

2023 IEEE International Conference on Multimedia and Expo (ICME 2023)

**Brisbane, Australia
10-14 July 2023**

Pages 1-587



**IEEE Catalog Number: CFP23ICM-POD
ISBN: 978-1-6654-6892-3**

**Copyright © 2023 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP23ICM-POD
ISBN (Print-On-Demand):	978-1-6654-6892-3
ISBN (Online):	978-1-6654-6891-6
ISSN:	1945-7871

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2023 IEEE International Conference on Multimedia and Expo (ICME) ICME 2023

Table of Contents

Welcome Message from the General Chairs	lxxii
Organising Committee	lxxiv
Reviewers	lxxviii
Meta Reviewers	ci

Advances in Language and Vision Research

Weakly Supervised Few-Shot and Zero-Shot Semantic Segmentation with Mean Instance Aware Prompt Learning	1
<i>Prashant Pandey (Indian Institute of Technology Delhi, India), Mustafa Chasmai (Indian Institute of Technology Delhi, India), Monish Natarajan (Indian Institute of Technology Kharagpur, India), and Brejesh Lall (Indian Institute of Technology Delhi, India)</i>	
Ada-SwinBERT: Adaptive Token Selection for Efficient Video Captioning with Online Self-Distillation	7
<i>Qianwen Cao (Beijing Institute of Technology, China), Heyan Huang (Beijing Institute of Technology, China), Minpeng Liao (Independent researcher, China), and Xianling Mao (Beijing Institute of Technology, China)</i>	
A Retriever-Reader Framework with Visual Entity Linking for Knowledge-Based Visual Question Answering	13
<i>Jiuxiang You (Guangdong University of Technology, China), Zhenguo Yang (Guangdong University of Technology, China), Qing Li (The Hong Kong Polytechnic University, China), and Wenyin Liu (Guangdong University of Technology, China)</i>	
2S-DFN: Dual-Semantic Decoding Fusion Networks for Fine-Grained Image Recognition	19
<i>Pufen Zhang (University of Science and Technology Beijing, China), Peng Shi (University of Science and Technology Beijing, China), and Song Zhang (University of Science and Technology Beijing, China)</i>	
MuDPT: Multi-Modal Deep-Symphysis Prompt Tuning for Large Pre-Trained Vision-Language Models	25
<i>Yongzhu Miao (National University of Defense Technology, China), Shasha Li (National University of Defense Technology, China), Jintao Tang (National University of Defense Technology, China), and Ting Wang (National University of Defense Technology, China)</i>	

Action-GPT: Leveraging Large-Scale Language Models for Improved and Generalized Action Generation	31
<i>Sai Shashank Kalakonda (IIIT Hyderabad, INDIA), Shubh Maheshwari (IIIT Hyderabad, INDIA), and Ravi Kiran Sarvadevabhatla (IIIT Hyderabad, INDIA)</i>	

Steganography & Watermarking

Protecting Intellectual Property of EEG-Based Model with Watermarking	37
<i>Tianhua Xu (Shenzhen University, China), Sheng-hua Zhong (Shenzhen University, China), and Zhijiao Xiao (Shenzhen University, China)</i>	
Making Adversarial Attack Imperceptible in Frequency Domain: A Watermark-Based Framework ..	43
<i>Hanxiu Zhang (East China Normal University, China), Guitao Cao (East China Normal University, China), Xinyue Zhang (East China Normal University, China), Jing Xiang (East China Normal University, China), and Chunwei Wu (East China Normal University, China)</i>	
Content-Adaptive Adversarial Embedding for Image Steganography using Deep Reinforcement Learning	49
<i>Jie Luo (Sichuan University, China), Peisong He (Sichuan University, China), Jiayong Liu (Sichuan University, China), Hongxia Wang (Sichuan University, China), Chunwang Wu (School of Cyber Science and Engineering, China), Yijing Chen (Sichuan University, China), Wanjie Li (Sichuan University, China), and Jiangchuan Li (Sichuan University, China)</i>	
A Robust Generative Image Steganography Method Based on Guidance Features in Image Synthesis	55
<i>Youqiang Sun (Beijing Univ Posts & Telecommun., China), Jianyi Liu (Beijing Univ Posts & Telecommun., China), and Ru Zhang (Beijing Univ Posts & Telecommun., China)</i>	
Adversarial Audio Watermarking: Embedding Watermark into Deep Feature	61
<i>Shiqiang Wu (University of Chinese Academy of Sciences; Institute of Automation, Chinese Academy of Sciences, China), Jie Liu (Institute of Automation, Chinese Academy of Sciences, China), Ying Huang (Beijing University of Posts and Telecommunications; Institute of Automation, Chinese Academy of Sciences, China), Hu Guan (Institute of Automation, Chinese Academy of Sciences, China), and Shuwu Zhang (Beijing University of Posts and Telecommunications; Institute of Automation, Chinese Academy of Sciences, China)</i>	
Deniable Diffusion Generative Steganography	67
<i>Tengjun Liu (Fudan University, China), Ying Chen (Fudan University, China), and Wanxuan Gu (NVIDIA, China)</i>	

