emotion Recognition: Exploring Model Portability across Speakers' Genders -- Lorenzo Vaiani (Politecnico di Torino), Alkis Koudounas (Politecnico di Torino), Moreno La Quatra (Politecnico di Torino), Luca Cagliero (Politecnico di Torino), Paolo Garza (Politecnico di Torino), Elena Baralis (Politecnico di Torino)
	14h40	muse005	On-site	Bridging the Gap: End-to-End Domain Adaptation for Emotional Vocalization Classification using Adversarial Learning -- Dominik Schiller (Augsburg University), Silvan Mertes (Augsburg University), Pol van Rijn (Augsburg University), Elisabeth André (Augsburg University)
	14h55	muse018		Leveraging Multi-modal Interactions among the Intermediate Representations of Deep Transformers for Emotion Recognition -- Yang Wu (Harbin Institute of Technology), Zhenyu Zhang (Harbin Institute of Technology), Pai Peng (Harbin Institute of Technology), Yanyan Zhao (Harbin Institute of Technology), Bing Qin (Harbin Institute of Technology)
Pav5 R5C	**9h-18h**			
SUMAC 2022: 4th ACM International workshop on Structuring and Understanding of Multimedia heritAge Contents	9h15			Opening & Welcome
	9h30	sumac25k	On-site	Building Blocks for a Virtual Time Machine -- Andreas Konrad Maier (Friedrich-Alexander-Universität Erlangen-Nürnberg)
	10h30	sumac0371	Virtual	Data-driven Automatic Attribution of Azerbaijani Flat Woven Carpets -- Rashid Bakirov (Bournemouth University), Roya Taghieva (Azerbaijani Carpet Makers' Union), Nigar Eyvazli (Independent Researcher), Umay Mammadzada (Azerbaijan State University of Culture and Arts)
	11h	sumac1323	On-site	Deep Level Annotation for Painter Attribution on Greek Vases utilizing Object Detection -- Marta Kipke (Georg-August-Universität Göttingen), Lukas Brinkmeyer (Universität Hildesheim), Souaybou Bagayoko (Universität Hildesheim), Lars Schmidt-Thieme (Universität Hildesheim), Martin Langner (Georg-August-Universität Göttingen)
	11h30	sumac7135	On-site	Approach to Identification of Changes from Local Surface Normal Analysis of RTI Data in Application to Cultural Heritage -- Sunita Saha (Warsaw University of Technology), David A Lewis (University of Burgundy), Robert Sitnik (Warsaw University of Technology)
	11h45	**Lunch Break**		
	14h	sumac22k	Virtual	Creating a Time Machine of Future Pasts -- Georgios Artopoulos (Cyprus Institute)
	15	sumac7423	Virtual	A Methodological Approach for Multi-temporal Tracking of Silver Tarnishing -- Amalia Siatou (University of Applied Sciences and Arts Western Switzerland : University of Burgundy), Yuly Castro (University of Burgundy), Marvin Nurit (University of Burgundy), Gaetan Le Goic (University of Burgundy), Hermine Chatoux (University of Burgundy : University of South Brittany), Christian Degrigny (University of Applied Sciences and Arts Western Switzerland), Laura Brambilla (University of Applied Sciences and Arts Western Switzerland), Alamin Mansouri (University of Burgundy)
	15h30	sumac0746	On-site	Contributions of Photometry to the 3D-digitization of Heritage -- Antoine Laurent (IRIT - UMR CNRS 5505), Jean Mélou (IRIT - UMR CNRS 5505), Thomas Sagory (Musée d' Archéologie Nationale), Carole Fritz (MSHS de Toulouse TRACES - UMR 5608), Jean-Denis Durou (IRIT - UMR CNRS 5505)
	15h45	**Technical Break for Best Paper Selection**		
	16h15			Wrap-up and Ceremony Award
Pav4 R1.07	**9h-13h**			
NarSUM 2022: 1st Workshop on User-centric Narrative Summarization of Long Videos	9h		On-site	Welcome and Opening Remarks
	9h05	narsum33k	On-site	Learning, Understanding and Interaction in Videos -- Manmohan Chandraker (NEC Laboratories America)
	9h45		On-site	Invited presentation: Compute to Tell the Tale: Goal-Driven Narrative Generation (Brave New Idea paper, ACM MM 2022)
	10h	narsum04	On-site	Narrative Dataset: Towards Goal-Driven Narrative Generation -- Karen Stephen (NEC Corporation), Rishabh Sheoran (NEC Corporation), Satoshi Yamazaki (NEC Corporation)
	10h15	narsum02	On-site	Soccer Game Summarization using Audio Commentary, Metadata, and Captions -- Sushant Gautam (Simula Metropolitan Center for Digital Engineering), Cise Midoglu (Simula Metropolitan Center for Digital Engineering), Saeed Shafiee Sabet (Simula Metropolitan Center for Digital Engineering), Dinesh Baniya Kshatri (Simula Metropolitan Center for Digital Engineering), Pal Halvorsen (Simula Metropolitan Center for Digital Engineering)
	10h30	narsum03	Virtual	Contrastive Representation Learning for Expression Recognition from Masked Face images -- Fanxing Luo (Ritsumeikan University), Longjiao Zhao (Nagoya University), Yu Wang (Hitotsubashi University), Jien Kato (Ritsumeikan University)
	10h45	**Break**		
	11h	narsum22k	On-site	NarSUM'22 Keynote Talk by Ioannis (Yiannis) Patras / Sheridan Will Update Title on Submission Page -- Ioannis (Yiannis) Patras (Queen Mary University of London)
	11h40	narsum11p	On-site	NarSUM Panel- User-Centric Narrative Summarization of Long Videos -- Karen Stephen (National University of Singapore), Mohan Kankanhalli (National University of Singapore), Jianquan Liu (National University of Singapore), Yongkang Wong (National University of Singapore)
	12h30		On-site	Closing Remarks
Pav4 R1.07	**14h-18h**			
FME 2022: 2nd Workshop on Facial Micro-Expression: Advanced Techniques for Multi-Modal Facial Expression Analysis	14h		On-site	Opening from the Chairs
	14h05	fme003	On-site	A Multi-Temporal Stream Network for Spotting Facial Macro- and Micro-Expression with Hard and Soft Pseudo-labels -- Gen Bing Liong (University of Malaya), Sze-Teng Liong (Feng Chia University), John See (Heriot-Watt University Malaysia), Chee-Seng Chan (University of Malaya)
	14h25	fme002	Virtual	A More Objective Quantification of Micro-Expression Intensity through Facial Electromyography -- Shaoyuan Lu (CAS Key Laboratory of Behavioral Science, Institute of Psychology : Department of Psychology, University of the Chinese Academy of Sciences), Jingting Li (CAS Key Laboratory of Behavioral Science, Institute of Psychology : Department of Psychology, University of the Chinese Academy of Sciences), Yan Wang (CAS Key Laboratory of Behavioral Science, Institute of Psychology : Department of Psychology, University of the Chinese Academy of Sciences), Zizhao Dong (CAS Key Laboratory of Behavioral Science, Institute of Psychology : Department of Psychology, University of the Chinese Academy of Sciences), Su-Jing Wang (CAS Key Laboratory of Behavioral Science, Institute of Psychology : Department of Psychology, University of the Chinese Academy of Sciences), Xiaolan Fu (State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences : Department of Psychology, University of the Chinese Academy of Sciences)
	14h45	fme022k	On-site	FME 2022 Keynote Talk by Hu Han -- Hu Han (Institute of Computing Technology, Chinese Academy of Sciences)
	15h25	**Break**		
	15h40		Virtual	MEGC2022: Spotting Task
			Virtual	Baseline: 3D-CNN for Facial Micro- and Macro-expression Spotting on Long Video Sequences using Temporal Oriented Reference Frame
			Virtual	1st place: Facial Expression Spotting Based on Optical Flow Features
			Virtual	2nd place: Rethinking Optical Flow Methods for Micro-Expression Spotting
			Virtual	3rd place: ABPN: Apex and Boundary Perception Network for Micro- and Macro-Expression Spotting
	16h10		Virtual	MEGC2022: Generation Task
			Virtual	1st place: Micro Expression Generation with Thin-plate Spline Motion Model and Face Parsing
			Virtual	2nd place: Fine-grained Micro-Expression Generation based on Thin-Plate Spline and Relative AU& Constraint
			Virtual	3rd place: Adaptive Dual Motion Model for Facial Micro-Expression Generation
	16h40		On-site	Awards and Summary from the Chairs
	16h50		On-site	Panel discussion and brainstorming for the next steps
	17h20		On-site	Closing

Tutorials

Pav5 R5A	**9h-13h**	**Deep Learning-based Point Cloud Coding for Immersive Experiences**
Pav5 R5A	**14h-18h**	**Advances In Quality Assessment Of Video Streaming Systems: Algorithms, Methods, Tools**
Pav5 R5B	**9h-13h**	**Multimedia Content Understanding in Harsh Environments**
Pav5 R5B	**14h-18h**	**Autonomous UAV Cinematography**
Pav4 R1.06	**14h-18h**	**Open Challenges of Interactive Video Search and Evaluation**
Main Foyer	**18h30**	**Reception**

Tuesday, October 11th (Main Conference)

Opening [Auditório I]	**8h45-9h15**	**Chairs: ACM Mutlimedia 2022 General Chairs, ACM Multimedia 2022 Program Chairs**		
Oral Session 1a [Auditório I]	**9h15-10h30**	**Chair:Jingkuan Song (University of Electronic Science and Technology of China)**		
Multimodal Fusion I	9h15	mmfp1804	Virtual	Non-Autoregressive Cross-Modal Coherence Modelling -- Yi Bin (University of Electronic Science and Technology of China, Tsinghua University), WENHAO SHI (University of Electronic Science and Technology of China), Jipeng Zhang (Department of Computer Science and Engineering, The Hong Kong University of Science and Technology), Yujuan Ding (Hong Kong Polytechnic University), Yang Yang (University of Electronic Science and Technology of China), Hengtao Shen (University of Electronic Science and Technology of China)
	9h33	mmfp2969	Virtual	Semantic Structure Enhanced Contrastive Adversarial Hash Attention Network for Cross-media Representation Learning -- Meiyu Liang (Beijing University of Posts and Telecommunications), Junping Du (Beijing University of Posts and Telecommunications), Xiaowen Cao (Beijing University of Posts and Telecommunications), Yang Yu (Beijing University of Posts and Telecommunications), Kangkang Lu (Beijing University of Posts and Telecommunications), Zhe Xue (Beijing University of Posts and Telecommunications), Min Zhang (Beijing University of Posts and Telecommunications)
	9h51	mmfp2446	Virtual	DHHN: Dual Hierarchical Hybrid Network for Weakly-Supervised Audio-Visual Video Parsing -- Xun Jiang (University of Electronic Science and Technology of China), Xing Xu (University of Electronic Science and Technology of China), Zhiguo Chen (University of Electronic Science and Technology of China), Jingran Zhang (University of Electronic Science and Technology of China), Jingkuan Song (University of Electronic Science and Technology of China, Tsinghua University), Fumin Shen (University of Electronic Science and Technology of China, Tsinghua University), Huimin Lu (Kyushu Institute of Technology), Hengtao Shen (University of Electronic Science and Technology of China)
	10h09	mmfp2491	Virtual	Span-based Audio-Visual Localization -- Yiling Wu (Peng Cheng Laboratory), Xinfeng Zhang (University of the Chinese Academy of Sciences), Yaowei Wang (Pengcheng Laboratory), Qingming Huang (Univ of Chinese Academy of Sciences)
Oral Session 1b [Auditório II]	**9h15-10h30**	**Chair: Yi Yu (National Institute of Informatics)**		
	9h15	mmfp1855	Virtual	Cross-Modal Retrieval with Heterogeneous Graph Embedding -- Dapeng Chen (Huawei Technologies Ltd.), Min Wang (Huawei Technologies Ltd.), Haobin Chen (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Chinese Academy of Sciences), Lin Wu (Hefei University of Technology), Jing Qin (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Wei Peng (Huawei Technologies Ltd.)

Cross-Modal Retrieval	9h33	mmfp2189	Virtual	C^3CMR: Cross-Modality Cross-Instance Contrastive Learning for Cross-Media Retrieval -- Junsheng Wang (Nanjing University of Science and Technology), Tiantian Gong (Nanjing University of Aeronautics and Astronautics), Zhixiong Zeng (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Changchang Sun (Illinois Institute of Technology), Yan Yan (Illinois Institute of Technology)
	9h51	mmfp2827	Virtual	A feature-space multimodal data augmentation technique for text-video retrieval -- Alex Falcon (Fondazione Bruno Kessler), Giuseppe Serra (University of Udine), Oswald Lanz (Free University of Bozen-Bolzano)
	10h09	mmfp2900	On-site	Machine Unlearning for Image Retrieval: A Generative Scrubbing Approach -- Peng-Fei Zhang (University of Queensland), Guangdong Bai (University of Queensland), Zi Huang (University of Queensland), Xin-Shun Xu (Shandong University)
Oral Session 1c [Auditório III]	**9h15-10h30**	**Chair: Hailin Shi (NIO)**		
Affective Multimedia Computing	9h15	mmfp1360	Virtual	Enlarging the Long-time Dependency via RL-based Memory Network in Movie Affective Analysis -- Jie Zhang (Alibaba Group), Yin Zhao (Alibaba Group), Kai Qian (University of the Chinese Academy of Sciences)
	9h33	mmfp1888	Virtual	Counterfactual Reasoning for Out-of-distribution Multimodal Sentiment Analysis -- Teng Sun (Shandong University), Wenjie Wang (National University of Singapore), Jing Qiang (Shandong University), Yiran Cui (Shandong University), Songxuemeng Song (Shandong University), Liqiang Nie (Harbin Institute of Technology (Shenzhen))
	9h51	mmfp1834	Virtual	MAFW: A Large-scale, Multi-modal, Compound Affective Database for Dynamic Facial Expression Recognition in the Wild -- Yuanyuan Liu (China University of Geosciences Wuhan), Wei Dai (China University of Geosciences Wuhan), Chuanxu Feng (China University of Geosciences Wuhan), Wenbin Wang (China University of Geosciences Wuhan), Guanghao Yin (China University of Geosciences Wuhan), Jiabei Zeng (Institute of Computing Technology, Chinese Academy of Sciences), Shiguang Shan (Institute of Computing Technology, Chinese Academy of Sciences)
	10h09	mmfp3080	Virtual	SER30K: A Large-Scale Dataset for Sticker Emotion Recognition -- Shengzhe Liu (Nankai University), Xin Zhang (Nankai University), Jufeng Yang (Nankai University)
Oral Session 1d [Auditório IV]	**9h15-10h30**	**Chair: Björn Þór Jónsson (Reykjavic University)**		
Weakly-Supervised Learning for Multimedia	9h15	mmfp1864	Virtual	Boat in the sky: Background Decoupling and Object-aware Pooling for Weakly Supervised Semantic Segmentation -- Jianjun Xu (University of Science and Technology of China), Hongtao Xie (University of Science and Technology of China), Hai Xu (University of Science and Technology of China), Yuxin Wang (University of Science and Technology of China), Sun-Ao Liu (University of Science and Technology of China), Yongdong Zhang (University of Science and Technology of China)
	9h33	mmfp0803	Virtual	Proxy Probing Decoder for Weakly Supervised Object Localization: A Baseline Investigation -- Jingyuan Xu (University of Science and Technology of China), Hongtao Xie (University of Science and Technology of China), Chuanbin Liu (University of Science and Technology of China), Yongdong Zhang (University of Science and Technology of China)
	9h51	mmfp1701	Virtual	Integrating Object-aware and Interaction-aware Knowledge for Weakly Supervised Scene Graph Generation -- Xingchen Li (Zhejiang University), Long Chen (Columbia University), Wenbo Ma (Zhejiang University), Yi Yang (Zhejiang University), Jun Xiao (Zhejiang University)
	10h09	mmfp1637	Virtual	ChebyLighter: Optimal Curve Estimation for Low-light Image Enhancement -- Jinwang Pan (Harbin Institute of Technology), Deming Zhai (Harbin Institute of Technology), Yuanchao Bai (Harbin Institute of Technology), Junjun Jiang (Harbin Institute of Technology), Debin Zhao (Harbin Institute of Technology), Xianming Liu (Harbin Institute of Technology)
Doctoral Symposium [Pav4 R1.07]	**9h15-10h30**			
Content Understanding	9h15	mmdc005	On-site	Unsupervised Multi-object Tracking via Dynamical VAE and Variational Inference -- Xiaoyu Lin (INRIA)
	9h30	mmdc008	On-site	Enabling Effective Low-Light Perception using Ubiquitous Low-Cost Visible-Light Cameras -- Igor Morawski (National Taiwan University)
Systems and Applications	9h45	mmdc001	On-site	Video Coding Enhancements for HTTP Adaptive Streaming -- Vignesh V Menon (Alpen-Adria Universität Klagenfurt)
	10h	mmdc010	On-site	Interaction with Immersive Cultural Heritage Environments: Using XR Technologies to represent multiple perspectives on Serralves Museum -- Manuel Silva (Universidade Catolica Portuguesa)
	10h15	mmdc012	On-site	Enriching existing educational video datasets to improve slide classification and analysis -- Travis Seng (IRIT - REVA)
Poster Session 1 [Pav. 4 - Hall]	**10h30-11h45**			*(Posters will be hung the whole day)*
Understanding Multimedia Content: Multimodal Fusion and Embeddings		mmfp0149	Virtual	Normalization-based Feature Selection and Restitution for Pan-sharpening -- man zhou (University of Science and Technology of China), Huang Jie (University of Science and Technology of China), Keyu Yan (University of Science and Technology of China), yang gang (University of Science and Technology of China), Aiping Liu (University of Science and Technology of China), Chongyi Li (Nanyang Technological University), Feng Zhao (University of Science and Technology of China)
		mmfp0737	Virtual	Adaptively Learning Low-high Frequency Information Integration for Pan-sharpening -- man zhou (University of Science and Technology of China), Huang Jie (University of Science and Technology of China), Chongyi Li (Nanyang Technological University), Hu Yu (University of Science and Technology of China), Keyu Yan (University of Science and Technology of China), Naishan Zheng (University of Science and Technology of China), Feng Zhao (University of Science and Technology of China)
		mmfp2657	Virtual	Complementary Graph Representation Learning for Functional Neuroimaging Identification -- Rongyao Hu (University of Electronic Science and Technology of China), Larry Peng (University of Electronic Science and Technology of China), Jiangzhang Gan (Massey University), Xiaoshuang Shi (University of Electronic Science and Technology of China, Tsinghua University), Xiaofeng Zhu (University of Electronic Science and Technology of China)
		mmfp1639	Virtual	Dynamically Adjust Word Representations Using Unaligned Multimodal Information -- 继伟 郭 (Hangzhou Dianzi University), Jiajia Tang (Hangzhou Dianzi University), Weichen Dai (Hangzhou Dianzi University), Yu Ding (Netease Fuxi AI Lab), Wanzeng Kong (Hangzhou Dianzi University)
		mmfp1661	On-site	Bipartite Graph-based Discriminative Feature Learning for Multi-View Clustering -- Weiqing Yan (Yantai University), jindong Xu (Yantai University), Jinglei Liu (Yantai University), Guanghui Yue (Shenzhen University), Chang Tang (China University of Geosciences)
		mmfp2046	Virtual	Dynamic Incomplete Multi-view Imputing and Clustering -- Li Feng (Nanjing University of Science and Technology), quan Sun (Nanjing University of Science and Technology), Zhenwen Ren (Southwest University Of Science And Technology), Hui Sun (Nanjing University of Science and Technology)
		mmfp2059	Virtual	Learning Smooth Representation for Multi-view Subspace Clustering -- Shudong Huang (Sichuan University), Yixi Liu (Sichuan University), Yazhou Ren (University of Electronic Science and Technology of China), Ivor W Tsang (University of Technology Sydney), Zenglin Xu (Harbin Institute of Technology, Shenzhen), Jiancheng Lv (Sichuan University)
		mmfp2259	Virtual	LFBCNet: Light Field Boundary-aware and Cascaded Interaction Network for Salient Object Detection -- Mianzhao Wang (Tianjin University of Technology), Fan Shi (Tianjin University of Technology), Xu Cheng (Norwegian University of Science and Technology), Meng Zhao (Tianjin University of Technology), Yao Zhang (Tianjin University), Chen Jia (Tianjin University of Technology), Weiwei Tian (Tianjin University of Technology), Shengyong Chen (Tianjin University of Technology)
		mmfp2623	Virtual	Multiple Kernel Clustering with Dual Noise Minimization -- Junpu Zhang (National University of Defense Technology), Liang Li (National University of Defense Technology), Siwei Wang (National University of Defense Technology), Jiyuan Liu (National University of Defense Technology), Yue Liu (National University of Defense Technology), Xinwang Liu (National University of Defense Technology), En Zhu (National University of Defense Technology)
		mmfp2670	Virtual	Webly Adaptive Image Hashing via Lightweight Semantic Transfer Network -- Hui Cui (Shandong Normal University), Lei Zhu (Shandong Normal University), Jingjing Li (University of Electronic Science and Technology of China), Zheng Zhang (Harbin Institute of Technology), WEILI GUAN (Monash University)
		mmfp0821	On-site	Rethinking Super-Resolution as Text-Guided Details Generation -- Chenxi Ma (Fudan University), Bo Yan (Fudan University), Qing Lin (Fudan University), Weimin Tan (Fudan University), Siming Chen (Fudan University)
		mmfp1106	Virtual	DEAL: An Unsupervised Domain Adaptive Framework for Graph-level Classification -- Nan Yin (National University of Defense Technology), Li Shen (JD Explore Academy), Baopu Li (The Chinese University of Hong Kong), Mengzhu Wang (National University of Defense Technology), Xiao Luo (Peking University), Chong Chen (Alibaba Group), Zhigang Luo (National University of Defense Technology, Tsinghua University), Xian-Sheng Hua (Zhejiang University)
		mmfp2343	On-site	AVQA: A Dataset for Audio-Visual Question Answering on Videos -- Pinci Yang (Tsinghua-Berkeley Shenzhen Institute,Tsinghua University), Xin Wang (Tsinghua University), Xuguang Duan (Tsinghua University, Tsinghua University), Hong Chen (Tsinghua University, Tsinghua University), Runze Hou (Tsinghua-Berkeley Shenzhen Institute, Tsinghua University), Cong Jin (Communication University of China), Wenwu Zhu (Tsinghua University, Tsinghua University)
		mmfp0447	Virtual	Prompting for Multi-Modal Tracking -- Jinyu Yang (University of Birmingham), Zhe Li (Southern University of Science and Technology), Feng Zheng (Southern University of Science and Technology), Ales Leonardis (University of Birmingham), Jingkuan Song (University of Electronic Science and Technology of China, Tsinghua University)
		mmfp2178	Virtual	mmBody Benchmark: 3D Body Reconstruction Dataset and Analysis for Millimeter Wave Radar -- Anjun Chen (Zhejiang University), Xiangyu Wang (Zhejiang University), Shaohao Zhu (Zhejiang University), Yanxu Li (Zhejiang University), Jiming Chen (Zhejiang University), Qi Ye (Zhejiang University)
		mmfp0618	Virtual	Eliminating Spatial Ambiguity for Weakly Supervised 3D Object Detection without Spatial Labels -- Haizhuang Liu (University of Science and Technology Beijing), Huimin Ma (University of Science and Technology Beijing), Yilin Wang (University of Science and Technology Beijing), Bochao Zou (University of Science and Technology Beijing), Tianyu Hu (University of Science and Technology Beijing), Rongquan Wang (University of Science and Technology Beijing), Jiansheng Chen (University of Science and Technology Beijing)
		mmfp0426	Virtual	Dynamic Graph Reasoning for Multi-person 3D Pose Estimation -- Zhongwei Qiu (University of Science and Technology Beijing), Yang Qiansheng (Baidu), Jian Wang (Baidu), Dongmei Fu (University of Science and Technology Beijing)
		mmfp0686	Virtual	DiT: Self-supervised Pre-training for Document Image Transformer -- Junlong Li (Shanghai Jiaotong University), Yiheng Xu (Harbin Institute of Technology), Tengchao Lv (Microsoft), Lei Cui (Microsoft Research Asia), Cha Zhang (Microsoft Research), Furu Wei (Microsoft Research)
		mmfp2889	On-site	Learning to Estimate External Forces of Human Motion in Video -- Nathan Louis (University of Michigan - Ann Arbor), Jason Corso (University of Michigan), Tylan Templin (Southwest Research Institute), Travis Eliason (Southwest Research Institute), Dan Nicolella (Southwest Research Institute)
		mmfp3216	Virtual	Query Prior Matters: A MRC Framework for Multimodal Named Entity Recognition -- Meihuizi Jia (Beijing Institute of Technology), Xin Shen (Australian National University), Lei Shen (Institute of Computing Technology, Chinese Academy of Sciences), Jinhui Pang (Beijing Institute of Technology), Lejian Liao (Beijing Institute of Technology, Tsinghua University), yang song (Algorithm engineer), Meng Chen (JD AI Research), Xiaodong He (JD AI Research)
		mmfp3127	On-site	Robust Multimodal Depth Estimation using Transformer based Generative Adversarial Network -- Md Fahim Faysal Khan (Pennsylvania State University), Anusha Devulapally (Pennsylvania State University), Siddharth Advani (Samsung Research America), Vijaykrishnan Narayanan (Pennsylvania State University)
		mmfp1574	On-site	Caption-Aware Medical VQA via Semantic Focusing and Progressive Cross-Modality Comprehension -- Fuze Cong (Beijing University of Posts and Telecommunications), Shibiao Xu (Beijing University of Posts and Telecommunications), Li Guo (Beijing University of Posts and Telecommunications), Yinbing Tian (Beijing University of Posts and Telecommunications)
		mmfp1857	On-site	Complementarity-Enhanced and Redundancy-Minimized Collaboration Network for Multi-agent Perception -- Guiyang Luo (Beijing University of Post and Telecommunication, Tsinghua University), Hui Zhang (Beijing Jiaotong University), Quan Yuan (Beijing University of Posts and Telecommunications), Jinglin Li (Beijing University of Posts and Telecommunications)
		mmfp2305	Virtual	Chunk-aware Alignment and Lexical Constraint for Visual Entailment with Natural Language Explanations -- Qian Yang (Harbin Institute of Technology), yunxin li (Harbin Institute of Technology), Baotian Hu (Harbin Institute of Technology, Shenzhen), Lin Ma (Meituan), YUXIN DING (Harbin Institute of Technology), Min zhang (Harbin Institute of Technology)
		mmfp0601	Virtual	Understanding Political Polarization via Jointly Modeling Users, Connections and Multimodal Contents on Heterogeneous Graphs -- Hanjia Lyu (University of Rochester), Jiebo Luo (University of Rochester)
		mmfp0407	Virtual	Hierarchical Hourglass Convolutional Network for Efficient Video Classification -- Yi Tan (University of Science and Technology of China), Yanbin Hao (University of Science and Technology of China), Hao Zhang (City University of Hong Kong), Shuo Wang (University of Science and Technology of China), Xiangnan He (University of Science and Technology of China)
		mmfp1262	Virtual	TextBlock: Towards Scene Text Spotting without Fine-grained Detection -- Jin Wei (Communication University of China), Yuan Zhang (Communication University of China), Yu Zhou (Institute of Information Engineering, CAS), Gangyan Zeng (Communication University of China), Zhi Qiao (University of the Chinese Academy of Sciences), Youhui Guo (University of the Chinese Academy of Sciences), haiying wu (Mashang Consumer Finance Co, Ltd), HONGBIN WANG (Mashang Consumer Finance), Weipinng Wang (IIE)
		mmfp2032	Virtual	Progressive Cross-modal Knowledge Distillation for Human Action Recognition -- Jianyuan Ni (Texas State University), Anne Ngu (Texas State University), Yan Yan (Illinois Institute of Technology)
		mmfp1850	Virtual	Finding the Host from the Lesion by Iteratively Mining the Registration Graph -- Zijie Yang (, Chinese Academy of Sciences), Lingxi Xie (Huawei Technologies Ltd.), Xinyue Huo (University of Science and Technology of China), Sheng Tang (Institute of Computing Technology, Chinese Academy of Sciences), Qi Tian (Huawei Technologies Ltd.), Yongdong Zhang (University of Science and Technology of China)
		mmfp0408	On-site	3D Body Reconstruction Revisited: Exploring the Test-time 3D Body Mesh Refinement Strategy via Surrogate Adaptation -- Jonathan Lumentut (Seoul National University), In Kyu Park (Inha University)
		mmfp1720	On-site	Domain Adaptation for Time-Series Classification to Mitigate Covariate Shift -- Felix Ott (LMU Munich), David Rügamer (Rheinisch Westfälische Technische Hochschule Aachen), Lucas Heublein (Fraunhofer-Institut für Integrierte Schaltungen IIS), Bernd Bischl (LMU), Christopher Mutschler (Fraunhofer IIS)
		mmfp3262	On-site	Face Anthropometry Aware Audio-visual Age Verification -- Pavel Korshunov (Idiap research institute), Sébastien Marcel (Idiap Research Institute)
		mmfp2932	Virtual	PDD-GAN: Prior-based GAN Network with Decoupling Ability for Single Image Dehazing -- Xiaoxuan Chai (Tsinghua University, Tsinghua University), Junchi Zhou (Tsinghua University, Tsinghua University), Hang Zhou (University of Idaho), Juihsin Lai (Pingan Technology)
		mmfp0137	Virtual	Active Patterns Perceived for Stochastic Video Prediction -- Yechao Xu (Nanjing University), Zhengxing Sun (Nanjing University), Li Qian (National University of Defense Technology), Yunhan Sun (Nanjing University), 守桐 骆 (Nanjing University)

Understanding Multimedia Content: Media Interpretation		mmfp0761	Virtual	Few-shot Open-set Recognition Using Background as Unknowns -- Nan Song (Nanyang Technological University), Chi Zhang (Tencent), Guosheng Lin (Nanyang Technological University)
		mmfp2481	Virtual	Self-supervised Scene Text Segmentation with Object-centric Layered Representations Augmented by Text Regions -- Yibo Wang (University of Science and Technology of China), Yunhu Ye (University of Science and Technology of China), Yuanpeng Mao (University of Science and Technology of China), Yanwei Yu (University of Science and Technology of China), Song Yuanping (University of Science and Technology of China)
		mmfp0338	Virtual	Self-Supervised Representation Learning for Skeleton-Based Group Activity Recognition -- Cunling Bian (Tianjin University), Wei Feng (Tianjin University), Song Wang (University of South Carolina)
		mmfp0471	Virtual	Graph-DETR3D: Rethinking Overlapping Regions for Multi-View 3D Object Detection -- Zehui Chen (University of Science and Technology of China), Zhenyu Li (Harbin Institute of Technology), Shiquan Zhang (senseauto), Liangji Fang (Shanghai Jiaotong University), Qinhong Jiang (Zhejiang University, Tsinghua University), Feng Zhao (University of Science and Technology of China)
		mmfp0136	Virtual	Adaptive Mixture of Experts Learning for Generalizable Face Anti-Spoofing -- Qianyu Zhou (Shanghai Jiao Tong University,), Keyue Zhang (East China Normal University), Taiping Yao (Tencent Youtu Lab), Ran Yi (Shanghai Jiao Tong University), Shouhong Ding (Tencent Youtu Lab), Lizhuang Ma (Dept. of Computer Sci. & Eng., Shanghai Jiao Tong University)
		mmfp0725	Virtual	Multi-Granular Semantic Mining for Weakly Supervised Semantic Segmentation -- Zhang Mj (Beijing Institute of Technology), Jianwu Li (Beijing Institute of Technology, China), Tianfei Zhou (Swiss Federal Institute of Technology)
		mmfp1877	Virtual	Consistency Learning based on Class-Aware Style Variation for Domain Generalizable Semantic Segmentation -- Siwei Su (SUN YAT-SEN UNIVERSITY), Haijian Wang (SUN YAT-SEN UNIVERSITY), Meng Yang (SUN YAT-SEN UNIVERSITY, School of Computer Science and Engineering)
		mmfp1936	Virtual	Delving into the Continuous Domain Adaptation -- Yinsong Xu (Beijing University of Post and Telecommunication), Zhuqing Jiang (Beijing University of Posts and Telecommunications), Aidong Men (Beijing University of Posts and Telecommunications), Yang Liu (Peking University), Qingchao Chen (Peking University)
		mmfp2456	On-site	Digging Into Normal Incorporated Stereo Matching -- Zihua Liu (Tokyo Institute of Technology, Tokyo Institute of Technology), Songyan Zhang (Tongji University), Zhicheng Wang (Tongji University), Masatoshi Okutomi (Tokyo Institute of Technology)
		mmfp2993	Virtual	Box-FaceS: A Bidirectional Method for Box-Guided Face Component Editing -- Wenjing Huang (Shanghai Jiao Tong University), Shikui Tu (Shanghai Jiao Tong University), Lei Xu (Shanghai Jiao Tong University)
		mmfp0989	Virtual	Learning Parallax Transformer Network for Stereo Image JPEG Artifacts Removal -- Xuhao Jiang (Fudan University), Weimin Tan (Fudan University), Ri Cheng (Fudan University), Shili Zhou (Fudan University), Bo Yan (Fudan University)
		mmfp0812	Virtual	Geometry-Aware Reference Synthesis for Multi-View Image Super-Resolution -- Ri Cheng (Fudan University), Yuqi Sun (Fudan University), Bo Yan (Fudan University), Weimin Tan (Fudan University), Chenxi Ma (Fudan University)
		mmfp0357	Virtual	Chinese Character Recognition with Augmented Character Profile Matching -- Xinyan Zu (Fudan University), Haiyang Yu (Fudan University), Bin Li (Fudan University), Xiangyang Xue (Fudan University)
		mmfp1861	Virtual	Hierarchical Scene Normality-Binding Modeling for Anomaly Detection in Surveillance Videos -- Qianyue Bao (Xidian University), Fang Liu (Xidian University), Yang Liu (Xidian University), Licheng Jiao (Xidian University), Xu Liu (Xidian University), Lingling Li (Xidian University)
		mmfp0726	Virtual	ParseMVS: Learning primitive-aware surface representations for sparse multi-view stereopsis -- Haiyang Ying (Tsinghua University, Tsinghua University), Jinzhi Zhang (Electronic Engineering, Tsinghua University, Tsinghua University), Yuzhe Chen (University of Electronic Science and Technology of China), Zheng Cao (BirenTech Research), Jing Xiao (Pingan Group), Ruqi Huang (Tsinghua Shenzhen International Graduate School/Tsinghua Berkeley Shenzhen Institute), Lu Fang (Tsinghua University, Tsinghua University)
		mmfp0248	On-site	Set-Based Face Recognition Beyond Disentanglement: Burstiness Suppression With Variance Vocabulary -- Jiong Wang (Zhejiang University), Zhou Zhao (Zhejiang University), Fei Wu (Zhejiang University)
		mmfp0596	Virtual	Gait Recognition in the Wild with Multi-hop Temporal Switch -- Jinkai Zheng (Hangzhou Dianzi University), Xinchen Liu (JD Explore Academy), Xiaoyan Gu (Institute of Information Engineering, Chinese Academy of Sciences), Yaoqi Sun (Hangzhou Dianzi University), Chuang Gan (MIT-IBM Watson AI Lab), Jiyong Zhang (Hangzhou Dianzi University), Wu Liu (JD Explore Academy), Chenggang Yan (Hangzhou Dianzi University, Tsinghua University)
		mmfp1469	Virtual	Generic Image Manipulation Localization through the Lens of Multi-scale Spatial Inconsistence -- Zan Gao (Qilu University of Technology), Shenghao Chen (Qilu University of Technology), Yangyang Guo (National University of Singapore), WEILI GUAN (Monash University), Jie Nie (Ocean University of China), Anan Liu (university of tianjin, china)
		mmfp1476	On-site	Beyond Geo-localization: Fine-grained Orientation of Street-view Images by Cross-view Matching with Satellite Imagery -- Wenmiao Hu (National University of Singapore), Yichen Zhang (National University of Singapore), Yuxuan Liang (National University of Singapore), Yifang Yin (I2R, A*STAR), Andrei Georgescu (Grab Chronos S.R.L.), An Tran (Grab Taxi Holdings), Hannes Kruppa (GrabTaxi Holdings Pte. Ltd.), See-Kiong Ng (National University of Singapore), Roger Zimmermann (National University of Singapore)
		mmfp1655	Virtual	Region-based Pixels Integration Mechanism for Weakly Supervised Semantic Segmentation -- Chen Qian (Tsinghua University, Tsinghua University), Hui Zhang (Tsinghua University, Tsinghua University)
		mmfp2812	Virtual	Point Cloud Completion via Multi-Scale Edge Convolution and Attention -- Rui Cao (Fudan University), Kaiyi Zhang (Fudan University), Yang Chen (Fudan University), Ximing Yang (Fudan University), Cheng Jin (Fudan University)
		mmfp0873	Virtual	CRNet: Unsupervised Color Retention Network for Blind Motion Deblurring -- Suiyi Zhao (Hefei University of Technology), Zhao Zhang (Hefei University of Technology), Richang Hong (Hefei University of Technology), Mingliang Xu (Zhengzhou University), Haijun Zhang (Harbin Institute of Technology, Shenzhen), Meng Wang (Hefei University of Technology), Shuicheng YAN (National University of Singapore)
Understanding Multimedia Content: Vision and Language		mmfp2668	Virtual	MVPTR: Multi-Level Semantic Alignment for Vision-Language Pre-Training via Multi-Stage Learning -- Zejun Li (Fudan University), Zhihao Fan (Fudan University), Huaixiao Tou (Fudan University), Jingjing Chen (Fudan University), zhongyu wei (Fudan University), Xuanjing Huang (Fudan University)
		mmfp2368	On-site	Combining Vision and Language Representations for Patch-based Identification of Lexico-Semantic Relations -- Prince Jha (Indian Institute of Technology, Patna), Gaël Dias (University of Caen Normandy), Alexis Lechervy (Université de Caen Basse Normandie), Jose Moreno (Université Paul Sabatier / Université de Toulouse III), Anubhav Jangra (Google), Sebastião Pais (Universidade da Beira Interior), Sriparna Saha (Indian Institute of Technology Patna, India)
		mmfp0084	Virtual	Multi-Attention Network for Compressed Referring Video Object Segmentation -- Weidong Chen (University of Chinese Academy of Sciences), Dexiang Hong (, Chinese Academy of Sciences), Yuankai Qi (The University of Adelaide), Zhenjun Han (University of the Chinese Academy of Sciences), Shuhui Wang (Institute of Computing Technology, Chinese Academy of Sciences), Laiyun Qing (University of the Chinese Academy of Sciences), Qingming Huang (University of Chinese Academy of Sciences), Guorong Li (University of Chinese Academy of Sciences)
		mmfp0046	Virtual	Cross-modal Co-occurrence Attributes Alignments for Person Search by Language -- Kai Niu (Northwest Polytechnical University Xi'an), Linjiang Huang (The Chinese University of Hong Kong), Yan Huang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), PENG WANG (Northwestern Polytechnical University), Liang Wang (CASIA), Yanning Zhang (Northwestern Polytechnical University)
		mmfp0118	On-site	RefCrowd: Grounding the Target in Crowd with Referring Expressions -- Heqian Qiu (University of Electronic Science and Technology of China), Hongliang Li (University of Electronic Science and Technology of China, Tsinghua University), Taijin Zhao (University of Electronic Science and Technology of China), Lanxiao Wang (University of Electronic Science and Technology of China), Qingbo Wu (University of Electronic Science and Technology of China), Fanman Meng (University of Electronic Science and Technology of China, Tsinghua University)
		mmfp0472	Virtual	Unified Normalization for Accelerating and Stabilizing Transformers -- Qiming Yang (Hikvision Research Institute), Kai Zhang (Hikvision Research Institute), Chaoxiang Lan (Zhejiang University), Zhi Yang (University of Science and Technology of China), Zheyang Li (Shanghai Jiao Tong University), Wenming Tan (Hikvision Research Institute), Jun Xiao (Zhejiang University), Shiliang Pu (Zhejiang University)
		mmfp1164	Virtual	Enhancing Semi-Supervised Learning with Cross-Modal Knowledge -- Hui Zhu (Institute of Computing Technology, Chinese Academy of Sciences), Yongchun Lu (Mashang Consumer Finance Co., Ltd.), HONGBIN WANG (Mashang Consumer Finance), Xunyi Zhou (Mashang Consumer Finance Co., Ltd.), Qin Ma (China Agricultural University), Yanhong Liu (Mashang Consumer Finance), ning jiang (Mashang Consumer Finance Co, Ltd), xin wei (Mashang Consumer Finance Co, Ltd), Linchengxi Zeng (Mashang Consumer Finance Co., Ltd.), Xiaofang Zhao (Institute of Computing Technology, Chinese Academy of Sciences)
		mmfp1298	Virtual	Dynamic Spatio-Temporal Modular Network for Video Question Answering -- Zi Qian (Tsinghua Univ.), Xin Wang (Tsinghua University), Xuguang Duan (Tsinghua University, Tsinghua University), Hong Chen (Tsinghua University, Tsinghua University), Wenwu Zhu (Tsinghua University, Tsinghua University)
		mmfp1465	Virtual	Micro-video Tagging via Jointly Modeling Social Influence and Tag Relation -- Xiao Wang (Shandong University), Tian Gan (Shandong University), Yinwei Wei (National University of Singapore), Jianlong Wu (Shandong University), dai meng (Baidu), Liqiang Nie (Harbin Institute of Technology (Shenzhen))
		mmfp1744	Virtual	MimCo: Masked Image Modeling Pre-training with Contrastive Teacher -- Qiang Zhou (Alibaba Group), Chaohui Yu (Alibaba Group), Hao Luo (Alibaba Group), Zhibin Wang (Alibaba Group), Li Hao (Alibaba Group)
		mmfp2127	Virtual	Multimodal Hate Speech Detection via Cross-Domain Knowledge Transfer -- Chuanpeng Yang (University of Chinese Academy of Sciences), Fuqing Zhu (Institute of Information Engineering, Chinese Academy of Sciences), Guihua Liu (University of Chinese Academy of Sciences), Jizhong Han (Institute of Information Engineering), Songiln Hu (Institute of Information Engeering, Chinese Academy of Sciences)
		mmfp1985	On-site	ARMANI: Part-level Garment-Text Alignment for Unified Cross-Modal Fashion Design -- xujie zhang (SUN YAT-SEN UNIVERSITY), Michael Kampffmeyer (UiT The Arctic University of Norway), Yu Sha (SUN YAT-SEN UNIVERSITY), Zhenyu Xie (SUN YAT-SEN UNIVERSITY), ZEQUN JIE (Meituan), Huang Chengwen (Shidi Inc.), Jianqing Peng (SUN YAT-SEN UNIVERSITY), Xiaodan Liang (SUN YAT-SEN UNIVERSITY)
		mmfp0197	Virtual	Skimming, Locating, then Perusing: A Human-Like Framework for Natural Language Video Localization -- Daizong Liu (Peking University), Wei Hu (Peking University)
		mmfp0245	On-site	Distance Matters in Human-Object Interaction Detection -- Guangzhi Wang (National University of Singapore), Yangyang Guo (National University of Singapore), Yongkang Wong (National University of Singapore), Mohan Kankanhalli (National University of Singapore)
		mmfp1510	Virtual	Token Embeddings Alignment For Cross-Modal Retrieval -- Chen-Wei Xie (Alibaba Group), wu Min (Northwest Polytechnical University Xi'an), Yun Zheng (SUN YAT-SEN UNIVERSITY), Pan Pan (Alibaba Group), Xian-Sheng Hua (Alibaba)
		mmfp0936	Virtual	From Token to Word: OCR Token Evolution via Contrastive Learning and Semantic Matching for Text-VQA -- Zan-Xia Jin (University of Science and Technology Beijing), Zheng Shou (National University of Singapore), Fang Zhou (University of Science and Technology Beijing), Satoshi Tsutsui (Indiana University), Jingyan Qin (University of Science and Technology Beijing), Xu-cheng Yin (University of Science and Technology Beijing)
		mmfp1512	On-site	IDEA: Increasing Text Diversity via Online Multi-Label Recognition for Vision-Language Pre-training -- Xinyu Huang (Fudan University), Youcai Zhang (Fudan University), Ying Cheng (Fudan University), Weiwei Tian (Fudan University), Rui-Wei Zhao (Fudan University), Rui Feng (Fudan University), Yuejie Zhang (Fudan University), Yaqian Li (OPPO Research Institute), Yandong Guo (OPPO Research Institute), Xiaobo Zhang (Fudan University)
		mmfp2709	Virtual	CLOP: Video-and-Language Pre-Training with Knowledge Regularizations -- Guohao Li (Tsinghua University, Tsinghua University), Hu Yang (Beijing University of Posts and Telecommunications), Feng He (Sichuan University), Zhifan Feng (Baidu), Yajuan Lyu (Baidu), hua wu (Baidu), Haifeng Wang (Baidu)
		mmfp1871	Virtual	Talk2Face: A Unified Sequence-based Framework for Diverse Face Generation and Analysis Tasks -- Yudong Li (Department of Software Engineering, Shenzhen University), Xianxu Hou (Shenzhen University), Zhe Zhao (Tencent AI Lab), Linlin Shen (Shenzhen University), Xuefeng Yang (Nanyang Technological University), Kimmo Yan (Tencent)
		mmfp1002	Virtual	TxVAD: Improved Video Action Detection by Transformers -- Zhenyu Wu (Texas A&M), Zhou Ren (Wormpex AI Research), Yi Wu (Wormpex AI Research), Zhangyang Wang (University of Texas, Austin), Gang Hua (Wormpex AI Research)
		mmfp0031	Virtual	Relational Representation Learning in Visually-Rich Documents -- Xin Li (Tencent Youtu Lab), Yan Zheng (Tencent Youtu Lab), Yiqing Hu (Tencent Youtu Lab), Haoyu Cao (Tencent Youtu Lab), Yunfei Wu (Tencent YouTu Lab), Deqiang Jiang (Tencent YouTu Lab), YINSONG LIU (Tencent Youtu Lab), Bo Ren (Tencent Youtu Lab)
		mmfp0926	Virtual	Unified Multimodal Model with Unlikelihood Training for Visual Dialog -- Zihao Wang (Tongji University), Junli Wang (Tongji University), changjun jiang (Tongji University)
		mmfp0982	Virtual	Tackling Instance-Dependent Label Noise with Dynamic Distribution Calibration -- Manyi Zhang (Tsinghua University), Yuxin Ren (Tsinghua University), Zihao Wang (Tsinghua University), Chun Yuan (Tsinghua University, Tsinghua University)
		mmfp0257	On-site	On Leveraging Variational Graph Embeddings for Open World Compositional Zero-Shot Learning -- Muhammad Umer Anwaar (Technical University Munich), Zhihui Pan (Department of Informatics, Technische Universität München), Martin Kleinsteuber (Technical University of Munich)
		mmfp1586	Virtual	Comprehensive Relationship Reasoning for Composed Query Based Image Retrieval -- Feifei Zhang (Tianjin University of Technology), Ming Yan (Alibaba Group), Ji Zhang (Alibaba Group), Changsheng Xu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
		mmfp1675	Virtual	Image Understanding by Captioning with Differentiable Architecture Search -- Ramtin Hosseini (University of California, San Diego), Pengtao Xie (University of California, San Diego)
		mmfp2721	Virtual	Atrous Pyramid Transformer with Spectral Convolution for Image Inpainting -- Muqi Huang (Wuhan University), Lefei Zhang (Wuhan University)
		mmfp2025	Virtual	QuadTreeCapsule: QuadTree Capsules for Deep Regression Tracking -- Ding Ma (Harbin Institute of Technology), Xiangqian Wu (Harbin Institute of Technology)
		mmfp0267	Virtual	End-to-End 3D Face Reconstruction with Expressions and Specular Albedos from Single In-the-wild Images -- Qixin Deng (University of Houston), Binh Le (University of Houston), Aobo Jin (University of Houston-Victoria), Zhigang Deng (University of Houston)
		mmfp2754	Virtual	Heterogeneous Learning for Scene Graph Generation -- Yunqing He (Nanjing University), Tongwei Ren (Nanjing University), Jinhui Tang (Nanjing University of Science and Technology), Gangshan Wu (Nanjing University)
		mmfp1199	Virtual	Equivariant and Invariant Grounding for Interpretable Video Question Answering -- Yicong Li (National University of Singapore), Xiang Wang (National University of Singapore), Junbin Xiao (National University of Singapore), Tat-seng Chua (National University of Singapore)
		mmfp0925	Virtual	Align and Adapt: A Two-stage Adaptation Framework for Unsupervised Domain Adaptation -- Yan YU (Zhejiang University), Yuchen Zhai (Zhejiang University, Tsinghua University), Yin Zhang (Zhejiang University)

		mmfp2354	On-site	Detach and Attach: Stylized Image Captioning without Paired Stylized Dataset -- Yutong Tan (University of Chinese Academy of Sciences), zheng Lin (Institute of Information Engineering, Chinese Academy of Sciences), Peng Fu (Institute of Information Engineering, Chinese Academy of Sciences), mingyu Zheng (University of Chinese Academy of Sciences), Lanrui Wang (Beijing University of Posts and Telecommunications), Yanan Cao (Institute of Information Engineering, Chinese Academy of Sciences), Weipinng Wang (IIE)
		mmfp1294	Virtual	PixelSeg: Pixel-by-Pixel Stochastic Semantic Segmentation for Ambiguous Medical Images -- Wei Zhang (Chongqing University), Xiaohong Zhang (Chongqing University), Sheng Huang (Chongqing University), Lu Yuting (Chongqing University), Kun Wang (Chongqing University)
		mmfp1549	Virtual	A Probabilistic Model for Controlling Diversity and Accuracy of Ambiguous Medical Image Segmentation -- Wei Zhang (Chongqing University), Xiaohong Zhang (Chongqing University), Sheng Huang (Chongqing University), Lu Yuting (Chongqing University), Kun Wang (Chongqing University)
		mmfp2092	Virtual	Crossmodal Few-shot 3D Point Cloud Semantic Segmentation -- Ziyu Zhao (UofSC), Zhenyao Wu (University of South Carolina), Xinyi Wu (University of South Carolina), Canyu Zhang (University of South Carolina), Song Wang (University of South Carolina)
		mmfp1795	Virtual	VQ-DcTr: Vector-Quantized Autoencoder With Dual-channel Transformer Points Splitting for 3D Point Cloud Completion -- Ben Fei (Fudan University), Weidong Yang (Fudan University), Wen-Ming Chen (Fudan University), Lipeng Ma (Fudan University)
		mmfp1239	Virtual	Fine-grained Action Recognition with Robust Motion Representation Decoupling and Concentration -- Baoli Sun (Dalian University of Technology), Xinchen Ye (Dalian University of Technology), Tiantian Yan (Dalian University of Technology), Zhihui Wang (Dalian University of Technology), Haojie Li (Dalian University of Technology), Zhiyong Wang (The University of Sydney)
		mmfp0207	Virtual	Concept Propagation via Attentional Knowledge Graph Reasoning for Video-Text Retrieval -- Sheng Fang (, Chinese Academy of Sciences), Shuhui Wang (Institute of Computing Technology, Chinese Academy of Sciences), Junbao Zhuo (, Chinese Academy of Sciences), Qingming Huang (Univ of Chinese Academy of Sciences), Bin MA (Meituan), Wei Xiaoming (Beijing University of Aeronautics and Astronautics), Xiaolin Wei (Meituan)
		mmfp2217	On-site	Domain Generalization via Frequency-domain-based Feature Disentanglement and Interaction -- Jingye Wang (Beijing University of Posts and Telecommunications), Ruoyi Du (Beijing University of Posts and Telecommunications), Chang Liang (Beijing University of Post and Telecommunication), Kongming Liang (Beijing University of Posts and Telecommunications), Zhanyu Ma (Beijing University of Post and Telecommunication)
		mmfp3226	Virtual	Immunofluorescence Capillary Imaging Segmentation: Cases Study -- runpeng hou (Southern University of Science and Technology), Ziyuan Ye (Southern University of Science and Technology), Chengyu Yang (Southern University of Science and Technology), Linhao Fu (Southern University of Science and Technology), Chao Liu (New York University), Quanying Liu (Southern University of Science and Technology)
		mmfp3119	Virtual	Imitated Detectors: Stealing Knowledge of Black-box Object Detectors -- Siyuan Liang (Institute of Information Engineering, Chinese Academy of Sciences (iie.ac.cn)), Aishan Liu (Beijing University of Aeronautics and Astronautics), Jiawei Liang (SUN YAT-SEN UNIVERSITY), Longkang Li (the Chinese University of Hong Kong (Shenzhen)), Yang Bai (University of Electronic Science and Technology of China), Xiaochun Cao (Institute of Information Engineering)
Engaging Users with Multimedia: Multimedia Search and Recommendation		mmfp0880	Virtual	Video Moment Retrieval with Hierarchical Contrastive Learning -- Bolin Zhang (Hunan University), Chao Yang (Hunan University), Bin Jiang (Hunan University), Xiaokang Zhou (Shiga University)
		mmfp2816	On-site	Learning to Retrieve Videos by Asking Questions -- Avinash Madasu (University of North Carolina at Chapel Hill), Junier Oliva (University of North Carolina, Chapel Hill), Gedas Bertasius (University of North Carolina, Chapel Hill)
		mmfp1588	Virtual	HEART:A Statistical Perspective for Learning to Hash with Label Noise -- Jinan Sun (Peking University), Haixin Wang (Peking University), Xiao Luo (Peking University), Shikun Zhang (Peking University), Wei Xiang (University of Texas, Arlington), Chong Chen (Alibaba Group), Xian-Sheng Hua (Zhejiang University)
		mmfp1564	Virtual	Learning Hybrid Behavior Patterns for Multimedia Recommendation -- Zongshen Mu (Zhejiang University, Tsinghua University), Yueting Zhuang (Zhejiang University), Jie Tan (The Chinese University of Hong Kong), Jun Xiao (Zhejiang University), Siliang Tang (Zhejiang University)
		mmfp1289	Virtual	Fine-Grained Relational Dependency and Bidirectional Generative Networks]{Image-Text Matching with Fine-Grained Relational Dependency and Bidirectional Attention-Based Generative Networks -- Jianwei Zhu (Guangxi Normal University), Zhixin Li (Guangxi Normal University), Yufei Zeng (Guangxi Normal University), Jiahui Wei (Guangxi Normal University), Huifang Ma (Northwest Normal University Lanzhou)
		mmfp2485	Virtual	Visual Grounding in Remote Sensing Images -- Yuxi Sun (Harbin Institute of Technology, Shenzhen), Shanshan Feng (Harbin Institute of Technology), Xutao Li (Harbin Institute of Technology, Shenzhen), Yunming Ye (Harbin Institute of Technology, Shenzhen), Jian Kang (Suzhou University), Xu Huang (Harbin Institute of Technology, Shenzhen)
		mmfp1060	Virtual	Prompt-based Zero-shot Video Moment Retrieval -- Guolong Wang (University of International Business and Economics), Xun Wu (Tsinghua University, Tsinghua University), Zhaoyuan Liu (Qilu University of Technology (Shandong Academy of Sciences)), Junchi Yan (Shanghai Jiao Tong University)
		mmfp1057	Virtual	Cross-Lingual Cross-Modal Retrieval with Noise-Robust Learning -- Yabing Wang (Zhejiang Gongshang University), Jianfeng Dong (Zhejiang Gongshang University), Tianxiang Liang (Zhejiang Gongshang University), Z MS (Zhejiang Gongshang University), Rui Cai (Zhejiang Gongshang University), Xun Wang (Zhejiang Gongshang University)
		mmfp0907	Virtual	Learn to Understand Negation in Video Retrieval -- Ziyue Wang (Renmin University of China), Aozhu Chen (Renmin University of China), Fan Hu (Renmin University of China), Xirong Li (Renmin University of China)
		mmfp1962	Virtual	AdsCVLR: Commercial Visual-Linguistic Representation Modeling in Sponsored Search -- Yongjie Zhu (Beijing University of Posts and Telecommunications), Chunhui Han (Microsoft), Yuefeng Zhan (Microsoft), Bochen Pang (Peking University), Zhaoju Li (Microsoft), Hao Sun (Microsoft), Si Li (Beijing University of Posts and Telecommunications), Boxin Shi (Peking University), Nan Duan (Microsoft Research Asia), Weiwei Deng (South China University of Technology), Ruofei Zhang (Microsoft), Liangjie Zhang (Microsoft), Qi Zhang (Washington University, St. Louis)
		mmfp1814	Virtual	Differentiable Cross-modal Hashing via Multimodal Transformers -- Junfeng Tu (Hefei University of Technology), Xueliang Liu (Hefei University of Technology), Zongxiang Lin (College of Computer Science and Technology, Zhejiang University), Richang Hong (Hefei University of Technology), Meng Wang (Hefei University of Technology)
Engaging Users with Multimedia: Emotional and Social Signals		mmfp1134	On-site	Representation Learning through Multimodal Attention and Time-Sync Comments for Affective Video Content Analysis -- Jicai Pan (University of Science and Technology of China), Shangfei Wang (University of Science and Technology of China,), Lin Fang (University of Science and Technology of China)
		mmfp2222	Virtual	TFF-Former: Temporal-Frequency Fusion Transformer for Zero-training Decoding of Two BCI Tasks -- Xujin Li (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Wei Wei (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Shuang Qiu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Huiguang He (Institute of Automation Chinese Academy of Sciences)
		mmfp0767	Virtual	Towards Unbiased Visual Emotion Recognition via Causal Intervention -- Yuedong Chen (Monash University), Xu Yang (Nanyang Technological University), Tat-Jen Cham (Nanyang Technological University), Jianfei Cai (Monash University)
		mmfp2821	On-site	A Dataset of Bodily Behaviors in Social Interaction -- Michal Balazia (INRIA), Philipp Müller (German Research Center for AI), Ákos Tánczos (INRIA), August Liechtenstein (German Research Center for AI), Francois Bremond (inria)
		mmfp2871	On-site	Learning from label relationships in human affect -- Niki Foteinopoulou (Queen Mary, University of London), Ioannis Patras (Queen Mary University of London)
		mmfp2148	Virtual	Brain Topography Adaptive Network for Satisfaction Modeling in Interactive Information Access System -- Ziyi Ye (Tsinghua University, Tsinghua University), Xiaohui Xie (Tsinghua University, Tsinghua University), Yiqun LIU (Tsinghua University), Zhihong Wang (Tsinghua University, Tsinghua University), Xuesong Chen (Tsinghua University, Tsinghua University), Min Zhang (Tsinghua University, Tsinghua University), Shaoping Ma (Tsinghua University)
Engaging Users with Multimedia: Summarization, Analytics, and Storytelling		mmfp1306	Virtual	Weakly-Supervised Temporal Action Alignment Driven by Unbalanced Spectral Fused Gromov-Wasserstein Distance -- Dixin Luo (Beijing Institute of Technology), Yutong Wang (Beijing Institute of Technology), Angxiao Yue (Beijing Institute of Technology), Hongteng Xu (Renmin University of China)
		mmfp1409	Virtual	A Knowledge Augmented and Multimodal-Based Framework for Video Summarization -- Jiehang Xie (Nankai University), Xuanbai Chen (CMU, Carnegie Mellon University), Shao-Ping Lu (Nankai University), Yulu Yang (Nankai University)
		mmfp0093	Virtual	FedMed-ATL: Misaligned Unpaired Cross-Modality Neuroimage Synthesis via Affine Transform Loss -- Jinbao Wang (South University of Science and Technology of China), Guoyang Xie (University of Surrey), Yawen Huang (Tencent), Yefeng Zheng (Tencent), Yaochu Jin (Universität Bielefeld), Feng Zheng (Southern University of Science and Technology)
		mmfp0818	Virtual	Towards Blind Watermarking: Combining Invertible and Non-invertible Mechanisms -- Rui Ma (Peking University), Mengxi Guo (Bytedance), Yi Hou (Peking University), Fan Yang (Peking University), Yuan Li (Peking University), Huizhu Jia (Peking University), Xiaodong Xie (Peking University)
		mmfp2052	Virtual	Improving Transferability for Domain Adaptive Detection Transformers -- Kaixiong Gong (Beijing Institute of Technology, Tsinghua University), Shuang Li (Beijing Institute of Technology), Li Gang (Beijing Institute of Technology), Rui Zhang (Beijing Institute of Technology), Chi Liu (Beijing Institute of Technology), Qiang Chen (Baidu)
		mmfp3209	Virtual	Support for Teaching Mathematics of the Blind by Sighted Tutors Through Multisensual Access to Formulas with Braille Converters and Speech -- Dariusz Mikulowski (Siedlce University of Natural Sciences and Humanities)
		mmfp2587	Virtual	Geometry Aligned Variational Transformer for Image-conditioned Layout Generation -- Yunning Cao (University of Science and Technology of China), Ye Ma (Alibaba Group), Min Zhou (Beijing University of Aeronautics and Astronautics), Chuanbin Liu (University of Science and Technology of China), Hongtao Xie (University of Science and Technology of China), Tiezheng Ge (Alibaba Group), Yuning Jiang (University of Science and Technology of China, Tsinghua University)
		mmfp0588	Virtual	PVSeRF: Joint Pixel-, Voxel- and Surface-Aligned Radiance Field for Single-Image Novel View Synthesis -- Xianggang Yu (The Chinese University of Hong Kong, Shenzhen), Jiapeng Tang (Technische Universität München), Yipeng Qin (Cardiff University), Chenghong Li (The Chinese University of Hong Kong, Shenzhen), Linchao Bao (Tencent AI Lab), Xiaoguang Han (The Chinese University of Hong Kong, Shenzhen), Shuguang Cui (The Chinese University of Hong Kong, Shenzhen)
		mmfp0277	Virtual	Cross-Modality High-Frequency Transformer for MR Image Super-Resolution -- Chaowei Fang (Xidian University), Dingwen Zhang (Northwestern Polytechnical University), liang wang (iRobot), Yulun Zhang (Swiss Federal Institute of Technology), Lechao Cheng (Zhejiang Lab), Junwei Han (Northwestern Polytechnical University, Tsinghua University)
		mmfp0337	Virtual	Adma-GAN: Attribute-Driven Memory Augmented GANs for Text-to-Image Generation. -- Xintian Wu (Zhejiang University), Hanbin Zhao (Zhejiang University), Liangli Zheng (Zhejiang University), Shouhong Ding (Tencent Youtu Lab), Xi Li (Zhejiang University)
		mmfp1678	Virtual	Efficient Multiple Kernel Clustering via Spectral Perturbation -- Chang Tang (China University of Geosciences), Zhenglai Li (China University of Geosciences Wuhan), Weiqing Yan (Yantai University), Guanghui Yue (Shenzhen University), Wei Zhang (Qilu University of Technology)
		mmfp1869	Virtual	DOMFN: A Divergence-Orientated Multi-Modal Fusion Network for Resume Assessment -- Yang Yang (Nanjing University of Science and Technology), Jingshuai Zhang (Baidu), Fan Gao (Tokyo Institute of Technology, Tokyo Institute of Technology), Xiaoru Gao (Rutgers University, Newark), Hengshu Zhu (Baidu)
		mmfp1912	Virtual	Generative Steganography Network -- Ping Wei (Fudan University), Sheng Li (Fudan University), Xinpeng Zhang (Fudan University), Ge Luo (Fudan University), Zhenxing Qian (Fudan University), Qing Zhou (Fudan University)
		mmfp1161	Virtual	You Only Hypothesize Once: Point Cloud Registration with Rotation-equivariant Descriptors -- Haiping Wang (Wuhan University), Yuan Liu (The University of Hong Kong), Zhen Dong (Wuhan University), Wenping Wang (HKU)
		mmfp0048	Virtual	Disentangled Representation Learning for Multimodal Emotion Recognition -- Dingkang Yang (Fudan University), Shuai Huang (Fudan University), Haopeng Kuang (Fudan University), Yangtao Du (Fudan University), Lihua Zhang (Tsinghua University, Tsinghua University)
		mmfp1086	Virtual	Relative Alignment Network for Source-Free Multimodal Video Domain Adaptation -- Yi Huang (Institute of automation, Chinese academy of science), Xiaoshan Yang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Ji Zhang (Alibaba Group), Changsheng Xu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
		mmfp1866	Virtual	PRO-Face: a generic framework for Privacy-preserving Recognizable Obfuscation of Face images -- Lin Yuan (Chongqing University of Post and Telecommunications), linguo liu (Chongqing University of Post and Telecommunications), Xiao Pu (Chongqing University of Post and Telecommunications), li zhao (Chongqing University of Post and Telecommunications), Hongbo Li (Chongqing University of Post and Telecommunications), Xinbo Gao (Chongqing University of Post and Telecommunications)
		mmfp0300	Virtual	Skeleton-Based Action Recognition via Adaptive Cross-Form Learning -- Xuanhan Wang (University of Electronic Science and Technology of China), Yan Dai (University of Electronic Science and Technology of China), Lianli Gao (University of Electronic Science and Technology of China, Tsinghua University), Jingkuan Song (University of Electronic Science and Technology of China, Tsinghua University)
		mmfp0713	On-site	Sample Weighted Multiple Kernel K-means via min-max optimization -- Yi Zhang (National University of Defense Technology), Weixuan Liang (National University of Defense Technology), Xinwang Liu (National University of Defense Technology), Sisi Dai (National University of Defense Technology), Siwei Wang (National University of Defense Technology), Liyang Xu (National University of Defense Technology), En Zhu (National University of Defense Technology)
		mmfp0641	On-site	MIntRec: A New Dataset for Multimodal Intent Recognition -- Hanlei Zhang (Tsinghua University), Hua Xu (Tsinghua University, Tsinghua University), Xin Wang (Hebei University of Science & Technology), Qianrui Zhou (Tsinghua University, Tsinghua University), Shaojie Zhao (Hebei University of Science & Technology), Jiayan Teng (Tsinghua University, Tsinghua University)
		mmfp1610	Virtual	Adaptive Transformer-Based Conditioned Variational Autoencoder for Incomplete Social Event Classification -- Zhangming Li (University of the Chinese Academy of Sciences), Shengsheng Qian (, Institute of automation, Chinese academy of science), Cao Jie (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Quan Fang (, Institute of automation, Chinese academy of science), Changsheng Xu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
		mmfp0050	Virtual	Learning Modality-Specific and -Agnostic Representations for Asynchronous Multimodal Language Sequences -- Dingkang Yang (Fudan University), Haopeng Kuang (Fudan University), Shuai Huang (Fudan University), Lihua Zhang (Tsinghua University, Tsinghua University)

Experience: Multimedia Applications		mmfp1407	Virtual	DoF-NeRF: Depth-of-Field Meets Neural Radiance Fields -- Zijin Wu (Huazhong University of Science and Technology), Xingyi Li (Huazhong University of Science and Technology), Juewen Peng (Huazhong University of Science and Technology), Hao Lu (Huazhong University of Science and Technology), Zhiguo Cao (Huazhong University of Science and Technology), Weicai Zhong (Xidian University)
		mmfp0321	Virtual	RKformer: Runge-Kutta Transformer with Random-Connection Attention for Infrared Small Target Detection -- Mingjin Zhang (Xidian University), Haichen Bai (Xidian University), Jing Zhang (The University of Sydney), 睿 张 (Xidian University), Chaoyue Wang (JD Explore Academy), Jie Guo (Xidian University), Xinbo Gao (Xidian University)
		mmfp0123	Virtual	Self-Supervised Human Pose based Multi-Camera Video Synchronization -- liqiang yin (Tianjin University), Ruize Han (Tianjin University), Wei Feng (Tianjin University), Song Wang (University of South Carolina)
		mmfp1345	Virtual	Energy-Based Domain Generalization for Face Anti-Spoofing -- Zhekai Du (University of Electronic Science and Technology of China), Jingjing Li (University of Electronic Science and Technology of China), Lin Zuo (University of Electronic Science and Technology of China), Lei Zhu (Shandong Normal University), Ke Lu (University of Electronic Science and Technology of China)
		mmfp0302	Virtual	Revisiting Stochastic Learning for Generalizable Person Re-identification -- Jiajian Zhao (Beihang University), Yifan Zhao (Peking University), Xiaowu Chen (Beihang University), Jia Li (Beihang University)
		mmfp1054	Virtual	D^2Animator: Dual Distillation of StyleGAN For High-Resolution Face Animation -- Zhuo Chen (Tsinghua University, Tsinghua University), Chaoyue Wang (JD Explore Academy), Haimei Zhao (University of Sydney), Bo Yuan (Tsinghua University), Xiu Li (Tsinghua University, Tsinghua University)
		mmfp1461	On-site	Adaptive Hierarchical Pooling for Weakly-supervised Sound Event Detection -- Lijian Gao (Jiangsu University), Ling Zhou (Computer Science), Qirong Mao (Jiangsu University), Ming Dong (Wayne State University)
		mmfp1666	Virtual	Mutual Adaptive Reasoning for Monocular 3D Multi-Person Pose Estimation -- Juze Zhang (ShanghaiTech University), Jingya Wang (ShanghaiTech University), Ye Shi (ShanghaiTech University), Fei Gao (ShanghaiTech University), Lan Xu (ShanghaiTech University), Jingyi Yu (ShanghaiTech University)
		mmfp2263	Virtual	Learning Generalizable Latent Representations for Novel Degradations in Super-Resolution -- fengjun li (Harbin Institute of Technology, Shenzhen), Xin Feng (Harbin Institute of Technology), Fanglin Chen (Harbin Institute of Technology (Shenzhen)), Guangming Lu (Harbin Institute of Technology), Wenjie Pei (Harbin Institute of Technology, Shenzhen)
		mmfp2962	Virtual	Rethinking the Vulnerability of DNN Watermarking: Are Watermarks Robust against Naturalness-aware Perturbations? -- Run Wang (Wuhan University), Haoxuan Li (Wuhan University), Lingzhou Mu (Wuhan University), Jixing Ren (Wuhan University), Shangwei Guo (Chongqing University), Li Liu (Fudan University), Liming Fang (Nanjing University of Aeronautics and Astronautics), Jing Chen (Wuhan University), Lina Wang (Wuhan University)
		mmfp0666	Virtual	In-N-Out Generative Learning for Dense Unsupervised Video Segmentation -- Xiao Pan (Zhejiang University), Peike Li (University of Technology Sydney), Zongxin Yang (University of Technology Sydney), Huiling Zhou (Hong Kong Polytechnic University), Chang Zhou (Peking University), Hongxia Yang (Duke University), Jingren Zhou (Columbia University), Yi Yang (Zhejiang University)
		mmfp0916	On-site	Everything is there in the Latent Space: Image Editing by Latent Space Manipulation -- Rishubh Parihar (Indian Institute of Science, Bangalore), Ankit Dhiman (Indian Institute of Science, Indian institute of science, Bangalore), Tejan Karmali (Indian Institute of Science, Bengaluru), Venkatesh R (Indian Institute of Science)
		mmfp2538	Virtual	An Image-to-video Model for Real-Time Video Enhancement -- Dongyu She (Tsinghua University, Tsinghua University), Kun Xu (Tsinghua University, Tsinghua University)
		mmfp0957	Virtual	Learning an Inference-accelerated Network from a Pre-trained Model with Frequency-enhanced Feature Distillation -- Xuesong Niu (Kuaishou Technology), Jili Gu (Kwai Inc), Guoxin Zhang (Kwai Inc.), Pengfei Wan (Kuaishou Technology), Zhongyuan Wang (Kuaishou Inc.)
		mmfp2194	Virtual	Exploring Feature Compensation and Cross-level Correlation for Infrared Small Target Detection -- Mingjin Zhang (Xidian University), Ke Yue (Xidian University), Jing Zhang (The University of Sydney), Yunsong Li (Xidian University), Xinbo Gao (Xidian University)
		mmfp1376	Virtual	Pixel Exclusion: Uncertainty-aware Boundary Discovery for Active Cross-Domain Semantic Segmentation -- fuming you (University of Electronic Science and Technology of China), Jingjing Li (University of Electronic Science and Technology of China), Zhi Chen (University of Queensland), Lei Zhu (Shandong Normal University)
		mmfp0461	Virtual	Deep Flexible Structure Preserving Image Smoothing -- Mingjia Li (Tianjin University), Yuanbin Fu (Tianjin University), Xinhui Li (Tianjin University), Xiaojie Guo (Tianjin University)
		mmfp2818	On-site	Defending Physical Adversarial Attack on Object Detection via Adversarial Patch-Feature Energy -- Taeheon Kim (Korea Advanced Institute of Science & Technology), Youngjoon Yu (Korea Advanced Institute of Science and Technology), Yong Ro (Korea Advanced Institute of Science and Technology)
		mmfp1301	On-site	Multiview Contrastive Learning for Completely Blind Video Quality Assessment of User Generated Content -- Shankhanil Mitra (Indian Institute of Science), Rajiv Soundararajan (University of Texas at Austin)
		mmfp0278	Virtual	Compound Batch Normalization for Long-tailed Image Classification -- Lechao Cheng (Zhejiang Lab), Chaowei Fang (Xidian University), Dingwen Zhang (Northwestern Polytechnical University), Guanbin Li (SUN YAT-SEN UNIVERSITY), Gang Huang (Zhejiang Lab)
		mmfp3213	Virtual	Alleviating Style Sensitivity then Adapting: Source-free Domain Adaptation for Medical Image Segmentation -- Yalan Ye (University of Electronic Science and Technology of China), Ziqi Liu (University of Electronic Science and Technology of China), Yangwuyong Zhang (University of Electronic Science and Technology of China), Jingjing Li (University of Electronic Science and Technology of China), Hengtao Shen (University of Electronic Science and Technology of China)
		mmfp1627	On-site	Multimedia Event Extraction From News With a Unified Contrastive Learning Framework -- Jian Liu (Beijing Jiaotong University), Yufeng Chen (Beijing jiaotong univercity), Xu Jinan (Beijing Jiaotong University)
		mmfp0343	On-site	DomainPlus: Cross Transform Domain Learning towards Efficient High Dynamic Range Imaging -- bolun zheng (Hangzhou Dianzi University), Xiaokai Pan (Hangzhou Dianzi University), Hua Zhang (Hangzhou Dianzi University), Xiaofei Zhou (Hangzhou Dianzi University), Greg Slabaugh (Queen Mary University London), Chenggang Yan (Hangzhou Dianzi University, Tsinghua University), Shanxin Yuan (Huawei Technologies Ltd.)
		mmfp1988	Virtual	Tracking Game: Self-adaptative Agent based Multi-object Tracking -- Shuai Wang (Beijing University of Aeronautics and Astronautics), Yubin Wu (Beijing University of Aeronautics and Astronautics), Da Yang (Beihang University), Yang Liu (Beihang University), Hao Sheng (Beijing University of Aeronautics and Astronautics)
		mmfp0630	On-site	Self-Supervised Text Erasing with Controllable Image Synthesis -- Gangwei Jiang (University of Science and Technology of China), Shiyao Wang (Alibaba Group), Tiezheng Ge (Alibaba Group), Yuning Jiang (University of Science and Technology of China, Tsinghua University), Ying Wei (City University of Hong Kong), Defu Lian (University of Science and Technology of China)
		mmfp1708	Virtual	Look Before You Leap: Improving Text-based Person Retrieval by Learning A Consistent Cross-modal Common Manifold -- Zijie Wang (Nanjing Tech University), Aichun Zhu (Nanjing Tech University), Jingyi Xue (Nanjing University), Xili Wan (Nanjing Tech University), Liu Chao (Jinling institute of technology), Tian Wang (Beihang University, China), Yifeng Li (Nanjing Tech University)
		mmfp2907	On-site	MONOPOLY: Financial Prediction from MONetary POLicY Conference Videos Using Multimodal Cues -- Puneet Mathur (University of Maryland, College Park), Atula Neerkaje (Manipal Institute of Technology), Malika Chhibber (Birla Institute of Technology and Science, K. K. Birla Goa Campus), Ramit Sawhney (Indraprastha Institute of Information Technology, Delhi, Dhirubhai Ambani Institute Of Information and Communication Technology), Fuming Guo (Fidelity Investments), Franck Dernoncourt (Adobe Systems), Sanghamitra Dutta (J.P. Morgan Chase), Dinesh Manocha (University of Maryland, College Park)
Experience: Art and Culture		mmfp2296	Virtual	Draw Your Art Dream: Diverse Digital Art Synthesis with Multimodal Guided Diffusion -- Nisha Huang (University of the Chinese Academy of Sciences), Fan Tang (Jilin University), Weiming Dong (Institute of Automation, Chinese Academy of Sciences), Changsheng Xu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
		mmfp0778	Virtual	AesUST: Towards Aesthetic-Enhanced Universal Style Transfer -- Zhizhong Wang (Zhejiang University), Zhanjie Zhang (Zhejiang University), Lei Zhao (Zhejiang University), Zhiwen Zuo (Zhejiang University), Ailin Li (College of computer science, Zhejiang University), Wei Xing (Zhejiang University), Dongming Lu (College of Computer Science and Technology, Zhejiang University)
		mmfp2849	On-site	Semi-supervised Human Pose Estimation in Art-historical Images -- Matthias Springstein (Leibniz Information Centre for Science and Technology), Stefanie Schneider (Ludwig-Maximilians-Universität München), Christian Althaus (University of Bern), Ralph Ewerth (L3S Research Center)
		mmfp3260	On-site	Understanding and Identifying Artwork Plagiarism with the Wisdom of Designers: A Case Study on Poster Artworks -- Shenglan Cui (Hunan University), FANG LIU (Hunan University), Tongqing Zhou (National University of Defense Technology), Mohan Zhang (Hunan University)
		mmfp0592	Virtual	REMOT: A Region-to-Whole Framework for Realistic Human Motion Transfer -- Quanwei Yang (University of Science and Technology of China), Xinchen Liu (JD Explore Academy), Wu Liu (JD Explore Academy), Hongtao Xie (University of Science and Technology of China), Xiaoyan Gu (Institute of Information Engineering, Chinese Academy of Sciences), Lingyun Yu (University of Science and Technology of China), Yongdong Zhang (University of Science and Technology of China)
Experience: Interactions and Quality of Experience		mmfp2906	Virtual	Improving meeting inclusiveness using speech interruption analysis -- Szu-Wei Fu (Microsoft), Yaran Fan (Microsoft), Yasaman Hosseinkashi (Microsoft), Jayant Gupchup (Johns Hopkins University), Ross Cutler (Microsoft)
		mmfp2044	Virtual	Transductive Aesthetic Preference Propagation for Personalized Image Aesthetics Assessment -- Yaohui Li (Department of Control Science and Intelligence Engineering, Nanjing University), Yuzhe Yang (OPPO Research Institute), Huaxiong Li (Nanjing University), Haoxing Chen (Nanjing University), LiWu Xu (Oppo Research Institute), Leida Li (Xidian University), Yaqian Li (OPPO Research Institute), Yandong Guo (OPPO Research Institute)
		mmfp1460	Virtual	Multi-Mode Interactive Image Segmentation -- Zheng Lin (Nankai University), Zhao Zhang (Nankai University), Ling-Hao Han (Nankai University), Shao-Ping Lu (Nankai University)
		mmfp2876	On-site	Deep-BVQM: A Deep-learning Bitstream-based Video Quality Model -- Nasim Jamshidiavanaki (Technische Universität Berlin), Steven Schmidt (Sony Interactive Entertainment), Thilo Michael (Technische Universität Berlin), Saman Zadtootaghaj (Dolby Laboratories), Sebastian Möller (Technische Universität Berlin)
Multimedia Systems: Transport and Delivery		mmfp2477	Virtual	Learning-Based Video Coding with Joint Deep Compression and Enhancement -- Tiesong Zhao (Fuzhou University), WeiZe Feng (Fuzhou University), HongJi Zeng (Fuzhou University), Yiwen Xu (Fuzhou University), Yuzhen Niu (Fuzhou University), Jiaying Liu (Peking University)
		mmfp1683	Virtual	Structure-Preserving Motion Estimation for Learned Video Compression -- Gao Han (University of Electronic Science and Technology of China), JINZHONG CUI (University of Electronic Science and Technology of China), Mao Ye (University of Electronic Science and Technology of China), Shuai Li (Shandong University), Yu Zhao (University of Electronic Science and Technology of China), Xiatian Zhu (University of Surrey)
		mmfp3267	Virtual	Learned Internet Congestion Control for Short Video Uploading -- Tianchi Huang (Tsinghua University), Chao Zhou (Peking University), Lianchen Jia (Tsinghua University, Tsinghua University), Ruixiao Zhang (Tsinghua University, Tsinghua University), Lifeng Sun (Tsinghua University, Tsinghua University)
Multimedia Systems: Data Systems Management and Indexing		mmfp1188	On-site	HyPS^2$ Loss: Beyond Hypersphere Metric Space for Multi-label Image Retrieval -- Chengyin Xu (Tsinghua University, Tsinghua University), Zenghao Chai (Tsinghua University, Tsinghua University), Zhengzhuo Xu (Tsinghua University, Tsinghua University), Chun Yuan (Tsinghua University, Tsinghua University), Yanbo Fan (Tencent AI Lab), Jue Wang (Tencent AI Lab)
Multimedia Systems: Systems and Middleware		mmfp0891	On-site	OISSR: Optical Image Stabilization Based Super Resolution on Smartphone Cameras -- Hao Pan (Shanghai Jiaotong University), Yi-Chao Chen (Shanghai Jiaotong University), Feitong Tan (Simon Fraser University), Wenhao Li (Shanghai Jiaotong University), Guangtao Xue (Shanghai Jiao Tong University)
		mmfp2505	Virtual	Improving Scalability, Sustainability and Availability via Workload Distribution in Edge-Cloud Gaming -- Iryanto Jaya (Nanyang Technological University), Yusen Li (Nankai University), Wentong Cai (Nanyang Technological University)
Demo Session 1 [Main Foyer]	**10h30-11h45**			*(Demos might be available the whole day)*
Audio/speech		mmde003	Virtual	SingMaster: A Sight-singing Evaluation System of "Shoot and Sing" Based on Smartphone -- Wei Xu (School of Electronic Information and Communications, Huazhong University of Science and Technology), Bowen Tian (School of Electronic Information and Communications, Huazhong University of Science and Technology), Lijie Luo (School of Electronic Information and Communications, Huazhong University of Science and Technology), Weiming Yang (Huazhong University of Science and Technology), Xianke Wang (School of Electronic Information and Communications, Huazhong University of Science and Technology), Lei Wu (School of Mechanical Science and Engineering, Huazhong University of Science and Technology)
		mmde006	Virtual	Seeing Speech: Magnetic Resonance Imaging-Based Vocal Tract Deformation Visualization Using Cross-Modal Transformer -- Kele XU (Computer Science Lab - Pierre and Marie Curie University, Paris, France), Ming Feng (Tongji University), weiquan Huang (Tongji University)
		mmde011	Virtual	A Platform for Deploying the TFE Ecosystem of Automatic Speech Recognition -- Yuanfeng SONG (AI Group, WeBank), Rongzhong Lian (WEBANK), Yixin Chen (WEBANK), Di Jiang (Hong Kong University of Science and Technology), Xuefang Zhao (WeBank Co., Ltd.), Conghui Tan (Tencent), Qian XU (WEBANK), Raymond Wong (HKUST)
		mmde020	On-site	Sync Sofa: Sofa-type Side-by-side Communication Experience Based on Multimodal Expression -- Yuki Tajima (KDDI Research, inc.), Shota Okubo (KDDI Research, inc.), Tomoaki Konno (KDDI Research, Inc.), Toshiharu Horiuchi (KDDI Research, Inc.), Tatsuya Kobayashi (KDDI Research, Inc.)
		mmde032	Virtual	Interpretable Melody Generation from Lyrics with Discrete-Valued Adversarial Training -- Wei Duan (NII, Tokyo Institute of Technology), Zhe Zhang (NII, SOKENDAI), Yi Yu (NII, National Institute of Informatics), Keizo Oyama (NII, National Institute of Informatics)
		mmde044	On-site	TWIZ: The Multimodal Conversational Task Wizard -- Rafael Ferreira (Universidade NOVA de Lisboa), Diogo Silva (Universidade NOVA de Lisboa), Diogo Tavares (Universidade NOVA de Lisboa), Frederico Vicente (Universidade Nova de Lisboa), Mariana Bonito (Universidade Nova de Lisboa), Gustavo Gonçalves (Carnegie Mellon University), Rui Margarido (Universidade NOVA de Lisboa), Paula Figueiredo (Universidade NOVA de Lisboa), Helder Rodrigues (Universidade NOVA de Lisboa), David Semedo (Universidade NOVA de Lisboa), Joao Magalhaes (Universidade Nova de Lisboa)
Art Exhibition [Pav 4]	**All day**			*(Exhibition available also Wednesday and Thursday)*
		mmia02	On-site	All is noise: In Search of Enlightenment, a VR experience -- Manuel Silva, Luana Santos, Luís Teixeira, José Carvalho
		mmia05	On-site	Beauty: Machine Microbial Interface as Artistic Experimentation -- Johnny DiBlasi, Carlos Castellanos, Bello Bello