Semantic Processing I

Sea Surface Object Detection Based on Background Dynamic Perception and Cross-Layer Semantic Interaction	72
<i>Songbin Li (University of Chinese Academy of Sciences, China), Xiangzhi Yang (University of Chinese Academy of Sciences, China), and Jingang Wang (University of Chinese Academy of Sciences, China)</i>	
Addressing Predicate Overlap in Scene Graph Generation with Semantic Granularity Controller	78
<i>Guikun Chen (Zhejiang University, China), Lin Li (Zhejiang University, China), Yawei Luo (Zhejiang University, China), and Jun Xiao (Zhejiang University, China)</i>	
Towards Discriminative Semantic Relationship for Fine-Grained Crowd Counting	84
<i>Shiqi Ren (University of Science and Technology Beijing, China), Chao Zhu (University of Science and Technology Beijing, China), Mengyin Liu (University of Science and Technology Beijing, China), and Xu-Cheng Yin (University of Science and Technology Beijing, China)</i>	
Region-Aware Semantic Consistency for Unsupervised Domain-Adaptive Semantic Segmentation .	90
<i>Jun Xie (University of Electronic Science and Technology of China, China), Yixuan Zhou (University of Electronic Science and Technology of China, China), Xing Xu (University of Electronic Science and Technology of China, China), Guoqing Wang (University of Electronic Science and Technology of China, China), Fumin Shen (University of Electronic Science and Technology of China, China), and Yang Yang (University of Electronic Science and Technology of China, China)</i>	
Deep Unsupervised Hashing with Selective Semantic Mining	96
<i>Chuang Zhao (School of Computer Science and Technology, HUST, China), Hefei Ling (School of Computer Science and Technology, HUST, China), Yuxuan Shi (School of Computer Science and Technology, HUST, China), Chengxin Zhao (School of Computer Science and Technology, HUST, China), Jiazhong Chen (School of Computer Science and Technology, HUST, China), and Qiang Cao (Wuhan National Laboratory for Optoelectronics, HUST, China)</i>	
Boosting Interactive Image Segmentation by Exploiting Semantic Clues	102
<i>Qiaoqiao Wei (Tsinghua University, China), Hui Zhang (Tsinghua University, China), and Jun-Hai Yong (Tsinghua University, China)</i>	

Media Retrieval I

Visual-Linguistic Alignment and Composition for Image Retrieval with Text Feedback	108
<i>Dafeng Li (Shenzhen University, China) and Yingying Zhu (Shenzhen University, China)</i>	
Scene Text Involved "Text"-to-Image Retrieval Through Logically Hierarchical Matching	114
<i>Xinyu Zhou (Wuhan University of Technology, China), Huen Chen (Wuhan University of Technology, China), Anna Zhu (Wuhan University of Technology, China), and Wei Pan (Wuhan University of Technology, China)</i>	

Federating Hashing Networks Adaptively for Privacy-Preserving Retrieval	120
<i>Yi Li (Dalian University of Technology, China), Meihua Yu (Dalian University of Technology, China), Xin Xie (Dalian University of Technology, China), Haiyan Fu (Dalian University of Technology, China), Hao He (InsightOne Tech Co., Ltd., China), and Yanqing Guo (Dalian University of Technology, China)</i>	
Deep Unsupervised Momentum Contrastive Hashing for Cross-Modal Retrieval	126
<i>Kangkang Lu (Beijing University of Posts and Telecommunications, China), Yanhua Yu (Beijing University of Posts and Telecommunications, China), Meiyu Liang (Beijing University of Posts and Telecommunications, China), Min Zhang (Beijing University of Posts and Telecommunications, China), Xiaowen Cao (Beijing University of Posts and Telecommunications, China), Zehua Zhao (Beijing University of Posts and Telecommunications, China), Mengran Yin (Beijing University of Posts and Telecommunications, China), and Zhe Xue (Beijing University of Posts and Telecommunications, China)</i>	
Uncertainty-Aware Cross-Modal Transfer Network for Sketch-Based 3D Shape Retrieval	132
<i>Yiyang Cai (Tongji University, China), Jiaming Lu (Tongji University, China), Jiewen Wang (Tongji University, China), and Shuang Liang (Tongji University, China)</i>	
Scene Graph Based Fusion Network for Image-Text Retrieval	138
<i>Guoliang Wang (Beijing University of Posts and Telecommunications, China), Yanlei Shang (Beijing University of Posts and Telecommunications, China), Yong Chen (Beijing University of Posts and Telecommunications, China), Chaoqi Zhen (Beijing University of Posts and Telecommunications, China), and Dequan Cheng (Beijing University of Posts and Telecommunications, China)</i>	

Detection I

Compact Intertemporal Coupling Network for Remote Sensing Change Detection	144
<i>Yuchao Feng (Zhejiang University of Technology, China), Honghui Xu (Zhejiang University of Technology, China), Jiawei Jiang (Zhejiang University of Technology, China), and Jianwei Zheng (Zhejiang University of Technology, China)</i>	
Boundary-Aware Shadow Detection via Mask Decoupling and Feature Correction	150
<i>Jueyu Chen (Sichuan University, China), Guanyu Xing (Sichuan University, China), Jingwei Liao (George Mason University, USA), Housheng Wei (Sichuan University, China), and Yanli Liu (Sichuan University, China)</i>	
Explore Faster Localization Learning For Scene Text Detection	156
<i>Yuzhong Zhao (University of Chinese Academy of Sciences, China), Yuanqiang Cai (Beijing University of Posts and Telecommunications, China), Weijia Wu (Zhejiang University, China), and Weiqiang Wang (University of Chinese Academy of Sciences, China)</i>	
Counterfactual Inference for Visual Relationship Detection in Videos	162
<i>Xiaofeng Ji (Beijing Institute of Technology, China), Jin Chen (Beijing Institute of Technology, China), and Xinxiao Wu (Beijing Institute of Technology, China)</i>	

Body-Part Joint Detection and Association via Extended Object Representation	168
<i>Huayi Zhou (Shanghai Jiao Tong University, China), Fei Jiang (East China Normal University, China), and Hongtao Lu (Shanghai Jiao Tong University, China)</i>	
Be-or-Not Prompt Enhanced Hard Negatives Generating for Memes Category Detection	174
<i>Jian Cui (Wuhan University of Technology, China), Lin Li (Wuhan University of Technology, China), and Xiaohui Tao (The University of Southern Queensland, Australia)</i>	

Health Application

Automatic Retinal Nerve Fiber Trajectory Simulation and Quasi-Polar Transformation for Detecting Retinal Nerve Fiber Layer Defect in Fundus Images	180
<i>Yanni Wang (Renmin University of China, China), Gang Yang (Renmin University of China), Dayong Ding (Visionary Intelligence Ltd., China), and Jianchun Zhao (Visionary Intelligence Ltd., China)</i>	
GA-HQS: MRI Reconstruction via a Generically Accelerated Unfolding Approach	186
<i>Jiawei Jiang (Zhejiang University of Technology, China), Jiacheng Chen (Zhejiang University of Technology, China), Honghui Xu (Zhejiang University of Technology, China), Yuchao Feng (Zhejiang University of Technology, China), and Jianwei Zheng (Zhejiang University of Technology, China)</i>	
Early Diagnosis of Alzheimer's Disease Based on Multimodal Hypergraph Attention Network .	192
<i>Yi Li (Guangdong University of Technology, China), BaoYao Yang (Guangdong University of Technology, China), Dan Pan (Guangdong Polytechnic Normal University, China), An Zeng (Guangdong University of Technology, China), Long Wu (Guangdong Provincial People's Hospital, China), and Yang Yang (Guangdong Provincial People's Hospital, China)</i>	
Score-Based Causal Feature Selection for Cancer Risk Prediction	198
<i>Shanshan Huang (Chongqing University, China), Qingsong Li (Chongqing University, China), Lei Wang (Chongqing University, China), Yuanhao Wang (Chongqing University, China), and Li Liu (Chongqing University, China)</i>	
Depression Diagnosis and Analysis via Multimodal Multi-Order Factor Fusion	204
<i>Chengbo Yuan (Wuhan University, China; Hubei LuoJia Laboratory, China), Qianhui Xu (Wuhan University, China; Hubei LuoJia Laboratory, China), and Yong Luo (Wuhan University, China; Hubei LuoJia Laboratory, China)</i>	

A Dual-Path Supplemental Information Learning Architecture for Breast Cancer Ki-67 Status Prediction in T2w MRI 210

Wentian Cai (South China University of Technology; Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, China), Yulin Cheng (South China University of Technology, China), Ying Gao (South China University of Technology; Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, China), Weixiao Liu (The First Affiliated Hospital of Jinan University, China), Xinyan Xie (South China University of Technology, China), Xiongwen Luo (South China University of Technology, China), Weixian Yang (South China University of Technology; Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, China), Zaiyi Liu (Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, China), and Changhong Liang (Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, China)

Expression & Face Computing

Expression-Guided Attention GAN for Fine-Grained Facial Expression Editing 216

Hui Zhang (East China Normal University, China), Shiqi Shen (South China Normal University, China), and Jinhua Xu (East China Normal University, China)

RMES: Real-Time Micro-Expression Spotting using Phase From Riesz Pyramid 222

Yini Fang (Hong Kong University of Science and Technology; Ydentity Organization), Didan Deng (Hong Kong University of Science and Technology), Liang Wu (Hong Kong University of Science and Technology), Frederic Jumelle (Bright Nation Limited; Ydentity Organization), and Bertram Shi (Hong Kong University of Science and Technology)

AU-aware graph convolutional network for Macroand Micro-expression spotting 228

Shukang Yin (University of Science and Technology of China, China), Shiwei Wu (University of Science and Technology of China, China), Tong Xu (University of Science and Technology of China, China), Shifeng Liu (University of Science and Technology of China, China), Sirui Zhao (University of Science and Technology of China, China), and Enhong Chen (University of Science and Technology of China, China)