[Pavillion 4 - Hall + R1.02, R1.03 & R1. 04]		mmia08	On-site	Being's Spread: Mirror of Life Interconnection -- Xinrui Wang, Yulu Song, Xiaohui Wang
		mmia11		CAPTCHA the Flag -- Tiago Rorke
		mmia14	On-site	Cellular Trending: Fragmented Information Dissemination on Social Media Through Generative Lens -- Bo Shui, Xiaohui Wang
		mmia17	On-site	Collaboration superpowers: the process of crafting an interactive storytelling animation -- Sofia Dias, Sara Silva, Beatriz Silva, Rui Nóbrega
		mmia20	On-site	Dream Painter: An Interactive Art Installation Bridging Audience Interaction, Robotics, and Creative AI -- Varvara Guljajeva, Mar Canet Sola
		mmia23	On-site	Emotional Machines -- Jorge Forero, Mónica Mendes, Gilberto Bernardes
		mmia26	On-site	Fragrance In Sight :Personalized Perfume Production Based on Style Recognition -- Jiaxiang You, Yinyu Chen, Xiaohui Wang
		mmia29	On-site	Meditation in Motion: Interactive Media Art Visualization Based on Ancient Tai Chi Chuan -- Ze Gao, Anqi Wang, Pan Hui, Tristan Braud
		mmia32	On-site	Read Your Voice -- Hugo Pauget Ballesteros, Gilles Azzaro, Jean Mélou, Yvain Quéau, Jean-Denis Durou
		mmia35	On-site	StimulusLoop: Game-Actuated Mutuality Artwork for Evoking Affective State -- Tai-Chen Tsai, Tse-Yu Pan, Min-Chun Hu, Ya-Lun Tao
		mmia38	On-site	Viva Contemporary! Mobile Music Laboratory -- Emily Graber, Charles Picasso, Elaine Chew
		mmia41		Wander: an AI-driven Chatbot to Visit the Future Earth -- Yuqian Sun, Chenhang Cheng, Ying Xu, Yihua Li, Chang Hee Lee, Ali Asadipour
Doctoral Symposium [Pav. 4 R1.07]	**10h50-11hh45**			
PhD School Talk 1	10h50		On-site	Title: Multimodal Computing for a Healthy Society -- Ramesh Jain (University California Irvine)
Best Paper Session [Auditório I]	**11h45-13h15**	**Chair: Cees Snoek (University of Amsterdam)**		
	11h45	mmfp1786	Remote	Search-oriented Micro-video Captioning -- Liqiang Nie (Harbin Institute of Technology (Shenzhen)), Leigang Qu (Shandong University), Dai Meng (Kuaishou), Min zhang (Harbin Institute of Technology (Shenzhen)), Qi Tian (Huawei Technologies Ltd.), Alberto del Bimbo (Universita di Firenze)
	12h03	mmfp3008	On-site	ConceptBeam: Concept driven target speech extraction -- Yasunori Ohishi (NTT Corporation), Marc Delcroix (NTT), Tsubasa Ochiai (NTT), Shoko Araki (The University of Tokyo, Tokyo University), Daiki Takeuchi (NTT), Daisuke Niizumi (NTT Communication Science Laboratories), Akisato Kimura (NTT Corporation), Noboru Harada (NTT Communication Science Laboratories), Kunio Kashino (NTT)
	12h21	mmfp0545	Virtual	DSE-GAN: Dynamic Semantic Evolution Generative Adversarial Network for Text-to-Image Generation -- Mengqi Huang (University of Science and Technology of China), Zhendong Mao (University of Science and Technology of China), Penghui Wang (University of Science and Technology of China), Quan Wang (Beijing University of Posts and Telecommunications), Yongdong Zhang (University of Science and Technology of China)
	12h39	mmfp1806	Virtual	Learnability Enhancement for Low-light Raw Denoising: Where Paired Real Data Meets Noise Modeling -- Hansen Feng (Beijing Institute of Technology), Lizhi Wang (Beijing Institute of Technology), Yuzhi Wang (Megvii Technology Inc.), Hua Huang (Beijing Normal University)
	12h57	mmfp1662	Virtual	Maze: A Cost-Efficient Video Deduplication System at Web-scale -- An Qin (Baidu), Mengbai Xiao (Shandong University), Huang Ben (Baidu), Xiaodong Zhang (Ohio State University)
Virtual Only Poster Session 1 [Gathertown]	**13h-14h30**			*(See Poster Session 1)*
Keynote Talk: Douwe Kiela [Auditório I]	**14h30-15h30**			Grounding, Meaning and Foundation Models: Adventures in Multimodal Machine Learning – Douwe Kiela
Oral Session 2a [Auditório I]	**15h30-16h45**	**Chair: Yan Yan (Illinois Institute of Technology)**		
Multimodal Modeling: Image, Text and Voice	15h30	mmfp0801	Virtual	GSRFormer: Grounded Situation Recognition Transformer with Alternate Semantic Attention Refinement -- Zhi-Qi Cheng (Carnegie Mellon University), Qi Dai (Microsoft Research Asia), Siyao Li (Carnegie Mellon University), Teruko Mitamura (Carnegie Mellon University), Alexander Hauptmann (Carnegie-Mellon University)
	15h48	mmfp1694	Virtual	Ordered Attention for Coherent Visual Storytelling -- Tom Braude (IDC), Idan Schwartz (Technion), Alex Schwing (University of Illinois, Urbana Champaign), Ariel Shamir (IDC)
	16h06	mmfp3228	On-site	Understanding News Text and Images Connection with Context-enriched Multimodal Transformers -- Claudio Bartolomeu (Universidade Nova de Lisboa), Rui Nóbrega (NOVA School of Science and Technology), David Semedo (Universidade NOVA de Lisboa)
	16h24	mmfp0228	Virtual	You Can even Annotate Text with Voice: Transcription-only-Supervised Text Spotting -- Jingqun Tang (Netease), Su Qiao (Zhejiang Gongshang University), Benlei Cui (Zhejiang Gongshang University), Yuhang Ma (University College London, University of London), Sheng Zhang (Zhejiang University), Dimitrios Kanoulas (University College London)
Oral Session 2b [Auditório II]	**15h30-16h45**	**Chair: Miriam Redi (King's College)**		
Top Papers Session I	15h30	mmfp2265	On-site	Interactive Video Corpus Moment Retrieval using Reinforcement Learning -- Zhixin Ma (Singapore Management University), Chong Wah Ngo (Singapore Management University)
	15h48	mmfp2042	Virtual	A Multi-view Spectral-Spatial-Temporal Masked Autoencoder for Decoding Emotions with Self-supervised Learning -- Rui Li (Shanghai Jiaotong University), Yiting Wang (Cornell University), Wei-Long Zheng (Shanghai Jiao Tong University), Bao-liang Lu (Shanghai Jiaotong University)
	16h06	mmfp0016	Virtual	TVFormer: Trajectory-guided visual quality assessment on 360° images with Transformers -- Li Yang (Beihang University), Mai Xu (Beihang University, Tsinghua University), Tie Liu (Beihang University), Liangyu Huo (Beijing University of Aeronautics and Astronautics), Xinbo Gao (Chongqing University of Post and Telecommunications)
	16h24	mmfp2586	Virtual	Improving Generalization for Neural Adaptive Video Streaming via Meta Reinforcement Learning -- Nuowen Kan (Shanghai Jiaotong University), Yuankun Jiang (Shanghai Jiao Tong University), Chenglin Li (Shanghai Jiao Tong University), Wenrui Dai (Shanghai Jiao Tong University), Junni Zou (Shanghai Jiao Tong University), Hongkai Xiong (Shanghai Jiao Tong University)
Oral Session 2c [Auditório III]	**15h30-16h45**	**Chair: Chengcui Zhang (Autonomous University of Barcelona)**		
Multimodal Fusion II	15h30	mmfp0368	Virtual	Inferring Speaking Styles from Multi-modal Conversational Context by Multi-scale Relational Graph Convolutional Networks -- Jingbei Li (Tsinghua University, Tsinghua University), Yi Meng (Tsinghua University, Tsinghua University), Xixin Wu (The Chinese University of Hong Kong), Zhiyong Wu (Tsinghua University, Tsinghua University), Jia Jia (Department of Computer Science and Technology, Tsinghua University), Helen Meng (The Chinese University of Hong Kong), Qiao Tian (ByteDance), Wang Yuping (University of Science and Technology of China), Yuxuan Wang (ByteDance)
	15h48	mmfp0166	On-site	Modality Eigen-Encodings Are Keys to Open Modality Informative Containers -- Yiyuan Zhang (Beijing Institute of Technology), Yuqi Ji (Beijing Institute of Technology)
	16h06	mmfp0797	Virtual	SPTS: Single-Point Text Spotting -- Dezhi Peng (South China University of Technology), Xinyu Wang (The University of Adelaide), Yuliang Liu (Huazhong University of Science and Technology), Jiaxin Zhang (South China University of Technology, South China University of Technology), Mingxin Huang (South China University of Technology), Songxuan Lai (South China University of Technology), Jing Li (National University of Singapore), Shenggao Zhu (Huawei Cloud), Dahua Lin (Toyota Technnological Institute), Chunhua Shen (Zhejiang University), Xiang Bai (Huazhong University of Science and Technology), Lianwen Jin (South China University of Technology)
	16h24	mmrp2022	Virtual	Reproducibility Companion Paper: Focusing on Persons: Colorizing Old Images Learning from Modern Historical Movies -- Xin Jin (Department of Cyber Security, Beijing Electronic Science and Technology Institute, Fengtai District), Ke Liu (Department of Cyber Security, Beijing Electronic Science and Technology Institute, Fengtai District), Zhonglan Li (Department of Cyber Security, Beijing Electronic Science and Technology Institute, Fengtai District), Heng Huang (Department of Cyber Security, Beijing Electronic Science and Technology Institute, Fengtai District), Dongqing Zou (SenseTime Research and Tetras.AI Beijing Beijing 100080, Beijing, China, Qing Yuan Research Institute, Shanghai Jiao Tong University), Vajira Thambawita (SimulaMet)
Oral Session 2d [Auditório IV]	**15h30-16h45**	**Chair: Chenglin Li (Shanghai Jiao Tong University)**		
Re-Identification	15h30	mmfp1955	Virtual	Cross-Compatible Embedding and Semantic Consistent Feature Construction for Sketch Re-identification -- Yafei Zhang (Kunmimg University of Science and Technology), Yongzeng Wang (Kunmimg University of Science and Technology), Huafeng Li (Kunmimg University of Science and Technology), Li Shuang (Kunmimg University of Science and Technology)
	15h48	mmfp3053	Virtual	Hierarchical Walking Transformer for Object Re-Identification -- Florence Philex (East China Normal University), Jun Liu (Tencent YouTu Lab), zhizhong zhang (East China Normal University), Chengjie Wang (Tencent YouTu Lab), Yanyun Qu (Xiamen University), Yuan Xie (East China Normal University), Lizhuang Ma (Dept. of Computer Sci. & Eng., Shanghai Jiao Tong University)
	16h06	mmfp2629	Virtual	Progressive Attribute Embedding for Accurate Cross-modality Person Re-ID -- AIHUA ZHENG (Anhui University), Peng Pan (Anhui University), Hongchao Li (Anhui University), Chenglong Li (Anhui University), Bin Luo (Anhui University), Chang Tan (University of Science and Technology of China, Tsinghua University), Ruoran Jia (IFLYTEK CO)
Doctoral Symposium [Pav. 4 R1.07]	**15h30-17h30**			
Multimodal Analysis	15h30	mmdc014	On-site	Zero-shot Generalization of Multimodal Dialogue Agents -- Diogo Tavares (Universidade NOVA de Lisboa)
	15h45	mmdc011	On-site	Multi-modal Learning Algorithms and Network Architectures for Information Extraction and Retrieval -- Maurits Bleeker (University of Amsterdam)
	16h	mmdc015	On-site	The First Impression: Understanding the impact of Multimodal System Responses on User Behavior in Task-oriented Agents -- Diogo Silva (Universidade NOVA de Lisboa)
PhD School Talk 2	16h35		Virtual	The Importance of Iteration on the Path from Idea to Research Paper -- Martha Larson (Radboud University)
Panel 1 [Auditório I]	**16h45-18h**			Groundbreaking Multimedia Research Directions (Moderator: Alan Smeaton, Panelists: Ramesh Jain, Alex Hauptmann, Fernando Pereira, Xirong Li)
Poster Session 1bis [Pav. 4 - Hall]	**16h45-18h**			*(See Poster Session 1)*
Demo Session 1bis [Main Foyer]	**16h45-18h**			*(See Demo Session 1)*

Wednesday, October 12th (Main Conference)

Oral Session 3a [Auditório I]	**9h15-10h30**	**Chair: Jianbo Jiao (University of Birmingham)**		
Self-Supervised Learning for Multimedia	9h15	mmfp0817	Virtual	Simple Self-supervised Multiplex Graph Representation Learning -- Mo Yujie (University of Electronic Science and Technology of China), Yuhuan Chen (University of Electronic Science and Technology of China), Larry Peng (University of Electronic Science and Technology of China), Xiaoshuang Shi (University of Electronic Science and Technology of China, Tsinghua University), Xiaofeng Zhu (University of Electronic Science and Technology of China)
	9h33	mmfp0990	Virtual	Self-supervised Exclusive Learning for 3D Segmentation with Cross-Modal Unsupervised Domain Adaptation -- Yachao Zhang (Xiamen University,), MiaoYu Li (Xiamen University), Yuan Xie (East China Normal University), Cuihua Li (XMU), Cong Wang (Huawei Technologies Ltd.), zhizhong zhang (East China Normal University), Yanyun Qu (Xiamen University)
	9h51	mmfp0159	Virtual	Unsupervised and Pseudo-Supervised Vision-Language Alignment in Visual Dialog -- Feilong Chen (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Duzhen Zhang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Xiuyi Chen (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Jing Shi (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Shuang Xu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Bo XU (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
	10h09	mmfp0205	Virtual	Reading and Writing: Discriminative and Generative Modeling for Self-Supervised Text Recognition -- Mingkun Yang (Huazhong University of Science and Technology), Minghui Liao (Huawei Cloud), lu pu (Huawei Technologies Ltd.), Jing Wang (Huawei Technologies Ltd.), Shenggao Zhu (Huawei Cloud), Hualin Luo (Huawei Technologies Ltd.), Qi Tian (Huawei Technologies Ltd.), Xiang Bai (Huazhong University of Science and Technology)

Oral Session 3b [Auditório II]	**9h15-10h30**	**Chair: Lorenzo Baraldi (University of Modena and Regio Emilia)**		
Captioning	9h15	mmfp1280	Virtual	Inferential Visual Question Generation -- Chao Bi (University of the Chinese Academy of Sciences), Shuhui Wang (Institute of Computing Technology, Chinese Academy of Sciences), Zhe Xue (Beijing University of Posts and Telecommunications), Shengbo Chen (Henan University), Qingming Huang (University of Chinese Academy of Sciences)
	9h33	mmfp2665	Virtual	A Baseline for Detecting Out-of-Distribution Examples in Image Captioning -- Gal Shalev (Bar Ilan University), Gabi Shalev (Bar Ilan University), Joseph Keshet (Bar-Ilan University)
	9h51	mmfp0548	Virtual	Fine-tuning with Multi-modal Entity Prompts for News Image Captioning -- Jingjing Zhang (University of Science and Technology of China), Shancheng Fang (University of Science and Technology of China), Zhendong Mao (University of Science and Technology of China), Zhiwei Zhang (Search), Yongdong Zhang (University of Science and Technology of China)
	10h09	mmfp2785	Virtual	Rethinking the Reference-based Distinctive Image Captioning -- Yangjun Mao (Zhejiang University), Long Chen (Columbia University), Zhihong Jiang (Zhejiang University), Zhang Dong (The Hong Kong University of Science and Technology), Zhimeng Zhang (Zhejiang University), Jian Shao (Zhejiang University), Jun Xiao (Zhejiang University)
Oral Session 3c [Auditório III]	**9h15-10h30**	**Chair: Qin Jin (Renmin University of China)**		
Document Analysis	9h15	mmfp2513	Virtual	CALM: Commen-Sense Knowledge Augmentation for Document Image Understanding -- Qinyi Du (Shanghai Jiaotong University), Qingqing Wang (Shanghai Jiaotong University), Keqian Li (University of California, Santa Barbara), Jidong Tian (Shanghai Jiao Tong University), Liqiang Xiao (Shanghai Jiao Tong University), Yaohui Jin (Shanghai Jiaotong University)
	9h33	mmfp1542	Virtual	LayoutLMv3: Pre-training for Document AI with Unified Text and Image Masking -- Yupan Huang (SUN YAT-SEN UNIVERSITY), Tengchao Lv (Microsoft), Lei Cui (Microsoft Research Asia), Yutong Lu (SUN YAT-SEN UNIVERSITY), Furu Wei (Microsoft Research)
	9h51	mmfp0553	Virtual	End-to-End Compound Table Understanding with Multi-Modal Modeling -- Zaisheng Li (University of the Chinese Academy of Sciences), Yi Li (ShanghaiTech University), Liang Qiao (Zhejiang University), Pengfei Li (Xi'an Jiaotong University), Zhanzhan Cheng (Zhejiang University), Yi Niu (Shanghai Jiao Tong University), Shiliang Pu (Zhejiang University), Xi Li (Zhejiang University)
	10h09	mmfp0534	Virtual	Query-driven Generative Network for Document Information Extraction in the Wild -- Haoyu Cao (Tencent Youtu Lab), Xin Li (Tencent Youtu Lab), Jiefeng Ma (University of Science and Technology of China), Deqiang Jiang (Tencent YouTu Lab), Antony Guo (Tencent Youtu Lab), Yiqing Hu (Tencent Youtu Lab), Hao Liu (Tencent Youtu Lab), YINSONG LIU (Tencent Youtu Lab), Bo Ren (Tencent Youtu Lab)
Oral Session 3d [Auditório IV]	**9h15-10h30**	**Chair: Irene Viola (Centrum Wiskunde & Informatica)**		
Art & Multimedia	9h15	mmfp1467	Virtual	PC-Dance: Posture-controllable Music-driven Dance Synthesis -- Jibin Gao (SUN YAT-SEN UNIVERSITY), Junfu Pu (Tencent ARC Lab), Honglun Zhang (Tencent PCG ARC Lab), Ying Shan (Tencent), Wei-Shi Zheng (SUN YAT-SEN UNIVERSITY)
	9h33	mmfp2758	On-site	ReLyMe: Improving Lyric-to-Melody Generation by Incorporating Lyric-Melody Relationships -- Chen Zhang (Zhejiang University), Luchin Chang (Zhejiang University), Songruoyao Wu (Zhejiang University), Xu Tan (Microsoft), Tao Qin (Microsoft Research Asia), Tie-Yan Liu (Microsoft), Kejun Zhang (Zhejiang University)
	9h51	mmfp2843	On-site	SongDriver: Real-time Music Accompaniment Generation without Logical Latency nor Exposure Bias -- Zihao Wang (Zhejiang University), Kejun Zhang (Zhejiang University), Yuxing Wang (Zhejiang University), Chen Zhang (Zhejiang University), Qihao LIANG (Zhejiang University), Pengfei Yu (Jingchu University of Technology), Feng Yongsheng (Shandong University), Wenbo Liu (College of Computer Science and Technology, Zhejiang University), Yikai Wang (Zhejiang University), BAO YUNTAO (Zhejiang University), Yiheng Yang (Zhejiang University)
	10h09	mmfp1676	On-site	EuglPollock: Rethinking Interspecies Collaboration through Art Making -- Kyungwon Lee (Yonsei University), Yu-Kyung Jang (Yonsei University), Jaewoo Jung (Yonsei University), DongHwan Kim (Yonsei University), Hyun Jean Lee (Yonsei University), Seung Ah Lee (Yonsei University)
Grand Challenges Session I [Pav. 4 R1.07]	**9h15-10h30**			
Overview	9h15	mmgc67	On-site	Overview of the ACM Multimedia Grand Challenges 2022 -- Miriam Redi (Wikimedia Foundation, Bell Labs Cambridge), Georges Quénot (LIG-CNRS), Maria Eskevich (CLARIN ERIC)
Deep Video Understanding Challenge	9h21	mmgc18	On-site	The ACM Multimedia 2022 Deep Video Understanding Grand Challenge -- Keith Curtis (Dublin City University), George Awad (National Institute of Standards and Technology), Shahzad Rajput (Georgetown University), Ian Soboroff (National Institute of Standards and Technology)
	9h27	mmgc12	Virtual	Deep Video Understanding with a Unified Multi-Modal Retrieval Framework -- Chen-Wei Xie (Alibaba Group), Sun Siyang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Liming Zhao (Alibaba Group), Wu Min (Northwest Polytechnical University Xi'an), Dangwei Li (Alibaba Group), Yun Zheng (SUN YAT-SEN UNIVERSITY)
	9h33	mmgc52	On-site	Multimodal Analysis for Deep Video Understanding with Video Language Transformer -- Beibei Zhang (Nanjing University), Fang Yaqun (Nanjing University), Tongwei Ren (Nanjing University), Gangshan Wu (Nanjing University)
Computational Paralinguistics ChallengE	9h39	mmgc32	On-site	The ACM Multimedia 2022 Computational Paralinguistics Challenge: Vocalisations, Stuttering, Activity, \& Mosquitoes -- Björn Schuller (Imperial College London)
	9h45	mmgc23	Virtual	VecNet: A Deep Learning Based Approach to Detect Mosquitoes -- Devesh Khandelwal (University of California, Berkeley), Sean Campos (University of California, Berkeley), Shwetha Nagaraj (University of California, Berkeley), Fred Nugen (University of California, Berkeley), Alberto Todeschini (University of California, Berkeley)
	9h51	mmgc06	On-site	Wav2vec2-based Paralinguistic Systems to Recognise Vocalised Emotions and Stuttering -- Tamás Grósz (Aalto University), Dejan Porjazovski (Aalto University), Yaroslav Getman (Aalto University), Sudarsana Kadiri (Aalto University), Mikko Kurimo (Aalto University)
Facial Micro-Expression Grand Challenge	9h57	mmgc54	Virtual	MEGC2022: ACM Multimedia 2022 Micro-Expression Grand Challenge -- Jingting Li (Institute of Psychology, Chinese Academy of Sciences), Moi Hoon Yap (Manchester Metropolitan University), Wen-Huang Cheng (National Yang Ming Chiao Tung University), John See (Heriot-Watt University), Xiaopeng Hong (Harbin Institute of Technology), Xiaobai Li (University of Oulu), Su-Jing Wang (Institute of Psychology, Chinese Academy of Sciences), Adrian Davison (University of Manchester), Yante Li (University of Oulu), Zizhao Dong (Institute of Psychology, Chinese Academy of Sciences)
	10h03	mmgc62	Virtual	Facial Expression Spotting Based on Optical Flow Features -- Jun Yu (University of Science and Technology of China), Zhongpeng Cai (University of Science and Technology of China), Zepeng Liu (University of Science and Technology of China), Guochen Xie (University of Science and Technology of China), Peng He (University of Science and Technology of China)
	10h09	mmgc63	Virtual	Micro Expression Generation with Thin-plate Spline Motion Model and Face Parsing -- Jun Yu (University of Science and Technology of China), Guochen Xie (University of Science and Technology of China), Zhongpeng Cai (University of Science and Technology of China), Peng He (University of Science and Technology of China), Fang Gao (GuangXi University), Qiang Ling (University of Science and Technology of China)
Social Media Prediction (SMP) Challenge	10h15	mmgc02	Virtual	Title-and-Tag Contrastive Vision-and-Language Transformer for Social Media Popularity Prediction -- weilong chen (University of Electronic Science and Technology of China), Chenghao Huang (University of Electronic Science and Technology of China), Yuan min (University of Electronic Science and Technology of China), Xiaolu Chen (University of Electronic Science and Technology of China), Wenhao Hu (University of Electronic Science and Technology of China), Xinran Zhang (University of Electronic Science and Technology of China), Yanru Zhang (University of Electronic Science and Technology of China)
	10h21	mmgc61	Virtual	An Efficient Multi-View Multimodal Data Processing Framework for Social Media Popularity Prediction -- YunPeng Tan (Beijing University of Posts and Telecommunications), Liu Fangyu (Beijing University of Posts and Telecommunications), Bowei Li (Beijing University of Posts and Telecommunications), Zheng Zhang (Beijing University of Posts and Telecommunications), Bo Zhang (Beijing University of Posts and Telecommunications)
Poster Session 2 [Pav. 4 - Hall]	**10h30-11h45**			*(Posters will be hung the whole day)*
Understanding Multimedia Content: Multimodal Fusion and Embeddings		mmfp2671	Virtual	Two-Stream Transformer for Multi-Label Image Classification -- Xuelin Zhu (Southeast University), Jiuxin Cao (Southeast University), Jiawei Ge (Southeast University), Weijia Liu (Southeast University), Bo Liu (Southeast University)
		mmfp3243	On-site	SoftSkip: Empowering Multi-Modal Dynamic Pruning for Single-Stage Referring Comprehension -- Dulanga Weerakoon (Singapore Management University), Vigneshwaran Subbaraju (Institute of High Performance Computing, Singapore, A*STAR), Tuan Tran (Singapore Management University), Archan Misra (Singapore Management University)
		mmfp0448	Virtual	Unbiased Directed Object Attention Graph for Object Navigation -- Ronghao Dang (Tongji University), Zhuofan Shi (ETHZ - ETH Zurich), Liuyi Wang (Tongji University), Zongtao He (Tongji University), Chengju Liu (Tongji University), Qijun Chen (Tongji University)
		mmfp2256	Virtual	FastPR: One-stage Semantic Person Retrieval via Self-supervised Learning -- Meng Sun (Tsinghua University, Tsinghua University), Ju Ren (Tsinghua University, Tsinghua University), Xin Wang (Tsinghua University), Wenwu Zhu (Tsinghua University, Tsinghua University), Yaoxue Zhang (Tsinghua University, Tsinghua University)
		mmfp0766	Virtual	Towards Counterfactual Image Manipulation via CLIP -- Yingchen Yu (Nanyang Technological University), Fangneng Zhan (Max Planck Institute for Informatics), Rongliang WU (Nanyang Technological University), Jiahui Zhang (Nanyang Technological University), Shijian Lu (Nanyang Technological University), miaomiao cui (Alibaba Group), Xuansong Xie (Electronic Engineering), Xian-Sheng Hua (Alibaba), Chunyan Miao (Nanyang Technological University)
		mmfp1209	Virtual	Bidirectionally Learning Dense Spatio-temporal Feature Propagation Network for Unsupervised Video Object Segmentation -- Jiaqing Fan (Nanjing University of Aeronautics and Astronautics), Tiankang Su TianKang Su (Nanjing University of Information Science and Technology), Kaihua Zhang (NUIST), Qingshan Liu (Nanjing University of Information Science and Technology)
		mmfp0688	Virtual	Weakly Supervised Video Salient Object Detection via Point Supervision -- Shuyong Gao (Fudan University), Hao Xing (FuDan University), Yan Wang (Fudan University), Qianyu Guo (Fudan University), Wei Zhang (Fudan University), Wenqiang Zhang (Fudan University)
		mmfp0476	Virtual	Look Less Think More: Rethinking Compositional Action Recognition -- Rui Yan (Nanjing University of Science and Technology), Peng Huang (Nanjing University of Science and Technology), Xiangbo Shu (Nanjing University of Science and Technology), David Junhao Zhang (National University of Singapore), Yonghua Pan (Nanjing University of Science and Technology), Jinhui Tang (Nanjing University of Science and Technology)
		mmfp0483	Virtual	Continual Multi-view Clustering -- Xinhang Wan (National University of Defense Technology), Jiyuan Liu (National University of Defense Technology), Weixuan Liang (National University of Defense Technology), Xinwang Liu (National University of Defense Technology), Yi Wen (National University of Defense Technology), En Zhu (National University of Defense Technology)
		mmfp1579	Virtual	Efficient Anchor Learning-based Multi-view Clustering -- A Late Fusion Method -- Tiejian Zhang (National University of Defense Technology), Xinwang Liu (National University of Defense Technology), En Zhu (National University of Defense Technology), sihang zhou (National University of Defense Technology), dong bin (National University of Defense Technology)
		mmfp2248	Virtual	Cross-modal Knowledge Graph Contrastive Learning for Machine Learning Method Recommendation -- Xianshuai Cao (Shandong University), Yuliang Shi (Shandong University), Jihu Wang (Shandong University), Han Yu (Nanyang Technological University), Xinjun Wang (Shandong University), Zhongmin Yan (Shandong University)
		mmfp0549	Virtual	Multigranular Visual-Semantic Embedding for Cloth-Changing Person Re-identification -- Zan Gao (Qilu University of Technology), Hongwei Wei (Qilu University of Technology), WEILI GUAN (Monash University), Weizhi Nie (Tianjin University), Meng Liu (Shandong Jianzhu University), Meng Wang (Hefei University of Technology)
		mmfp3242	Virtual	Adaptive Structural Similarity Preserving for Unsupervised Cross Modal Hashing -- Liang Li (Fudan University), Baihua Zheng (Singapore Management University), Weiwei Sun (Fudan University)
		mmfp1163	Virtual	CubeMLP: An MLP-based Model for Multimodal Sentiment Analysis and Depression Estimation -- Hao Sun (Zhejiang University), Hongyi Wang (Zhejiang University), Jiaqing LIU (Ritsumeikan University), Yen-wei Chen (Ritsumeikan University), Lanfen Lin (Zhejiang University)
		mmfp1668	On-site	Generalized Global Ranking-Aware Neural Architecture Ranker for Efficient Image Classifier Search -- Bicheng Guo (Zhejiang University), Tao Chen (Fudan University), Shibo He (Zhejiang University), Haoyu Liu (NetEase), Lilin Xu (Zhejiang University), Peng Ye (Fudan University), Jiming Chen (Zhejiang University)
		mmfp2486	Virtual	Exploiting Transformation Invariance and Equivariance for Self-supervised Sound Localisation -- Jinxiang Liu (Shanghai Jiaotong University), Chen Ju (Shanghai Jiao Tong University), Weidi Xie (University of Oxford), Ya Zhang (Shanghai Jiao Tong University)
		mmfp0387	On-site	Unsupervised Video Hashing with Multi-granularity Contextualization and Multi-structure Preservation -- Yanbin Hao (University of Science and Technology of China), Jingru Duan (University of Science and Technology of China), Hao Zhang (City University of Hong Kong), Bin Zhu (City University of Hong Kong), Pengyuan Zhou (University of Science and Technology of China), Xiangnan He (University of Science and Technology of China)
		mmfp3051	Virtual	DisCo: Disentangled Implicit Content and Rhythm Learning for Diverse Co-Speech Gestures Synthesis -- Haiyang Liu (the university of tokyo), Naoya Iwamoto (Huawei Technologies Japan K.K.), Zihao Zhu (Keio University), Zhengqing Li (Huawei Technologies Ltd.), YOU ZHOU (Huawei Technologies Ltd.), Elif Bozkurt (Huawei Turkey R&D Center), Bo Zheng (Huawei Technologies Ltd.)
		mmfp0783	Virtual	Adaptively-weighted Integral Space for Fast Multiview Clustering -- Mansheng Chen (SUN YAT-SEN UNIVERSITY), Tuo Liu (SUN YAT-SEN UNIVERSITY), Chang-dong Wang (SUN YAT-SEN UNIVERSITY), Dong Huang (South China Agricultural University), Jianhuang Lai (SUN YAT-SEN UNIVERSITY)
		mmfp0900	Virtual	Towards All Weather and Unobstructed Multi-Spectral Image Stitching: Algorithm and Benchmark -- Zhiying Jiang (Dalian University of Technology), ZengXi Zhang (Dalian University of Technology), Xin Fan (Dalian University of Technology), Risheng Liu (Dalian University of Technology)
		mmfp0984	Virtual	A Parameter-free Multi-view Information Bottleneck Clustering Method by Cross-view Weighting -- Shizhe Hu (Zhengzhou University), Ruilin Geng (Zhengzhou University), Zhaoxu Cheng (Zhengzhou University), Chaoyang Zhang (Zhengzhou University), Guoliang Zou (Zhengzhou University), Zhengzheng Lou (Zhengzhou University), Yangdong Ye (Zhengzhou University)

		mmfp2617	Virtual	HERO: HiErarchical spatio-tempoRal reasOning with Contrastive Action Correspondence for End-to-End Video Object Grounding -- Mengze Li (Zhejiang University, Tsinghua University), Tianbao Wang (Zhejiang University), Haoyu Zhang (Zhejiang University), Shengyu Zhang (Zhejiang University), Zhou Zhao (Zhejiang University), Wenqiao Zhang (National University of Singapore), Jiaxu Miao (Zhejiang University), Shiliang Pu (Zhejiang University), Fei Wu (Zhejiang University)
		mmfp2926	Virtual	MAVT-FG: Multimodal Audio-Visual Transformer for Weakly-supervised Fine-Grained Recognition -- Xiaoyu Zhou (University of Electronic Science and Technology of China), Xiaotong Song (University of Electronic Science and Technology of China), Hao Wu (University of Electronic Science and Technology of China), Jingran Zhang (University of Electronic Science and Technology of China), Xing Xu (University of Electronic Science and Technology of China)
		mmfp1362	Virtual	Dynamic Graph Modeling for Weakly-Supervised Temporal Action Localization -- Haichao Shi (Institute of Information Engineering, Chinese Academy of Sciences), Xiao-Yu Zhang (Institute of Information Engineering, Chinese Academy of Sciences), Changsheng Li (Beijing Institute of Technology), Lixing Gong (JingDong), Yong Li (, Institute of automation, Chinese academy of science), Yongjun Bao (Peking University)
Understanding Multimedia Content: Media Interpretation		mmfp2037	Virtual	SGINet: Toward Sufficient Interaction Between Single Image Deraining and Semantic Segmentation -- Yanyan Wei (Hefei University of Technology), Zhao Zhang (Hefei University of Technology), Huan Zheng (Hefei University of Technology), Richang Hong (Hefei University of Technology), Yi Yang (Zhejiang University), Meng Wang (Hefei University of Technology)
		mmfp1546	Virtual	FCL-GAN: A Lightweight and Real-Time Baseline for Unsupervised Blind Image Deblurring -- Suiyi Zhao (Hefei University of Technology), Zhao Zhang (Hefei University of Technology), Richang Hong (Hefei University of Technology), Mingliang Xu (Zhengzhou University), Yi Yang (Zhejiang University), Meng Wang (Hefei University of Technology)
		mmfp0777	On-site	Task-adaptive Spatial-Temporal Video Sampler for Few-shot Action Recognition -- Huabin Liu (Shanghai Jiao Tong University), Weixian L (Shanghai Jiaotong University), John See (Heriot-Watt University), Weiyao Lin (Shanghai Jiao Tong University)
		mmfp0496	Virtual	MM-Pyramid: Multimodal Pyramid Attentional Network for Audio-Visual Event Localization and Video Parsing -- Jiashuo Yu (Fudan University), Ying Cheng (Fudan University), Rui-Wei Zhao (Fudan University), Rui Feng (Fudan University), Yuejie Zhang (Fudan University)
		mmfp1378	On-site	Lip-to-Speech Synthesis for Arbitrary Speakers in the Wild -- Sindhu Hegde (International Institute of Information Technology Hyderabad), Prajwal K R (University of Oxford), Rudrabha Mukhopadhyay (International Institute of Information Technology Hyderabad), Vinay Namboodiri (University of Bath), C.V. Jawahar (IIIT Hyderabad)
		mmfp2378	Virtual	Attribute-guided Dynamic Routing Graph Network for Transductive Few-shot Learning -- Chaofan Chen (University of Science and Technology of China), Xiaoshan Yang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Ming Yan (Alibaba Group), Changsheng Xu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
		mmfp0397	Virtual	OS-MSL: One Stage Multimodal Sequential Link Framework for Scene Segmentation and Classification -- Ye Liu (Shanghai Jiaotong University), Lingfeng Qiao (Tencent), di yin (Tencent Youtu), Zhuoxuan Jiang (Tencent Inc.), Xinghua Jiang (Tencent YouTu Lab), Deqiang Jiang (Tencent YouTu Lab), Bo Ren (Tencent Youtu Lab)
		mmfp0495	Virtual	Modality-aware Contrastive Instance Learning with Self-Distillation for Weakly-Supervised Audio-Visual Violence Detection -- Jiashuo Yu (Fudan University), Jinyu Liu (Fudan University), Ying Cheng (Fudan University), Rui Feng (Fudan University), Yuejie Zhang (Fudan University)
		mmfp0837	On-site	Parameterization of Cross-token Relations with Relative Positional Encoding for Vision MLP -- Zhicai Wang (University of Science and Technology of China), Yanbin Hao (University of Science and Technology of China), Xingyu Gao (Chinese Academy of Sciences), Hao Zhang (City University of Hong Kong), Shuo Wang (University of Science and Technology of China), Tingting Mu (University of Manchester), Xiangnan He (University of Science and Technology of China)
		mmfp0222	Virtual	Real-time Semantic Segmentation with Parallel Multiple Views Feature Augmentation -- Jianjun Qiao (Southwest Jiaotong University), Zhi-Qi Cheng (Carnegie Mellon University), Xiao Wu (Southwest Jiaotong University), Wei Li (Southwest Jiaotong University), Ji Zhang (Southwest Jiaotong University)
		mmfp0362	Virtual	Exposure-Consistency Representation Learning for Exposure Correction -- Huang Jie (University of Science and Technology of China), man zhou (University of Science and Technology of China), Yajing Liu (University of Science and Technology of China, Tsinghua University), Mingde Yao (University of Science and Technology of China), Feng Zhao (University of Science and Technology of China), Zhiwei Xiong (USTC)
		mmfp0376	Virtual	Global Meets Local: Effective Multi-Label Image Classification via Category-Aware Weak Supervision -- Jiawei Zhan (Tencent), Jun Liu (Tencent YouTu Lab), wei tang (Google), GUANNAN JIANG (University of New South Wales), Xi Wang (South China University of Technology), Bin-Bin Gao (Tencent YouTu Lab), Tianliang Zhang (Tencent Youtu Lab), Wenlong Wu (Tencent YouTu Lab), Wei Zhang (Technische Universität München), Chengjie Wang (Tencent YouTu Lab), Yuan Xie (East China Normal University)
		mmfp0568	Virtual	Domain-Specific Conditional Jigsaw Adaptation for Enhancing transferability and Discriminability -- Qi He (Southwest Jiaotong University), Zhaoquan Yuan (Southwest Jiaotong University), Xiao Wu (Southwest Jiaotong University), HE JUNYAN (Alibaba DAMO Academy)
		mmfp0802	Virtual	Effective Video Abnormal Event Detection by Learning A Consistency-Aware High-Level Feature Extractor -- Guang Yu (National University of Defense Technology), Siqi Wang (National University of Defense Technology), Zhiping Cai (National University of Defense Technology), Xinwang Liu (National University of Defense Technology), Chengkun Wu (University of Manchester)
		mmfp0945	Virtual	Less is More: Consistent Video Depth Estimation with Masked Frames Modeling -- Yiran Wang (Huazhong University of Science and Technology), Zhiyu Pan (Huazhong University of Science and Technology), Xingyi Li (Huazhong University of Science and Technology), Zhiguo Cao (Huazhong University of Science and Technology), Ke Xian (Nanyang Technological University), Jianming Zhang (Adobe Systems)
		mmfp1186	Virtual	Deep Multi-Resolution Mutual Learning for Image Inpainting -- Huan Zheng (Hefei University of Technology), Zhao Zhang (Hefei University of Technology), Haijun Zhang (Harbin Institute of Technology, Shenzhen), Yi Yang (Zhejiang University), Shuicheng YAN (National University of Singapore), Meng Wang (Hefei University of Technology)
		mmfp1264	Virtual	TGDM: Target Guided Dynamic Mixup for Cross-Domain Few-Shot Learning -- Linhai Zhuo (Fudan University), YuQian Fu (Fudan University), Jingjing Chen (Fudan University), Yixin Cao (Singapore Management University), Yu-Gang Jiang (Fudan University)
		mmfp1754	Virtual	SIR-Former: Stereo Image Restoration Using Transformer -- Zizheng Yang (University of Science and Technology of China), Mingde Yao (University of Science and Technology of China), Huang Jie (University of Science and Technology of China), man zhou (University of Science and Technology of China), Feng Zhao (University of Science and Technology of China)
		mmfp2911	Virtual	Learning Occlusion-aware Coarse-to-Fine Depth Map for Self-supervised Monocular Depth Estimation -- Zhengming Zhou (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Qiulei Dong (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
		mmfp2994	Virtual	Guess-It-Generator: Generating in a Lewis Signaling Framework through Logical Reasoning -- Arghya Pal (Harvard University), Sailaja Rajanala (Indian Institute of Technology, Hyderabad), Raphael Phan (Monash University), KokSheik Wong (Monash University Malaysia)
		mmfp2562	Virtual	Long-Term Person Re-identification with Dramatic Appearance Change: Algorithm and Benchmark -- Liu Mengmeng (Nankai University), Ma Zhi (Nankai University), Tao Liu (Nankai University), Yanfeng Jiang (Nankai University), Kai Wang (Nankai University)
		mmfp1018	Virtual	PaCL: Part-level Contrastive Learning for Fine-grained Few-shot Image Classification -- Chuanming Wang (Beijing University of Posts and Telecommunications), Huiyuan Fu (Beijing University of Posts and Telecommunications), Huadong Ma (Beijing University of Post and Telecommunication, Tsinghua University)
		mmfp1123	Virtual	FMNet: Frequency-Aware Modulation Network for SDR-to-HDR Translation -- Gang Xu (Nankai University), Qibin Hou (Nankai University), Le Zhang (University of Electronic Science and Technology of China), Ming-ming Cheng (Nankai University, Tsinghua University)
		mmfp0477	Virtual	CrossNet: Boosting Crowd Counting with Localization -- Ji Zhang (Southwest Jiaotong University), Zhi-Qi Cheng (Carnegie Mellon University), Xiao Wu (Southwest Jiaotong University), Wei Li (Southwest Jiaotong University), Jianjun Qiao (Southwest Jiaotong University)
		mmfp0288	Virtual	NeRF-SR: High Quality Neural Radiance Fields using Supersampling -- Chen Wang (Tsinghua University, Tsinghua University), Xian Wu (Kuaishou Technology), Yuanchen Guo (Tsinghua University), Song-Hai Zhang (Tsinghua University, Tsinghua University), Yu-wing Tai (Kuaishou Technology), Shi-min Hu (Tsinghua University, Tsinghua University)
		mmfp1256	On-site	Rail Detection: An Efficient Row-based Network and A New Benchmark -- Xinpeng Li (South China University of Technology), Xiaojiang Peng (Shenzhen Technology University)
		mmfp0759	Virtual	Robust Attention Deraining Network for Synchronous Rain Streaks and Raindrops Removal -- Yanyan Wei (Hefei University of Technology), Zhao Zhang (Hefei University of Technology), Mingliang Xu (Zhengzhou University), Richang Hong (Hefei University of Technology), Jicong Fan (The Chinese University of Hong Kong (Shenzhen)), Shuicheng Yan (National University of Singapore)
		mmfp1207	Virtual	TSRFormer: Table Structure Recognition with Transformers -- Weihong Lin (Microsoft), Zheng Sun (Institute of automation, Chinese academy of science), Chixiang Ma (University of Science and Technology of China), Mingze Li (Shanghai Jiaotong University), Jiawei Wang (University of Science and Technology of China), Lei Sun (Microsoft), Qiang Huo (Microsoft)
		mmfp2811	Virtual	Structure- and Texture-Aware Learning for Low-Light Image Enhancement -- JingHao Zhang (University of Science and Technology of China), Huang Jie (University of Science and Technology of China), Mingde Yao (University of Science and Technology of China), man zhou (University of Science and Technology of China), Feng Zhao (University of Science and Technology of China)
		mmfp0538	Virtual	CLUT-Net: Learning Adaptively Compressed Representations of 3DLUTs for Lightweight Image Enhancement -- Fengyi Zhang (Tongji University), Hui Zeng (The Hong Kong Polytechnic University), Tianjun Zhang (Tongji University), Lin Zhang (Tongji University)
		mmfp2858	On-site	Automatic piano fingering from partially annotated scores using autoregressive neural networks -- Pedro Ramoneda (Universitat Pompeu Fabra), Dasaem Jeong (Sogang University), Eita Nakamura (Kyoto University), Xavier Serra (Universitat Pompeu Fabra), Marius Miron (Universitat Pompeu Fabra)
		mmfp1575	Virtual	Boosting Single-Frame 3D Object Detection by Simulating Multi-Frame Point Clouds -- Wu Zheng (The Chinese University of Hong Kong), Li Jiang (Max-Planck Institute), Fanbin Lu (Department of Computer Science and Engineering, The Chinese University of Hong Kong), Yangyang Ye (Zhejiang University), Chi-Wing Fu (The Chinese University of Hong Kong)
		mmfp3171	On-site	Towards Complex Document Understanding By Discrete Reasoning -- Fengbin ZHU (National University of Singapore), Wenqiang Lei (Sichuan University), Fuli Feng (University of Science and Technology of China), chao wang (Tsinghua University, Tsinghua University), Haozhou Zhang (Sichuan University), Tat-seng Chua (National University of Singapore)
		mmfp0138	Virtual	RPPformer-Flow: Relative Position Guided Point Transformer for Scene Flow Estimation -- Hanlin Li (University of Electronic Science and Technology of China), Dong Guanting (University of Science and Technology of China), Yueyi Zhang (University of Science and Technology of China), Xiaoyan Sun (University of Science and Technology of China), Zhiwei Xiong (USTC)
		mmfp3074	Virtual	mmLayout: Multi-grained MultiModal Transformer for Document Understanding -- Wenjin Wang (Zhejiang University), Zhengjie Huang (Sun Yat-sen University), Bin Luo (Baidu), Qianglong Chen (Zhejiang University), Qiming Peng (Baidu), Yinxu Pan (Tsinghua University, Tsinghua University), Weichong Yin (Baidu), shikun feng (Baidu), Yu Sun (Baidu), dianhai yu (Baidu), Yin Zhang (Zhejiang University)
		mmfp1089	Virtual	Boosting Video-Text Retrieval with Explicit High-Level Semantics -- Haoran Wang (Baidu), Di Xu (University of the Chinese Academy of Sciences), Dongliang He (Baidu), Fu Li (Dalian University of Technology), Zhong Ji (Tianjin university), Jungong Han (Aberystwyth University), Errui Ding (Baidu)
		mmfp2338	Virtual	Rethinking the Mechanism of the Pattern Pruning and the Circle Importance Hypothesis -- Hengyi Zhou (Xi'an Jiaotong University), Longjun Liu (Xi'an Jiaotong University), Haonan Zhang (Xi'an Jiaotong university), Nanning Zheng (Xi'an Jiaotong university)
		mmfp1740	Virtual	A Region-based Document VQA -- Xinya Wu (Beijing University of Posts and Telecommunications), Duo Zheng (Beijing University of Posts and Telecommunications), RUONAN WANG (Beijing University of Posts and Telecommunications), Jiashen Sun (Beijing University of Posts and Telecommunications), Hu Minzhen (Beijing University of Posts and Telecommunications), Fangxiang Feng (Beijing University of Posts and Telecommunications), Xiaojie Wang (Beijing University of Post and Telecommunication), Huixing Jiang (Li Auto), Fan Yang (Institute of Automation, Chinese Academy of Sciences)
		mmfp1822	Virtual	CyclicShift: A Data Augmentation Method For Enriching Data Patterns -- Hui Lu (University of Electronic Science and Technology of China), Xuan Cheng (University of Electronic Science and Technology of China), Wentao Xia (University of Electronic Science and Technology of China), Pan Deng (University of Electronic Science and Technology of China), MingHui Liu (University of Electronic Science and Technology of China), Tianshu Xie (University of Electronic Science and Technology of China), XiaoMin Wang (University of Electronic Science and Technology of China), Ming Liu (University of Electronic Science and Technology of China)
		mmfp1682	On-site	Counterexample Contrastive Learning for Spurious Correlation Elimination -- Jinqiang Wang (Beijing Jiaotong University), Rui Hu (Beijing Jiaotong University), Chaoquan Jiang (Beijing Jiaotong University), Rui Hu (Beijing Jiaotong University), Jitao Sang (Beijing Jiaotong University)
		mmfp1316	Virtual	MC-SLT: Towards Low-Resource Signer-Adaptive Sign Language Translation -- Tao Jin (Zhejiang University), Zhou Zhao (Zhejiang University), Meng Zhang (Huawei Technologies Ltd.), Xingshan Zeng (Huawei Technologies Ltd.)
		mmfp0732	Virtual	Deep Evidential Learning with Noisy Correspondence for Cross-modal Retrieval -- Yang Qin (Sichuan University), Dezhong Peng (Sichuan University), Xi Peng (Sichuan University), Shawn Wang (Sichuan University), Peng Hu (Sichuan University)
		mmfp2506	Virtual	CAliC: Accurate and Efficient Image-Text Retrieval via Contrastive Alignment and Visual Contexts Modeling -- Hongyu Gao (University of Science and Technology Beijing), Chao Zhu (University of Science and Technology Beijing), Mengyin Liu (University of Science and Technology Beijing), Weibo Gu (Tencent), WANG HongFa (Chinese Academy of Sciences), Wei Liu (Tencent AI Lab), Xu-cheng Yin (University of Science and Technology Beijing)
		mmfp0066	Virtual	Correspondence Matters for Video Referring Expression Comprehension -- Meng Cao (Peking University), Jiang Ji (Peking University), Long Chen (Columbia University), Zou Yuexian (Peking University)
		mmfp2026	Virtual	Point to Rectangle Matching for Image Text Retrieval -- Zheng WANG (University of Electronic Science and Technology of China), Zhenwei Gao (University of Electronic Science and Technology of China), Xing Xu (University of Electronic Science and Technology of China), Yadan Luo (University of Queensland), Yang Yang (University of Electronic Science and Technology of China), Hengtao Shen (University of Electronic Science and Technology of China)
		mmfp1884	Virtual	Shifting Perspective to see difference: a novel multi-view method for skeleton based action recognition -- Ruijie Hou (Zhejiang University), Yanran Li (University of Birmingham), Ningyu Zhang (Zhejiang University), YULIN ZHOU (Zhejiang University), Xiaosong Yang (Bournemouth University), Zhao Wang (Zhejiang University)
		mmfp3003	Virtual	Counterfactually Measuring and Eliminating Social Bias in Vision-Language Pre-training Models -- Yi Zhang (Beijing Jiaotong University), Junyang Wang (Beijing Jiaotong University), Jitao Sang (Beijing Jiaotong University)
		mmfp0268	Virtual	Towards Adversarial Attack on Vision-Language Pre-training Models -- Jiaming Zhang (Beijing Jiaotong University), QI YI (Beijing Jiaotong University), Jitao Sang (Beijing Jiaotong University)
		mmfp0546	On-site	TPSNet: Reverse Thinking of Thin Plate Splines for Arbitrary Shape Scene Text Representation -- Wei Wang (Institute of Information Engineering, Chinese Academy of Sciences), Yu Zhou (Institute of Information Engineering, CAS), Jiahao Lv (Beijing Normal University), Dayan Wu (iie,cas), Guoqing Zhao (Mashang Consumer Finance Co., Ltd.), ning jiang (Mashang Consumer Finance Co, Ltd), Weipinng Wang (IIE)

Understanding Multimedia Content: Vision and Language		mmfp0406	On-site	Efficient Modeling of Future Context for Image Captioning -- Zhengcong Fei (Institute of Computing Technology, Chinese Academy of Sciences)
		mmfp0233	Virtual	Relative Pose Estimation for Multi-Camera Systems from Point Correspondences with Scale Ratio -- Banglei Guan (National University of Defense Technology), Ji Zhao (TuSimple)
		mmfp0073	Virtual	Towards Open-Ended Text-to-Face Generation, Combination and Manipulation -- Jun Peng (Xiamen University), Han Pan (Xiamen University), Yiyi Zhou (Xiamen University), Jing He (Xiamen University), Xiaoshuai Sun (Xiamen University), Yan Wang (Pinterest), YONGJIAN WU (Wuhan University), Rongrong Ji (Columbia University)
		mmfp1940	Virtual	Improving Fusion of Region Features and Grid Features via Two-Step Interaction for Image-Text Retrieval -- Dongqing Wu (Northwest Polytechnical University Xi'an), Huihui Li (Northwest Polytechnical University Xi'an), Cang Gu (Northwestern Polytechnical University, Northwest Polytechnical University Xi'an), Lei Guo (Northwest Polytechnical University Xi'an), Hang LIU (Northwest Polytechnical University Xi'an)
		mmfp0558	Virtual	A Numerical DEs Perspective on Unfolded Linearized ADMM Networks for Inverse Problems -- Weixin An (Xi'an University of Electronic Science and Technology), yue yingjie (Xi'an University of Electronic Science and Technology), Yuanyuan Liu (Xidian University), Fanhua Shang (The Chinese University of Hong Kong), Hongying Liu (Xidian University)
		mmfp0705	Virtual	UDoc-GAN: Unpaired Document Illumination Correction with Background Light Prior -- Yonghui Wang (University of Science and Technology of China), Wengang Zhou (University of Science and Technology of China), Zhenbo Lu (University of Science and Technology of China), Houqiang Li (University of Science and Technology of China)
		mmfp0557	On-site	Dilated Context Integrated Network with Cross-Modal Consensus for Temporal Emotion Localization in Videos -- Juncheng Li (Zhejiang University), Junlin Xie (Zhejiang University), Linchao Zhu (University of Technology Sydney), Long Qian (Zhejiang University), Siliang Tang (Zhejiang University), Wenqiao Zhang (National University of Singapore), Haochen Shi (Université de Montréal), Shengyu Zhang (Zhejiang University), Longhui Wei (Huawei Cloud Technologies Ltd.), Qi Tian (Huawei Technologies Ltd.), Yueting Zhuang (Zhejiang University)
		mmfp0108	Virtual	Balanced Gradient Penalty Improves Deep Long-Tailed Learning -- Dong Wang (Xidian University), Yicheng Liu (The Chinese University of Hong Kong), Liangji Fang (Shanghai Jiaotong University), Fanhua Shang (The Chinese University of Hong Kong), Yuanyuan Liu (Xidian University), Hongying Liu (Xidian University)
		mmfp0147	Virtual	Uncertainty-Aware 3D Human Pose Estimation from Monocular Video -- Jinlu Zhang (Wuhan University), Yujin Chen (Technische Universität München), Zhigang Tu (Wuhan University)
		mmfp0250	On-site	MVSPlenOctree: Fast and Generic Reconstruction of Radiance Fields in PlenOctree from Multi-view Stereo -- Wenpeng Xing (Hong Kong Baptist University), Jie Chen (Hong Kong Baptist University)
		mmfp0251	On-site	A Dual-Masked Auto-Encoder for Robust Motion Capture with Spatial-Temporal Skeletal Token Completion -- Junkun Jiang (Hong Kong Baptist University), Jie Chen (Hong Kong Baptist University), Yike Guo (Hong Kong Baptist University)
		mmfp0331	Virtual	Learning Dynamic Prior Knowledge for Text-to-Face Pixel Synthesis -- Jun Peng (Xiamen University), Xiaoxiong Du (Xiamen University), Yiyi Zhou (Xiamen University), Jing He (Xiamen University), Yunhang Shen (Xiamen University), Xiaoshuai Sun (Xiamen University), Rongrong Ji (Columbia University)
		mmfp0475	Virtual	Correct Twice at Once: Learning to Correct Noisy Labels for Robust Deep Learning -- Jingzheng Li (Beihang university), Hailong Sun (School of Software, Beihang University)
		mmfp0814	Virtual	Align, Reason and Learn: Enhancing Medical Vision-and-Language Pre-training with Knowledge -- Zhihong Chen (THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN), Guanbin Li (SUN YAT-SEN UNIVERSITY), Xiang Wan (Shenzhen Research Institute of Big Data)
		mmfp0856	Virtual	Diverse Human Motion Prediction via Gumbel-Softmax Sampling from an Auxiliary Space -- Lingwei Dang (South China University of Technology), Yongwei Nie (South China University of Technology), Chengjiang Long (Meta Reality Labs), Qing Zhang (SUN YAT-SEN UNIVERSITY), Guiqing Li (South China University of Technology)
		mmfp0865	Virtual	Towards High-Fidelity Face Normal Estimation -- Meng Wang (Tianjin University), Chaoyue Wang (JD Explore Academy), Xiaojie Guo (Tianjin University), Jiawan Zhang (Tianjin University)
		mmfp0992	On-site	Generating Transferable Adversarial Examples against Vision Transformers -- Yuxuan Wang (Beijing University of Aeronautics and Astronautics), Jiakai Wang (Beihang University), Zixin Yin (Beijing University of Aeronautics and Astronautics), Ruihao Gong (Beihang University), Jingyi Wang (Beihang University), Aishan Liu (Beijing University of Aeronautics and Astronautics), Xianglong Liu (Beihang University, Tsinghua University)
		mmfp1016	Virtual	Video-Guided Curriculum Learning for Spoken Video Grounding -- Yan Xia (Zhejiang University), Zhou Zhao (Zhejiang University), Ye Shangwei (Zhejiang University), Yang Zhao (Zhejiang University), Haoyuan Li (Zhejiang University), Yi Ren (Zhejiang University)
		mmfp1100	Virtual	Multi-Scale Coarse-to-Fine Transformer for Frame Interpolation -- Chen Li (Shanghai Jiao Tong University), Li Song (Shanghai Jiao Tong University), Xueyi Zou (Huawei Technologies Ltd.), Jeremie Guo (Huawei Technologies Ltd.), Youliang Yan (Huawei Technologies Ltd.), Wenjun Zhang (Shanghai Jiao Tong University)
		mmfp1380	Virtual	S-CCR: Super-Complete Comparative Representation for Low-Light Image Quality Inference In-the-wild -- Miaohui Wang (Shenzhen University), Zhuowei Xu (Shenzhen Key Laboratory of Media Security, Shenzhen University), Yuanhao Gong (Shenzhen University), Wuyuan Xie (Department of Software Engineering, Shenzhen University)
		mmfp1475	On-site	Talking Head Generation from Speech Audio using a Pre-trained Image Generator -- Mohammed Alghamdi (University of Leeds), Andrew Bulpitt (University of Leeds), He Wang (University of Leeds), David Hogg (University of Leeds)
		mmfp1547	Virtual	Exploring High-quality Target Domain Information for Unsupervised Domain Adaptive Semantic Segmentation -- Junjie Li (University of Science and Technology of China), Zilei Wang (University of Science and Technology of China), Yuan Gao (University of Science and Technology of China), Xiaoming Hu (University of Science and Technology of China)
		mmfp1692	On-site	Semantics-Driven Generative Replay for Few-Shot Class Incremental Learning -- Aishwarya Agarwal (Indian Institute of Technology, Bombay), Biplab Banerjee (Indian Institute of Technology Bombay, Dhirubhai Ambani Institute Of Information and Communication Technology), Fabio Cuzzolin (Oxford Brookes University), Subhasis Chaudhuri (University of California, San Diego)
		mmfp2290	Virtual	Global-local Cross-view Fisher Discrimination for View-Invariant Action Recognition -- Lingling Gao (University of Electronic Science and Technology of China), Yanli Ji (University of Electronic Science and Technology of China, Tsinghua University), Yang Yang (University of Electronic Science and Technology of China), Hengtao Shen (University of Electronic Science and Technology of China)
		mmfp2437	On-site	Reflecting on Experiences for Response Generation -- Chenchen Ye (National University of Singapore), Lizi Liao (Singapore Management University), SuYu Liu (Singapore Management University), Tat-seng Chua (National University of Singapore)
Engaging Users with Multimedia: Multimedia Search and Recommendation		mmfp1665	Virtual	Multi-Level Region Matching for Fine-Grained Sketch-Based Image Retrieval -- Zhixin Ling (Fudan University), Zhen Xing (Fudan University), jiangtong li (Shanghai Jiao Tong University), Li Niu (Shanghai Jiao Tong University)
		mmfp1344	Virtual	DDGHM: Dual Dynamic Graph with Hybrid Metric Training for Cross-Domain Sequential Recommendation -- Xiaolin Zheng (Zhejiang University), Jiajie Su (Zhejiang University), Liu Weiming (Zhejiang University), Chaochao Chen (Zhejiang University)
		mmfp2708	On-site	Spatial-Temporal Aligned Multi-Agent Learning for Visual Dialog Systems -- Yong Zhuang (Carnegie Mellon University), Tong Yu (Adobe Research), Junda Wu (New York University), Shiqu Wu (University of California, San Diego), Shuai Li (John Hopcroft Center, Shanghai Jiao Tong University)
		mmfp2385	Virtual	Learning Intrinsic and Extrinsic Intentions for Cold-start Recommendation with Neural Stochastic Processes -- Huafeng Liu (University of Hong Kong), Liping Jing (Beijing Jiaotong University), Dahai Yu (Northwestern Polytechnical University), Mingjie Zhou (The University of Hong Kong), Michael Ng (The University of Hong Kong)
		mmfp1330	Virtual	Camera-specific Informative Data Augmentation Module for Unbalanced Person Re-identification -- Pingting Hong (Institute of Information Engineering. Chinese Academy of Sciences), Dayan Wu (iie,cas), Bo Li (Chinese Academy of Sciences), Weipinng Wang (IIE)
		mmfp2351	Virtual	TopicVAE: Topic-aware Disentanglement Representation Learning for Enhanced Recommendation -- Zhiqiang Guo (Huazhong University of Science and Technology), Guohui Li (Huazhong University of Science and Technology), Jianjun Li (Huazhong University of Science and Technology), Huaicong Chen (Huazhong University of Science and Technology)
		mmfp1379	Virtual	Pixel-Level Anomaly Detection via Uncertainty-aware Prototypical Transformer -- Huang Chao (Harbin Institute of Technology), Chengliang Liu (Harbin Institute of Technology), Zheng Zhang (Harbin Institute of Technology), Zhihao Wu (Harbin Institute of Technology, Shenzhen), Jie Wen (Harbin Institute of Technology), Qiuping Jiang (Ningbo University), Yong Xu (Harbin Institute of Technology)
		mmfp0112	Virtual	Dynamic Prototype Mask for Occluded Person Re-Identification -- Lei Tan (Xiamen University), Pingyang Dai (Xiamen University), Rongrong Ji (Columbia University), YONGJIAN WU (Wuhan University)
		mmfp2017	On-site	Meta Reconciliation Normalization for Lifelong Person Re-Identification -- Nan Pu (Leiden University, Leiden University), Yu Liu (Department of Electrical Engineering, KU Leuven, Belgium, KU Leuven), Wei Chen (National University of Defense Technology), Erwin Bakker (Leiden University), Michael Lew (Leiden University)
		mmfp0861	Virtual	Attack is the Best Defense: Towards Preemptive-Protection Person Re-Identification -- Lin Wang (IIE,CAS), Wanqian Zhang (IIE, CAS), Dayan Wu (iie,cas), Fei Zhu (Institute of Information Engineering. Chinese Academy of Sciences), Bo Li (Chinese Academy of Sciences)
		mmfp1120	Virtual	TAGPerson: A Target-Aware Generation Pipeline for Person Re-identification -- Kai CHEN (Tsinghua University), Weihua Chen (Alibaba Group), Tao He (Tsinghua University), Rong Du (Alibaba Group), Fan Wang (Alibaba Group), Xiuyu Sun (Alibaba Group), Yuchen Guo (Tsinghua University, Tsinghua University), Guiguang Ding (Tsinghua University)
		mmfp1122	Virtual	Efficient Hash Code Expansion by Recycling Old Bits -- Dayan Wu (iie,cas), Qinghang Su (Institute of Information Engineering, CAS), Bo Li (Chinese Academy of Sciences), Weiping Wang (IIE, CAS)
Engaging Users with Multimedia: Emotional and Social Signals		mmfp0487	Virtual	DPCNet: Dual Path Multi-Excitation Collaborative Network for Facial Expression Representation Learning in Videos -- Yan Wang (Fudan University), Yixuan Sun (Fudan University), Wei Song (Shanghai Ocean University), Shuyong Gao (Fudan University), Yiwen Huang (Fudan University), Zhaoyu Chen (Fudan University), Weifeng Ge (Fudan University), Wenqiang Zhang (Fudan University)
		mmfp1550	Virtual	Pursuing Knowledge Consistency: Supervised Hierarchical Contrastive Learning for Facial Action Unit Recognition -- Yingjie Chen (Peking University), Chong Chen (Alibaba Group), Xiao Luo (Peking University), Jianqiang Huang (Nanyang Technological University), Xian-Sheng Hua (Zhejiang University), Tao Wang (Peking University), Yun Liang (Google)
		mmfp2566	Virtual	Unsupervised Domain Adaptation Integrating Transformer and Mutual Information for Cross-Corpus Speech Emotion Recognition -- Shiqing Zhang (Taizhou University), Ruixin Liu (Taizhou University), Yijiao Yang (Taizhou College), Xiaoming Zhao (Taizhou University), Jun Yu (Hangzhou Dianzi University)
		mmfp1802	Virtual	Co-Completion for Occluded Facial Expression Recognition -- Zhen Xing (Fudan University), Weimin Tan (Fudan University), Ruian He (Fudan University), Yangle Lin (Fudan University), Bo Yan (Fudan University)
		mmfp2453	On-site	Generalized Similarity Loss for Gait Recognition -- Weichen Yu (, Institute of automation, Chinese academy of science), Hongyuan Yu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Yan Huang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Liang Wang (CASIA)
		mmfp2275	Virtual	Feeling Without Sharing: A Federated Video Emotion Recognition Framework Via Privacy-Agnostic Hybrid Aggregation -- Fan Qi (Tianjin University of Technology), Zixin Zhang (Tianjin University of Technology), Xiaoshan Yang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), huaiwen zhang (Inner Mongolia University), Changsheng Xu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
Engaging Users with Multimedia: Summarization, Analytics, and Storytelling		mmfp1159	Virtual	MMT: Image-guided Story Ending Generation with Multimodal Memory Transformer -- Dizhan Xue (Institute of Automation, Chinese Academy of Sciences), Shengsheng Qian (, Institute of automation, Chinese academy of science), Quan Fang (, Institute of automation, Chinese academy of science), Changsheng Xu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
		mmfp2123	Virtual	An End-to-End Conditional Generative Adversarial Network Based on Depth Map for 3D Craniofacial Reconstruction -- Niankai Zhang (Qingdao University), Junli Zhao (Qingdao University), Fuqing Duan (Beijing normal university), Zhenkuan Pan (Qingdao University), Zhongke Wu (Beijing Normal University), Mingquan Zhou (Beijing Normal University), David Gu (Stony Brook University)
		mmfp2724	Virtual	The More, The Better? Active Silencing of Non-Positive Transfer for Efficient Multi-Domain Few-Shot Classification -- Xingxing Zhang (Tsinghua University, Tsinghua University), Zhizhe Liu (Beijing Jiaotong University), Weikai Yang (Softwarn Engineering, Tsinghua University, Tsinghua University), Liyuan Wang (Tsinghua University), Jun Zhu (Tsinghua University)
		mmfp3100	On-site	Hierarchical Few-Shot Object Detection: Problem, Benchmark and Method -- Lu Zhang (Fudan University), Yang Wang (Tongji University), Jiaogeng Zhou (Huaiyin Normal Univeristy), Chenbo Zhang (Fudan University), Yinglu Zhang (Fudan University), Jihong Guan (Tongji University), An Bian (Tencent AI Lab), Shuigeng Zhou (Fudan University)
		mmfp1357	On-site	Few-shot X-ray Prohibited Item Detection: A Benchmark and Weak-feature Enhancement Network -- Tao Renshuai (Beihang University), tianbo Wang (Beijing University of Aeronautics and Astronautics), Ziyang Wu (University of Science and Technology of China), Cong Liu (iFLYTEK), Aishan Liu (Beijing University of Aeronautics and Astronautics), Xianglong Liu (Beihang University)
		mmfp0539	Virtual	High-Fidelity Variable-Rate Image Compression via Invertible Activation Transformation -- Shilv Cai (Huazhong University of Science and Technology), Zhijun Zhang (Huazhong University of Science and Technology), Liqun Chen (Huazhong University of Science and Technology), Luxin Yan (Huazhong University of Science and Technology), Sheng Zhong (Huazhong University of Science and Technology), Xu Zou (Huazhong University of Science and Technology)
		mmfp1631	Virtual	Cycle Encoding of a StyleGAN Encoder for Improved Reconstruction and Editability -- Xudong Mao (SUN YAT-SEN UNIVERSITY), Liujuan Cao (Xiamen University), Aurele Gnanha (City University of Hong Kong), Zhenguo Yang (Guangdong University of Technology), Qing Li (City University of Hong Kong), Rongrong Ji (Columbia University)
		mmfp0446	Virtual	Speech Fusion to Face: Bridging the Gap Between Human's Vocal Characteristics and Facial Imaging -- Yeqi BAI (Nanyang Technological University), TAO MA (Shanghai AI Laboratory), Lipo Wang (Nanyang Technological University), Zhenjie Zhang (Neuron Mobility Pte. Ltd.)
		mmfp0350	Virtual	Learning Action-guided Spatio-temporal Transformer for Group Activity Recognition -- Wei Li (Southwest Jiaotong University), Tianzhao Yang (Southwest Jiaotong University), Xiao Wu (Southwest Jiaotong University), Xian-Jun Du (Southwest Jiaotong University), Jianjun Qiao (Southwest Jiaotong University)
		mmfp0504	On-site	A Unified End-to-End Retriever-Reader Framework for Knowledge-based VQA -- Yangyang Guo (National University of Singapore), Liqiang Nie (Harbin Institute of Technology (Shenzhen)), Yongkang Wong (National University of Singapore), Yibing Liu (City University of Hong Kong), Zhiyong Cheng (Shandong Artificial Intelligence Institute), Mohan Kankanhalli (National University of Singapore)
		mmfp1216	Virtual	PIA: Parallel Architecture with Illumination Allocator for Joint Enhancement and Detection in Low-Light -- Tengyu Ma (Dalian University of Technology), Long Ma (Dalian University of Technology), Xin Fan (Dalian University of Technology), Zhongxuan Luo (Dalian University of Technology), Risheng Liu (Dalian University of Technology)
		mmfp3084	On-site	Robust Actor Recognition in Entertainment Multimedia at Scale -- Abhinav Aggarwal (Amazon), Yash Pandya (Amazon), Lokesh Ravindranathan (Amazon), Laxmi Shivaji Ahire (Amazon), Manivel Sethu (Amazon), Kaustav Nandy (Department of Computer Science, University of Maryland, College Park)

Experience: Multimedia Applications		mmfp3115	Virtual	MF-Net: A Novel Few-shot Stylized Multilingual Font Generation Method -- Yufan Zhang (Duke University), Junkai Man (Duke University), Peng Sun (Duke University)
		mmfp0530	Virtual	Feature and Semantic Views Consensus Hashing for Image Set Classification -- Yuan Sun (Sichuan University), Dezhong Peng (Sichuan University), Haixiao Huang (Sichuan Provincial Commision of Politics and Law), Zhenwen Ren (Southwest University Of Science And Technology)
		mmfp1417	Virtual	Evidential Reasoning for Video Anomaly Detection -- Che Sun (Beijing Institute of Technology), Yunde Jia (Beijing Institute of Technology), Yuwei Wu (Beijing Institute of Technology)
		mmfp1873	Virtual	Gaze- and Spacing-flow Unveil Intentions: Hidden Follower Discovery -- danni xu (Wuhan University), Ruimin Hu (Wuhan University), Zheng Wang (Wuhan University, China), Linbo Luo (Xidian University), Dengshi Li (Jianghan University), Wenjun Zeng (Eastern Institute for Advanced Study)
		mmfp0949	Virtual	Semi-supervised Learning for Multi-label Video Action Detection -- Hongcheng Zhang (Shanghai Jiaotong University), Xu Zhao (Shanghai Jiao Tong University), Dongqi Wang (Shanghai Jiaotong University)
		mmfp0870	Virtual	Learning Cross-Image Object Semantic Relation in Transformer for Few-Shot Fine-Grained Image Classification -- Bo Zhang (Fudan University), Jiakang Yuan (Fudan University), Baopu Li (The Chinese University of Hong Kong), Tao Chen (Fudan University), Jiayuan Fan (Fudan University), Botian Shi (Shanghai AI Lab)
		mmfp0527	Virtual	Progressive Spatial-temporal Collaborative Network for Video Frame Interpolation -- Mengshun Hu (Wuhan University), Kui Jiang (Wuhan University), Liang Liao (Nanyang Technological University), Zhixiang Nie (Wuhan University), Jing Xiao (Wuhan University), Zheng Wang (Wuhan University, China)
		mmfp2165	Virtual	Best of Both Worlds: See and Understand Clearly in the Dark -- Xinwei Xue (Dalian University of Technology), Jia He (Dalian University of Technology), Long Ma (Dalian University of Technology), Yi Wang (Dalian University of Technology), Xin Fan (Dalian University of Technology), Risheng Liu (Dalian University of Technology)
		mmfp0617	Virtual	Meta Clustering Learning for Large-scale Unsupervised Person Re-identification -- Xin Jin (University of Science and Technology of China), Tianyu He (Alibaba Group), Xu Shen (University of Science and Technology of China), Tongliang Liu (University of Sydney), Xinchao Wang (National University of Singapore), Jianqiang Huang (Nanyang Technological University), Zhibo Chen (University of Science and Technology of China, Tsinghua University), Xian-Sheng Hua (Zhejiang University)
		mmfp0134	Virtual	Adjustable Memory-efficient Image Super-resolution via Individual Kernel Sparsity -- Luo Xiaotong (XMU), Mingliang Dai (Fudan University), Yulun Zhang (Swiss Federal Institute of Technology), Yuan Xie (East China Normal University), Ding Liu (University of Illinois, Urbana Champaign), Yanyun Qu (Xiamen University), Yun Fu (Northeastern University), Junping Zhang (Fudan University)
		mmfp0153	Virtual	GT-MUST: Gated Try-on by Learning the Mannequin-Specific Transformation -- Ning Wang (Wuhan University), Jing Zhang (The University of Sydney), Lefei Zhang (Wuhan University), Dacheng Tao (JD.com)
		mmfp0161	Virtual	PC^2-PU: Patch Correlation and Point Correlation for Effective Point Cloud Upsampling -- Chen Long (Wuhan University), Wenxiao Zhang (Singapore University of Technology and Design), Ruihui Li (Hunan University), Hao Wang (Wuhan University), Zhen Dong (Wuhan University), Bisheng Yang (Wuhan University)
		mmfp0750	Virtual	Self-Supervised Multi-view Stereo via Adjacent Geometry Guided Volume Completion -- Luoyuan Xu (Huazhong University of Science and Technology), Tao Guan (Huazhong University of Science and Technology), yuesong Wang (Huazhong University of Science and Technology), Yawei Luo (Zhejiang University, Tsinghua University), CHEN ZHUO (Huazhong University of Science and Technology), Wenkai Liu (Huazhong University of Science and Technology), Wei Yang (Huazhong University of Science and Technology)
		mmfp1691	Virtual	AtHom: Two Divergent Attentions Stimulated By Homomorphic Training in Text-to-Image Synthesis -- Zhenbo Shi (University of Science and Technology of China), Zhi Chen (University of Science and Technology of China), Zhenbo Xu (University of Science and Technology of China), Wei Yang (University of Science and Technology of China), Liusheng Huang (University of Science and Technology of China)
		mmfp2350	Virtual	One-step Low-Rank Representation for Clustering -- Zhiqiang Fu (Beijing Jiaotong University), Yao Zhao (Beijing Jiaotong University), Dongxia Chang (Beijing Jiaotong University), Yiming Wang (Beijing Jiaotong University), Jie Wen (Harbin Institute of Technology), Xingxing Zhang (Tsinghua University, Tsinghua University), Guodong Guo (West Virginia University)
		mmfp3116	Virtual	Customizing GAN Using Few-shot Sketches -- Syed Israr (University of Science and Technology of China), Feng Zhao (University of Science and Technology of China)
		mmfp1925	On-site	Video coding using learned latent GAN compression -- Mustafa Shukor (InterDigital, Inc), Bharath Damodaran (Interdigital R&D), Xu YAO (Télécom Paris), Pierre Hellier (Interdigital)
		mmfp2311	Virtual	Action-conditioned On-demand Motion Generation -- QIUJING LU (University of California, Los Angeles), Yipeng Zhang (University of California, Los Angeles), MINGJIAN LU (University of California, Los Angeles), vwani Roychowdhury (University of California, Los Angeles)
		mmfp0775	On-site	Universal Domain Adaptive Object Detector -- Wenxu Shi (Chongqing University), Lei Zhang (Chongqing University), Weijie Chen (Hikvision Research Institute), Shiliang Pu (Hikvision Research Institute)
		mmfp1250	On-site	PIMoG: An Effective Screen-shooting Noise-Layer Simulation for Deep-Learning-Based Watermarking Network -- Han Fang (National University of Singapore), Zhaoyang Jia (University of Science and Technology of China), Zehua Ma (University of Science and Technology of China), Ee-Chien Chang (National University of Singapore), Weiming Zhang (University of Science and Technology of China)
		mmfp1399	Virtual	Universal Domain Adaptive Object Detector -- Wenxu Shi(Chongqing University), Lei Zhang(Chongqing University), Weijie Chen(Hikvision Research Institute), Shiliang Pu (Hikvision Research Institute)
		mmfp1187	Virtual	Composite Photograph Harmonization with Complete Background Cues -- Yazhou Xing (The Hong Kong University of Science and Technology), Yu Li (International Digital Economy Academy), Xintao Wang (Applied Research Center, Tencent PCG), Ye Zhu (South China University of Technology), Qifeng Chen (Hong Kong University of Science and Technology)
		mmfp1890	Virtual	Self-supervised Multi-view Stereo via Inter and Intra Network Pseudo Depth -- Ke Qiu (Peking University), Yawen Lai (Peking University), Shiyi Liu (Peking University), Ronggang Wang (Peking University Shenzhen Graduate School)
		mmfp2022	Virtual	Delegate-based Utility Preserving Synthesis for Pedestrian Image Anonymization -- Zhenzhong Kuang (Hangzhou Dianzi University), Longbin Teng (Hangzhou Dianzi University), Zhou Yu (Hangzhou Dianzi University), Jun Yu (Hangzhou Dianzi University), Jianping Fan (AI Lab at Lenovo Research), Mingliang Xu (Zhengzhou University)
		mmfp0699	Virtual	Video Instance Lane Detection via Deep Temporal and Geometry Consistency Constraints -- Mingqian Wang (Tianjin University), Yujun Zhang (Tianjin University), Wei Feng (Tianjin University), Lei Zhu (Hong Kong University of Science and Technology), Song Wang (University of South Carolina)
		mmfp1131	Virtual	Learning Visible Surface Area Estimation for Irregular Objects -- Xu Liu (Harbin Institute of Technology), Jianing Li (Peking University), Xianqi Zhang (Harbin Institute of Technology), Jingyuan Sun (Harbin Institute of Technology), Xiaopeng Fan (Harbin Institute of Technology), Yonghong Tian (Peking University)
		mmfp1732	Virtual	Blind Robust Video Watermarking Based on Adaptive Region Selection and Channel Reference -- Qinwei Chang (Tencent), Leichao Huang (Huazhong University of Science and Technology), Shaoteng Liu (Northwest Polytechnical University Xi'an), Hualuo Liu (Tencent), Yang Tianshu (Tencent Video), Yexin Wang (Peking University)
		mmfp1638	Virtual	Disparity-based Stereo Image Compression with Aligned Cross-View Priors -- Yongqi Zhai (Peking University), Luyang Tang (Peking University), Yi Ma (Peking University), Rui Peng (Peking University), Ronggang Wang (Peking University Shenzhen Graduate School)
		mmfp1291	Virtual	Label-Efficient Domain Generalization via Collaborative Exploration and Generalization -- Junkun Yuan (Zhejiang University), Ma Xu (Zhejiang University), Defang Chen (Zhejiang University), Kun Kuang (Zhejiang University), Fei Wu (Zhejiang University), Lanfen Lin (Zhejiang University)
		mmfp0242	Virtual	Progressive Unsupervised Learning of Local Descriptors -- Wufan Wang (Beijing Institute of Technology), Lei Zhang (Beijing Institute of Technology), Hua Huang (Beijing Normal University)
		mmfp0463	Virtual	Graph Reasoning Transformer for Image Parsing -- Zhang Dong (The Hong Kong University of Science and Technology), Jinhui Tang (Nanjing University of Science and Technology), Kwang-Ting Cheng (HKUST)
		mmfp2175	Virtual	Opportunistic Backdoor Attacks: Exploring Human-imperceptible Vulnerabilities on Speech Recognition Systems -- Qiang Liu (National University of Defense Technology), Tongqing Zhou (National University of Defense Technology), Zhiping Cai (National University of Defense Technology), Yonghao Tang (National University of Defense Technology)
		mmfp2518	Virtual	Certifying Better Robust Generalization for Unsupervised Domain Adaptation -- Zhiqiang Gao (Xi'an Jiaotong-Liverpool University), Shufei Zhang (Shanghai AI Lab), Kaizhu Huang (Duke Kunshan University), Qiufeng Wang (Xi'an Jiaotong-Liverpool University), Rui Zhang (Xi'an Jiaotong-Liverpool University), Chaoliang Zhong (Fujitsu Research and Development Center Co. Ltd.)
		mmfp1300	Virtual	Multimodal In-bed Pose and Shape Estimation under the Blankets -- Yu Yin (Northeastern University), Joseph Robinson (Vicarious Surgical), Yun Fu (Northeastern University)
		mmfp1028	Virtual	Progressive Limb-Aware Virtual Try-On -- Xiaoyu Han (Harbin Institute of Technology), Shengping Zhang (Harbin Institute of Technology), Qinglin Liu (Harbin Institute Of Technology), Zonglin Li (Harbin Institute of Technology), Chenyang Wang (Harbin Institute of Technology)
		mmfp2033	On-site	Text Style Transfer based on Multi-factor Disentanglement and Mixture -- Anna Zhu (Wuhan University of Technology), Zhanhui Yin (Wuhan University of Technology), Brian Iwana (Kyushu University), Xinyu Zhou (Wuhan University of Technology), Shengwu Xiong (Wuhan University of Technology)
		mmfp0297	On-site	Cloud2Sketch: Augmenting Clouds with Imaginary Sketches -- Zhaoyi Wan (JQ Investments), Dejia Xu (University of Texas at Austin), Zhangyang Wang (University of Texas, Austin), Jian Wang (Snap Inc.), Jiebo Luo (University of Rochester)
Experience: Art and Culture		mmfp1874	On-site	CharFormer: A Glyph Fusion based Attentive Framework for High-precision Character Image Denoising -- daqian Shi (University of Trento), Xiaolei Diao (University of Trento), Lida Shi (Jilin University), Hao Tang (ETH Zurich), Yang Chi (Jilin University), Chuntao Li (Jilin University), Hao Xu (Jilin University)
		mmfp0195	Virtual	Delving into the Frequency: Temporally Consistent Human Motion Transfer in the Fourier Space -- Guang Yang (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Wu Liu (JD Explore Academy), Xinchen Liu (JD Explore Academy), Xiaoyan Gu (Institute of Information Engineering, Chinese Academy of Sciences), Juan Cao (Institute of Computing Technology, Chinese Academy of Sciences), Jintao Li (Institute of Computing Technology, Chinese Academy of Sciences)
		mmfp2577	Virtual	Adaptive Affine Transformation: A Simple and Effective Operation for Spatial Misaligned Image Generation -- Zhimeng Zhang (Shandong University), Yu Ding (Netease Fuxi AI Lab)
		mmfp2683	On-site	RCRN: Real-world Character Image Restoration Network via Skeleton Extraction -- daqian Shi (University of Trento), Xiaolei Diao (University of Trento), Hao Tang (ETH Zurich), Xiaomin Li (Jilin University), Hao Xing (Jilin University), Hao Xu (Jilin University)
Experience: Interactions and Quality of Experience		mmfp0606	Virtual	Quality Assessment of Image Super-Resolution: Balancing Deterministic and Statistical Fidelity -- Wei Zhou (University of Waterloo), Zhou Wang (University of Waterloo)
		mmfp1748	Virtual	No-reference Omnidirectional Image Quality Assessment Based on Joint Network -- C. Zhang (Tianjin University), Shiguang Liu (Tianjin University)
		mmfp2099	On-site	PassWalk: Spatial Authentication Leveraging Lateral Shift and Gaze on Mobile Headsets -- Abhishek Kumar (University of Helsinki), LIK-HANG LEE (Korea Advanced Institute of Science & Technology), J Ch (University of Southampton), Xiang Su (Norwegian University of Science and Technology), Mohammad Hoque (University of Helsinki), Susanna Pirttikangas (University of Oulu), Sasu Tarkoma (University of Helsinki), Pan Hui (The Hong Kong University of Science and Technology)
		mmfp2630	Virtual	Adaptive Hypergraph Convolutional Network for No-Reference 360-degree Image Quality Assessment -- Jun Fu (University of Science and Technology of China), Chen Hou (University of Science and Technology of China), Wei Zhou (University of Waterloo), Jiahua Xu (University of Science and Technology of China), Zhibo Chen (University of Science and Technology of China, Tsinghua University)
		mmfp1852	On-site	DeepWSD: Projecting Degradations in Perceptual Space to Wasserstein Distance in Deep Feature Space -- Xingran Liao (City University of Hong Kong), Baoliang Chen (City University of Hong Kong), Hanwei Zhu (City University of Hong Kong), Shiqi Wang (City University of Hong Kong), Mingliang Zhou (City University of Hong Kong), Sam Kwong (City University of Hong Kong)
Multimedia Systems: Transport and Delivery		mmfp1753	Virtual	PicT: A Slim Weakly Supervised Vision Transformer for Pavement Distress Classification -- Wenhao Tang (Chongqing University), Sheng Huang (Chongqing University), Zhang Xiaoxian (Chongqing University), Luwen Huangfu (Center for Human Dynamics in the Mobile Age)
		mmfp1900	Virtual	Rate-Distortion-Guided Learning Approach with Cross-Projection Information for V-PCC Fast CU Decision -- Hang Yuan (Peking University), Wei Gao (Peking University Shenzhen Graduate School), Ge Li (Peking University Shenzhen Graduate School), Zhu Li (University of Missouri-Kansas City)
		mmfp1927	On-site	Evaluating the Impact of Tiled User-Adaptive Real-Time Point Cloud Streaming on VR Remote Communication -- Shishir Subramanyam (Centrum voor Wiskunde en Informatica), Irene Viola (Centrum voor Wiskunde en Informatica), Jack Jansen (Centrum voor Wiskunde en Informatica), Evangelos Alexiou (Centrum voor Wiskunde en Informatica), Alan Hanjalic (Delft University of Technology), Pablo Cesar (Delft University of Technology)
		mmfp0460	On-site	Prism: Handling packet loss for ultra-low latency video. -- Devdeep Ray (Carnegie Mellon University), Vicente Bobadilla Riquelme (Carleton College), Srinivasan Seshan (Computer Science Department, School of Computer Science)
Multimedia Systems: Data Systems Management and Indexing		mmfp2752	Virtual	Online Deep Learning from Doubly-Streaming Data -- Heng Lian (Old Dominion University), John Atwood (Old Dominion University), Bo-Jian Hou (University of Pennsylvania), Jian Wu (Old Dominion University), Yi He (Old Dominion University)
Multimedia Systems: Systems and Middleware		mmfp2088	Virtual	Display of 3D Illuminations using Flying Light Specks -- Shahram Ghandeharizadeh (University of Southern California)
Demo Session 2 [Main Foyer]	**10h30-11h45**			*(Demos might be available the whole day)*