Semi-Supervised Facial Expression Recognition by Exploring False Pseudo-Labels 234

Hao Sun (Central China Normal University, China), Chenchen Pi (Central China Normal University, China), and Wei Xie (Central China Normal University, China)

CPNet: Exploiting CLIP-Based Attention Condenser and Probability Map Guidance for High-Fidelity Talking Face Generation 240

Jingning Xu (Tongji University; Department of AI Technology, China), Benlai Tang (Department of AI Technology, China), Mingjie Wang (Zhejiang Sci-Tech University, China), Minghao Li (Department of AI Technology, China), and Meirong Ma (Department of AI Technology, China)

ANYRES: Generating High-Resolution Visible-Face Images from Low-Resolution Thermal-Face Images	246
<i>David Anghelone (Université Côte d'Azur, France), Sarah Lannes (Thales, France), and Antitza Dantcheva (Université Côte d'Azur, France)</i>	
A Visually Interpretable Convolutional-Transformer Model for Assessing Depression from Facial Images	252
<i>Yutong Li (Lanzhou University, China), Zhenyu Liu (Lanzhou University, China), Gang Li (Third People's Hospital of Tianshui, China), Qiongqiong Chen (Northwest Minzu University, China), Zhijie Ding (Third People's Hospital of Tianshui, China), Xiping Hu (Lanzhou University, China), and Bin Hu (Lanzhou University, China)</i>	

Learning Techniques II

FreConv: Frequency Branch-and-Integration Convolutional Networks	258
<i>Zhaowen Li (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Xu Zhao (Chinese Academy of Sciences, China), Peigeng Ding (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Zongxing Gao (Beijing Institute of Graphic Communication, China), Yuting Yang (Chinese Academy of Sciences, China), Ming Tang (Chinese Academy of Sciences, China), and Jinqiao Wang (Chinese Academy of Sciences; Peng Cheng Laboratory, China)</i>	
Class-Aware Variational Auto-Encoder for Open Set Recognition	264
<i>Ruofan Wang (Fudan University, China), Jiayu Guo (Fudan University, China), Rui-Wei Zhao (Fudan University, China), Ling Su (Children's Hospital of Fudan University, China), Yingzi Ye (Children's Hospital of Fudan University, China), Xiaobo Zhang (Children's Hospital of Fudan University, China), Yuejie Zhang (Fudan University, China), and Rui Feng (Fudan University, China)</i>	
Repnas: Searching for Efficient Re-Parameterizing Blocks	270
<i>Mingyang Zhang (Zhejiang University of Technology), Xinyi Yu (Zhejiang University of Technology), Jingtao Rong (Zhejiang University of Technology), and Linlin Ou (Zhejiang University of Technology)</i>	
Difference-Aware Iterative Reasoning Network for Key Relation Detection	276
<i>Bowen Zhao (University of Science and Technology of China, China), Weidong Chen (University of Science and Technology of China, China), Bo Hu (University of Science and Technology of China, China), Hongtao Xie (University of Science and Technology of China, China), and Zhendong Mao (University of Science and Technology of China, China)</i>	
Injecting-Diffusion: Inject Domain-Independent Contents into Diffusion Models for Unpaired Image-to-Image Translation	282
<i>Luying Li (Shanghai Jiao Tong University, China) and Lizhuang Ma (Shanghai Jiao Tong University, China)</i>	

Semi-Supervised Top-k Feature Selection with a General Optimization Framework	288
<i>Lei Xu (Northwestern Polytechnical University, P.R. China), Rong Wang (Northwestern Polytechnical University, P.R. China), Feiping Nie (Northwestern Polytechnical University, P.R. China), Jun Wu (Huawei Technologies, co. Ltd, P.R. China), and Xuelong Li (Northwestern Polytechnical University, P.R. China)</i>	
Towards Boosting Black-box Attack via Sharpness-Aware	294
<i>Yukun Zhang (University of Electronic Science and Technology of China, China), Shengming Yuan (University of Electronic Science and Technology of China, China), Jingkuan Song (University of Electronic Science and Technology of China, China), Yixuan Zhou (University of Electronic Science and Technology of China, China), Lin Zhang (University of Electronic Science and Technology of China, China), and Yulan He (The University of Warwick, UK)</i>	
Learning Group Residual Representation for Group Activity Prediction	300
<i>Xiaolin Zhai (Nankai University, China), Zhengxi Hu (Nankai University, China), Dingye Yang (Nankai University, China), Shichao Wu (Nankai University, China), and Jingtai Liu (Nankai University, China)</i>	
SAFE: Simultaneous Alignment of Features and Predictions for Dense Object Detectors	306
<i>Xuesong Guo (University of Science and Technology of China, China), Shuo Wang (University of Science and Technology of China, China), Jiahao Chang (University of Science and Technology of China, China), Zehui Chen (University of Science and Technology of China, China), and Feng Zhao (University of Science and Technology of China, China)</i>	
MSG-CAM:Multi-Scale Inputs Make a Better Visual Interpretation of CNN Networks	312
<i>Xiaohong Xiang (Chongqing University of Posts and Telecommunications, China), Fuyuan Zhang (Chongqing University of Posts and Telecommunications, China), Xin Deng (Chongqing University of Posts and Telecommunications, China), and Ke Hu (Chongqing University of Posts and Telecommunications, China)</i>	