Cultural heritage, Art, and Aesthetics		mmde014	On-site	Mediascape XR: A cultural heritage experience in Social VR -- Ignacio Reimat (UPC - Barcelona School of Informatics, Universidad Politécnica de Cataluna), Yanni Mei (Centrum voor Wiskunde en Informatica), Evangelos Alexiou (Centrum voor Wiskunde en Informatica), Jack Jansen (Centrum voor Wiskunde en Informatica), Jie Li (Centrum voor Wiskunde en Informatica), Shishir Subramanyam (Centrum voor Wiskunde en Informatica), Irene Viola (Centrum voor Wiskunde en Informatica), Johan Oomen (N/A), Pablo Cesar (Delft University of Technology)
		mmde019	Virtual	AI Carpet: Automatic Generation of Aesthetic Carpet Pattern -- Ziyi Wang (Tsinghua University, Tsinghua University), Xingqi Wang (Tsinghua University, Tsinghua University), Zeyu Jin (Tsinghua University, Tsinghua University), Xiaohan Li (Tsinghua University, Tsinghua University), Shikun Sun (Tsinghua University, Tsinghua University), Jia Jia (Tsinghua University)
		mmde026	Virtual	Attribute Controllable Beautiful Caucasian Face Generation by Aesthetics Driven Reinforcement Learning -- Xin Jin (Beijing Electronic Science and Technology Institute), Shu Zhao (Beijing Electronic Science and Technology Institute), Le Zhang (Tsinghua University, Tsinghua University), Xin Zhao (Tsinghua University), Deng Qiang (Beijing Electronic Science and Technology Institute), Chaoen Xiao (Beijing Electronic Science and Technology Institute)
		mmde027	On-site	An AI Powered Re-Identification System for Real-time Contextual Multimedia Applications -- Giuseppe Becchi (University of Florence), Andrea Ferracani (MICC - University of Florence), Filippo Principi (MICC - University of Florence), Alberto Del Bimbo (MICC - University of Florence)
		mmde036	On-site	ALEGORIA: joint multimodal search and spatial navigation into the geographic iconographic heritage -- Florent Geniet (IGN), Valerie Gouet-Brunet (LaSTIG / IGN, Univ. Gustave Eiffel), Mathieu Brédif (IGN)
		mmde045	On-site	Engaging museum visitors with gamification of body and facial expressions -- Marco Bertini (Università degli Studi di Firenze), Maria Donadio (Università degli Studi di Firenze), Filippo Principi (MICC - University of Florence, IT), Andrea Ferracani (MICC - University of Florence, IT), Alberto Del Bimbo (Università degli Studi di Firenze)
		mmde039	On-site	GetWild: A VR Editing System with AI-Generated 3D Object and Terrain -- Shing Ming Wong (National Tsinghua University), Chien-Wen Chen (National Cheng Kung University), Tse-Yu Pan (National Tsing Hua University), Hung-Kuo Chu (National Tsing Hua University), Min-Chun Hu (Department of Computer Science, National Tsing Hua University, National Tsing Hua University)
Art Exhibition	**All day**			*(See Art Exhibition on Tuesday)*
Grand Challenges Session II [Pav. 4 R1.07]	**11h-12h15**			
Short Video Streaming Challenge	11h	mmgc20	Virtual	Bandwidth-Efficient Multi-video Prefetching for Short Video Streaming -- Xutong Zuo (Computer Science, Tsinghua University), Yishu Li (Tsinghua University), Mohan Xu (Tsinghua University), Wei Ooi (National University of Singapore), Jiangchuan Liu (Simon Fraser University), Junchen Jiang (University of Chicago), Xinggong ZHANG (Peking University), Kai Zheng (Tsinghua University, Tsinghua University), Yong Cui (Tsinghua University, Tsinghua University)
	11h06	mmgc05	Virtual	PDAS: Probability-Driven Adaptive Streaming for Short Video -- Chao Zhou (Peking University), Yixuan Ban (Peking University), Yangchao Zhao (Nanjing University), Dan Yang (Kuaishou), Liang Guo (Kuaishou), Bing Yu (Kuaishou)
	11h12	mmgc31	Virtual	QoE-aware Download Control and Bitrate Adaptation for Short Video Streaming -- Ximing Wu (Department of Software Engineering, Shenzhen University), Lei Zhang (Shenzhen University), Laizhong Cui (Shenzhen University)
Pre-training for Video Understanding Challenge	11h18	mmgc17	Virtual	Auto-captions on GIF: A Large-scale Video-sentence Dataset for Vision-language Pre-training -- Yingwei Pan (JD AI Research), Yehao Li (SUN YAT-SEN UNIVERSITY), Jianjie Luo (SUN YAT-SEN UNIVERSITY), Jun Xu (University of Science and Technology of China), Ting Yao (JD AI Research), Tao Mei (JD Explore Academy)
	11h24	mmgc22	On-site	Multiple Temporal Fusion based Weakly-supervised Pre-training Techniques for Video Categorization -- Xiaochen Cai (Nanjing University), Hengxing Cai (SUN YAT-SEN UNIVERSITY), Boqing Zhu (National University of Defense Technology), Kele XU (Computer Science Lab - Pierre and Marie Curie University, Paris, France), Wei-Wei Tu (4Paradigm Inc.), Dawei Feng (National University of Defense Technology)
MultiMediate challenge	11h30	mmgc28	On-site	MultiMediate'22: Backchannel Detection and Agreement Estimation in Group Interactions -- Philipp Müller (German Research Center for AI), Michael Dietz (Universität Augsburg), Dominik Schiller (Universität Augsburg), Dominike Thomas (University of Stuttgart, Universität Stuttgart), Hali Lindsay (German Research Center for AI), Patrick Gebhard (German Research Center for AI), Elisabeth Andre (University of Augsburg), Andreas Bulling (University of Stuttgart)
	11h36	mmgc26	Virtual	TA-CNN: A Unified Network for Human Behavior Analysis in Multi-Person Conversations -- Fuyan Ma (Hunan University), Ziyu Ma (Hunan University), Bin Sun (Hunan University), Shutao Li (Hunan University)
Detecting CheapFakes	11h42	mmgc45	Virtual	A Textual-Visual-Entailment-based Unsupervised Algorithm for Cheapfake Detection -- Tien Tran (Ho Chi Minh city University of Science, Vietnam National University), Phuc Tran (Ho Chi Minh city University of Science, Vietnam National University), Minh-Son Dao (National Institute of Information and Communications Technology (NICT)), La Vinh (University of Information and Technology at Viet Nam), Anh-Duy Tran (VNUHCM-University of Science), Duc Tien Dang Nguyen (University of Bergen)
	11h48	mmgc44	On-site	A Combination of Visual-Semantic Reasoning and Text Entailment-based Boosting Algorithm for Cheapfake Detection -- La Vinh (University of Information and Technology at Viet Nam), Minh-Son Dao (National Institute of Information and Communications Technology (NICT)), Tran Tien (Ho Chi Minh city University of Science, Vietnam National University), Phuc Tran (Ho Chi Minh city University of Science, Vietnam National University), Anh-Duy Tran (VNUHCM-University of Science), Duc Tien Dang Nguyen (University of Bergen)
Conversational Head Generation	11h54	mmgc08	Virtual	Audio-driven Talking Head Generation with Transformer and 3D Morphable Model -- Ricong Huang (SUN YAT-SEN UNIVERSITY), Weizhi Zhong (SUN YAT-SEN UNIVERSITY), Guanbin Li (SUN YAT-SEN UNIVERSITY)
	12h	mmgc16		Semantic-aware Responsive Listener Head Synthesis -- Wei Zhao (Hunan University), Peng Xiao (Hunan University), Rongju Zhang (Hunan University), Yijun Wang (Hunan University), Jianxin Lin (Hunan University, Tsinghua University)
Grand Challenge Poster Session [Pav. 4 - Hall]	**12h15-13h**			
Computational Paralinguistics ChallengE		mmgc14	Virtual	Masked Modeling-based Audio Representation for ACM Multimedia 2022 Computational Paralinguistics ChallengE -- kang you (Tongji University), Kele XU (Computer Science Lab - Pierre and Marie Curie University, Paris, France), Boqing Zhu (National University of Defense Technology), Ming Feng (Tongji University), Dawei Feng (National University of Defense Technology), Bo Liu (National University of Defense Technology), Tian Gao (iFlytek), Bo Ding (National University of Defense Technology)
		mmgc27	On-site	End-to-End and Self-Supervised Learning for ComParE 2022 Stuttering Sub-Challenge -- Shakeel Sheikh (INRIA), Md Sahidullah (Indian Institute of Technology Kharagpur), Fabrice Hirsch (INRIA), Slim Ouni (Loria)
		mmgc40	On-site	How Much Attention Should we Pay to Mosquitoes? -- Moreno La Quatra (Politecnico di Torino), Lorenzo Vaiani (Polytechnic Institute of Turin), Alkis Koudounas (Polytechnic Institute of Turin), Luca Cagliero (Politecnico di Torino), Paolo Garza (Politecnico di Torino), Elena Baralis (Polytechnic Institute of Turin)
		mmgc48	Virtual	A Transformer Based Approach for Activity Detection -- Gulshan Sharma (Indian Institute of Technology Ropar), Abhinav Dhall (Monash University), Ramanathan Subramanian (University of Canberra)
		mmgc60	On-site	Audio Features from the Wav2Vec 2.0 Embeddings for the ACM Multimedia 2022 Stuttering Challenge -- Claude Montacié (STIH Sens Texte Informatique Histoire), Marie-José Caraty (Université de Paris), Nikola LACKOVIC (Stih)
Conversational Head Generation		mmgc03	On-site	A Baseline for ViCo Conversational Head Generation Challenge -- Meng Liu (Shandong Jianzhu University), Shuyan Zhai (Shandong University), Yongqiang Li (Shandong University), WEILI GUAN (Monash University), Liqiang Nie (Harbin Institute of Technology (Shenzhen))
		mmgc11	Virtual	Perceptual Conversational Head Generation with Regularized Driver and Enhanced Renderer -- Ailin Huang (Megvii Technology Inc.), Zhewei Huang (Megvii Technology Inc.), Shuchang Zhou (Chinese Academy of Sciences)
		mmgc19	On-site	Generating Smooth and Facial-Details-Enhanced Talking Head Video: A Perspective of Pre and Post Processes -- Tian Lv (Tsinghua University, Tsinghua University), Yuhui Wen (Tsinghua University, Tsinghua University), Zhiyao Sun (Tsinghua University, Tsinghua University), Zipeng Ye (Tsinghua University, Tsinghua University), Yong-jin Liu (Tsinghua University, Tsinghua University)
Deep Video Understanding Challenge		mmgc01	On-site	A Multi-Stream Approach for Video Understanding -- Lutharsanen Kunam (University of Zurich), Luca Rossetto (Department of Informatics, University of Zurich), Abraham Bernstein (Department of Informatics, University of Zurich)
		mmgc09	Virtual	Two stage Multi-Modal Modeling for Video Interaction Analysis in Deep Video Understanding Challenge -- Sun Siyang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Xiong Xiong (Alibaba Group), Yun Zheng (SUN YAT-SEN UNIVERSITY)
		mmgc58	Virtual	Unified QA-aware Knowledge Graph Generation Based on Multi-modal Modeling -- Penggang Qin (University of Science and Technology of China), Jiarui Yu (University of Science and Technology of China), Yan Gao (University of Science and Technology of China), Derong Xu (University of Science and Technology of China), Yunkai Chen (University of Science and Technology of China), Shiwei Wu (University of Science and Technology of China), Tong Xu (University of Science and Technology of China), Yanbin Hao (University of Science and Technology of China), Enhong Chen (University of Science and Technology of China)
		mmgc65	Virtual	Leveraging Text Representation and Face-head Tracking for Long-form Multimodal Semantic Relation Understanding -- Raksha Ramesh (Columbia University), Vishal Anand (Microsoft), Zifan Chen (Columbia University), Yifei Dong (Columbia University), Yun Chen (Graphen), Ching-Yung Lin (Graphen)
Detecting CheapFakes		mmgc56	On-site	Sentiment-aware Classifier for Out-of-Context Caption Detection -- Muhannad Alkaddour (American University of Sharjah), Abhinav Dhall (Monash University), Usman Tariq (American University of Sharjah), Hasan Al Nashash (American University of Sharjah), Fares Al-Shargie (American University of Sharjah)
Facial Micro-Expression Grand Challenge		mmgc04	On-site	3D-CNN for Facial Micro- and Macro-expression Spotting on Long Video Sequences using Temporal Oriented Reference Frame -- Chuin Hong Yap (The Manchester Metropolitan University), Moi Hoon Yap (Manchester Metropolitan University), Adrian Davison (University of Manchester), Connah Kendrick (Manchester Metropoltian University), Jingting Li (Institute of Psychology, Chinese Academy of Sciences), Su-Jing Wang (Institute of Psychology, Chinese Academy of Sciences), Ryan Cunningham (Manchester Metropolitan University)
		mmgc36	On-site	Adaptive Dual Motion Model for Facial Micro-Expression Generation -- Xinqi Fan (City University of Hong Kong), Ali Shahid (City University of Hong Kong), Hong Yan (City University of Hong Kong)
		mmgc46	Virtual	Fine-grained Micro-Expression Generation based on Thin-Plate Spline and Relative AU Constraint -- Sirui Zhao (University of Science and Technology of China), Shukang Yin (University of Science and Technology of China), Huaying Tang (University of Science and Technology of China), Jin Rijin (University of Science and Technology of China), Yifan Xu (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)
		mmgc51	On-site	ABPN: Apex and Boundary Perception Network for Micro- and Macro-Expression Spotting -- Wenhao Leng (University of Science and Technology of China), Sirui Zhao (University of Science and Technology of China), Yiming Zhang (University of Science and Technology of China), Shiifeng Liu (University of Science and Technology of China), Xinglong Mao (University of Science and Technology of China), Hao Wang (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)
		mmgc55	On-site	Rethinking Optical Flow Methods for Micro-Expression Spotting -- Yuan Zhao (Chongqing University of Technology), Xin Tong (Hubei University of Technology), Zhu Chong (Central South University), Jianda Sheng (Pingan Technology), Lei Dai (Pingan Technology), Ling Xuling (Pingan Technology), Xuehai Xia (Pingan Technology), Jian Yu (Pingan Technology), Jiao Li (Pingan Technology)
MultiMediate challenge		mmgc59	On-site	Graph-based Group Modelling for Backchannel Detection -- Garima Sharma (Monash University), Kalin Stefanov (University of Southern California), Abhinav Dhall (Monash University), Jianfei Cai (Monash University)
Short Video Streaming Challenge		mmgc07	Virtual	DAM: Deep Reinforcement Learning based Preload Algorithm with Action Masking for Short Video Streaming -- Si-Ze Qian (Communication University of China), Yuhong Xie (Communication University of China), Zipeng Pan (Communication University of China), Yuan Zhang (Communication University of China), TAO LIN (Communication University of China)
Social Media Prediction (SMP) Challenge		mmgc10	Virtual	Deeply Exploit Visual and Language Information for Social Media Popularity Prediction -- wu Min (Northwest Polytechnical University Xi'an), Liming Zhao (Alibaba Group), Dangwei Li (Alibaba Group), Chen-Wei Xie (Alibaba Group), Sun Siyang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Yun Zheng (SUN YAT-SEN UNIVERSITY)
		mmgc37	Virtual	A Comprehensive Study of Spatiotemporal Feature Learning for Social Medial Popularity Prediction -- Chih-Chung Hsu (National Cheng Kung University), Pi-Ju Tsai (National Cheng Kung University), Ting-Chun Yeh (National Cheng Kung University), Xiu-Yu Hou (National University of Tainan)
Oral Session 4a [Auditório I]	**11h45-13h**	**Chair: Mea Wang (University of Calgary)**		
Recommendation and Retrieval	11h45	mmfp0929	Virtual	Partially Relevant Video Retrieval -- Jianfeng Dong (Zhejiang Gongshang University), Xianke Chen (Zhejiang Gongshang University), Z MS (Zhejiang Gongshang University), Xun Yang (University of Science and Technology of China), Shujie Chen (Zhejiang Gongshang University), Xirong Li (Renmin University of China), Xun Wang (Zhejiang Gongshang University)
	12h03	mmfp2829	Virtual	From Abstract to Details: A Generative Multimodal Fusion Framework for Recommendation -- Fangxiong Xiao (Institute of Computing Technology, Chinese Academy of Sciences), Lixi deng (Institute of Computing Technology, CAS), Jingjing Chen (Fudan University), Houye Ji (JD.COM), Xiaorui Yang (jd), Zhuoye Ding (Fudan University), Bo Long (State University of New York, Binghamton)
	12h21	mmfp1149	Virtual	Bi-directional Heterogeneous Graph Hashing towards Efficient Outfit Recommendation -- WEILI GUAN (Monash University), Songxuemeng Song (Shandong University), Haoyu Zhang (Shandong University), Meng Liu (Shandong Jianzhu University), Chung-Hsing Yeh (Monash University), Xiaojun Chang (University of Technology Sydney)
	12h39	mmfp3225	Virtual	DVR: Micro-Video Recommendation Optimizing Watch-Time-Gain under Duration Bias -- Yu Zheng (Electronic Engineering, Tsinghua University, Tsinghua University), Chen Gao (Tsinghua University, Tsinghua University), Jingtao Ding (Electronic Engineering, Tsinghua University), LINGLING YI (Tencent), Depeng Jin (Tsinghua University, Tsinghua University), Yong Li (Tsinghua University, Tsinghua University), Meng Wang (Hefei University of Technology)

BNI Session [Auditório II]	**11h45-13h15**	**Chairs: Hayley Hung (TU Delft) and Chong-Wah Ngo (Singapore Management University)**		
Brave New Ideas	11h45	mmbni17	On-site	Compute to Tell the Tale: Goal-Driven Narrative Generation -- Yongkang Wong (National University of Singapore), Shaojing Fan (National University of Singapore), Yangyang Guo (National University of Singapore), Ziwei Xu (National University of Singaore), Rishabh Sheoran (national university of singapore, National University of Singapore), Karen Stephen (NEC), Anusha Bhamidipati (NEC), Vivek Barsopia (Indian Institute of Technology, Bombay, Dhirubhai Ambani Institute Of Information and Communication Technology), Jianquan Liu (NEC Corporation), Mohan Kankanhalli (National University of Singapore)
	12h03	mmbni25	On-site	Demographic Feature Isolation for Bias Research Using Deepfakes -- Kurtis Haut (Department of Computer Science, University of Rochester), Caleb Wohn (University of Rochester), Victor Antony (University of Rochester), Aidan Goldfarb (University of Rochester), Melissa Welsh (University of Rochester), Dillanie Sumanthiran (University of Rochester), Mohammed Rafayet Ali (University of Rochester Department of Computer Science), Mohammed Hoque (University of Rochester)
	12h21	mmbni21	Virtual	Benign Adversarial Attack: Tricking Models for Goodness -- Jitao Sang (Beijing Jiaotong University), Xian Zhao (Beijing Jiaotong University), Jiaming Zhang (Beijing Jiaotong University), Zhiyu Lin (Beijing Jiaotong University)
	12h39	mmbni38	On-site	Recipe-oriented Food Logging for Nutritional Management -- Yoko Yamakata (The University of Tokyo), Akihisa Ishino (The University of Tokyo), Akiko Sunto (Kanagawa University of Human Services), Sosuke Amano (foo.log Inc.), Kiyoharu Aizawa (The University of Tokyo)
	12h57	mmbni02	Virtual	Can Language Understand Depth? -- Renrui Zhang (MMLab of CUHK & Shanghai AI Laboratory), Ziyao Zeng (ShanghaiTech University), Ziyu Guo (Peking University)
Oral Session 4c [Auditório III]	**11h45-13h**	**Chair: Yinqiang Zheng (The University of Tokyo)**		
Multimedia 3D Processing	11h45	mmfp0536	Virtual	LVI-ExC: A Target-free LiDAR-Visual-Inertial Extrinsic Calibration Framework -- Zhong Wang (Tongji University), Lin Zhang (Tongji University), Ying Shen (Tongji University), Yicong Zhou (University of Macau)
	12h03	mmfp1232	Virtual	Cross-Domain 3D Model Retrieval Based On Contrastive Learning And Label Propagation -- Dan Song (Tianjin University), Yue Yang (Tianjin University), Weizhi Nie (Tianjin University), Xuanya Li (Baidu), Anan Liu (university of tianjin, china)
	12h21	mmfp1653	Virtual	HMTN: Hierarchical Multi-scale Transformer Network for 3D Shape Recognition -- Yue Zhao (Tianjin University), Weizhi Nie (Tianjin University), Zan Gao (Qilu University of Technology), Anan Liu (Tianjin University)
	12h39	mmfp0587	Virtual	MAPLE: Masked Pseudo-Labeling autoEncoder for Semi-supervised Point Cloud Action Recognition -- Xiaodong Chen (University of Science and Technology of China), Wu Liu (JD Explore Academy), Xinchen Liu (JD Explore Academy), Yongdong Zhang (University of Science and Technology of China), Jungong Han (Aberystwyth University), Tao Mei (JD Explore Academy)
Oral Session 4d [Auditório IV]	**11h45-13h**	**Chair: Lin Wu (Hefei University of Technology)**		
Scene Understanding and Experience I	11h45	mmfp2443	Virtual	SIM-Trans: Structure Information Modeling Transformer for Fine-grained Visual Categorization -- Hongbo Sun (Peking University), Xiangteng He (Peking University), Yuxin Peng (Peking University)
	12h03	mmfp1707	Virtual	Rethinking Open-World Object Detection in Autonomous Driving Scenarios -- Zeyu Ma (University of Electronic Science and Technology of China), Yang Yang (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), Hengtao Shen (University of Electronic Science and Technology of China), Mingxing Zhang (University of Electronic Science and Technology of China)
	12h21	mmfp1342	Virtual	MVLayoutNet: 3D layout reconstruction with multi-view panoramas -- Zhihua Hu (Wuhan University), Bo Duan (Huawei Technologies Ltd.), Yanfeng Zhang (Huawei Technologies Ltd.), mingwei Sun (Huawei Technologies Ltd.), Jingwei Huang (Stanford University)
	12h39	mmfp2205	Virtual	Decoupling Recognition from Detection: Single Shot Self-Reliant Scene Text Spotter -- Jingjing Wu (Harbin Institute of Technology), Pengyuan Lyu (Baidu), Guangming Lu (Harbin Institute of Technology), Chengquan Zhang (Baidu), Kun Yao (Baidu), Wenjie Pei (Harbin Institute of Technology, Shenzhen)
Virtual Only Poster Session 2 [Gathertown]	**13h-14h30**			*(See Poster Session 2)*
Keynote Talk: Yoelle Maarek [Auditório I]	**14h30-15h30**			Alexa, let's work together! How Alexa helps customers complete tasks with verbal and visual guidance in the Alexa Prize TaskBot Challenge -- Yoelle Maarek
Technical Achievement Award [Auditório I]	**15h30-16h15**	**Chair: Phoebe Chen (La Trobe University)**		Where should we go with Multimedia ? -- Alan Smeaton
Rising Star Award [Auditório I]	**16h15-16h45**			Large-scale Multimodal Representation Learning for Image, Video, and Embodied Agent -- Jianlong Fu
Poster Session 2bis [Pav. 4 - Hall]	**15h30-16h45**			*(See Poster Session 2)*
Demo Session 2bis [Main Foyer]	**15h30-16h45**			*(See Demo Session 2)*
Oral Session 5a [Auditório I]	**16h45-18h**	**Chair: Xavier Alameda-Pineda (Inria)**		
Top Papers Session II	16h45	mmfp3095	On-site	MM-ALT: A Multimodal Automatic Lyric Transcription System -- Xiangming Gu (National University of Singapore), Longshen Ou (National University of Singapore), Danielle Ong (national university of singaore, National University of Singapore), Ye Wang (National University of Singapore)
	17h03	mmfp0493	Virtual	Semi-supervised Crowd Counting via Density Agency -- Hui LIN (Xi'an Jiaotong University), Zhiheng Ma (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Chinese Academy of Sciences), Xiaopeng Hong (Harbin Institute of Technology), Yaowei Wang (Pengcheng Laboratory), su zhou (Xi'an Jiaotong University)
	17h21	mmfp0079	Virtual	Im2Oil: Stroke-Based Oil Painting Rendering with Linearly Controllable Fineness Via Adaptive Sampling -- Zhengyan Tong (Shanghai Jiaotong University), Xiaohang Wang (Shanghai Jiaotong University), Shengchao Yuan (Shanghai Jiaotong University), Xuanhong Chen (Shanghai Jiao Tong University), Junjie Wang (Shanghai Jiaotong University), Xiangzhong Fang (Shanghai Jiao Tong University)
	17h39	mmfp0235	Virtual	CACOLIT: Cross-domain Adaptive Co-learning for Imbalanced Image-to-Image Translation -- Yijun Wang (Hunan University), Tao Liang (University College London, University of London), Jianxin Lin (Hunan University, Tsinghua University)
Oral Session 5b [Auditório II]	**16h45-18h**	**Chair: Silvia Rossi (Centrum Wiskunde & Informatica)**		
Advanced Learning for Multimedia I	16h45	mmfp1539	Virtual	Difference Residual Graph Neural Networks -- Liang Yang (Hebei University of Technology), Weihang Peng (Hebei University of Technology), Zhou Miao (Hebei University of Technology,), bingxin niu (Hebei University of Techonology), Junhua Gu (Hebei University of Technology), Chuan Wang (institute of information engineering), Yuanfang Guo (Beihang University), Dongxiao He (Jilin University, China), Xiaochun Cao (Institute of Information Engineering)
	17h03	mmfp0659	On-site	Long-term Leap Attention, Short-term Periodic Shift for Video Classification -- Hao Zhang (City University of Hong Kong), Lechao Cheng (Zhejiang Lab), Yanbin Hao (University of Science and Technology of China), Chong-wah Ngo (City University of Hong Kong)
	17h21	mmfp2536	Virtual	Dynamic Scene Graph Generation via Temporal Prior Inference -- Wang Shuang (University of Electronic Science and Technology of China), Xinyu Lyu (University of Electronic Science and Technology of China), Yuyu Guo (University of Electronic Science and Technology of China), Pengpeng Zeng (University of Electronic Science and Technology of China), Jingkuan Song (University of Electronic Science and Technology of China, Tsinghua University), Lianli Gao (University of Electronic Science and Technology of China, Tsinghua University)
	17h39	mmfp1684	Virtual	Confederated Learning: Going Beyond Centralization -- Zitai Wang (University of the Chinese Academy of Sciences), Qianqian Xu (Institute of Computing Technology, Chinese Academy of Sciences), Ke Ma (University of Chinese Academy of Sciences), Xiaochun Cao (Institute of Information Engineering), Qingming Huang (University of Chinese Academy of Sciences)
Oral Session 5c [Auditório III]	**16h45-18h**	**Chair: Michael Riegler (Simula)**		
Image Processing and Enhancement I	16h45	mmfp0129	Virtual	ELMformer: Efficient Raw Image Restoration with a Locally Multiplicative Transformer -- Jiaqi Ma (Wuhan University), Shengyuan Yan (Wuhan University), Lefei Zhang (Wuhan University), Guoli Wang (Tsinghua University, Tsinghua University), Qian Zhang (Horizon Robotics)
	17h03	mmfp0023	Virtual	JPEG Compression-aware Image Forgery Localization -- Menglu Wang (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), Jiawei Liu (University of Science and Technology of China), Zheng-Jun Zha (University of Science and Technology of China)
	17h21	mmfp2196	Virtual	Delving Globally into Texture and Structure for Image Inpainting -- Haipeng Liu (Hefei University of Technology), Yang Wang (Hefei University of Technology), Meng Wang (Hefei University of Technology), Yong Rui (Lenovo)
	17h39	mmfp0083	Virtual	Magic ELF: Image Deraining Meets Association Learning and Transformer -- Kui Jiang (Wuhan University), Zhongyuan Wang (Wuhan University), Chen Chen (University of Central Florida), Zheng Wang (Wuhan University, China), Laizhong Cui (Shenzhen University), Chia-Wen Lin (National Tsing Hua University)
Oral Session 5d [Auditório IV]	**16h45-18h**	**Chair: Yuewei Lin (Brookhaven National Laboratory)**		
Scene Understanding and Experience II	16h45	mmfp1581	Virtual	Factorized and Controllable Neural Re-Rendering of Outdoor Scene for Photo Extrapolation -- Boming Zhao (Zhejiang University), Bangbang Yang (Zhejiang University), Zhenyang Li (Baidu), Zuoyue Li (ETH Zürich), Guofeng Zhang (Zhejiang University), Jiashu Zhao (Wilfrid Laurier University), Dawei Yin (Lehigh University), Zhaopeng Cui (Zhejiang University), Hujun Bao (Zhejiang University)
	17h03	mmfp1428	Virtual	Skeleton2Humanoid: Animating Simulated Characters for Physically-plausible Motion In-betweening -- Yunhao Li (Shanghai Jiao Tong University), Zhenbo Yu (Shanghai Jiao Tong University), Yucheng Zhu (Shanghai Jiaotong University), Bingbing Ni (Shanghai Jiao Tong University), Guangtao Zhai (Shanghai Jiao Tong University), Wei Shen (Shanghai Jiao Tong University)
	17h21	mmfp2719	On-site	Multi-view Layout Design for VR Concert Experience -- Minju Kim (Korea Institute of Science and Technology Information), Yuhyun Lee (Soongsil University), Jungjin Lee (Soongsil University)
	17h39	mmfp0283	Virtual	AggCast: Practical Cost-effective Scheduling for Large-scale Cloud-edge Crowdsourced Live Streaming -- Ruixiao Zhang (Tsinghua University, Tsinghua University), changpeng yang (Huawei Cloud), Xiaochan Wang (Tsinghua University, Tsinghua University), Tianchi Huang (Tsinghua University), Chenglei Wu (Tsinghua University, Tsinghua University), Jiangchuan Liu (Simon Fraser University), Lifeng Sun (Tsinghua University, Tsinghua University)
Grand Challenges Poster Session bis	**16h45-18h**			*(See Grand Challenges Poster Session)*
CCB	**19h30**	**Social Event - Banquet**		

Thursday, October 13th (Main Conference)

Oral Session 6a [Auditório I]	**9h15-10h30**	**Chair: Vincent Oria (New Jersey Institute of Technology)**		
Text2Image	9h15	mmfp3195	Virtual	A Tree-Based Structure-Aware Transformer Decoder for Image-To-Markup Generation -- Shuhan Zhong (Department of Computer Science and Engineering, Hong Kong University of Science and Technology), Sizhe Song (The Hong Kong University of Science and Technology), Guanyao Li (Department of Computer Science and Engineering, Hong Kong University of Science and Technology), S.-H. Chan (The Hong Kong University of Science and Technology)
	9h33	mmfp0236	Virtual	AI Illustrator: Translating Raw Descriptions into Images by Prompt-based Cross-Modal Generation -- Yiyang Ma (Microsoft), Huan Yang (Microsoft), Bei Liu (Microsoft Research Asia), Jianlong Fu (Microsoft), Jiaying Liu (Peking University)
	9h51	mmfp1680	Virtual	Background Layout Generation and Object Knowledge Transfer for Text-to-Image Generation -- Zhuowei Chen (University of Science and Technology of China), Zhendong Mao (University of Science and Technology of China), Shancheng Fang (University of Science and Technology of China), Bo Hu (University of Science and Technology of China)

	10h09	mmfp0306	Virtual	Synthesizing Counterfactual Samples for Effective Image-Text Matching -- Hao Wei (, Chinese Academy of Sciences), Shuhui Wang (Institute of Computing Technology, Chinese Academy of Sciences), XINZHE HAN (, Chinese Academy of Sciences), Zhe Xue (Beijing University of Posts and Telecommunications), Bin MA (Meituan), Wei Xiaoming (Beijing University of Aeronautics and Astronautics), Xiaolin Wei (Meituan)
Oral Session 6b [Auditório II]	**9h15-10h30**	**Chair: Marcel Worring (University of Amsterdam)**		
Multimodal Domain Adaptation	9h15	mmfp0489	Virtual	Semantic Data Augmentation based Distance Metric Learning for Domain Generalization -- Mengzhu Wang (Alibaba Group), Jianlong Yuan (Alibaba Group), Qi Qian (Alibaba Group), Zhibin Wang (Alibaba Group), Li Hao (Alibaba Group)
	9h33	mmfp2468	Virtual	Mix-DANN and Dynamic-Modal-Distillation for Video Domain Adaptation -- Yuehao Yin (Fudan University), Bin Zhu (City University of Hong Kong), Jingjing Chen (Fudan University), Lechao Cheng (Zhejiang Lab), Yu-Gang Jiang (Fudan University)
	9h51	mmfp1677	Virtual	Source-Free Active Domain Adaptation via Energy-Based Locality Preserving Transfer -- Xinyao Li (University of Electronic Science and Technology of China), Zhekai Du (University of Electronic Science and Technology of China), Jingjing Li (University of Electronic Science and Technology of China), Lei Zhu (Shandong Normal University), Ke Lu (University of Electronic Science and Technology of China)
	10h09	mmfp1660	Virtual	Class Discriminative Adversarial Learning for Unsupervised Domain Adaptation -- Lihua Zhou (University of Electronic Science and Technology of China), Mao Ye (University of Electronic Science and Technology of China), Xiatian Zhu (University of Surrey), Li Shuaifeng (University of Electronic Science and Technology of China), Yiguang Liu (Sichuan University)
Oral Session 6c [Auditório III]	**9h15-10h30**	**Chair: Paulo Jorge Lourenço Nunes (Instituto Universitário de Lisboa)**		
Segmentation for Experience	9h15	mmfp3035	Virtual	Leveraging GAN Priors for Few-Shot Part Segmentation -- Mengya Han (Wuhan University), Heliang Zheng (USTC), Chaoyue Wang (JD Explore Academy), Yong Luo (Wuhan University), Han Hu (Beijing Institute of Technology), Bo Du (Wuhan University)
	9h33	mmfp0560	Virtual	MaMiCo: Macro-to-Micro Semantic Correspondence for Self-supervised Video Representation Learning -- Bo Fang (Institute Information of Engineer), Wenhao Wu (University of Sydney), Chang Liu (Tsinghua University, Tsinghua University), Yu Zhou (Institute of Information Engineering, CAS), Dongliang He (Baidu), Weipinng Wang (IIE)
	9h51	mmfp1959	Virtual	Bidirectional Self-Training with Multiple Anisotropic Prototypes for Domain Adaptive Semantic Segmentation -- Yulei Lu (Zhejiang University), Yawei Luo (Zhejiang University, Tsinghua University), Li Zhang (Jilin University), Zheyang Li (Shanghai Jiao Tong University), Yi Yang (Zhejiang University), Jun Xiao (Zhejiang University)
	10h09	mmfp0270	Virtual	KnifeCut: Refining Thin Part Segmentation with Cutting Lines -- Zheng Lin (Nankai University), Zheng-Peng Duan (Xi'an University of Electronic Science and Technology), Zhao Zhang (Nankai University), Chunle Guo (Nankai University), Ming-ming Cheng (Nankai University, Tsinghua University)
Oral Session 6d [Auditório IV]	**9h15-10h30**	**Chair: Chong-Wah Ngo (Singapore Management University)**		
Privacy and Secutiry in Multimedia	9h15	mmfp0280	Virtual	Profiling Channel Variance and Spatial Gradient for Robust Face Forgery Detection -- Luchuan Song (University of Science and Technology of China), Xiaodan Li (University of Science and Technology of China), Zheng Fang (Beijing University of Aeronautics and Astronautics), Jin Zhenchao (University of Science and Technology of China), YueFeng Chen (Alibaba Group), Chenliang Xu (University of Rochester)
	9h33	mmfp1302	Virtual	Purifier: Plug-and-play Backdoor Mitigation for Pre-trained Models Via Anomaly Activation Suppression -- Xiaoyu Zhang (Xidian University), Jin Yulin (Xidian University), Wang Tao (Xidian University), Jian Lou (Xidian University), Xiaofeng Chen (Xidian University, Tsinghua University)
	9h51	mmfp2845	Virtual	Hierarchical Graph Embedded Pose Regularity Learning via Spatio-Temporal Transformer for Abnormal Behavior Detection -- Huang Chao (Harbin Institute of Technology), Yabo Liu (Harbin Institute of Technology), Zheng Zhang (Harbin Institute of Technology), Chengliang Liu (Harbin Institute of Technology), Jie Wen (Harbin Institute of Technology), Yong Xu (Harbin Institute of Technology), Yaowei Wang (Pengcheng Laboratory)
	10h09	mmfp2732	Virtual	Keyword Spotting in the Homomorphic Encrypted Domain Using Deep Complex-Valued CNN -- Peijia Zheng (SUN YAT-SEN UNIVERSITY), Char Giway (SUN YAT-SEN UNIVERSITY), Zeng Huicong (SUN YAT-SEN UNIVERSITY), Jiwu Huang (Shenzhen Univerisity)
Industry Session 1 [Pav. 4 - Room 1.07]	**9h15-10h45**			
Keynote	9h15			Neural Photo Avatar for Communication and Content Creation -- Ming-Yu Liu
Paper Session Ind1: Multimedia Content Creation and Recom.	10h15	mmind003	Virtual	CreaGAN: An Automatic Creative Generation Framework for Display Advertising -- Shiyao Wang (Alibaba Group), Qi Liu (University of Science and Technology of China), Yicheng Zhong (Peking University), Zhilong Zhou (Alibaba Group), Tiezheng Ge (Alibaba Group), Defu Lian (University of Science and Technology of China), Yuning Jiang (University of Science and Technology of China, Tsinghua University)
	10h30	mmind009	Virtual	Personality-Driven Social Multimedia Content Recommendation -- Qi Yang (Somin.ai), Sergey Nikolenko (N/A), Alfred Huang (Nanyang Technological University), Aleksandr Farseev (National University of Singapore)
Poster Session 3 [Pav. 4 - Hall]	**10h30-11h45**			*(Posters will be hung the whole day)*
Understanding Multimedia Content: Multimodal Fusion and Embeddings		mmfp0995	Virtual	Cross-Domain and Cross-Modal Knowledge Distillation in Domain Adaptation for 3D Semantic Segmentation -- MiaoYu Li (Xiamen University), Yachao Zhang (Xiamen University,), Yuan Xie (East China Normal University), Zuodong Gao (Xiamen University), Cuihua Li (XMU), zhizhong zhang (East China Normal University), Yanyun Qu (Xiamen University)
		mmfp1169	Virtual	AVA-AVD: Audio-visual Speaker Diarization in the Wild -- Eric Zhongcong Xu (national university of singaore, National University of Singapore), Zeyang Song (national university of singaore, National University of Singapore), Satoshi Tsutsui (Indiana University), Chao Feng (University of Michigan - Ann Arbor), Mang Ye (Wuhan University), Zheng Shou (National University of Singapore)
		mmfp2448	Virtual	Image-Signal Correlation Network for Textile Fiber Identification -- Bo Peng (Fudan University), Liren He (Fudan University), Yining Qiu (Fudan University), Wu Dong (Fudan University), Mingmin Chi (Fudan University)
		mmfp2947	Virtual	Relation-enhanced Negative Sampling for Multimodal Knowledge Graph Completion -- Derong Xu (University of Science and Technology of China), Tong Xu (University of Science and Technology of China), Shiwei Wu (University of Science and Technology of China), Jingbo Zhou (Baidu Research), Enhong Chen (University of Science and Technology of China)
		mmfp0520	Virtual	Symmetric Uncertainty-Aware Feature Transmission for Depth Super-Resolution -- Wuxuan Shi (Wuhan University), Mang Ye (Wuhan University), Bo Du (Wuhan University)
		mmfp2548	Virtual	DTR: An Information Bottleneck Based Regularization Framework for Video Action Recognition -- Jiawei Fan (Beijing University of Posts and Telecommunications), Yu Zhao (Harbin Institute of Technology), Xie Yu (Beijing University of Posts and Telecommunications), Lihua Ma (Xidian University), Junqi Liu (Beihang University), Fangqiu Yi (Peking University), Boxun Li (Megvii Technology Inc.)
		mmfp1572	Virtual	Self-Supervised Graph Neural Network for Multi-Source Domain Adaptation -- Jin Yuan (Southeast University, Tsinghua University), Feng Hou (Institute of Computing Technology, Chinese Academy of Sciences), Yangzhou Du (Tsinghua University, Tsinghua University), zhongchao shi (Lenovo Research), Xin Geng (Southeast University), Jianping Fan (AI Lab at Lenovo Research), Yong Rui (Lenovo)
		mmfp0254	On-site	ChoreoGraph: Music-conditioned Automatic Dance Choreography over a Style and Tempo Consistent Dynamic Graph -- Ho Yin Au (Hong Kong Baptist University), Jie Chen (Hong Kong Baptist University), Junkun Jiang (Hong Kong Baptist University), Yike Guo (Hong Kong Baptist University)
		mmfp1148	On-site	Pixelwise Adaptive Discretization with Uncertainty Sampling for Depth Completion -- Rui Peng (Peking University), Tord Zhang (Peking University), Bing Li (KAUST), Yitong Wang (Peking University)
		mmfp0589	Virtual	Robust Diversified Graph Contrastive Network for Incomplete Multi-view Clustering -- Zhe Xue (Beijing University of Posts and Telecommunications), Junping Du (Beijing University of Posts and Telecommunications), Hai Zhu (Beijing University of Posts and Telecommunications), Andrew Guan (Beijing University of Posts and Telecommunications), 龙 云飞 (Beijing University of Posts and Telecommunications), Yu Zang (Beijing University of Posts and Telecommunications), Meiyu Liang (Beijing University of Posts and Telecommunications)
		mmfp1457	On-site	Calibrating Class Weights with Multi-Modal Information for Partial Video Domain Adaptation -- Xiyu Wang (Nanyang Technological University), Yuecong Xu (Institute for Infocomm Research, A*STAR), JIANFEI YANG (Nanyang Technological University), Kezhi Mao (Nanyang Technological University)
		mmfp0757	Virtual	Cyclical Fusion: Accurate 3D Reconstruction via Cyclical Monotonicity -- Duo Chen (Sichuan University), Zixin Tang (Sichuan University), Yiguang Liu (Sichuan University)
		mmfp0928	Virtual	Keypoint-Guided Modality-Invariant Discriminative Learning for Visible-Infrared Person Re-identification -- Tengfei Liang (Beijing Jiaotong University), Yi Jin (Beijing Jiaotong University), Wu Liu (JD Explore Academy), Songhe Feng (Beijing Jiaotong University), Tao Wang (Beijing Jiaotong University), Yidong Li (Beijing Jiaotong University)
		mmfp1308	Virtual	Model-Guided Multi-Contrast Deep Unfolding Network for MRI Super-resolution Reconstruction -- yang gang (University of Science and Technology of China), Li Zhang (University of Science and Technology of China), man zhou (University of Science and Technology of China), Aiping Liu (University of Science and Technology of China), Xun Chen (University of Science and Technology of China), Zhiwei Xiong (USTC), Feng Wu (University of Science and Technology of China)
		mmfp1971	Virtual	Learning from Different text-image Pairs: A Relation-enhanced Graph Convolutional Network for Multimodal NER -- Fei Zhao (Nanjing University), Chunhui li (Nanjing University), Zhen Wu (Nanjing University), Shangyu Xing (Nanjing University), Xinyu Dai (Nanjing University)
		mmfp0389	Virtual	Multi-directional Knowledge Transfer for Few-Shot Learning -- Shuo Wang (University of Science and Technology of China), XinYu Zhang (University of Science and Technology of China), Yanbin Hao (University of Science and Technology of China), Chengbing Wang (University of Science and Technology of China), Xiangnan He (University of Science and Technology of China)
		mmfp0620	Virtual	DetFusion: A Detection-driven Infrared and Visible Image Fusion Network -- Yiming Sun (Tianjin University), Bing Cao (University of Tianjin of China), Pengfei Zhu (Tianjin University), Qinghua Hu (Tianjin University)
		mmfp1009	Virtual	Sketch Transformer: Asymmetrical Disentanglement Learning from Dynamic Synthesis -- Cuiqun Chen (Wuhan University), Mang Ye (Wuhan University), Meibin Qi (Hefei University of Technology), Bo Du (Wuhan University)
		mmfp0450	Virtual	Rethinking the Metric in Few-shot Learning: From an Adaptive Multi-Distance Perspective -- Jinxiang Lai (Tencent Youtu Lab), Siqian Yang (Tongji University), GUANNAN JIANG (University of New South Wales), Xi Wang (South China University of Technology), Yuxi Li (Tencent Youtu Lab), Zihui Jia (Peking University), Xiaochen Chen (Tencent YouTu Lab), Jun Liu (Tencent YouTu Lab), Bin-Bin Gao (Tencent YouTu Lab), Wei Zhang (Technische Universität München), Yuan Xie (East China Normal University), Chengjie Wang (Tencent YouTu Lab)
		mmfp0033	Virtual	Cross-Modality Domain Adaptation for Freespace Detection: A Simple yet Effective Baseline -- Yuanbin Wang (Beijing University of Aeronautics and Astronautics), Leyan Zhu (Beihang University), Shaofei Huang (Sensetime Technology), Tianrui Hui (Institute of Information Engineering, Chinese Academy of Sciences), xiaojie li (Sensetime), Fei Wang (Sensetime), Si Liu (Beihang University)
		mmfp0590	On-site	Learning a Dynamic Cross-modal Network for Multi-spectral Pedestrian Detection -- Jin Xie (Chongqing University), Rao Anwer (Mohamed bin Zayed University of Artificial Intelligence), Hisham Cholakkal (MBZUAI), Jing Nie (university of tianjin of china, Tianjin University), Jiale Cao (Tianjin University), Jorma Laaksonen (Aalto University), Fahad Khan (Inception Institute of Artificial Intelligence)
		mmfp1856	Virtual	Two-Stage Multi-Scale Resolution-Adaptive Network for Low-Resolution Face Recognition -- Haihan Wang (University of Science and Technology of China), Shangfei Wang (University of Science and Technology of China,), Lin Fang (University of Science and Technology of China)
		mmfp0951	Virtual	When True Becomes False: Few-Shot Link Prediction beyond Binary Relations through Mining False Positive Entities -- Xuan Zhang (Renmin University of China), Xun Liang (Renmin University of China), XiangPing ZHENG (Renmin University of China), Bo-Wen Wu (Renmin University of China), Yuhui Guo (Renmin University of China)
		mmfp1377	On-site	Extreme-scale Talking-Face Video Upsampling with Audio-Visual Priors -- Sindhu Hegde (International Institute of Information Technology Hyderabad), Rudrabha Mukhopadhyay (International Institute of Information Technology Hyderabad), Vinay Namboodiri (University of Bath), C.V. Jawahar (IIIT Hyderabad)
		mmfp0826	Virtual	Enhancement by Your Aesthetic: An Intelligible Unsupervised Personalized Enhancer for Low-Light Images -- Naishan Zheng (University of Science and Technology of China), Huang Jie (University of Science and Technology of China), Qi Zhu (University of Science and Technology of China), man zhou (University of Science and Technology of China), Feng Zhao (University of Science and Technology of China), Zheng-Jun Zha (University of Science and Technology of China)
		mmfp0946	Virtual	Scale-flow: Estimating 3D Motion from Video -- ling han (Nanjing University of Science and Technology), Sun Quansen (Nanjing University of Science and Technology), Zhenwen Ren (Southwest University Of Science And Technology), ya liu (Nanjing University of Science and Technology), Hongyuan Wang (Nanjing University of Science and Technology), Zichen Wang (Suzhou Zhito Technology Co., Ltd)
		mmfp1843	On-site	SlimSeg: Slimmable Semantic Segmentation with Boundary Supervision -- Danna Xue (Northwestern Polytechnical University), Fei Yang (Computer Vision Center, Universitat Autònoma de Barcelona), Pei Wang (Northwestern Polytechnical University), Luis Herranz (Computer Vision Center, Universitat Autònoma de Barcelona), Jinqiu Sun (Northwest Polytechnical University Xi'an), Yu Zhu (Northwest Polytechnical University Xi'an), Yanning Zhang (Northwestern Polytechnical University)
		mmfp0846	Virtual	Saliency in Augmented Reality -- Huiyu Duan (Shanghai Jiao Tong University), Wei Shen (Shanghai Jiao Tong University), Xiongkuo Min (Shanghai Jiao Tong University), Danyang Tu (Shanghai Jiao Tong University), Jing Li (Alibaba Group), Guangtao Zhai (Shanghai Jiao Tong University)
		mmfp2626	Virtual	T-former: An Efficient Transformer for Image Inpainting -- Ye Deng (Xi'an Jiaotong University), Siqi Hui ('Xi'an Jiaotong University), Sanping Zhou (Xi'an Jiaotong University), Deyu Meng, Jinjun Wang (Xi'an Jiaotong University)
		mmfp1213	On-site	Cycle Self-Training for Semi-Supervised Object Detection with Distribution Consistency Reweighting -- Hao Liu (Institute of Computing Technology, Chinese Academy of Sciences), Bin Chen (Institute of Computing Technology, Chinese Academy of Sciences), Bo Wang (Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Chinese Academy of Sciences), Chunpeng Wu (Duke University), Feng Dai (ICT, Chinese Academy of Sciences), Peng Wu (Beijing Institute of Technology)