Trustworthy Federated Learning for Multimedia I

Personalization Disentanglement for Federated Learning	318
<i>Peng Yan (University of Technology Sydney, Australia) and Guodong Long (University of Technology Sydney, Australia)</i>	
Fairness-Aware Client Selection for Federated Learning	324
<i>Yuxin Shi (Nanyang Technological University (NTU); Alibaba-NTU Singapore Joint Research Institute, Singapore; Alibaba Group, China), Zelei Liu (Nanyang Technological University (NTU), Singapore), Zhuan Shi (University of Science and Technology of China, China), and Han Yu (Nanyang Technological University (NTU), Singapore)</i>	
Utility-Maximizing Bidding Strategy for Data Consumers in Auction-Based Federated Learning.....	330
<i>Xiaoli Tang (Nanyang Technological University, Singapore) and Han Yu (Nanyang Technological University, Singapore)</i>	

Federated Learning for Personalized Image Aesthetics Assessment	336
<i>Zhiwei Xiong (Nanyang Technological University, Singapore), Han Yu (Nanyang Technological University, Singapore), and Zhiqi Shen (Nanyang Technological University, Singapore)</i>	
Decentralized Federated Learning Via Mutual Knowledge Distillation	342
<i>Yue Huang (Shandong University), Lanju Kong (Shandong University; Dareway Software Co.), Qingzhong Li (Shandong University; Dareway Software Co.), and Baochen Zhang (Shandong University)</i>	
Fedward: Flexible Federated Backdoor Defense Framework with Non-IID Data	348
<i>Zekai Chen (Fuzhou university, China), Fuyi Wang (Deakin University), Zhiwei Zheng (Fuzhou university, China), Ximeng Liu (Fuzhou university, China), and Yujie Lin (Fuzhou university, China)</i>	

Quality Assessment I

Contrastive Fusion Representation: Mitigating Adversarial Attacks on VQA Models	354
<i>Jialing He (Chongqing University, China), Zhen Qin (Chongqing University, China), Hangcheng Liu (Nanyang Technological University, Singapore), Shangwei Guo (Chongqing University, China), Biwen Chen (Chongqing University, China), Ning Wang (Chongqing University, China), and Tao Xiang (Chongqing University, China)</i>	
Improving Point Cloud Quality Metrics with Noticeable Possibility Maps	360
<i>Zhengyu Wang (Shanghai Jiao Tong University, China), Yujie Zhang (Shanghai Jiao Tong University, China), Qi Yang (MediaLab, Tencent, China), Yiling Xu (Shanghai Jiao Tong University, China), Yifei Zhou (Shanghai Maritime University, China), Jun Sun (Shanghai Jiao Tong University, China), and Shan Liu (MediaLab, Tencent, United States)</i>	
Exploring Opinion-Unaware Video Quality Assessment with Semantic Affinity Criterion	366
<i>Haoning Wu (Nanyang Technological University), Liang Liao (Nanyang Technological University), Jingwen Hou (Nanyang Technological University), Chaofeng Chen (Nanyang Technological University), Erli Zhang (Nanyang Technological University), Annan Wang (Nanyang Technological University), Wenxiu Sun (Sensetime Research), Qiong Yan (Sensetime Research), and Weisi Lin (Nanyang Technological University)</i>	
Just Noticeable Difference Estimation for Screen Content Images: A Content Uncertainty-Guided Approach	372
<i>Lirong Huang (Shenzhen University, China), Rong Zhang (Shenzhen University, China), and Miaohui Wang (Shenzhen University, China)</i>	
Intermediate-Task Learning with Pretrained Model for Synthesized Speech Mos Prediction	378
<i>Hui Wang (Nankai University, China), Xiguang Zheng (Kuaishou Technology, China), and Yong Qin (Nankai University, China)</i>	
Cross-Modal-Aware Representation Learning with Syntactic Hypergraph Convolutional Network for VideoQA	384
<i>Zenan Xu (Sun Yat-sen University, China), Wanjun Zhong (Sun Yat-sen University, China), Qinliang Su (Sun Yat-sen University; Guangdong Key Laboratory of Big Data Analysis and Processing, China), and Fuwei Zhang (Sun Yat-sen University, China)</i>	

Semantic Segmentation

SASFormer: Transformers for Sparsely Annotated Semantic Segmentation	390
<i>Hui Su (Zhejiang Lab, China), Yue Ye (Zhejiang Lab, China), Wei Hua (Zhejiang Lab, China), Lechao Cheng (Zhejiang Lab, China), and Mingli Song (Zhejiang University, China)</i>	
Holistic Weighted Distillation for Semantic Segmentation	396
<i>Wujie Sun (Zhejiang University, China), Defang Chen (Zhejiang University, China), Can Wang (Zhejiang University, China), Deshi Ye (Zhejiang University, China), Yan Feng (Zhejiang University, China), and Chun Chen (Zhejiang University, China)</i>	
Knowledge Distillation from 3D to Bird's-Eye-View for LiDAR Semantic Segmentation	402
<i>Feng Jiang (Fudan University, China), Heng Gao (Fudan University, China), Shoumeng Qiu (Fudan University, China), Haiqiang Zhang (Mogo Auto, China), Ru Wan (Mogo Auto, China), and Jian Pu (Fudan University, China)</i>	
A Pseudo-Dual Self-Rectification Framework for Semantic Segmentation	408
<i>Huazheng Hao (Ningbo University, China), Hui Xiao (Ningbo University, China), Li Dong (Ningbo University, China), Diqun Yan (Ningbo University, China), Dongtai Liang (Ningbo University, China), Jiayan Zhuang (Ningbo Institute of Materials Technology & Engineering, Chinese Academy of Sciences, China), and Chengbin Peng (Ningbo University; Key Laboratory of Mobile Network Application Technology of Zhejiang Province, China)</i>	
ICANet: A Lightweight Increasing Context Aided Network for Real-Time Image Semantic Segmentation	414
<i>Lei Chen (Inner Mongolia University, China), Huhe Dai (Inner Mongolia University, China), and Yuan Zheng (Inner Mongolia University, China)</i>	
Dual-Level Consistency Learning for Unsupervised Domain Adaptive Night-Time Semantic Segmentation	420
<i>Feifei Ding (Hangzhou Dianzi University, China), Jianjun Li (Hangzhou Dianzi University, China), and Wanyong Tian (The Key laboratory of Data Link of China Electronics Technology Group Corporation, China)</i>	

Zero-Shot Learning

Modality-Fusion Spiking Transformer Network for Audio-Visual Zero-Shot Learning	426
<i>Wenrui Li (Harbin Institute of Technology, China), Zhengyu Ma (Peng Cheng Laboratory, China), Liang-Jian Deng (University of Electronic Science and Technology of China, China), Hengyu Man (Harbin Institute of Technology, China), and Xiaopeng Fan (Harbin Institute of Technology, China)</i>	
Privacy-Enhanced Zero-Shot Learning via Data-Free Knowledge Transfer	432
<i>Rui Gao (Xi'an Jiaotong University, China), Fan Wan (Durham University, UK), Daniel Organisciak (Northumbria University, UK), Jiyao Pu (Durham University, UK), Haoran Duan (Durham University, UK), Peng Zhang (Durham University, UK), Xingsong Hou (Xi'an Jiaotong University, China), and Yang Long (Durham University, UK)</i>	

Swap-Reconstruction Autoencoder for Compositional Zero-Shot Learning	438
<i>Ting Guo (Shanxi University, China), Jiye Liang (Shanxi University, China), and Guo-Sen Xie (Nanjing University of Science and Technology, China)</i>	
Synthetic Feature Assessment for Zero-Shot Object Detection	444
<i>Xinmiao Dai (Ningbo University, China), Chong Wang (Ningbo University, China), Haohe Li (Ningbo University, China), Sunqi Lin (Ningbo University, China), Li Dong (Ningbo University, China), Jiafei Wu (SenseTime Research, China), and Jun Wang (China University of Mining and Technology, China)</i>	
Audio-Visual Generalized Zero-Shot Learning Based on Variational Information Bottleneck	450
<i>Yapeng Li (Wuhan University, China; Hubei LuoJia Laboratory, China), Yong Luo (Wuhan University, China; Hubei LuoJia Laboratory, China), and Bo Du (Wuhan University, China; Hubei LuoJia Laboratory, China)</i>	
Fine-Grained Primitive Representation Learning for Compositional Zero-Shot Classification ...	456
<i>Han Jiang (University of Science and Technology of China), Xiaoshan Yang (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Peng Cheng Laboratory), Chaofan Chen (University of Science and Technology of China), and Changsheng Xu (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences; Peng Cheng Laboratory)</i>	

Super Resolution & Inpainting I

Medical Image Super-Resolution via Diagnosis-Guided Attention	462
<i>Jingwei Wang (Anhui University, China), Peng Zhou (Anhui University, China), Xianjun Han (Anhui University, China), and Yanming Chen (Anhui University, China)</i>	
Denser is Better:cost Distribution Super-Resolution Network for More Accurate sub-Pixel Disparity	468
<i>Hong Zhang (Dalian university of technology), Shenglun Chen (Dalian university of technology), Zhihui Wang (Dalian university of technology), Haojie Li (Dalian university of technology), and Wanli Ouyang (The University of Sydney)</i>	
DSP-Net: Diverse Structure Prior Network for Image Inpainting	474
<i>Lin Sun (Hunan University, China), Chao Yang (Hunan University, China), and Bin Jiang (Hunan University, China)</i>	
Joint Feature Aggregation for Stereo Image Super-Resolution	480
<i>Zekun Ai (Xiamen University, China), Xiaotong Luo (Xiamen University, China), and Yanyun Qu (Xiamen University, China)</i>	
Joint Super-Resolution and Classification Based on Bidirectional Mapping and Multiple Constraints	486
<i>Zijian Yuan (Guangxi University, China), Kan Chang (Guangxi University, China), Zhiquan Liu (Guangxi University, China), Xinjie Wei (Guangxi University, China), and Boning Chen (The University of Melbourne, Australia)</i>	