Understanding Multimedia Content: Media Interpretation		mmfp1374	Virtual	VMRF: View Matching Neural Radiance Fields -- Jiahui Zhang (Nanyang Technological University), Fangneng Zhan (Max Planck Institute for Informatics), Rongliang WU (Nanyang Technological University), Yingchen Yu (Nanyang Technological University), Wenqing Zhang (Huazhong University of Science and Technology), Song Bai (University of Oxford), Xiaoqin Zhang (Wenzhou University), Shijian Lu (Nanyang Technological University)
		mmfp1014	Virtual	ME-D2N: Multi-Expert Domain Decompositional Network for Cross-Domain Few-Shot Learning -- YuQian Fu (Fudan University), YU XIE (Fudan University), Yanwei Fu (Fudan University), Jingjing Chen (Fudan University), Yu-Gang Jiang (Fudan University)
		mmfp1121	Virtual	Towards Causality Inference for Very Important Person Localization -- Xiao Wang (University of Science and Technology), Zheng Wang (Wuhan University, China), Wu Liu (JD Explore Academy), Xin Xu (Wuhan University of Science and Technology), Qijun Zhao (Sichuan University), Shin'ichi Satoh (National Institute of Informatics)
		mmfp2171	Virtual	MMDV: Interpreting DNNs via Building Evaluation Metrics, Manual Manipulation and Decision Visualization -- keyang Cheng (Jiangsu University), Yu Si (Jiangsu University), Zhou Hao (Jiangsu university), Rabia Tahir (Jiangsu University Zhenjiang China)
		mmfp1558	Virtual	Learning Dual Convolutional Dictionaries for Image De-raining -- Chengjie Ge (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), Zheng-Jun Zha (University of Science and Technology of China)
		mmfp3090	Virtual	Source-Free Domain Adaptation for Real-world Image Dehazing -- Hu Yu (University of Science and Technology of China), Huang Jie (University of Science and Technology of China), Yajing Liu (University of Science and Technology of China, Tsinghua University), Qi Zhu (University of Science and Technology of China), man zhou (University of Science and Technology of China), Feng Zhao (University of Science and Technology of China)
		mmfp1745	On-site	Knowledge Guided Representation Disentangle for Face Recognition from Low Illumination Images -- miao xiangyu (University of Science and Technology of China), Shangfei Wang (University of Science and Technology of China,)
		mmfp1699	Virtual	Tracking Objects in Low Frame Rate Videos -- Tao Zhou (Zhejiang University), Wenhan Luo (SUN YAT-SEN UNIVERSITY), Zhiguo Shi (Zhejiang University), Jiming Chen (Zhejiang University), Qi Ye (Zhejiang University)
		mmfp1010	Virtual	ICNet: Joint Alignment and Reconstruction via Iterative Collaboration for Video Super-Resolution -- Jiaxu Leng (Chongqing University of Post and Telecommunications), Jia Wang (Chongqing University of Post and Telecommunications), Xinbo Gao (Chongqing University of Post and Telecommunications), Bo Hu (Chongqing University of Post and Telecommunications), Ji Gan (Chongqing University of Post and Telecommunications), CHENQIANG GAO (School of Computer Science, Carnegie Mellon University)
		mmfp1185	Virtual	Estimation of Reliable Proposal Quality for Temporal Action Detection -- Junshan Hu (University of Science and Technology of China), Chaoxu Guo (Alibaba Group), Liansheng Zhuang (University of Science and Technology of China), Biao Wang (, Chinese Academy of Sciences), Tiezheng Ge (Alibaba Group), Yuning Jiang (University of Science and Technology of China, Tsinghua University), Houqiang Li (University of Science and Technology of China)
		mmfp2744	Virtual	Semi-supervised Semantic Segmentation via Prototypical Contrastive Learning -- Zenggui Chen (Peking University), Zhouhui Lian (Peking University)
		mmfp3055	Virtual	Towards Understanding Cross Resolution Feature Matching for Surveillance Face Recognition -- Chiawei Kuo (National Yang Ming Chiao Tung University), Yi-Ting Tsai (National Yang Ming Chiao Tung University), Hong-Han Shuai (National Chiao Tung University), Yi-ren Yeh (National Kaohsiung Normal University), Ching-Chun Huang (National Yang Ming Chiao Tung University)
		mmfp0626	Virtual	Single Image Shadow Detection via Complementary Mechanism -- Yurui Zhu (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), Chengzhi Cao (University of Science and Technology of China), 曦 汪 (University of Science and Technology of China), Qibin Sun (University of Science and Technology of China, Tsinghua University), Zheng-Jun Zha (University of Science and Technology of China)
		mmfp3271	Virtual	Distilling Resolution-robust Identity Knowledge for Texture-Enhanced Face Hallucination -- Qiqi Bao (Tsinghua University, Tsinghua University), Rui Zhu (City University), Bowen Gang (Fudan University), Pengyang Zhao (Electronic Engineering, Tsinghua University.), Wenming Yang (Tsinghua University, Tsinghua University), Qingmin Liao (Tsinghua University, Tsinghua University)
		mmfp2036	Virtual	Phoneme-Aware Adaptation with Discrepancy Minimization and Dynamically-Classified Vector for Text-independent Speaker Verification -- Jia Wang (Shenzhen University), Tianhao Lan (Shenzhen University), Jie Chen (Shenzhen University), Chengwen Luo (Shenzhen University), chao wu (Department of Software Engineering, Shenzhen University), Jianqiang Li (Shenzhen University)
		mmfp1042	Virtual	Anomaly Warning: Learning and Memorizing Future Semantic Patterns for Unsupervised Ex-ante Potential Anomaly Prediction -- Jiaxu Leng (Chongqing University of Post and Telecommunications), Mingpi Tan (Chongqing University of Post and Telecommunications), Xinbo Gao (Chongqing University of Post and Telecommunications), LU WEN (Xi'an University of Electronic Science and Technology), Zongyi Xu (Chongqing University of Post and Telecommunications)
		mmfp2391	On-site	DuetFace: Collaborative Privacy-Preserving Face Recognition via Channel Splitting in the Frequency Domain -- Yuxi Mi (Fudan University), Yuge Huang (Zhejiang University), Jiazhen Ji (Tencent Youtu Lab), Liu Hongquan (Fudan University), Xingkun Xu (Beijing University of Posts and Telecommunications), Shouhong Ding (Tencent Youtu Lab), Shuigeng Zhou (Fudan University)
		mmfp1630	Virtual	3D Human Mesh Reconstruction by Learning to Sample Joint Adaptive Tokens for Transformers -- YOUZE XUE (Tsinghua University, Tsinghua University), Jiansheng Chen (University of Science and Technology Beijing), Yudong Zhang (Tsinghua University, Tsinghua University), Cheng Yu (Tsinghua University, Tsinghua University), Huimin Ma (University of Science and Technology Beijing), Hongbing Ma (Tsinghua University)
		mmfp0175	Virtual	Grouped Adaptive Loss Weighting for Person Search -- Yanling TIAN (Nanjing University of Science and Technology), Di Chen (Nanjing University of Science and Technology), Yunan Liu (Nanjing University of Science and Technology), Shanshan Zhang (Nanjing University of Science and Technology), Jian Yang (Nanjing University of Science and Technology)
		mmfp0789	Virtual	Multi-view Gait Video Synthesis -- Weilai Xiang (Beihang University), Hongyu Yang (Beihang University), Di Huang (Beihang University), Yunhong Wang (Beihang University)
		mmfp2228	Virtual	Curriculum-NAS: Curriculum Weight-Sharing Neural Architecture Search -- Yuwei Zhou (Tsinghua University, Tsinghua University), Xin Wang (Tsinghua University), Hong Chen (Tsinghua University, Tsinghua University), Xuguang Duan (Tsinghua University, Tsinghua University), Chaoyu Guan (Tsinghua University, Tsinghua University), Wenwu Zhu (Tsinghua University, Tsinghua University)
		mmfp0723	Virtual	Content and Gradient Model-driven Deep Network for Single Image Reflection Removal -- Ya-Nan Zhang (Shenzhen University), Linlin Shen (Shenzhen University), Qiufu Li (Department of Software Engineering, Shenzhen University)
		mmfp0432	Virtual	TransCNN-HAE: Transformer-CNN Hybrid AutoEncoder for Blind Image Inpainting -- Haoru Zhao (Ocean University of China), Zhaorui Gu (Ocean University of China), Bing Zheng (Ocean University of China), Haiyong Zheng (Ocean University of China)
		mmfp1422	On-site	Trajectory Prediction from Hierarchical Perspective -- Tangwen Qian (Institute of Computing Technology, Chinese Academy of Sciences), Yongjun Xu (Institute of Computing Technology, Chinese Academy of Sciences), Zhao Zhang (Institute of Computing Technology, Chinese Academy of Sciences), Fei Wang (Institute of Computig Technology, Chinese Academy of Sciences)
		mmfp1299	Virtual	Exploring Effective Knowledge Transfer for Few-shot Object Detection -- Zhiyuan Zhao (Beijing University of Aeronautics and Astronautics), Qingjie LIU (Beihang University), Yunhong Wang (Beihang University)
		mmfp0332	Virtual	More is better: Multi-source Dynamic Parsing Attention for Occluded Person Re-identification -- Xinhua Cheng (Peking University), Mengxi Jia (Peking University ShenZhen Graduate School), Qian Wang (Sichuan University), Jian Zhang (Peking University)
		mmfp0915	On-site	ReFu: Refine and Fuse the Unobserved View for Detail-Preserving Single-Image 3D Human Reconstruction -- GYUMIN SHIM (Korea Advanced Institute of Science & Technology), Minsoo Lee (Korea Advanced Institute of Science and Technology), Jaegul Choo (Korea Advanced Institute of Science and Technology)
		mmfp2016	On-site	Transformers in Spectral Domain for Estimating Image Geometric Transformation -- Min-Gi Choi (Hallym University), Sangyeong Lee (Hallym University), Heesun Jung (Hallym University), Jong-Uk Hou (Hallym University)
Understanding Multimedia Content: Vision and Language		mmfp2946	Virtual	AI-VQA: Visual Question Answering based on Agent Interaction with Interpretability -- Rengang Li (Inspur Electronic Information Industry Co.,Ltd), Cong Xu (Inspur (Beijing) Electronic Information Industry Co., Ltd), Zhenhua Guo (Inspur (Beijing) Electronic Information Industry Co., Ltd), Baoyu Fan (Inspur Electronic Information Industry Co.,Ltd.), Runze Zhang (Inspur Electronic Information Industry Co.,Ltd), Wei Liu (Inspur Electronic Information Industry Co.,Ltd), Yaqian Zhao (Inspur Electronic Information Industry Co.,Ltd), Endong Wang (Inspur Electronic Information Industry Co.,Ltd), Endong Wang (Inspur Group)
		mmfp1204	On-site	Situational Perception Guided Image Matting -- Bo Xu (OPPO), Jiake Xie (Winroad Holdings Limited), Han Huang (OPPO Research Institute), Ziwen Li (University of California, San Diego), Cheng Lu (Xpeng Motors), Yong Tang (Winroad Holdings Ltd), Yandong Guo (OPPO Research Institute)
		mmfp1934	Virtual	ROMA: Cross-Domain Region Similarity Matching for Unpaired Nighttime Infrared to Daytime Visible Video Translation -- yu Jie (Beijing Institute of Technology), Kai Chen (Henan Univeristy), Shuang Li (Beijing Institute of Technology), Bingfeng Han (Beijing Institute of Technology), Chi Liu (Beijing Institute of Technology), Shuigen Wang (Beijing Institute of Technology)
		mmfp0068	Virtual	A^{3}GAN: Attribute-Aware Anonymization Networks for Face De-identification -- Liming Zhai (Nanyang Technological University), Qing Guo (Nanyang Technological University), Xiaofei Xie (Singapore Management University), Lei Ma (University of Alberta), Yi Wang (Nanyang Technological University), Yang Liu (Nanyang Technological University)
		mmfp1288	Virtual	CAIBC: Capturing All-round Information Beyond Color for Text-based Person Retrieval -- Zijie Wang (Nanjing Tech University), Aichun Zhu (Nanjing Tech University), Jingyi Xue (Nanjing University), Xili Wan (Nanjing Tech University), Liu Chao (Jinling institute of technology), Tian Wang (Beihang University, China), Yifeng Li (Nanjing Tech University)
		mmfp1760	Virtual	PreyNet: Preying on Camouflaged Objects -- Miao Zhang (Dalian University of Technology), Shuang Xu (Dalian University of Technology), Yongri Piao (Dalian University of Technology), Dongxiang Shi (Dalian University of Technology), Shusen Lin (Dalian University of Technology), Huchuan Lu (Dalian University of Technology)
		mmfp0914	On-site	Not All Pixels Are Matched: Dense Contrastive Learning for Cross-Modality Person Re-Identification -- Hanzhe Sun (East China Normal University), Jun Liu (Tencent YouTu Lab), zhizhong zhang (East China Normal University), Chengjie Wang (Tencent YouTu Lab), Yanyun Qu (Xiamen University), Yuan Xie (East China Normal University), Lizhuang Ma (Dept. of Computer Sci. & Eng., Shanghai Jiao Tong University)
		mmfp3265	Virtual	Asymmetric Adversarial-based Feature Disentanglement Learning for Cross-Database Micro-Expression Recognition -- Shiting XU (South China University of Technology), Zhiheng Zhou (South China University of Technology), Junyuan Shang (South China University of Technology)
		mmfp2358	Virtual	Backdoor Attacks on Crowd Counting -- Yuhua Sun (Huazhong University of Science and Technology), Tailai Zhang (Huazhong University of Science and Technology), Xingjun Ma (Deakin University), Pan Zhou (Huazhong University of Science and Technology), Jian Lou (Xidian University), Zichuan Xu (Dalian University of Technology), Xing Di (Protagolabs Inc.), Yu Cheng (Microsoft Research), Lichao Sun (Lehigh University)
		mmfp2417	Virtual	Robust Industrial UAV/UGV-Based Unsupervised Domain Adaptive Crack Recognitions with Depth and Edge Awareness: From System and Database Constructions to Real-Site Inspections -- Kangcheng Liu (The Chinese University of Hong Kong)
		mmfp2371	On-site	$ Forcing the whole video as background: an adversarial learning strategy for weakly temporal action localization $ -- Ziqiang Li (Chongqing University), Yongxin Ge (Chongqing University), Yu Jiaruo (Chongqing University), Zhongming Chen (Chongqing University)
		mmfp0356	Virtual	Towards Accurate Post-Training Quantization for Vision Transformer -- Yifu Ding (Beihang University), Haotong Qin (Beihang University), Qinghua YAN (Beihang University), Zhenhua Chai (Meituan), Junjie Liu (Meituan), Xiaolin Wei (Meituan), Xianglong Liu (Beihang University, Tsinghua University)
		mmfp1700	Virtual	Neighbor Correspondence Matching for Flow-based Video Frame Synthesis -- Zhaoyang Jia (University of Science and Technology of China), Yan Lu (Microsoft Research Asia), Houqiang Li (University of Science and Technology of China)
		mmfp3088	Virtual	ReFormer: The Relational Transformer for Image Captioning -- Xuewen Yang (State University of New York, Stony Brook), Xin Wang (SUNY at Stony Brook), Yingru Liu (Facebook Reality Lab)
		mmfp2221	Virtual	Transcript to Video: Efficient Clip Sequencing from Texts -- Yu Xiong (The Chinese University of Hong Kong), Fabian heilbron (Adobe Systems), Dahua Lin (Toyota Technnological Institute)
		mmfp0751	Virtual	Domain Reconstruction and Resampling for Robust Salient Object Detection -- Senbo Yan (Zhejiang University), Liang Peng (FABU Inc), Chuer Yu (Zhejiang University), Zheng Yang (Fabu Inc), Haifeng Liu (Zhejiang University), Deng Cai (Zhejiang University)
		mmfp1023	Virtual	Phase-based Memory Network for Video Dehazing -- Ye Liu (Tianjin University), Liang Wan (Tianjin University), Huazhu Fu (Institute of High Performance Computing, Singapore, A*STAR), Jing Qin (Hong Kong Polytechnic University), Lei Zhu (Hong Kong University of Science and Technology)
		mmfp0146	Virtual	UConNet: Unsupervised Controllable Network for Image and Video Deraining -- Jun-hao Zhuang (University of Electronic Science and Technology of China), Yisi Luo (University of Electronic Science and Technology of China), Xi-Le Zhao (University of Electronic Science and Technology of China), Tai-Xiang Jiang (Southwestern University of Finance and Economics), Bichuan Guo (Tencent)
		mmfp2643	Virtual	AGTGAN: Unpaired Image Translation for Photographic Ancient Character Generation -- Hongxiang Huang (South China University of Technology), Daihui Yang (South China University of Technology), Gang Dai (South China University of Technology), Zhen Han (Ludwig Maximilian University of Munich), Yuyi Wang (Swiss Federal Institute of Technology), Kin-man Lam (The Hong Kong Polytechnic University), Fan Yang (South China University of Technology), Shuangping Huang (South China University of Technology), Yongge Liu (Anyang Normal University), Mengchao He (Alibaba Group)
		mmfp1663	Virtual	CLIPTexture: Text-Driven Texture Synthesis -- Yiren Song (Shanghai Jiaotong University)
		mmfp0178	Virtual	OCR-Pose: Occlusion-aware Contrastive Representation for Unsupervised 3D Human Pose Estimation -- Junjie Wang (Shanghai Jiaotong University), Zhenbo Yu (Shanghai Jiao Tong University), Zhengyan Tong (Shanghai Jiaotong University), Hang Wang (Huawei Hisilicon), Jinxian Liu (Shanghai Jiao Tong University), Wenjun Zhang (Shanghai Jiao Tong University), Xiaoyan Wu (Shanghai Jiaotong University)
		mmfp0857	Virtual	DualSign: Semi-supervised Sign Language Production with Balanced Multi-modal Multi-task Dual Transformation -- Wencan Huang (Zhejiang University), Zhou Zhao (Zhejiang University), Jinzheng He (Zhejiang University), Mingmin Zhang (Zhejiang University)
		mmfp0420	Virtual	A Lightweight Graph Transformer Network for Human Mesh Reconstruction from 2D Human Pose -- Ce Zheng (University of Central Florida), Matias Mendieta (University of Central Florida), Pu Wang (University of North Carolina at Charlotte), Aidong Lu (University of North Carolina at Charlotte), Chen Chen (University of Central Florida)
		mmfp1220	Virtual	Repainting and Imitating Learning for Lane Detection -- Yue He (Hunan University), Minyue Jiang (Baidu), Xiaoqing Ye (Baidu), Liang Du (Tencent), Kang Zou (Huazhong University of Science and Technology, Tsinghua University), Wei Zhang (Baidu), Xiao Tan (Baidu), Errui Ding (Baidu)

		mmfp0294	Virtual	Paired Cross-Modal Data Augmentation for Fine-Grained Image-to-Text Retrieval -- Hao Wang (Nanyang Technological University), Guosheng Lin (Nanyang Technological University), Steven HOI (Singapore Management University), Chunyan Miao (Nanyang Technological University)
		mmfp0312	On-site	BlumNet: Graph Component Detection for Object Skeleton Extraction -- Yulu Zhang (Shanghai Huipainter Software Technology Co., Ltd.), Liang Sang (University of the Chinese Academy of Sciences), Marcin Grzegorzek (Lubkeck University), John See (Heriot-Watt University), Cong Yang (Soochow University)
		mmfp1398	Virtual	PPMN: Pixel-Phrase Matching Network for One-Stage Panoptic Narrative Grounding -- Zihan Ding (Beihang University), Zihan Ding (Beihang University), Tianrui Hui (Institute of Information Engineering, Chinese Academy of Sciences), Junshi Huang (National University of Singapore), Wei Xiaoming (Beijing University of Aeronautics and Astronautics), Xiaolin Wei (Meituan), Si Liu (Beihang University)
		mmfp1922	Virtual	Incremental Few-Shot Semantic Segmentation via Embedding Adaptive-Update and Hyper-class Representation -- Guangchen Shi (Hohai University), Yirui Wu (hohai University), Jun Liu (Singapore University of Technology and Design), SHAOHUA WAN (Wuhan University), Wenhai Wang (Shanghai AI Laboratory), Tong Lu (Nanjing University)
		mmfp0755	Virtual	Synthetic Data Supervised Salient Object Detection -- Zhenyu Wu (Beihang University), Lin Wang (Beihang University), wei wang (Harbin Institute of Technology), Tengfei Shi (beihang University), Chenglizhao Chen (China University of Petroleum), Aimin Hao (Beihang University), Shuo Li (Western University)
		mmfp1181	Virtual	Learning Granularity-Unified Representations for Text-to-Image Person Re-identification -- Zhiyin Shao (South China University of Technology), Xinyu Zhang (Baidu), Meng Fang (Eindhoven University of Technology), Zhifeng Lin (South China University of Technology), Jian Wang (Baidu), Changxing Ding (South China University of Technology)
		mmfp1283	Virtual	Flexible Hybrid Lenses Light Field Super-Resolution using Layered Refinement -- Song ChangSong (Beijing jiaotong univercity), Youfang Lin (Beijing jiaotong univercity), Shuo Zhang (Beijing jiaotong university)
		mmfp2739	On-site	DS-MVSNet: Unsupervised Multi-view Stereo via Depth Synthesis -- Jingliang Li (University of the Chinese Academy of Sciences), Zhengda Lu (University of the Chinese Academy of Sciences), Yiqun Wang (King Abdullah University of Science and Technology), Ying WANG (University of the Chinese Academy of Sciences), Jun Xiao (University of Chinese Academy of Sciences)
		mmfp2887	Virtual	Enhancing Image Rescaling using Dual Latent Variables in Invertible Neural Network -- Min Zhang (University of Southern California), Zhihong Pan (Baidu USA LLC), Xin Zhou (University of Michigan - Ann Arbor), C.-C. Kuo (University of Southern California)
		mmfp2747	Virtual	ScatterNet: Point Cloud Learning via Scatters -- Qi Liu (Shanghai Jiaotong University), Nianjuan Jiang (SmartMore), Jiangbo Lu (SmartMore Corporation), Mingang Chen (Shanghai Development Center of Computer Software Technology), Ran Yi (Shanghai Jiao Tong University), Lizhuang Ma (Dept. of Computer Sci. & Eng., Shanghai Jiao Tong University)
		mmfp1983	Virtual	Making The Best of Both Worlds: A Domain-Oriented Transformer for Unsupervised Domain Adaptation -- Wenxuan Ma (Beijing Institute of Technology), jinming Zhang (Beijing Institute of Technology), Shuang Li (Beijing Institute of Technology), Chi Liu (Beijing Institute of Technology), Yulin Wang (Tsinghua University, Tsinghua University), Wei Li (Peking University)
		mmfp0365	Virtual	Gloss Semantic-Enhanced Network with Online Back-Translation for Sign Language Production -- Shengeng Tang (Hefei University of Technology), Richang Hong (Hefei University of Technology), Dan Guo (Hefei University of Technology), Meng Wang (Hefei University of Technology)
		mmfp0585	Virtual	Paint and Distill: Boosting 3D Object Detection with Semantic Passing Network -- BO JU (Baidu), Kang Zou (Huazhong University of Science and Technology, Tsinghua University), Xiaoqing Ye (Baidu), Minyue Jiang (Baidu), Xiao Tan (Baidu), Errui Ding (Baidu), Jingdong Wang (Baidu)
		mmfp0200	Virtual	Dual Contrastive Learning for Spatio-temporal Representation -- Shuangrui Ding (Shanghai Jiaotong University), Rui Qian (The Chinese University of Hong Kong), Hongkai Xiong (Shanghai Jiao Tong University)
		mmfp2367	Virtual	Fine-Grained Fragment Diffusion for Cross Domain Crowd Counting -- Huilin Zhu (Wuhan University of Technology), Jingling Yuan (Wuhan University of Technology), Zhengwei Yang (Wuhan University), Xian Zhong (Wuhan University of Technology), Zheng Wang (Wuhan University, China)
		mmfp1205	Virtual	Depth-inspired Label Mining for Unsupervised RGB-D Salient Object Detection -- Teng Yang (Dalian University of Technology), Yue Wang (Dalian University of Technology), Lu Zhang (Dalian University of Technology), Jinqing Qi (Dalian University of Technology, Tsinghua University), Huchuan Lu (Dalian University of Technology)
		mmfp1854	Virtual	FastLTS: Non-autoregressive End-To-End Unconstrained Lip-To-Speech Synthesis -- Yongqi Wang (Zhejiang University), Zhou Zhao (Zhejiang University)
		mmfp1620	Virtual	Interact with Open Scenes: A Life-long Evolution Framework for Interactive Segmentation Models -- Ruitong Gan (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Junsong Fan (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Yuxi Wang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Zhaoxiang Zhang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
		mmfp1738	Virtual	Visual Dialog for Spotting the Differences between Pairs of Similar Images -- Duo Zheng (Beijing University of Posts and Telecommunications), Fandong Meng (WeChat AI, Tencent Inc.), Qingyi Si (Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China), Hairun Fan (Beijing University of Posts and Telecommunications), Zipeng Xu (University of Trento), Jie Zhou (WeChat AI, Tencent Inc.), Fangxiang Feng (Beijing University of Posts and Telecommunications), Xiaojie Wang (Beijing University of Post and Telecommunication)
		mmfp0991	Virtual	Time and Memory Efficient Large-Scale Canonical Correlation Analysis in Fourier Domain -- Xiang-Jun Shen (school of computer science), Zhaorui Xu (Jiangsu university), Liangjun Wang (Jiangsu University), Zechao Li (Nanjing University of Science and Techonolgy)
Engaging Users with Multimedia: Multimedia Search and Recommendation		mmfp3150	Virtual	Adaptive Anti-Bottleneck Multi-Modal Graph Learning Network for Personalized Micro-video Recommendation -- desheng cai (Hefei University of Technology), Shengsheng Qian (, Institute of automation, Chinese academy of science), Quan Fang (, Institute of automation, Chinese academy of science), Jun Hu (National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences), Changsheng Xu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
		mmfp0415	On-site	Show Me What I Like: Detecting User-Specific Video Highlights Using Content-Based Multi-Head Attention -- Uttaran Bhattacharya (University of Maryland, College Park), Gang Wu (Adobe Research), Stefano Petrangeli (Adobe Systems), Viswanathan Swaminathan (Adobe Systems), Dinesh Manocha (University of Maryland, College Park)
		mmfp2922	Virtual	Prototype-based Selective Knowledge Distillation for Zero-Shot Sketch Based Image Retrieval -- Kai Wang (University of Electronic Science and Technology of China), Yifan Wang (University of Electronic Science and Technology of China), Xing Xu (University of Electronic Science and Technology of China), Xin Liu (Huaqiao University), Weihua Ou (Guizhou Normal University), Huimin Lu (Kyushu Institute of Technology)
		mmfp1641	Virtual	ARRA: Absolute-Relative Ranking Attack against Image Retrieval -- Siyuan Li (University of Electronic Science and Technology of China), Xing Xu (University of Electronic Science and Technology of China), Zailei Zhou (University of Electronic Science and Technology of China), Yang Yang (University of Electronic Science and Technology of China), Guoqing Wang (University of Electronic Science and Technology of China), Hengtao Shen (University of Electronic Science and Technology of China)
		mmfp3071	Virtual	Invariant Representation Learning for Multimedia Recommendation -- Xiaoyu Du (Nanjing University of Science and Technology), Zike Wu (South China University of Technology), Fuli Feng (University of Science and Technology of China), Xiangnan He (University of Science and Technology of China), Jinhui Tang (Nanjing University of Science and Technology)
		mmfp0685	Virtual	X-CLIP: End-to-End Multi-grained Contrastive Learning for Video-Text Retrieval -- Yiwei Ma (Xiamen University), Guohai Xu (Alibaba Group), Xiaoshuai Sun (Xiamen University), Ming Yan (Alibaba Group), Ji Zhang (Alibaba Group), Rongrong Ji (Columbia University)
		mmfp2041	Virtual	Mixed Supervision for Instance Learning in Object Detection with Few-shot Annotation -- yi zhong (SUN YAT-SEN UNIVERSITY), Chengyao Wang (SUN YAT-SEN UNIVERSITY), Shiyong Li (Huawei Technologies Ltd.), Zhu Zhou (Huawei Technologies Ltd.), Yaowei Wang (Pengcheng Laboratory), Wei-Shi Zheng (SUN YAT-SEN UNIVERSITY)
		mmfp3065	Virtual	Improved Deep Unsupervised Hashing via Prototypical Learning -- Zeyu Ma (Harbin Institute of Technology), Wei Ju (Peking University), Xiao Luo (Peking University), Chong Chen (Alibaba Group), Xian-Sheng Hua (Zhejiang University), Guangming Lu (Harbin Institute of Technology)
		mmfp1902	Virtual	Adaptive Camera Margin for Mask-guided Domain Adaptive Person Re-identification -- Wang Rui (Anhui University), Feng Chen (Anhui University of Technology), Jun Tang (Anhui University), Pu Yan (anhui jianzhu university)
		mmfp2232	Virtual	BadHash: Invisible Backdoor Attacks against Deep Hashing with Clean Label -- Shengshan Hu (Huazhong University of Science and Technology), Ziqi Zhou (Huazhong University of Science and Technology), Yechao Zhang (Huazhong University of Science and Technology), Leo Zhang (Deakin University), Yifeng Zheng (Harbin Institute of Technology), Yuanyuan HE (Huazhong University of Science and Technology), Hai Jin (Huazhong University of Science and Technology)
		mmfp3069	Virtual	EliMRec: Eliminating Single-modal Bias in Multimedia Recommendation -- Xiaohao Liu (University of the Chinese Academy of Sciences), Zhulin Tao (Communication University of China), Jiahong Shao (Communication University of China), Yang Lifang (Communication University of China), Xianglin Huang (Communication University of China)
		mmfp1784	Virtual	Patch-based Knowledge Distillation for Lifelong Person Re-Identification -- Zhicheng Sun (Peking University), Yadong MU (Peking University)
Engaging Users with Multimedia: Emotional and Social Signals		mmfp0868	Virtual	Self-Paced Label Distribution Learning for In-The-Wild Facial Expression Recognition -- Jianjian Shao (University of Electronic Science and Technology of China), Zhenqian Wu (University of Electronic Science and Technology of China), Yuanyan Luo (University of Electronic Science and Technology of China), Shudong Huang (Sichuan University), Xiaorong Pu (University of Electronic Science and Technology of China), Yazhou Ren (University of Electronic Science and Technology of China)
		mmfp2307	Virtual	Uncertainty-Aware Semi-Supervised Learning of 3D Face Rigging from Single Image -- Yong Zhao (Vrije Universiteit Brussel), Haifeng Chen (Northwestern Polytechnical University), Hichem Sahli (ETRO), Ke Lu (University of Chinese Academy of Sciences), Dongmei Jiang (Northwestern Polytechnical University)
		mmfp1566	Virtual	A Unified Framework against Topology and Class Imbalance -- Junyu Chen (Institute of Information Engineering, Chinese Academy of Sciences), Qianqian Xu (Institute of Computing Technology, Chinese Academy of Sciences), Zhiyong Yang (Chinese Academy of Sciences), Xiaochun Cao (Institute of Information Engineering), Qingming Huang (University of Chinese Academy of Sciences)
		mmfp2438	Virtual	Unified Multi-modal Pre-training for Few-shot Sentiment Analysis with Prompt-based Learning -- Yang Yu (Suzhou University), Dong Zhang (Soochow University, China), Shoushan Li (Soochow University)
		mmfp1067	Virtual	Temporal Sentiment Localization: Listen and Look in Untrimmed Videos -- Zhicheng Zhang (Nankai University), Jufeng Yang (Nankai University)
		mmfp2837	Virtual	VigilanceNet: Decouple Intra- and Inter-Modality Learning for Multimodal Vigilance Estimation in RSVP-Based BCI -- Xinyu Cheng (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Wei Wei (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Changde Du (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Shuang Qiu (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Sanli Tian (University of the Chinese Academy of Sciences), Xiaojun Ma (Institute of Automation, Chinese Academy of Sciences), Huiguang He (Institute of Automation Chinese Academy of Sciences)
		mmfp1062	Virtual	EASE: Robust Facial Expression Recognition via Emotion Ambiguity-SEnsitive Cooperative Networks -- Lijuan Wang (Nankai University), Guoli Jia (Nankai University), ning jiang (Mashang Consumer Finance Co, Ltd), haiying wu (Mashang Consumer Finance Co, Ltd), Jufeng Yang (Nankai University)
		mmfp1805	On-site	Mimicking the Annotation Process for Recognizing the Micro Expressions -- Bo-Kai Ruan (National Chiao Tung University), Ling Lo (National Chiao Tung University), Hong-Han Shuai (National Chiao Tung University), Wen-Huang Cheng (Academia Sinica)
Engaging Users with Multimedia: Summarization, Analytics, and Storytelling		mmfp1198	Virtual	Clustering Generative Adversarial Networks for Story Visualization -- Bowen Li (Department of Computer Science, University of Oxford), Philip Torr (University of Oxford), Thomas Lukasiewicz (Institute of Logic and Computation, Technische Universität Wien)
		mmfp2997	Virtual	DeViT: Deformed Vision Transformers in Video Inpainting -- Jiayin Cai (Kuaishou), Changlin Li (Kuaishou), Xin Tao (Kuaishou), Chun Yuan (Tsinghua University, Tsinghua University), Yu-wing Tai (Kuaishou Technology)
		mmfp1505	Virtual	Multi-Level Spatiotemporal Network for Video Summarization -- Ming Yao (Alibaba Group), Yu Bai (Alibaba Group), Wei Du (Alibaba Group), Xuejun Zhang (Alibaba Group), Heng Quan (Beijing University of Posts and Telecommunications), Fuli Cai (Haerbin Engineering University), Hongwei Kang (Alibaba Group)
		mmfp0358	Virtual	CycleHand: Increasing 3D Pose Estimation Ability on In-the-wild Monocular Image through Cyclic Flow -- Daiheng Gao (Alibaba Group), Xindi Zhang (Queen Mary University London), Xingyu Chen (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Andong Tan (Siemens), Bang Zhang (Alibaba Group), Pan Pan (Alibaba Group), Ping Tan (Simon Fraser University)
		mmfp0736	Virtual	Defeating DeepFakes via Adversarial Visual Reconstruction -- Ziwen He (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Wei Wang (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Weinan Guan (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Jing Dong (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Tieniu Tan (NLPR China)
		mmfp1736	Virtual	Content based User Preference Modeling in Music Generation -- Xichu Ma (national university of singaore, National University of Singapore), Yuchen Wang (national university of singaore, National University of Singapore), Ye Wang (National University of Singapore)
		mmfp2735	Virtual	CrossHuman: learning cross-guidance from multi-frame images for Human Reconstruction -- Liliang Chen (University of Pittsburgh), Jiaqi Li (School of Instrumentation and Optoelectronic Engineering Beihang University), Han Huang (OPPO Research Institute), Yandong Guo (OPPO Research Institute)
		mmfp3159	Virtual	High-Quality 3D Face Reconstruction with Affine Convolutional Networks -- Zhiqian Lin (Zhejiang University), Jiangke Lin (NetEase Fuxi AI Lab), Lincheng Li (NetEase Fuxi AI Lab), Yi Yuan (Wuhan University), Zhengxia Zou (Beihang University)
		mmfp3141	On-site	xCloth: Extracting template-free textured 3D clothes from a monocular image -- Astitva Srivastava (International Institute of Information Technology Hyderabad, Dhirubhai Ambani Institute Of Information and Communication Technology), Chandradeep Pokhariya (International Institute of Information Technology, Hyderabad, International Institute of Information Technology Hyderabad), Sai Jinka (International Institute of Information Technology Hyderabad, Dhirubhai Ambani Institute Of Information and Communication Technology), Avinash Sharma (International Institute of Information Technology Hyderabad, Dhirubhai Ambani Institute Of Information and Communication Technology)
		mmfp0237	Virtual	SD-GAN: Semantic Decomposition for Face Image Synthesis with Discrete Attribute -- Kangneng Zhou (University of Science and Technology Beijing), Xiaobin Zhu (University of Science and Technology Beijing), Daiheng Gao (Alibaba Group), Kai Lee (University of Science and Technology Beijing), Xinjie Li (Pennsylvania State University), Xu-cheng Yin (University of Science and Technology Beijing)
		mmfp0456	Virtual	SingGAN: Generative Adversarial Network For High-Fidelity Singing Voice Generation -- Rongjie Huang (Zhejiang University), Chenye Cui (Zhejiang University), 陈 飞扬 (College of Computer Science and Technology, Zhejiang University), Yi Ren (Zhejiang University), Jinglin Liu (Zhejiang University), Zhou Zhao (Zhejiang University), Baoxing Huai (Huawei Cloud Computing Technology Co., Ltd), Zhefeng Wang (Huawei Technologies Ltd.)
		mmfp1517	Virtual	Design What You Desire: Icon Generation from Orthogonal Application and Theme Labels -- Yinpeng Chen (Huazhong University of Science and Technology), Zhiyu Pan (Huazhong University of Science and Technology), Min Shi (Huazhong University of Science and Technology), Hao Lu (Huazhong University of Science and Technology), Zhiguo Cao (Huazhong University of Science and Technology), Weicai Zhong (Xidian University)