Inpainting of Remote Sensing Sea Surface Temperature Image with Multi-scale Physical Constraints	492
<i>Qichen Wei (Ocean University of China, China), Zijie Zuo (Ocean University of China, China), Jie Nie (Ocean University of China, China), Jiahao Du (Ocean University of China, China), Yaning Diao (Ocean University of China, China), Min Ye (Ocean University of China, China), and Xinyue Liang (Ocean University of China, China)</i>	

Enhancement, Restoration, Deblurring

Generalized Compressed Video Restoration by Multi-Scale Temporal Fusion and Hierarchical Quality Score Estimation	498
<i>Zhijie Huang (Peking University, China), Tianyi Sun (Peking University, China), Xiaopeng Guo (Peking University, China), Yanze Wang (Peking University, China), and Jun Sun (Peking University, China)</i>	
Edgeformer: Edge-Enhanced Transformer for High-Quality Image Deblurring	504
<i>Yuan Zou (University of Science and Technology of China, China) and Yinyao Ma (University of Science and Technology of China, China)</i>	
Generative Iris Prior Embedded Transformer for Iris Restoration	510
<i>Yubo Huang (Beijing University of Posts and Telecommunications, China), Jia Wang (Beijing University of Posts and Telecommunications, China), Peipei Li (Beijing University of Posts and Telecommunications, China), Liuyu Xiang (Beijing University of Posts and Telecommunications, China), Peigang Li (Beijing University of Posts and Telecommunications, China), and Zhaofeng He (Beijing University of Posts and Telecommunications, China)</i>	
Industrial Image Enhancement Method Based on Multi-exposure Image Sequence Fusion	516
<i>Yangliu Ding (Jiangsu University, China), Keyang Cheng (Jiangsu University, China), Yongzhao Zhan (Jiangsu University, China), and Liuyang Yan (Jiangsu University, China)</i>	
MBDFNet: Multi-Scale Bidirectional Dynamic Feature Fusion Network for Efficient Image Deblurring	522
<i>Zhongbao Yang (Nanjing University of Science and Technology) and Jinshan Pan (Nanjing University of Science and Technology)</i>	
Multiple Degraded Image Restoration via Degradation History Estimation	528
<i>Minhua Liu (Shenzhen University, China), Yuanman Li (Shenzhen University, China), Rongqin Liang (Shenzhen University, China), Jiaxiang You (Shenzhen University, China), and Xia Li (Shenzhen University, China)</i>	

Learning Techniques III

Gradual Migration and Style Consistency for Unsupervised Domain Adaptation	534
<i>Jintao Zhang (Hunan University, China) and Guangyi Xiao (Hunan University, China)</i>	

Adapt then Generalize: A Simple Two-Stage Framework for Semi-Supervised Domain Generalization	540
<i>Han Xie (Fuzhou University, China), Zhifeng Shen (Fuzhou University, China), Shicai Yang (Hikvision Research Institute, China), Weijie Chen (Hikvision Research Institute, China), and LuoJun Lin (Fuzhou University, China)</i>	
Rethinking Overfitting of Multiple Instance Learning for Whole Slide Image Classification	546
<i>Hongjian Song (South China University of Technology, China), Jie Tang (South China University of Technology, China), Hongzhao Xiao (South China University of Technology, China), and Juncheng Hu (Guilin Medical University, China)</i>	
A Unified MRC Framework with Multi-query for Multi-modal Relation Triplets Extraction	552
<i>Qiang Chen (Soochow University, China), Dong Zhang (Soochow University, China), Shoushan Li (Soochow University, China), and Guodong Zhou (Soochow University, China)</i>	
Feature Bias Correction: A Feature Augmentation Method for Long-Tailed Recognition	558
<i>Jiaxin Yang (National university of defense technology, China), Xiaofei Li (National university of defense technology, China), Jun Zhang (National university of defense technology, China), and Shuohao Li (National university of defense technology, China)</i>	
Recombination Samples Training for Robust Natural Language Visual Reasoning	564
<i>Yuling Jiang (Nanjing University of Posts and Telecommunications, China), Yingyuan Zhao (Nanjing University of Posts and Telecommunications, China), and Bing-kun Bao (Nanjing University of Posts and Telecommunications, China)</i>	
SG-NeRF: Semantic-Guided Point-Based Neural Radiance Fields	570
<i>Yansong Qu (Beihang University, China), Yuze Wang (Beihang University, China), and Yue Qi (Beihang University, China; Peng Cheng Laboratory, China)</i>	
RTMC: A Robust Trusted Multi-view Classification Framework	576
<i>Hai Zhou (Beijing University of Posts and Telecommunications, China), Zhe Xue (Beijing University of Posts and Telecommunications, China), Ying Liu (Beijing University of Posts and Telecommunications, China), Boang Li (Beijing University of Posts and Telecommunications, China), Junping Du (Beijing University of Posts and Telecommunications, China), and Meiyu Liang (Beijing University of Posts and Telecommunications, China)</i>	
DF-CLIP: Towards Disentangled and Fine-Grained Image Editing from Text	582
<i>Xinjiao Zhou (Hunan University, China), Bin Jiang (Hunan University, China), Chao Yang (Hunan University, China), Haotian Hu (Hunan University, China), and Xiaofei Huo (Hunan University, China)</i>	
Letter Embedding Guidance Diffusion Model for Scene Text Editing	588
<i>Changshuo Wang (Shandong University, China), Lei Wu (Shandong University, China), Xu Chen (Shandong University, China), Xiang Li (Shandong University, China), Lei Meng (Shandong University, China), and Xiangxu Meng (Shandong University, China)</i>	