Experience: Multimedia Applications		mmfp1508	Virtual	Semantically-Consistent Dynamic Blurry Image Generation for Image Deblurring -- Jing Zhaohui (Northwest Polytechnical University Xi'an), Youjian Zhang (The University of Sydney, University of Sydney), Chaoyue Wang (JD Explore Academy), Daqing Liu (JD.com Inc.), Yong Xia (Northwestern Polytechnical University)
		mmfp0701	Virtual	RepSR: Training Efficient VGG-style Super-Resolution Networks with Structural Re-Parameterization and Batch Normalization -- Xintao Wang (Applied Research Center, Tencent PCG), Chao Dong (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Chinese Academy of Sciences), Ying Shan (Tencent)
		mmfp0266	Virtual	Rotation Invariant Transformer for Recognizing Object in UAVs -- Shuoyi Chen (Wuhan University), Mang Ye (Wuhan University), Bo Du (Wuhan University)
		mmfp0335	Virtual	Active Learning for Point Cloud Semantic Segmentation via Spatial-Structural Diversity Reasoning -- Feifei Shao (Zhejiang University), Yawei Luo (Zhejiang University, Tsinghua University), Ping Liu (Institute of High Performance Computing, Singapore, A*STAR), jie Chen (Zhejiang University), Yi Yang (Zhejiang University), Yulei Lu (Zhejiang University), Jun Xiao (Zhejiang University)
		mmfp0381	On-site	Free-Lunch for Cross-Domain Few-Shot Learning: Style-Aware Episodic Training with Robust Contrastive Learning -- Ji Zhang (University of Electronic Science and Technology of China), Jingkuan Song (University of Electronic Science and Technology of China, Tsinghua University), Lianli Gao (University of Electronic Science and Technology of China, Tsinghua University), Hengtao Shen (University of Electronic Science and Technology of China)
		mmfp0457	Virtual	ProDiff: Progressive Fast Diffusion Model for High-Quality Text-to-Speech -- Rongjie Huang (Zhejiang University), Zhou Zhao (Zhejiang University), Huadai Liu (Zhejiang University), Jinglin Liu (Zhejiang University), Chenye Cui (Zhejiang University), Yi Ren (Zhejiang University)
		mmfp0654	Virtual	Joint Learning Content and Degradation Aware Feature for Blind Super-Resolution -- Yifeng Zhou (University of Electronic Science and Technology of China), Chuming Lin (Fudan University), Donghao Luo (Tencent YouTu Lab), Yong Liu (, Chinese Academy of Sciences), Ying Tai (Tencent), Chengjie Wang (Tencent YouTu Lab), Mingang Chen (Shanghai Development Center of Computer Software Technology)
		mmfp0996	Virtual	Self-Aligned Concave Curve: Illumination Enhancement for Unsupervised Adaptation -- Wenjing Wang (Peking University), Zhengbo Xu (Peking University), Huang Haofeng (Peking University), Jiaying Liu (Peking University)
		mmfp1498	Virtual	Photorealistic Style Transfer via Adaptive Filtering and Channel Seperation -- Hong Ding (Guangxi University of Finance and Economics), Fei LUO (Wuhan University), Caoqing Jiang (Guangxi University of Finance and Economics), Gang Fu (Wuhan University), Zipei Chen (Wuhan University), Shenghong Hu (Hubei University of Economics), Chunxia Xiao (Wuhan University)
		mmfp1563	Virtual	Recurrent Meta-Learning against Generalized Cold-start problem in CTR Prediction -- Junyu Chen (Institute of Information Engineering, Chinese Academy of Sciences), Qianqian Xu (Institute of Computing Technology, Chinese Academy of Sciences), Zhiyong Yang (Chinese Academy of Sciences), Ke Ma (University of Chinese Academy of Sciences), Xiaochun Cao (Institute of Information Engineering), Qingming Huang (University of Chinese Academy of Sciences)
		mmfp1870	Virtual	Learning Projection Views for Sparse-View CT Reconstruction -- Liutao Yang (Nanjing University of Aeronautics and Astronautics), Rongjun Ge (Nanjing University of Aeronautics and Astronautics), Shichang Feng (Nanjing University of Aeronautics and Astronautics), Daoqiang Zhang (Nanjing University of Aeronautics and Astronautics, Tsinghua University)
		mmfp2364	Virtual	Unsupervised Textured Terrain Generation via Differentiable Rendering -- Peichi Zhou (East China Normal University), Dingbo Lu (East China Normal University), Chen Li (East China Normal University), Jian Zhang (East China Normal University), Long Liu (East China Normal University), Changbo Wang (East China Normal University)
		mmfp0394	On-site	MegaPortraits: One-shot Megapixel Neural Head Avatars -- Nikita Drobyshev (Skolkovo Institute of Science and Technology), Jenya Chelishev (Technion - Israel Institute of Technology, Technion), Taras Khakhulin (skolkovo university of science and technology), Aleksei Ivakhnenko (Samsung), Victor Lempitsky (Skolkovo Institute of Science and Technology), Egor Zakharov (Skolkovo Institute of Science and Technology)
		mmfp1657	Virtual	Event-guided Video Clip Generation from Blurry Images -- Xin Ding (NingboTech University), Tsuyoshi Takatani (University of Tsukuba), Zhongyuan Wang (Wuhan University), Ying Fu (Beijing Institute of Technology), Yinqiang Zheng (The University of Tokyo, Tokyo Institute of Technology)
		mmfp0753	Virtual	Consistency-Contrast Learning for Conceptual Coding -- Jianhui Chang (Peking University), Jian Zhang (Peking University), Youmin Xu (Peking University), Jiguo Li (Peking University), Siwei Ma (Peking University), Wen Gao (Peking University)
		mmfp0624	Virtual	Order-aware Human Interaction Manipulation -- Mandi Luo (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Jie Cao (Institute of automation, Chinese academy of science, Chinese Academy of Sciences), Ran He (Institute of automation, Chinese academy of science, Chinese Academy of Sciences)
		mmfp1353	Virtual	Semi-supervised Video Shadow Detection via Image-assisted Pseudo-label Generation -- Zipei Chen (Wuhan University), Xiao Lu (Hunan Normal University), Ling Zhang (Wuhan University), Chunxia Xiao (Wuhan University)
		mmfp0345	Virtual	Towards Robust Video Object Segmentation with Adaptive Object Calibration -- Xiaohao Xu (Huazhong University of Science and Technology), Jinglu Wang (Microsoft Research Asia), Xiang Ming (Xi'an Jiaotong University), Yan Lu (Microsoft Research Asia)
		mmfp1237	On-site	Split-PU: Hardness-aware Training Strategy for Positive-Unlabeled Learning -- Chengming Xuu (Fudan University), Chen Liu (The Hong Kong University of Science and Technology), Siqian Yang (Tongji University), Yabiao Wang (Tencent Youtu Lab), Shijie Zhang (Fudan University), Lijie Jia (Shanghai Jiaotong University), Yanwei Fu (Fudan University)
		mmfp2995	Virtual	Multi-Camera Collaborative Depth Prediction via Consistent Structure Estimation -- Jialei Xu (Harbin Institute of Technology), Xianming Liu (Harbin Institute of Technology), Yuanchao Bai (Harbin Institute of Technology), Junjun Jiang (Harbin Institute of Technology), Kaixuan Wang (Hong Kong University of Science and Technology), Xiaozhi Chen (Electronic Engineering, Tsinghua University, Tsinghua University), Xiangyang Ji (Tsinghua University, Tsinghua University)
		mmfp2955	Virtual	Fast Hierarchical Deep Unfolding Network for Image Compressed Sensing -- Wenxue Cui (Harbin Institute of Technology), Shaohui Liu (Harbin Institute of Technology), Debin Zhao (Harbin Institute of Technology)
		mmfp0431	Virtual	Restoration of User Videos Shared on Social Media -- Hongming Luo (Department of Electronics and information engineering, Shenzhen University), Fei Zhou (Shenzhen University), Kin-man Lam (The Hong Kong Polytechnic University), Guoping Qiu (University of Nottingham)
		mmfp0764	On-site	Real-time Streaming Video Denoising with Bidirectional Buffers -- Chenyang Qi (Hong Kong University of Science and Technology), Junming Chen (Hong Kong University of Science and Technology), Xin Yang (Hong Kong University of Science and Technology), Qifeng Chen (Hong Kong University of Science and Technology)
		mmfp2516	Virtual	Learning Hierarchical Dynamics with Spatial Adjacency for Image Enhancement -- Yudong Liang (School of Computer and Information Technology, Shanxi University), bin wang (Shanxi University), Wenqi Ren (Chinese acadamy of sciences), Jiaying Liu (Peking University), Wenjian Wang (Shanxi University), Wangmeng Zuo (Harbin Institute of Technology)
		mmfp1488	Virtual	Text's Armor: Optimized Local Adversarial Perturbation Against Scene Text Editing Attacks -- Tao Xiang (Chongqing University), Hangcheng Liu (Chongqing University), Shangwei Guo (Chongqing University), Hantao Liu (Cardiff University), Tianwei Zhang (Nanyang Technological University)
		mmfp2309	Virtual	ChartStamp: Robust Chart Encoding for Real-World Applications -- Jiayun Fu (Huazhong University of Science and Technology), Bin Zhu (Microsoft Research), Haidong Zhang (Microsoft Research Asia), Yayi Zou (Huazhong University of Science and Technology), Song Ge (Research, Microsoft), Weiwei Cui (Research, Microsoft), Yun Wang (Microsoft Research), Dongmei Zhang (Microsoft), Xiaojing Ma (Huazhong University of Science and Technology), Hai Jin (Huazhong University of Science and Technology)
		mmfp1685	Virtual	Few-shot Image Generation Using Discrete Content Representation -- Yan Hong (Shanghai Jiao Tong University), Li Niu (Shanghai Jiao Tong University), Jianfu Zhang (RIKEN), Liqing Zhang (Shanghai Jiao Tong University)
		mmfp1899	Virtual	Marior: Margin Removal and Iterative Content Rectification for Document Dewarping in the Wild -- Jiaxin Zhang (South China University of Technology), Canjie Luo (South China University of Technology), Lianwen Jin (South China University of Technology), Fengjun Guo (Shanghai Jiaotong University), Kai Ding (INTSIG Information)
		mmfp0727	On-site	Image Inpainting Detection via Enriched Attentive Pattern with Near Original Image Augmentation -- Wenhan Yang (Nanyang Technological University), RIZHAO CAI (Nanyang Technological University), Alex Kot (Nanyang Technological University)
		mmfp2316	On-site	Searching Lightweight Neural Network for Image Signal Processing -- Haojia Lin (Xiamen University), lijiang Li (Xiamen University), Xiawu Zheng (PengCheng Lab), Fei Chao (Xiamen University), Rongrong Ji (Columbia University)
		mmfp1652	Virtual	Image Generation Network for Covert Transmission in Online Social Network -- Zhengxin You (Fudan University), Ying Qichao (Fudan University), Sheng Li (Fudan University), Zhenxing Qian (Fudan University), Xinpeng Zhang (Fudan University)
		mmfp1858	Virtual	Augmented Dual-Contrastive Aggregation Learning for Unsupervised Visible-Infrared Person Re-Identification -- Bin Yang (Wuhan University), Mang Ye (Wuhan University), Jun Chen (Wuhan University), Zesen Wu (Wuhan University)
		mmfp0011	On-site	DrawMon: A Distributed System for Detection of Atypical Sketch Content in Concurrent Pictionary Games -- Nikhil Bansal (International Institute of Information Technology Hyderabad, Dhirubhai Ambani Institute Of Information and Communication Technology), Kartik Gupta (International Institute of Information Technology, Hyderabad), Kiruthika Kannan (International Institute of Information Technology Hyderabad), Sivani Pentapati (International Institute of Information Technology Hyderabad), Ravi Sarvadevabhatla (International Institute of Information Technology Hyderabad)
		mmfp2440	Virtual	Approximate Shifted Laplacian Reconstruction for Multiple Kernel Clustering -- Jiali You (Southwest University of Science and Technology), Zhenwen Ren (Southwest University Of Science And Technology), Quansen Sun (Nanjing University of Science and Technology), Yuan Sun (Sichuan University), Li Feng (Nanjing University of Science and Technology)
		mmfp1989	Virtual	Towards Continual Adaptation in Industrial Anomaly Detection -- Wujin Li (Tsinghua University, Tsinghua University), Jiawei Zhan (Tencent), Jinbao Wang (South University of Science and Technology of China), Bizhong Xia (Tsinghua University, Tsinghua University), Bin-Bin Gao (Tencent YouTu Lab), Jun Liu (Tencent YouTu Lab), Chengjie Wang (Tencent YouTu Lab), Feng Zheng (Southern University of Science and Technology)
		mmfp1222	On-site	SDRTV-to-HDRTV via Hierarchical Dynamic Context Feature Mapping -- Gang He (Xidian University), Kepeng Xu (Xidian University), Li Xu (School of Telecommunications Engineering, Xidian University), Chang Wu (Xidian University), Ming Sun (Kuaishou Tech), Xing Wen (Kuaishou Technology), Yu-wing Tai (Kuaishou Technology)
		mmfp1052	Virtual	Arbitrary Bit-width Network: A Joint Layer-Wise Quantization and Adaptive Inference Approach -- Chen Tang (Tsinghua University), Haoyu Zhai (Tsinghua University), Kai Ouyang (Tsinghua University, Tsinghua University), Zhi Wang (SIGS, Tsinghua University), Yifei Zhu (Shanghai Jiaotong University), Wenwu Zhu (Tsinghua University, Tsinghua University)
		mmfp2933	On-site	Privacy-preserving Reflection Rendering for Augmented Reality -- Yiqin Zhao (Worcester Polytechnic Institute), Sheng Wei (Rutgers University), Tian Guo (Worcester Polytechnic Institute)
Experience: Art and Culture		mmfp0269	Virtual	Exploring Negatives in Contrastive Learning for Unpaired Image-to-Image Translation -- Yupei Lin (Guangdong University of Technology), Sen Zhang (The University of Sydney, University of Sydney), Tianshui Chen (Guangdong University of Technology), Yongyi Lu (Johns Hopkins University), Guangping Li (Guangdong University of Technology), Yukai Shi (Guangdong University of Technology)
		mmfp0741	On-site	Sundial-GAN: A Cascade Generative Adversarial Networks Framework for Deciphering Oracle Bone Inscriptions -- Xiang Chang (University of Wales, Aberystwyth), Fei Chao (Xiamen University), Changjing Shang (Aberystwyth University), Qiang Shen (Aberystwyth University)
		mmfp1383	On-site	Structure-Enhanced Pop Music Generation via Harmony-Aware Learning -- Xueyao Zhang (The Chinese University of Hong Kong, Shenzhen), Jinchao Zhang (WeChat AI), Yao Qiu (Tencent WeChat AI), Li Wang (Communication University of China), Jie Zhou (WeChat AI, Tencent Inc.)
		mmfp2227	Virtual	Dynamic Weighted Semantic Correspondence for Few-Shot Image Generative Adaptation -- Xingzhong Hou (State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences), Boxiao Liu (State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences), Shuai Zhang (, Chinese Academy of Sciences), Lulin Shi (Hong Kong University of Science and Technology), zite Jiang (Hangzhou Dianzi University), Haihang You (Institute of Computing Technology, Chinese Academy of Sciences)
		mmfp1618	On-site	The Beauty of Repetition in Machine Composition Scenarios -- Zhejing Hu (Hong Kong Polytechnic University), Xiao MA (Hong Kong Polytechnic University), Yan Liu (The Hong Kong Polytechnic University), Gong Chen (Hong Kong Polytechnic University), Yongxu Liu (Hong Kong Polytechnic University)
		mmfp1391	Virtual	CariPainter: Sketch Guided Interactive Caricature Generation -- Xin Huang (Tongji University), Dong Liang (Tongji University), Hongrui Cai (University of Science and Technology of China), Juyong Zhang (University of Science and Technology of China), Jinyuan Jia (Tongji University)
		mmfp1447	Virtual	Cartoon-Flow: A Flow-Based Generative Adversarial Network for Arbitrary-Style Photo Cartoonization -- Jieun Lee (Korea University), Hyeonwoo Kim (Korea University), Jonghwa Shim (Korea University), Eenjun Hwang (Korea University)
Experience: Interactions and Quality of Experience		mmfp2324	On-site	Angular Gap: reducing the uncertainty of image difficulty through model calibration -- Bohua Peng (Imperial College London, Imperial College London), Mobarakol Islam (Imperial College London), peng zhang (Xidian University), Mei Tu (Samsung)
		mmfp3179	Virtual	GCL: Graph Calibration Loss for Trustworthy Graph Neural Network -- Min Wang (National University of Defense Technology), Hao Yang (National University of Defense Technology), Qing Cheng (National University of Defense Technology)
		mmfp0513	Virtual	Image Quality Assessment: From Mean Opinion Score to Opinion Score Distribution -- Yixuan Gao (Shanghai Jiaotong University), Xiongkuo Min (Shanghai Jiao Tong University), Yucheng Zhu (Shanghai Jiaotong University), Jing Li (Alibaba Group), Xiao-Ping Zhang (Ryerson University), Guangtao Zhai (Shanghai Jiao Tong University)
		mmfp0953	Virtual	No-Reference Image Quality Assessment using Dynamic Complex-valued Neural Network -- Zihan Zhou (South China University of Technology), Yong Xu (Peng Cheng Laboratory), Ruotao Xu (South China University of Technology), Yuhui Quan (South China University of Technology)
		mmfp1612	Virtual	Hybrid Conditional Deep Inverse Tone Mapping -- Tong Shao (Harbin Institute of Technology), Deming Zhai (Harbin Institute of Technology), Junjun Jiang (Harbin Institute of Technology), Xianming Liu (Harbin Institute of Technology)
		mmfp1862	Virtual	Where Are You Looking?: A Large-Scale Dataset of Head and Gaze Behavior for 360-Degree Videos and a Pilot Study -- Yili Jin (The Chinese University of Hong Kong, Shenzhen), Junhua Liu (The Chinese University of Hong Kong,shenzhen), Fangxin Wang (The Chinese University of Hong Kong, Shenzhen), Shuguang Cui (The Chinese University of Hong Kong, Shenzhen)
Multimedia Systems: Transport and Delivery		mmfp2823	Virtual	Exploring Spherical Autoencoder for Spherical Video Content Processing -- Jin Zhou (George Mason University), Na Li (Rutgers University, New Brunswick), Yao Liu (Rutgers University), Shuochao Yao (George Mason University), Songqing Chen (George Mason University)
		mmfp0026	Virtual	Sophon: Super-Resolution Enhanced 360° Video Streaming with Visual Saliency-aware Prefetch -- Jianxin Shi (Nankai University), Lingjun Pu (Nankai University), Xinjing Yuan (Nankai University), Qianyun Gong (Nankai University), Jingdong Xu (Nankai University)
		mmfp2931	On-site	Error Concealment of Dynamic 3D Point Cloud Streaming -- Tzu-Kuan Hung (National Tsinghua University), I-Chun Huang (Department of Computer Science, National Tsing Hua University, National Tsinghua University), Samuel Cox (national university of singaore, National University of Singapore), Wei Ooi (National University of Singapore), Cheng-Hsin Hsu (Department of Computer Science, National Tsing Hua University, National Tsinghua University)
		mmfp1246	Virtual	Personalized 360-Degree Video Streaming: A Meta-Learning Approach -- Yiyun Lu (Shanghai Jiaotong University), Yifei Zhu (Shanghai Jiaotong University), Zhi Wang (SIGS, Tsinghua University)

Systems: Data Systems Management and		mmfp2281	On-site	Re-ordered Micro Image based High Efficient Residual Coding in Light Field Compression -- Hyunmin Jung (Seoul National University), Hyuk-Jae Lee (Seoul National University), Chae Eun Rhee (Inha University)
		mmfp3103	On-site	Accelerating general-purpose lossless compression via simple and scalable parameterization -- YU MAO (City University of Hong Kong), Yufei CUI (, McGill University), Tei-Wei Kuo (City University of Hong Kong), Chun Jason Xue (City University of Hong Kong)
Multimedia Systems: Systems and Middleware		mmfp1739	On-site	Physical Backdoor Attacks to Lane Detection Systems in Autonomous Driving -- XINGSHUO HAN (Nanyang Technological University), Guowen Xu (Nanyang Technological University), Yuan Zhou (Nanyang Technological University), Xuehuan Yang (Nanyang Technological University), Jiwei Li (Stanford University), Tianwei Zhang (Nanyang Technological University)
		mmfp1967	Virtual	Dynamic Transformer for Few-shot Instance Segmentation -- Haochen Wang (University of Amsterdam), Jie Liu (University of Amsterdam), Yongtuo Liu (University of Amsterdam), Subhransu Maji (Department of Computer Science, University of Massachusetts, Amherst), Jan-jakob Sonke (Netherlands Cancer Institute), Efstratios Gavves (University of Amsterdam)
Demo Session 3 [Main Foyer]	**10h30-11h45**			*(Demos might be available the whole day)*
E-commerce, Video, and Sports		mmde037	On-site	Restoration of Analog Videos Using Swin-UNet -- Marco Bertini (Università degli Studi di Firenze), Lorenzo Agnolucci (University of Florence), Leonardo Galteri (University of Florence), Alberto Del Bimbo (Università degli Studi di Firenze)
		mmde007	On-site	Developing Embodied Conversational Agents in the Unreal Engine: the FANTASIA plugin -- Antonio Origlia (University of Naples Federico II), Martina Di Bratto (University of Naples Federico II), Maria Di Maro (University of Naples Federico II), Sabrina Mennella (University of Catania)
		mmde030	Virtual	A High-resolution Image-based Virtual Try-on System in Taobao E-commerce Scenario -- Zhilong Zhou (Alibaba Group), Shiyao Wang (Alibaba Group), Tiezheng Ge (Alibaba Group), Yuning Jiang (University of Science and Technology of China, Tsinghua University)
		mmde033	Virtual	WOC: A Handy Webcam-based 3D Online Chatroom -- Chuanhang Yan (Beijing Institute of Technology), Yu Sun (Harbin Institute of Technology), Qian Bao (JD Explore Academy), Jinhui Pang (Beijing Institute of Technology), Wu Liu (JD Explore Academy), Tao Mei (JD Explore Academy)
		mmde034	On-site	BetterSight: Immersive Vision Training for Basketball Players -- PinXuan Liu (Department of Computer Science, National Tsing Hua University, National Tsinghua University), Tse-Yu Pan (National Tsing Hua University), Hsin-Shih Lin (National Cheng Kung University), Hung-Kuo Chu (National Tsing Hua University), Min-Chun Hu (Department of Computer Science, National Tsing Hua University, National Tsing Hua University)
		mmde040	On-site	ScoreActuary: Hoop-Centric Trajectory-Aware Network for Fine-Grained Basketball Shot Analysis -- Ting-Yang Kao (National Tsing Hua University), Tse-Yu Pan (National Tsing Hua University), Chen-Ni Chen (National Tsing Hua University), Tsung-Hsun Tsai (National Tsing Hua University), Hung-Kuo Chu (National Tsing Hua University), Min-Chun Hu (Department of Computer Science, National Tsing Hua University, National Tsing Hua University)
		mmde043	On-site	A Conversational Shopping Assistant for Online Virtual Stores -- Diogo Tavares (Universidade NOVA de Lisboa), Pedro Valente (Universidade NOVA de Lisboa), Rafael Ferreira (Universidade NOVA de Lisboa), Diogo Silva (Universidade NOVA de Lisboa), David Semedo (Universidade NOVA de Lisboa), Joao Magalhaes (Universidade Nova de Lisboa), Nuno Correia (Universidade Nova de Lisboa), Tiago Fornelos (Universidade NOVA de Lisboa)
Art Exhibition	**All day**			*(See Art Exhibition on Tuesday)*
Industry Session 2 [Pav. 4 - Room 1.07]	**11h30-13h**			
Paper session Ind2: User Modeling, Indexing, and Information Extraction	11h30	mmind006	Virtual	Learning Interest-oriented Universal User Representation via Self-supervision -- Qinghui Sun (South China University of Technology), Jie Gu (Alibaba Group), Xiaoxiao Xu (University of Science and Technology of China), Renjun Xu (Zhejiang University), Ke Liu (Zhejiang University), Bei Yang (Alibaba Group), Hong Liu (Alibaba Group), huan xu (Georgia Institute of Technology)
	11h45	mmind008	Virtual	MMH-index: Enhancing Apache Lucene with High-Performance Multi-Modal Indexing and Searching -- Ruicheng Liu (N/A), Liang Jialing (University of Science and Technology of China), Peiquan Jin (University of Science and Technology of China), Yi Wang (N/A)
	12h	mmind002	On-site	Layout-Aware Information Extraction for Document-Grounded Dialogue: Dataset, Method and Demonstration -- Zhenyu Zhang (Chinese Academy of Sciences), Bowen Yu (Chinese Academy of Sciences), Haiyang Yu (Xiamen University), Tingwen Liu (Institute of Information Engineering, Chinese Academy of Sciences), Cheng Fu (N/A), Jingyang Li (Alibaba Group), chengguang tang (Alibaba Group), Jian SUN (Alibaba Group, DAMO Academy)
Paper session Ind3: Multimedia Search and Privacy	12h15	mmind013	Virtual	Learnable Privacy-Preserving Anonymization for Pedestrian Images -- Junwu Zhang (Wuhan University), Mang Ye (Wuhan University), Yao Yang (Zhejiang Lab China)
	12h30	mmind015	Virtual	Pyramidal Transformer with Conv-Patchify for Person Re-identification -- He Li (Wuhan University), Mang Ye (Wuhan University), Cong Wang (Huawei Technologies Ltd.), Bo Du (Wuhan University)
	12h45	mmind014	Virtual	Few-Shot Model Agnostic Federated Learning -- Wenke Huang (Wuhan University), Mang Ye (Wuhan University), Bo Du (Wuhan University), Xiang Gao (NSFOCUS Technologies Group Co Ltd)
Oral Session 7a [Auditório I]	**11h45-13h**	**Chair: Xavier Giro-i-Nieto (Politechnical University of Catalonia)**		
Multimedia in the Real World	11h45	mmfp2141	Virtual	Visual Knowledge Graph for Human Action Reasoning in Videos -- ma yue (Tsinghua University, Tsinghua University), Yali Wang (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Chinese Academy of Sciences), Yue WU (Chongqing University of Post and Telecommunications), Ziyu Lyu (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Chinese Academy of Sciences), siran Chen (University of Chinese Academy of Sciences), Xiu Li (Tsinghua University, Tsinghua University), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences)
	12h03	mmfp2293	Virtual	Target-Driven Structured Transformer Planner for Vision-Language Navigation -- Yusheng Zhao (Beijing University of Aeronautics and Astronautics), Jinyu Chen (Beihang University), Chen Gao (Beihang University), Wenguan Wang (University of Technology Sydney), Lirong Yang (Meituan Inc.), Haibing Ren (Meituan Inc.), Huaxia Xia (Meituan), Si Liu (Beihang University)
	12h21	mmfp2298	Virtual	Cross-modal Semantic Alignment Pre-training for Vision-and-Language Navigation -- Siying Wu (University of Science and Technology of China), Xueyang Fu (University of Science and Technology of China), Feng Wu (University of Science and Technology of China), Zheng-Jun Zha (University of Science and Technology of China)
	12h39	mmfp3126	Virtual	Towards Further Comprehension on Referring Expression with Rationale -- Rengang Li (Inspur Electronic Information Industry Co.,Ltd), Baoyu Fan (Inspur Electronic Information Industry Co.,Ltd.), Xiaochuan Li (Inspur Electronic Information Industry Co.,Ltd), Runze Zhang (Inspur Electronic Information Industry Co.,Ltd), Zhenhua Guo (Inspur Electronic Information Industry Co., Ltd), Kun Zhao (Inspur Electronic Information Industry Co.,Ltd), Yaqian Zhao (Inspur Electronic Information Industry Co.,Ltd), Weifeng Gong (Inspur Electronic Information Industry Co.,Ltd), Endong Wang (Inspur Electronic Information Industry Co.,Ltd)
Oral Session 7b [Auditório II]	**11h45-13h**	**Chair: Bhojan Anand (National University of Singapore)**		
Content Generation and Experience	11h45	mmfp0370	Virtual	Wavelet-enhanced Weakly Supervised Local Feature Learning for Face Forgery Detection -- Jiaming Li (University of Science and Technology of China), Hongtao Xie (University of Science and Technology of China), Lingyun Yu (University of Science and Technology of China), Yongdong Zhang (University of Science and Technology of China)
	12h03	mmfp1823	Virtual	Multi-Modal Experience Inspired AI Creation -- Qian Cao (Renmin University of China), Xu Chen (Renmin University of China), Ruihua Song (Renmin University of China), Hao Jiang (Huawei Technologies Ltd.), Guang Yang (Huawei Technologies Ltd.), Zhao Cao (Huawei Technologies Ltd.)
	12h21	mmfp1523	Virtual	On Generating Identifiable Virtual Faces -- Zhuowen Yuan (University of Illinois at Urbana-Champaign), Zhengxin You (Fudan University), Sheng Li (Fudan University), Zhenxing Qian (Fudan University), Xinpeng Zhang (Fudan University), Alex Kot (Nanyang Technological University)
	12h39	mmfp0810	Virtual	Geometric warping error aware CNN for DIBR oriented view synthesis -- Shuai Li (Shandong University), KaiXin Wang (Shandong University), Yanbo Gao (Shandong University), Xun Cai (Shandong University), Mao Ye (University of Electronic Science and Technology of China)
Oral Session 7c [Auditório III]	**11h45-13h**	**Chair: Laura Toni (University College London)**		
Spatio-temporal Priors for Multimedia	11h45	mmfp0698	Virtual	Deepfake Video Detection with Spatiotemporal Dropout Transformer -- DaiChi Zhang (University of the Chinese Academy of Sciences), Fanzhao Lin (Chinese Academy of Sciences), Yingying Hua (Institute of Information Engineering, Chinese Academy of Sciences), Pengju Wang (University of the Chinese Academy of Sciences), Dan Zeng (Shanghai University), Shiming Ge (Institute of Information Engineering, Chinese Academy of Sciences)
	12h03	mmfp1247	Virtual	AEDNet: Asynchronous Event Denoising with Spatial-Temporal Correlation among Irregular Data -- Huachen Fang (Xi'an University of Electronic Science and Technology), Jinjian Wu (Xidian University), Leida Li (Xidian University), Junhui Hou (City University of Hong Kong), Weisheng Dong (Xidian University), Guangming Shi (Xidian University)
	12h21	mmfp0422	Virtual	Hybrid Spatial-Temporal Entropy Modelling for Neural Video Compression -- Jiahao Li (Microsoft Research Asia), Bin Li (Microsoft), Yan Lu (Microsoft Research Asia)
	12h39	mmfp0525	Virtual	You Only Align Once: Bidirectional Interaction for Spatial-Temporal Video Super-Resolution -- Mengshun Hu (Wuhan University), Kui Jiang (Wuhan University), Zhixiang Nie (Wuhan University), Zheng Wang (Wuhan University, China)
Oral Session 7d [Auditório IV]	**11h45-13h**	**Chair: Stefano Petrangeli (Adobe)**		
Video Streaming and Quality	11h45	mmfp2082	On-site	DAO: Dynamic Adaptive Offloading for Video Analytics -- taslim murad (Georgia State University), Anh Nguyen (George Mason University), Zhisheng Yan (George Mason University)
	12h03	mmfp1192	On-site	AdaMask: Enabling Machine-Centric Video Streaming with Adaptive Frame Masking for DNN Inference Offloading -- Shengzhong Liu (University of Illinois at Urbana-Champaign), Tianshi Wang (Department of Computer Science, University of Illinois at Urbana-Champaign), Jinyang Li (University of Illinois at Urbana-Champaign), Dachun Sun (Department of Computer Science, University of Illinois at Urbana-Champaign), Mani Srivastava (University of California, Los Angeles), Tarek Abdelzaher (University of Illinois, Urbana Champaign)
	12h21	mmfp1799	On-site	InDiD: Instant Disorder Detection via a Principled Neural Network -- Evgenia Romanenkova (Skolkovo Institute of Science and Technology), Alexander Stepikin (Moscow Institute of Physics and Technology), Matvey Morozov (Moscow Institute of Physics and Technology), Alexey Zaytsev (Skoltech)
	12h39	mmfp2848	Virtual	R-FEC: RL-based FEC Adjustment for better QoE in WebRTC -- Insoo Lee (University of Colorado at Boulder), Seyeon Kim (KAIST, Korea Advanced Institute of Science & Technology), Sandesh Sathyanarayana (University of Colorado at Boulder), Kyungmin Bin (Seoul National University), Song Chong (Korea Advanced Institute of Science and Technology), Kyunghan Lee (Seoul National University), Dirk Grunwald (University of Colorado at Boulder), Sangtae Ha (University of Colorado at Boulder)
Virtual Only Poster Session 3 [Gathertown]	**13h-14h30**			*(See Poster Session 3)*
Oral Session 8a [Auditório I]	**14h30-15h45**	**Chair: Svebor Karaman (Dataminr)**		
Avdanced Learning for Multimedia II	14h30	mmfp2140	Virtual	Pay Attention to Your Positive Pairs: Positive Pair Aware Contrastive Knowledge Distillation -- Zhipeng Yu (University of the Chinese Academy of Sciences), Qianqian Xu (Institute of Computing Technology, Chinese Academy of Sciences), Yangbangyan Jiang (Institute of Information Engineering, Chinese Academy of Sciences), Haoyu Qin (SenseTime Ltd.), Qingming Huang (University of Chinese Academy of Sciences)
	14h48	mmfp0310	Virtual	RONF: Reliable Outlier Synthesis under Noisy Feature Space for Out-of-Distribution Detection -- Rundong He (Shandong University), Zhongyi Han (Shandong University), Xiankai Lu (Shandong University, Tsinghua University), Yilong Yin (Shandong University)
	15h06	mmfp1271	On-site	IDEAL: High-Order-Ensemble Adaptation Network for Learning with Noisy Labels -- Peng-Fei Zhang (University of Queensland), Zi Huang (University of Queensland), Guangdong Bai (University of Queensland), Xin-Shun Xu (Shandong University)
	15h24	mmfp0895	Virtual	Bayesian based Re-parameterization for DNN Model Pruning -- Xiaotong Lu (Xi'an University of Electronic Science and Technology), TENG XI (Baidu), Baopu Li (The Chinese University of Hong Kong), Gang Zhang (Baidu), Weisheng Dong (Xidian University)
Oral Session 8b [Auditório II]	**14h30-15h45**	**Chair: Huijuan Xu (Penn State University)**		
Zero/Few-shot Learning for Multimedia	14h30	mmfp0565	Virtual	Dual Part Discovery Network for Zero-Shot Learning -- Jiannan Ge (University of Science and Technology of China), Hongtao Xie (University of Science and Technology of China), Shaobo Min (University of Science and Technology of China), Pandeng Li (University of Science and Technology of China), Yongdong Zhang (University of Science and Technology of China)
	14h48	mmfp1152	Virtual	CoHOZ: Contrastive Multimodal Prompt Tuning for Hierarchical Open-set Zero-shot Recognition -- Ning Liao (Shanghai Jiaotong University), Yifeng Liu (Alibaba Group), LI XIAOBO (Alibaba Group), Chenyi Lei (Alibaba Group), guoxin wang (Zhejiang University), Xian-Sheng Hua (Alibaba Group), Junchi Yan (Shanghai Jiao Tong University)
	15h06	mmfp1074	Virtual	Zero-shot Video Classification with Appropriate Web and Task Knowledge Transfer -- Junbao Zhuo (, Chinese Academy of Sciences), Yan Zhu (University of the Chinese Academy of Sciences), Shuhao Cui (Meituan Group), Shuhui Wang (Institute of Computing Technology, Chinese Academy of Sciences), Bin MA (Meituan), Qingming Huang (Univ of Chinese Academy of Sciences), Wei Xiaoming (Beijing University of Aeronautics and Astronautics), Xiaolin Wei (Meituan)
	15h24	mmfp0909	Virtual	Reducing the Vision and Language Bias for Temporal Sentence Grounding -- Daizong Liu (Peking University), Xiaoye Qu (Huawei Technologies Ltd.), Wei Hu (Peking University)

Oral Session 8c [Auditório III]	**14h30-15h45**	**Chair: Mylene Farias (University of Brasilia)**		
Quality Assessment	14h30	mmfp0754	Virtual	ADGNet: Attention Discrepancy Guided Deep Neural Network for Blind Image Quality Assessment -- Xiaoyu Ma (Communication University of Zhejiang), YAQI WANG (Communication University of Zhejiang), Chang Liu (Communication University of Zhejiang), Suiyu Zhang (Communication University of Zhejiang), Dingguo Yu (Communication University of Zhejiang)
	14h48	mmfp2886	On-site	Domain-Specific Fusion Of Objective Video Quality Metrics -- Aaron Chadha (iSize), Ioannis Katsavounidis (Facebook), Ayan Kumar Bhunia (iSIZE), Cosmin Stejerean (Facebook), Mohammad Umar Khan (iSIZE), Yiannis Andreopoulos (University College London)
	15h06	mmfp0439	On-site	Exploring the Effectiveness of Video Perceptual Representation in Blind Video Quality Assessment -- Liang Liao (Nanyang Technological University), Kangmin Xu (Wuhan University), Haoning Wu (Nanyang Technological University), Chaofeng Chen (Nanyang Technological University), Wenxiu Sun (SenseTime Group Limited), Qiong Yan (SenseTime Research), Weisi Lin (Nanyang Technological University)
	15h24	mmfp2574	Virtual	A Deep Learning based No-reference Quality Assessment Model for UGC Videos -- Wei Sun (Shanghai Jiaotong University), Xiongkuo Min (Shanghai Jiao Tong University), Wei Lu (Shanghai Jiaotong University), Guangtao Zhai (Shanghai Jiao Tong University)
Oral Session 8d [Auditório IV]	**14h30-15h45**	**Chair: Lei Zhu (Hong Kong University of Science and Technology)**		
Image Processing and Enhancement II	14h30	mmfp0373	Virtual	Real-World Blind Super-Resolution via Feature Matching with Implicit High-Resolution Priors -- Chaofeng Chen (Nanyang Technological University), Xinyu Shi (University of Waterloo), Yipeng Qin (Cardiff University), Xiaoming Li (Harbin institute of Technology), Xiaoguang Han (The Chinese University of Hong Kong, Shenzhen), Yang Tao (Tsinghua University, Tsinghua University), Shihui Guo (Xiamen University)
	14h48	mmfp0303	Virtual	ReCoRo: Region-Controllable Robust Light Enhancement by User-Specified Imprecise Masks -- Dejia Xu (University of Texas at Austin), Hayk Poghosyan (Picsart Inc), Shant Navasardyan (Yerevan State University), Yifan Jiang (University of Texas, Austin), Honghui Shi (University of Oregon), Zhangyang Wang (University of Texas, Austin)
	15h06	mmfp0906	Virtual	Learning for Motion Deblurring with Hybrid Frames and Events -- Wen Yang (Xidian University), Jinjian Wu (Xidian University), Jupo Ma (Xidian University), Leida Li (Xidian University), Weisheng Dong (Xidian University), Guangming Shi (Xidian University)
	15h24	mmfp1063	Virtual	Cycle-Interactive Generative Adversarial Network for Robust Unsupervised Low-Light Enhancement -- Zhangkai NI (Tongji University), Wenhan Yang (Nanyang Technological University), Hanli Wang (Tongji University), Shiqi Wang (City University of Hong Kong), Lin Ma (Meituan), Sam Kwong (City University of Hong Kong)
Open Source Competition [Pav4 R1.06]	**15h30-18h**			
Oral Session	15h30	mmos03	On-site	CVNets: High Performance Library for Computer Vision -- Sachin Mehta (Apple), Farzad Abdolhosseini (Apple), Mohammad Rastegari (Apple)
	15h48	mmos06	On-site	MMRotate: A Rotated Object Detection Benchmark using PyTorch -- Yue Zhou (Shanghai Jiao Tong University), Xue Yang (Shanghai Jiao Tong University), Gefan Zhang (Shanghai Jiao Tong University), Jiabao Wang (Northwestern Polytechnical University), Yanyi Liu (Northeastern University), Liping Hou (University of Chinese Academy of Sciences), Xue Jiang (Shanghai Jiao Tong University), Xingzhao Liu (Shanghai Jiao Tong University), Junchi Yan (Shanghai Jiao Tong University), Chengqi Lyu (Shanghai AI Laboratory), Wenwei Zhang (Nanyang Technological University), Kai Chen (SenseTime Research)
	16h04	mmos09	On-site	MoZuMa: a Model Zoo for Multimedia Applications -- Stéphane Massonnet (EPFL), Marco Romanelli (EPFL), Rémi Lebret (EPFL), Niels Poulsen (EPFL), Karl Aberer (EPFL)
Spotlight Session	16h22	mmos12	Virtual	OpenHardwareVC: An Open Source Library for 8K UHD Video Coding Hardware Implementation -- Wei Gao (Peking University; Peng Cheng laboratory), Hang Yuan (Peking University; Peng Cheng laboratory), Yang Guo (Peking University; Peng Cheng laboratory), Lvfang Tao (Peking University; Peng Cheng laboratory), Zhanyuan Cai (Peking University; Peng Cheng laboratory), Ge Li (Peking University)
	16h27	mmos15	Virtual	Low Latency Live Streaming Implementation in DASH and HLS -- Abdelhak Bentaleb (National University of Singapore), Zhengdao Zhan (National University of Singapore), Farzad Tashtarian (Alpen-Adria-Universität Klagenfurt), May Lim (National University of Singapore), Saad Harous (University of Sharjah), Christian Timmerer (Alpen-Adria-Universität Klagenfurt), Hermann Hellwagner (Alpen-Adria-Universität Klagenfurt), Roger Zimmermann (National University of Singapore)
	16h32	mmos18	Virtual	OpenPointCloud: An Open-Source Algorithm Library of Deep Learning Based Point Cloud Compression -- Hua Ye (Peng Cheng Laboratory, China), Wei Gao (School of Electronic and Computer Engineering, Peking University, China), Ge Li (School of Electronic and Computer Engineering, Peking University, China), Huiming Zheng (School of Electronic and Computer Engineering, Peking University, China), Yuyang Wu (School of Electronic and Computer Engineering, Peking University, China), Liang Xie (School of Electronic and Computer Engineering, Peking University, China)
	16h37	mmos21	Virtual	PYSKL: Towards Good Practices for Skeleton Action Recognition -- Haodong Duan (The Chinese University of Hong Kong), Jiaqi Wang (Shanghai AI Laboratory), Kai Chen (Shanghai AI Laboratory), Dahua Lin (The Chinese University of Hong Kong)
	16h42	mmos24	Virtual	DavarOCR: A Toolbox for OCR and Multi-Modal Document Understanding -- Liang Qiao (Hikvision Research Institute), Hui Jiang (Hikvision Research Institute), Ying Chen (Hikvision Research Institute), Can Li (Hikvision Research Institute), Pengfei Li (Hikvision Research Institute), Zaisheng Li (Hikvision Research Institute), Baorui Zou (Hikvision Research Institute), Dashan Guo (Hikvision Research Institute), Yingda Xu (Hikvision Research Institute), Yunlu Xu (Hikvision Research Institute), Zhanzhan Cheng (Hikvision Research Institute), Yi Niu (Hikvision Research Institute)
	16h47	mmos30	Virtual	CurML: A Curriculum Machine Learning Library -- Yuwei Zhou (Tsinghua University), Hong Chen (Tsinghua University), Zirui Pan (Tsinghua University), Chuanhao Yan (Tsinghua University), Fanqi Lin (Tsinghua University), Xin Wang (Tsinghua University), Wenwu Zhu (Tsinghua University)
Poster Session	17h02	ALL Papers		
Panel 2 [Auditório I]	**15h45-17h**			Ethics and Responsible Multimedia AI Research -- Moderator: TBA, Panelists: TBA.
Oral Session 9b [Auditório II]	**15h45-17h**	**Chair: Florian Metze (Meta)**		
Avdanced Learning for Multimedia III	15h45	mmfp0842	Virtual	Robust Low-Rank Convolution Network for Image Denoising -- Ren JiaHuan (Hefei University of Technology), Zhao Zhang (Hefei University of Technology), Richang Hong (Hefei University of Technology), Mingliang Xu (Zhengzhou University), Haijun Zhang (Harbin Institute of Technology, Shenzhen), Mingbo Zhao (Donghua University, Shanghai), Meng Wang (Hefei University of Technology)
	16h03	mmfp1893	Virtual	Weakly-supervised Disentanglement Network for Video Fingerspelling Detection -- Ziqi Jiang (Zhejiang University), Shengyu Zhang (Zhejiang University), Siyuan Yao (Zhejiang University), Wenqiao Zhang (National University of Singapore), Sihan Zhang (The Chinese University of Hong Kong), Juncheng Li (Zhejiang University), Zhou Zhao (Zhejiang University), Fei Wu (Zhejiang University)
	16h21	mmfp3047	Virtual	Breaking Isolation: Multimodal Graph Fusion for Multimedia Recommendation by Edge-wise Modulation -- Feiyu Chen (University of Electronic Science and Technology of China), Junjie Wang (Waseda University), Yinwei Wei (National University of Singapore), Hai-Tao Zheng (Tsinghua University, Tsinghua University), Jie Shao (University of Electronic Science and Technology of China)
	16h39	mmfp2057	Virtual	Neural Network Model Protection with Piracy Identification and Tampering Localization Capability -- Cheng Xiong (University of Shanghai for Science and Technology), Guorui Feng (Shanghai University), Xinran Li (University of Shanghai for Science and Technology), Xinpeng Zhang (Fudan University), Chuan Qin (University of Shanghai for Science and Technology)
Oral Session 9c [Auditório III]	**15h45-17h**	**Chair: Xavier Alameda-Pineda (Inria)**		
Multimodal Fusion III	15h45	mmfp1872	Virtual	LS-GAN: Iterative Language-based Image Manipulation via Long and Short Term Consistency Reasoning -- Gaoxiang Cong (Shandong University), Liang Li (Institute of Computing Technology, Chinese Academy of Sciences), Zhenhuan Liu (ICT, Chinese Academy of Sciences), Yunbin Tu (University of Chinese Academy of Sciences), Weijun Qin (Kuaishou Technology), zhang shenyuan (China People's Daily), Chenggang Yan (Hangzhou Dianzi University, Tsinghua University), Wenyu Wang (Shandong University), Bin. Jiang (Shandong University)
	16h03	mmfp2345	Virtual	CMAL: A Novel Cross-Modal Associative Learning Framework for Vision-Language Pre-Training -- Zhiyuan Ma (Huazhong University of Science and Technology), Jianjun Li (Huazhong University of Science and Technology), Guohui Li (Huazhong University of Science and Technology), kaiyan Huang (Huazhong University of Science and Technology)
	16h21	mmfp1162	Virtual	Progressive Tree-Structured Prototype Network for End-to-End Image Captioning -- Pengpeng Zeng (University of Electronic Science and Technology of China), Jinkuan Zhu (University of Electronic Science and Technology of China), Jingkuan Song (University of Electronic Science and Technology of China, Tsinghua University), Lianli Gao (University of Electronic Science and Technology of China, Tsinghua University)
	16h39	mmfp1304	Virtual	Early-Learning regularized Contrastive Learning for Cross-Modal Retrieval with Noisy Labels -- Tianyuan Xu (Hefei University of Technology), Xueliang Liu (Hefei University of Technology), Zhen Huang (National University of Defense Technology), Dan Guo (Hefei University of Technology), Richang Hong (Hefei University of Technology), Meng Wang (Hefei University of Technology)
Oral Session 9d [Auditório IV]	**15h45-17h**	**Chair: Valérie Gouet-Brunet (National Institute for Geographical and Forest Information)**		
Geometry, 3D and Multimedia	15h45	mmfp0510	Virtual	IVT: An End-to-End Instance-guided Video Transformer for 3D Pose Estimation -- Zhongwei Qiu (University of Science and Technology Beijing), Yang Qiansheng (Baidu), Jian Wang (Baidu), Dongmei Fu (University of Science and Technology Beijing)
	16h03	mmfp1278	Virtual	Class Gradient Projection For Continual Learning -- Chen Cheng (University of Electronic Science and Technology of China), Ji Zhang (University of Electronic Science and Technology of China), Jingkuan Song (University of Electronic Science and Technology of China, Tsinghua University), Lianli Gao (University of Electronic Science and Technology of China, Tsinghua University)
	16h21	mmfp1414	Virtual	GroupDancer: Music to Multi-People Dance Synthesis with Style Collaboration -- Zixuan Wang (Tsinghua University, Tsinghua University), Jia Jia (Department of Computer Science and Technology, Tsinghua University), Haozhe Wu (Tsinghua University, Tsinghua University), Junliang Xing (Department of Computer Science and Technology, Tsinghua University), Jinghe Cai (Tsinghua University, Tsinghua University), Fanbo Meng (Tencent), Guowen Chen (Tencent PCG), YANFENG WANG (Tencent PCG AIBU)
	16h39	mmfp2106	Virtual	MESH2IR: Neural Acoustic Impulse Response Generator for Complex 3D Scenes -- Anton Ratnarajah (University of Maryland, College Park), Zhenyu Tang (University of Maryland, College Park), Rohith Aralikatti (University of Maryland, College Park), Dinesh Manocha (University of Maryland, College Park)
Poster Session 3bis [Pav. 4 - Hall]	**15h45-17h**			*(See Poster Session 3)*
Demo Session 3bis [Main Foyer]	**15h45-17h**			*(See Demo Session 3)*
Keynote Talk: Nuria Oliver [Auditório I]	**17h-18h**			Data Science against COVID-19 -- Nuria Oliver
Closing [Auditório I]	**18h-18h30**			

Friday, October 14th (Workshops and Tutorials)

Workshops

Auditorio II	**9h-13h**			
IMuR 2022: 2nd International Workshop on Interactive Multimedia Retrieval	9h	imur2103	On-site	Video Instructions Linking for Complex Tasks -- Sophie Fischer (University of Glasgow), Carlos Gemmell (University of Glasgow), Iain Mackie (University of Glasgow), Jeffrey Dalton (University of Glasgow)
	9h20	imur3871	On-site	Influence of Late Fusion of High-Level Features on User Relevance Feedback for Videos -- Omar Shahbaz Khan (IT University of Copenhagen), Jan Zahálka (Technical University in Prague), Björn Þór Jónsson (Reykjavik University)
	9h40	imur4349	On-site	Combining Semantic and Visual Image Graphs for Efficient Search and Exploration of Large Dynamic Image Collections -- Kai Uwe Barthel (HTW Berlin), Nico Hezel (HTW Berlin), Konstantin Schall (HTW Berlin), Klaus Jung (HTW Berlin)
	10h	**Break**		
	10h30	imur4490	On-site	An Asynchronous Scheme for the Distributed Evaluation of Interactive Multimedia Retrieval -- Loris Sauter (University of Basel), Ralph Gasser (University of Basel), Abraham Bernstein (University of Zurich), Heiko Schuldt (University of Basel), Luca Rossetto (University of Zurich)
	10h50	imur7986	On-site	Impact of Blind Image Quality Assessment on the Retrieval of Lifelog Images -- Ricardo Ribeiro (University of Aveiro), Alina Trifan (University of Aveiro), António J. R. Neves (University of Aveiro)
	11h10	imur8568	On-site	Graph-Based Implementation of the M -- Nikolaj Mertz (IT University of Copenhagen), Björn Þór Jónsson (Reykjavik University), Aaron Duane (IT University of Copenhagen)

Auditorio III	**9h-18h**			
APCCPA 2022: 1st International Workshop on Advances in Point Cloud Compression, Processing and Analysis		apccpa03	Virtual	IPDAE: Improved Patch-Based Deep Autoencoder for Lossy Point Cloud Geometry Compression -- Kang You (Nanjing University of Aeronautics and Astronautics), Pan Gao (Nanjing University of Aeronautics and Astronautics), Qing Li (Southern University of Science and Technology)
		apccpa04	On-site	GRASP-Net: Geometric Residual Analysis and Synthesis for Point Cloud Compression -- Jiahao Pang (InterDigital), Muhammad Asad Lodhi (InterDigital), Dong Tian (InterDigital)
		apccpa05		Wiener Filter-Based Point Cloud Adaptive Denoising for Video-based Point Cloud Compression -- Jinrui Xing (Shandong University), Hui Yuan (Shandong University), Chen Chen (Shandong University), Tian Guo (Shandong University)
		apccpa06		OpenPointCloud-V2: A Deep Learning Based Open-Source Algorithm Library of Point Cloud Processing -- Yongchi Zhang (Peng Cheng Laboratory), Wei Gao (Peking University Shenzhen Graduate School), Ge Li (Peking University Shenzhen Graduate School)
		apccpa07	Virtual	End-to-End Point Cloud Geometry Compression and Analysis with Sparse Tensor -- Liang Xie (Peking University Shenzhen Graduate School : Peng Cheng Laboratory), Wei Gao (Peking University Shenzhen Graduate School : Peng Cheng Laboratory), Hui Ming Zheng (Peking University Shenzhen Graduate School : Peng Cheng Laboratory)
		apccpa08	Virtual	Transformer and Upsampling-Based Point Cloud Compression -- Junteng Zhang (Hangzhou Normal University), Gexin Liu (Hangzhou Normal University), Dandan Ding (Hangzhou Normal University), Zhan Ma (Nanjing University)
		apccpa09	On-site	Quality Evaluation of Machine Learning-based Point Cloud Coding Solutions -- Joao Prazeres (Universidade da Beira Interior : Instituto de Telecomunicacoes), Rafael Rodrigues (Universidade da Beira Interior : Instituto de Telecomunicacoes), Manuela Pereira (Universidade da Beira Interior : Instituto de Telecomunicacoes), Antonio M.G. Pinheiro (Universidade da Beira Interior : Instituto de Telecomunicacoes)
		apccpa10	On-site	View-Adaptive Streaming of Point Cloud Scenes through combined Decomposition and Video-based Coding -- Michael Rudolph (University of Duisburg-Essen), Amr Rizk (University of Duisburg-Essen)
Auditorio IV	**9h-13h**			
Meet4MM: Methodologies and Tools for Multimedia	9h	m4mm0025k	On-site	Tools for collecting, synchronising and annotating ecologically valid social behaviour in the wild -- Hayley Hung (Technical University of Delft)
	10h	m4mm0121		Playing Lottery Tickets in Style Transfer Models -- Meihao Kong (Nanjing University), Jing Huo (Nanjing University), Wenbin Li (Nanjing University), Jing Wu (Cardiff University), Yu-Kun Lai (Cardiff University), Yang Gao (Nanjing University)
	10h18	m4mm2231		Optimal Tensor Bipartite Graph Learning -- Haizhou Yang (Xidian University), Wenhui Zhao (Xidian University), Quanxue Gao (Xidian University), Xiangdong Zhang (Xidian University), Wei Xia (Xidian University)
	11h	m4mm0022k	On-site	A Multimodal Dynamical Variational Autoencoder for Audiovisual Speech Representation Learning -- Simon Leglaive (CentraleSupélec, IETR)
	12h	m4mm2415	On-site	HPFL: Federated Learning by Fusing Multiple Sensor Modalities with Heterogeneous Privacy Sensitivity Levels -- Yuanjie Chen (National Tsing Hua University), Chih-Fan Hsu (National Yang Ming Chiao Tung University), Chung-Chi Tsai (Qualcomm Technologies, Inc.), Cheng-Hsin Hsu (National Tsing Hua University)
	12h18	m4mm2692	On-site	Boosting Few-shot Learning by Self-calibration in Feature Space -- Kaipeng Zheng (University of Electronic Science and Technology of China), Liu Cheng (Southwest University), Jie Shen (University of Electronic Science and Technology of China)
Pav5 R5A	**9h-13h**			
QoEVMA 2022: 2nd Workshop on Quality of Experience in Visual Multimedia Applications	9h15	qoevma22k	On-site	Estimating the Quality of Experience of Immersive Contents -- Mylène C.Q. Farias (University of Brasilia)
	10h	qoevma01	On-site	Adversarial Attacks Against Blind Image Quality Assessment Models -- Jari Korhonen (Shenzhen University), Junyong You (Norwegian Research Centre)
	10h15	qoevma02	On-site	Simulating Visual Mechanisms by Sequential Spatial-Channel Attention for Image Quality Assessment -- Junyong You (NORCE Norwegian Research Centre), Jari Korhonen (Shenzhen University : University of Aberdeen)
	10h30	qoevma04		No-reference Point Clouds Quality Assessment using Transformer and Visual Saliency -- Salima Bourbia (Mohammed V University in Rabat), Ayoub Karine (ISEN Yncréa Ouest), Aladine Chetouani (Université d'Orléans), Mohammed El Hassouni (Mohammed V University in Rabat), Maher Jridi (ISEN Yncréa Ouest, 33 Quater Chemin du Champ de Manœuvre 44470, Carquefou)
	11h45	**Break**		
	11h	qoevma08	On-site	From Just Noticeable Differences to Image Quality -- Ali Ak (Nantes University), Andreas Pastor (Nantes University), Patrick Le Callet (Nantes University)
	11h15	qoevma10		No-Reference Quality Assessment of Stereoscopic Video Based on Deep Frequency Perception -- Shuai Xiao (Tianjin University), Jiabao Wen (Tianjin University), Jiachen Yang (Tianjin University), Yanshuang Zhou (Tianjin University)
	11h30	qoevma12		Point Cloud Quality Assessment Using Cross-correlation of Deep Features -- Marouane Tliba (University of Orleans), Aladine Chetouani (University of Orleans), Giuseppe Valenzise (University of Paris-Saclay), Frederic Dufaux (University of Paris-Saclay)
	11h45	qoevma13	Virtual	On Objective and Subjective Quality of 6DoF Synthesized Live Immersive Videos -- Yuan-Chun Sun (National Tsing Hua University), Sheng-Ming Tang (National Tsing Hua University), Ching-Ting Wang (National Tsing Hua University), Cheng-Hsin Hsu (National Tsing Hua University)
	12h	qoevma15	On-site	Impact of Content on Subjective Quality of Experience Assessment for 3D Video -- Dawid Juszka (AGH University of Science and Technology), Zdzisław Papir (AGH University of Science and Technology)
Pav4 R1.02	**9h-18h**			
IXR 2022: 1st Workshop on Interactive eXtended Reality		ixr001		Optimal Camera Placement for 6 Degree-of-Freedom Immersive Video Streaming Without Accessing 3D Scenes -- Sheng-Ming Tang (National Tsing Hua University), Yuan-Chun Sun (National Tsing Hua University), Jia-Wei Fang (National Tsing Hua University), Kuan-You Lee (National Tsing Hua University), Ching-Ting Wang (National Tsing Hua University), Cheng-Hsin Hsu (National Tsing Hua University)
		ixr002	On-site	Tele-Robotics VR with Holographic Vision in Immersive Video -- Gauthier Lafruit (UPM), Laurie Van Bogaert (UPM), Jaime Sancho Aragon (UPM), Michael Panzirsch (UPM), Gregoire Hirt (UPM), Klaus Strobl (UPM), Eduardo Juarez Martinez (UPM)
		ixr003	On-site	Partially Reliable Transport Layer for QUICker Interactive Immersive Media Delivery -- Hemanth Kumar Ravuri (Ghent University - imec), Maria Torres Vega (Ghent University - imec), Jeroen van der Hooft (Ghent University - imec), Tim Wauters (Ghent University - imec), Filip De Turck (Ghent University - imec)
		ixr004	On-site	Virtual Visits: UX Evaluation of a Photorealistic AR-based Video Communication Tool -- Marina Alvarez (TNO), Alexander Toet (TNO), Sylvie Dijkstra-Soudarissanane (TNO)
		ixr005	On-site	Engagement and Quality of Experience in Remote Business Meetings: A Social VR Study -- Simardeep Singh (TNO), Sylvie Dijkstra-Soudarissanane (TNO), Simon Gunkel (TNO)
		ixr006	On-site	Behavioural Analysis in a 6-DoF VR System: Influence of Content, Quality and User Disposition -- Silvia Rossi (CWI), Irene Viola (CWI), Pablo Cesar (CWI)
		ixr007	On-site	Effects of Haptic Feedback on User Perception and Performance in Interactive Projected Augmented Reality -- Sam Van Damme (Ghent University - imec), Nicolas Legrand (Ghent University - imec), Joris Heyse (Ghent University - imec), Femke De Backere (Ghent University - imec), Filip De Turck (Ghent University - imec), Maria Torres Vega (Ghent University : KULeuven)
		ixr008	On-site	DHR: Distributed Hybrid Rendering for Metaverse Experiences -- Yu Wei Tan (National University of Singapore), Alden Tan (National University of Singapore), Nicholas Nge (National University of Singapore), Anand Bhojan (National University of Singapore)
		ixr009	On-site	Generating Realistic Synthetic Head Rotation Data for Extended Reality using Deep Learning -- Jakob Struye (University of Antwerp - imec), Filip Lemic (Universitat Politècnica de Catalunya), Jeroen Famaey (University of Antwerp - imec)
		ixr022k	On-site	Adventures in the Realverse: Experiences and Challenges in XR Communications -- Pablo Perez (Nokia)
		ixr025k	On-site	Deep Learning-based Extended Reality: Making Humans and Machines Speak the Same Visual Language -- Fernando Pereira (Instituto Superior Tecnico - Instituto de Telecomunicacoes)
Pav4 R1.03	**9h-13h**			
UoLMM 2022: 2nd Robust Understanding of Low-quality Multimedia Data: Unitive Enhancement, Analysis and Evaluation		uolmm10		Hybrid GAN Exposure Interpolation for Two Large-Exposure-Ratio Images -- Chaobing Zheng (Wuhan University of Science and Technology), Wei Wang (Wuhan University of Science and Technology), Wenjian Ying (Wuhan University of Science and Technology), Yuwen Li (Wuhan University of Science and Technology), Shiqian Wu (Wuhan University of Science and Technology)
	9h			Welcome and Opening Remarks
	9h10			Keynote Speech: Unlocking the Potential of Disentangled Representation for Robust Media Understanding -- Wenjun Zeng, IEEE Fellow, EIT Institute for Advanced Study
	9h50			Keynote Speech: Visual Signal Assessment, Analysis and Enhancement for Low-resolution or Varying-illumination Environment -- Weisi Lin, IEEE Fellow, Nanyang Technological University
	10h30			Keynote Speech: Computational Imaging on Mobile Phones -- Jinwei Gu, Executive R&D Director, SenseBrain
	11h10	uolmm20		In-training Restoration Models Matter: Data Augmentation for Degraded-reference Image Quality Assessment -- Jiazhi Du (Harbin Institute of Technology), Dongwei Ren (Harbin Institute of Technology), Yue Cao (Harbin Institute of Technology), Wangmeng Zuo (Harbin Institute of Technology)
Pav4 R1.03	**14h-18h**			
PIC 2022: 4th Person in Context workshop and Challenge		pic01	Virtual	STVGFormer: Spatio-Temporal Video Grounding with Static-Dynamic Cross-Modal Understanding -- Zihang Lin (Sun Yat-sen University), Chaolei Tan (Sun Yat-sen University), Jian-Fang Hu (Sun Yat-sen University : Guangdong Province Key Laboratory of Information Security Technology), Zhi Jin (Sun Yat-sen University), Tiancai Ye (Tencent), Wei-Shi Zheng (Sun Yat-sen University)
		pic02	Virtual	Cascaded Decoding and Multi-Stage Inference for Spatio-Temporal Video Grounding -- Li Yang (Institute of Automation, Chinese Academy of Sciences : University of Chinese Academy of Sciences), Peixuan Wu (Institute of Automation, Chinese Academy of Sciences : University of Chinese Academy of Sciences), Chunfeng Yuan (Institute of Automation, Chinese Academy of Sciences), Bing Li (Institute of Automation, Chinese Academy of Sciences), Weiming Hu (Institute of Automation, Chinese Academy of Sciences : University of Chinese Academy of Sciences : CAS Center for Excellence in Brain Science and Intelligence Technology)
		pic03	On-site	Human-centric Spatio-Temporal Video Grounding via the Combination of Mutual Matching Network and TubeDETR -- Fan Yu (Nanjing University), Zhixiang Zhao (Nanjing University), Yuchen Wang (Nanjing University), Yi Xu (Nanjing University), Tongwei Ren (Nanjing University), Gangshan Wu (Nanjing University)
		pic04	Virtual	PIC 4th Challenge: Semantic-Assisted Multi-Feature Encoding and Multi-Head Decoding for Dense Video Captioning -- Yifan Lu (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Ziqi Zhang (Institute of Automation, Chinese Academy of Sciences), Yuxin Chen (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Chunfeng Yuan (Institute of Automation, Chinese Academy of Sciences), Bing Li (Institute of Automation, Chinese Academy of Sciences), Weiming Hu (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; : CAS Center for Excellence in Brain Science and Intelligence Technology)
		pic05		Fine-grained Video Captioning via Precise Key Point Positioning -- Yunjie Zhang (Jiangnan University), Tiangyang Xu (Jiangnan University), Xiaoning Song (Jiangnan University), Zhenghua Feng (University of Surrey), Xiao-Jun Wu (Jiangnan University)
		pic06	On-site	Exploiting Feature Diversity for Make-up Temporal Video Grounding -- Xiujun Shu (Tencent Youtu Lab), Wei Wen (Tencent Youtu Lab), Taian Guo (Tencent Youtu Lab), Sunan He (Tencent Youtu Lab), Chen Wu (Tencent Youtu Lab), Ruizhi Qiao (Tencent Youtu Lab)
Pav4 R1.04	**9h-18h**			
DDAM 2022: 1st	9h10	ddam022k	Virtual	Lessons Learned from ASVSpoof and Remaining Challenges -- Junichi Yamagishi (National Institute of Informatics)
	10h30	ddam003		Detection of Synthetic Speech Based on Spectrum Defects -- JiaCheng Deng (Ningbo University), Terui Mao (Ningbo University), Diqun Yan (Ningbo University), Li Dong (Ningbo University), Mingyu Dong (Ningbo University)
	10h45	ddam004	On-site	Low-quality Fake Audio Detection through Frequency Feature Masking -- Il-Youp Kwak (Korea University), Sunmook Choi (Korea University), Jonghoon Yang (Korea University), Yerin Lee (Korea University), Soyul Han (Korea University), Seungsang Oh (Korea University)
	11h	ddam006		Audio Deepfake Detection Based on a Combination of F0 Information and Real Plus Imaginary Spectrogram Features -- Jun Xue (School of Computer Science and Technology, Anhui University), Cunhang Fan (School of Computer Science and Technology, Anhui University), Zhao Lv (School of Computer Science and Technology, Anhui University), Jianhua Tao (NLPR, Institute of Automation, Chinese Academy of Sciences), Jiangyan Yi (NLPR, Institute of Automation, Chinese Academy of Sciences), Chengshi Zheng (Institute of Acoustics, Chinese Academy of Sciences), Zhengqi Wen (Qiyuan Laboratory), Minmin Yuan (National Environmental Protection Engineering and Technology Center for Road Traffic Noise Control, Research Institute of Highway Ministry of Transport), Shegang Shao (National Environmental Protection Engineering and Technology Center for Road Traffic Noise Control, Research Institute of Highway Ministry of Transport)
	11h15	ddam007		Fully Automated End-to-End Fake Audio Detection -- Chenglong Wang (University of Science and Technology of China), Jiangyan Yi (Institute of Automation, Chinese Academy of Sciences), Jianhua Tao (Institute of Automation, Chinese Academy of Sciences), Xun Chen (University of Science and Technology of China), Zhengkun Tian (Institute of Automation, Chinese Academy of Sciences), Haoxin Ma (Institute of Automation, Chinese Academy of Sciences), Cunhang Fan (Anhui University), Ruibo Fu (Institute of Automation, Chinese Academy of Sciences)
	11h30	ddam014	On-site	A Comparative Study on Physical and Perceptual Features for Deepfake Audio Detection -- Menglu Li (Toronto Metropolitan University), Yasaman Ahmadiadli (Toronto Metropolitan University), Xiao-Ping Zhang (Toronto Metropolitan University)

International Workshop on Deepfake Detection for Audio Multimedia	11h45	ddam016		Deepfake Detection System for the ADD Challenge Track 3.2 Based on Score Fusion -- Yuxiang Zhang (Institute of Acoustics, Chinese Academy of Sciences), Jingze Lu (Institute of Acoustics, Chinese Academy of Sciences), Xingming Wang (Duke Kunshan University), Zhuo Li (Institute of Acoustics, Chinese Academy of Sciences), Runqiu Xiao (Institute of Acoustics, Chinese Academy of Sciences), Wenchao Wang (Institute of Acoustics, Chinese Academy of Sciences), Ming Li (Duke Kunshan University), Pengyuan Zhang (Institute of Acoustics, Chinese Academy of Sciences)
	12h	**Lunch Break**		
	14h30	ddam005		Singing-Tacotron: Global Duration Control Attention and Dynamic Filter for End-to-end Singing Voice Synthesis -- Tao Wang (Institute of Automation, Chinese Academy of Sciences), Ruibo Fu (Institute of Automation, Chinese Academy of Sciences), Jiangyan Yi (Institute of Automation, Chinese Academy of Sciences), Zhengqi Wen (Institute of Automation, Chinese Academy of Sciences), Jianhua Tao (Institute of Automation, Chinese Academy of Sciences)
	14h45	ddam009		An Initial Investigation for Detecting Vocoder Fingerprints of Fake Audio -- Xinrui Yan (Institute of Automation, Chinese Academy of Sciences), Jiangyan Yi (Institute of Automation, Chinese Academy of Sciences), Jianhua Tao (Institute of Automation, Chinese Academy of Sciences), Chenglong Wang (Institute of Automation, Chinese Academy of Sciences), Haoxin Ma (Institute of Automation, Chinese Academy of Sciences), Tao Wang (Institute of Automation, Chinese Academy of Sciences), Shiming Wang (Institute of Automation, Chinese Academy of Sciences), Ruibo Fu (Institute of Automation, Chinese Academy of Sciences)
	15h	ddam011		Deep Spectro-temporal Artifacts for Detecting Synthesized Speech -- Xiaohui Liu (Tianjin University), Meng Liu (Tianjin University), Lin Zhang (National Institute of Informatics), Linjuan Zhang (Taiyuan University of Technology), Chang Zeng (National Institute of Informatics), Kai Li (Japan Advanced Institute of Science and Technology), Nan Li (Tianjin University), Kong Aik Lee (Institute for Infocomm Research, A?STAR), Longbiao Wang (Tianjin University), Jianwu Dang (Japan Advanced Institute of Science and Technology)
	15h15	ddam012		Acoustic or Pattern? Speech Spoofing Countermeasure based on Image Pre-training Models -- Jingze Lu (Institute of Acoustics, Chinese Academy of Sciences), Zhuo Li (Institute of Acoustics, Chinese Academy of Sciences), Yuxiang Zhang (Institute of Acoustics, Chinese Academy of Sciences), Wenchao Wang (Institute of Acoustics, Chinese Academy of Sciences), Pengyuan Zhang (Institute of Acoustics, Chinese Academy of Sciences)
	15h30	ddam013	On-site	Human Perception of Audio Deepfakes -- Nicolas M. Müller (Fraunhofer AISEC), Karla Pizzi (Fraunhofer AISEC), Jennifer Williams (University of Southampton)
	15h45	ddam015		Improving Spoofing Capability for End-to-end Any-to-many Voice Conversion -- Hua Hua (Institute of Acoustics, Chinese Academy of Sciences), Ziyi Chen (Institute of Acoustics, Chinese Academy of Sciences), Yuxiang Zhang (Institute of Acoustics, Chinese Academy of Sciences), Ming Li (Duke Kunshan University), Pengyuan Zhang (Institute of Acoustics, Chinese Academy of Sciences)
Pav4 R1.05	**9h-13h**			
HCMA 2022: 3rd International Workshop on Human-Centric Multimedia Analysis	9h			Workshop Opening
	9h10	hcma31k	Virtual	Bridging Isolated Islands in Human Activity Understanding -- Cewu Lu (Shanghai Jiao Tong University)
	9h50	hcma34k	Virtual	Learning to Anticipate Human Actions from Videos -- Basura Fernando (Agency for Science, Technology and Research (A*STAR) Singapore)
	10h30	hcma37k	Virtual	Neural Digital Human Avatar: Motion Capture, Reconstruction and Rendering -- Yebin Liu (Tsinghua University)
	11h10	**Tea Break**		
	11h20	hcma06	Virtual	Anonym-Recognizer: Relationship-preserving Face Anonymization and Recognition -- Chunlei Peng (Xidian University), Shuang Wan (Xidian University), Zimin Miao (Xidian University), Decheng Liu (Xidian University), Yu Zheng (Xidian University), Nannan Wang (Xidian University)
	12h	hcma18		Cycle-Consistent Learning for Weakly Supervised Semantic Segmentation -- Bin Wang (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Yu Qiao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences), Dahua Lin (SenseTime Research), Stephen D.H. Yang (SenseTime Research), Weijia Li (School of Geospatial Engineering and Science, Sun Yat-sen University)
	12h20	hcma28		Dual Domain-Adversarial Learning for Audio-Visual Saliency Prediction -- Yingzi Fan (Xidian University), Longfei Han (Beijing Technology and Business University), Yue Zhang (Xi'an Jiaotong University), Lechao Cheng (Zhejiang Lab), Chen Xia (Northwestern Polytechnical University), Di Hu (Renmin University of China)
	12h40	hcma21	Virtual	Multi-level Multi-modal Feature Fusion for Action Recognition in Videos -- Xinghang Hu (University of Electronic Science and Technology of China), Yanli Ji (University of Electronic Science and Technology of China), Gedamu Alemu Kumie (Sichuan Artificial intelligence Research Institute)
	13h	hcma09		Two-branch Objectness-centric Open World Detection -- Yan Wu (State Key Lab of Software Development Environment, Beihang University), Xiaowei Zhao (State Key Lab of Software Development Environment, Beihang University), Yuqing Ma (State Key Lab of Software Development Environment, Beihang University), Duorui Wang (State Key Lab of Software Development Environment, Beihang University), Xianglong Liu (State Key Lab of Software Development Environment, Beihang University)
	11h20	hcma03p	On-site	Augmented Transformer with Adaptive Graph for Temporal Action Proposal Generation -- Shuning Chang (National University of Singapore), Pichao WANG (Alibaba Group), Fan Wang (Alibaba Group), Li Hao (National University of Singapore), Zheng Shou (National University of Singapore)
		hcma07p	On-site	Domain Camera Adaptation and Collaborative Multiple Feature Clustering for Unsupervised Person Re-ID -- Yuanpeng Tu (Tongji University)
		hcma12p	Virtual	PSINet: Progressive Saliency Iteration Network for RGB-D Salient Object Detection -- Songsong Duan (Anhui University of Science and Technology), Chenxing Xia (Anhui University of Science and Technology), Xianjin Fang (Anhui University of Science and Technology), Bin Ge (Anhui University of Science and Technology), Xiuju Gao (Anhui University of Science and Technology), Kuan-Ching Li (Providence University)
		hcma13p	Virtual	Multimodal Network with Cross-Modal Attention for Audio-Visual Event Localization -- Qianchao Tan (University of Electronic Science and Technology of China)
		hcma22p	On-site	Real-time Embedded Demo System for Fall Detection under 15W Power -- Junyi Lu (School of Artificial Intelligence and Automation, Huazhong University of Science and Technology), Wenxiang Jiang (ArcSoft Corporation Limited), Yang Xiao (School of Artificial Intelligence and Automation, Huazhong University of Science and Technology), Tingbing Yan (School of Artificial Intelligence and Automation, Huazhong University of Science and Technology), Zhiguo Cao (School of Artificial Intelligence and Automation, Huazhong University of Science and Technology), Zhiwen Fang (School of Biomedical Engineering, Southern Medical University), Joey Tianyi Zhou (A*STAR Centre for Frontier AI Research (CFAR))
		hcma23p	On-site	Face Clustering via Adaptive Aggregation of Clean Neighbors -- Shiyong Hong (Inspur Electronic Information Industry Co., Ltd.), Yaobin Zhang (Inspur Electronic Information Industry Co., Ltd.), Xu Ling (Inspur Electronic Information Industry Co., Ltd.), Weihong Deng (Inspur Electronic Information Industry Co., Ltd.), Yunfeng Yin (Inspur Electronic Information Industry Co., Ltd.), Yingjie Zhang (Inspur Electronic Information Industry Co., Ltd.), Hongzhi Shi (Inspur Electronic Information Industry Co., Ltd.), Dongchao Wen (Inspur Electronic Information Industry Co., Ltd.)
		hcma26p	Virtual	Cross-modal Token Selection for Video Understanding -- Liyong Pan (Nanjing University of Science and Technology), Zechao Li (Nanjing University of Science and Technology), Henghao Zhao (Nanjing University of Science and Technology), Rui Yan (Nanjing University of Science and Technology)
Pav4 R1.05	**14h-18h**			
MCFR 2022: 1st Workshop on Multimedia Computing towards Fashion Recommendation	14h10	mcfr022k	Virtual	Fashion Meets Computer Vision -- Wen-Huang Cheng (National Chiao Tung University)
	14h55	mcfr011	Virtual	On Leveraging the Metapath and Entity Aware Subgraphs for Recommendation -- Muhammad Umer Anwaar (Unite Network SE), Zhiwei Han (fortiss GmbH), Shyam Arumugaswamy (Unite Network SE), Rayyan Ahmad Khan (Technical University of Munich), Thomas Weber (Ludwig Maximilian Universität), Tianming Qiu (fortiss GmbH), Hao Shen (fortiss GmbH), Yuanting Liu (Fortiss GmbH), Martin Kleinsteuber (Unite Network SE)
	15h15	mcfr017	On-site	I-MALL An Effective Framework for Personalized Visits. Improving the Customer Experience in Stores -- Federico Becattini (Università degli Studi di Firenze), Giuseppe Becchi (Università degli Studi di Firenze), Andrea Ferracani (Università degli Studi di Firenze), Alberto Del Bimbo (Università degli Studi di Firenze), Liliana Lo Presti (Università degli Studi di Palermo), Giuseppe Mazzola (Università degli Studi di Palermo), Marco La Cascia (Università degli Studi di Palermo), Federico Cunico (Università degli Studi di Verona), Andrea Toaiari (Università degli Studi di Verona), Marco Cristani (Università degli Studi di Verona), Antonio Greco (Università degli Studi di Salerno), Alessia Saggese (Università degli Studi di Salerno), Mario Vento (Università degli Studi di Salerno)
	15h50	mcfr012	On-site	Orthogonal Vector-Decomposed Disentanglement Network of Interactive Image Retrieval for Fashion Outfit Recommendation -- Chen Chen (Xidian University), Jie Guo (Xidian University), Bin Song (Xidian University), Tong Zhang (Xidian University)
	16h10	mcfr014	Virtual	Counterfactural Inference towards Unbiased Outfit Compatibility Modeling -- Liqiang Jing (Shandong University), Minghui Tian (Shandong University), Xiaolin Chen (Shandong University), Teng Sun (Shandong University), Weili Guan (Monash University), Xuemeng Song (Shandong University)
Pav4 R1.06	**9h-18h**			
PIES-ME 2022: 1st Workshop on Photorealistic Image and Environment Synthesis for Multimedia Experiments	9h15			Welcom and Opening Remarks
	9h30	pies01	On-site	Delving into Light-Dark Semantic Segmentation for Indoor Scenes Understanding -- Xiaowen Ying (Lehigh University), Bo Lang (Lehigh University), Zhihao Zheng (Lehigh University), Mooi Choo Chuah (Lehigh University)
		pies02		Language-guided Semantic Style Transfer of 3D Indoor Scenes -- Bu Jin (University of Chinese Academy of Sciences), Beiwen Tian (Tsinghua University), Hao Zhao (Peking University), Guyue Zhou (Tsinghua University)
	10h	pies022k	On-site	Hyper-Realistic and Immersive Imaging for Enhanced Quality of Experience -- Frederic Dufaux (Université Paris-Saclay, CNRS, CentraleSupélec)
	10h30	**Coffee Break**		
	11h	pies03	On-site	Towards a Calibrated 360 Stereoscopic HDR Image Dataset for Architectural Lighting Studies -- Michèle Atié (Nantes Université), Toinon Vigier (Nantes Université), François Eymond (Université de Lyon), Céline Drozd (Nantes Université), Raphaël Labayrade (Université de Lyon), Daniel Siret (Nantes Université), Yannick Sutter (Nantes Université)
	11h30			I-SPIES Dataset Reveal
	12h			Panel Discussion: "Photorealistic datasets to enable multimedia research: creation, curation and use"
	13h	**Lunch Break**		
	14h			Keynote Speech
	15h	pies04	On-site	Subjective Study of the Impact of Compression, Framerate, and Navigation Trajectories on the Quality of Free-Viewpoint Video -- Jesús Gutiérrez (Universidad Politécnica de Madrid), Adriana Galán (Universidad Politécnica de Madrid), Pablo Pérez (Nokia), Daniel Corregidor (Universidad Politécnica de Madrid), Teresa Hernando (Universidad Politécnica de Madrid), Javier Usón (Universidad Politécnica de Madrid), Daniel Berjón (Universidad Politécnica de Madrid), Julián Cabrera (Universidad Politécnica de Madrid), Narciso García (Universidad Politécnica de Madrid)
	15h30	pies05		Comparative Evaluation of Temporal Pooling Methods for No-Reference Quality Assessment of Dynamic Point Clouds -- Pedro G. Freitas (University of Brasília), Giovani D. Lucafo (University of São Paulo), Mateus Gonçalves (University of Brasília), Johann Homonnai (University of Brasília), Rafael Diniz (University of Brasília), Mylène C.Q. Farias (University of Brasília)
	16h			Closing Remarks
Tutorials				
Pav5 R5A	14h-18h	**Memory Networks**		
Pav4 R1.07	9h-13h	**Video Grounding and Its Generalization**		