Applications

Cluster-Driven GNN-Based Federated Recommendation with Biased Message Dropout	594
<i>Rongyu Zhang (The Chinese University of Hong Kong, China), Yun Chen (The Chinese University of Hong Kong, China), Chenrui Wu (The Chinese University of Hong Kong, China), and Fangxin Wang (The Chinese University of Hong Kong; The Guangdong Provincial Key Laboratory of Future Networks of Intelligence; Peng Cheng Laboratory, China)</i>	
SQT: Debiased Visual Question Answering via Shuffling Question Types	600
<i>Tianyu Huai (East China Normal University, China), Shuwen Yang (East China Normal University, China), Junhang Zhang (East China Normal University, China), Guoan Wang (East China Normal University, China), Xinru Yu (East China Normal University, China), Tianlong Ma (East China Normal University, China), and Liang He (East China Normal University, China)</i>	
Fast Personalized Human Activity Recognition on Heuristic Parameter Estimation	606
<i>Shizhuo Deng (Northeastern University, China), Chuangui Yang (Baidu Inc., China), Zhubao Guo (Northeastern University, China), Boqian Lin (Northeastern University, China), Dongyue Chen (Northeastern University, China), Tong Jia (Northeastern University, China), and Botao Wang (Northeastern University, China)</i>	
Improving Automatic Singing Skill Evaluation with Timbral Features, Attention, and Singing Voice Separation	612
<i>Yaolong Ju (Huawei Technologies Co., Ltd., China), Chunyang Xu (Huawei Technologies Co., Ltd., China), Yichen Guo (Huawei Technologies Co., Ltd., China), Jinhu Li (Huawei Technologies Co., Ltd., China), and Simon Lui (Huawei Technologies Co., Ltd., China)</i>	
Learning High Frequency Surface Functions In Shells	618
<i>Han Guo (Fuzhou University), Yuanlong Yu (Fuzhou University), Yujie Wang (Shandong University), Xuelin Chen (Tencent AI Lab), and Yixin Zhuang (Fuzhou University)</i>	
Multi-template Tracker Driven by Cache Manager Algorithm, Towards Multi-distractor Scenarios	624
<i>Eli Lei (Wuhan University, China), Jia Shao (Huazhong Agricultural University, China), Youfa Liu (Huazhong Agricultural University, China), and Bo Du (Wuhan University, China)</i>	
Material-Aware Self-Supervised Network for Dynamic 3D Garment Simulation	630
<i>Aoran Liu (The University of Sydney, Australia), Kun Hu (The University of Sydney, Australia), Wenxi Yue (The University of Sydney, Australia), Qiuxia Wu (South China University of Technology, China), and Zhiyong Wang (The University of Sydney, Australia)</i>	
Multi-Speaker Direction of Arrival Estimation using Audio and Visual Modalities with Convolutional Neural Network	636
<i>Yulin Wu (Wuhan University, China), Ruimin Hu (Wuhan University, China), and Xiaochen Wang (Wuhan University, China)</i>	
Multi-scale Hybrid Fusion Network for Mandarin Audio-Visual Speech Recognition	642
<i>Jinxin Wang (Ocean University of China, China), Zhongwen Guo (Ocean University of China, China), Chao Yang (University of Technology Sydney, Australia), Xiaomei Li (Ocean University of China, China), and Ziyuan Cui (Ocean University of China, China)</i>	

Trustworthy Federated Learning for Multimedia II

Cross-Training with Prototypical Distillation for Improving the Generalization of Federated Learning	648
<i>Tianhan Liu (Shandong University, China), Zhuang Qi (Shandong University, China), Zitan Chen (Shandong University, China), Xiangxu Meng (Shandong University, China), and Lei Meng (Shandong University; Shandong Research Institute of Industrial Technology, China)</i>	
A Content-based Viewport Prediction Framework for 360Video Using Personalized Federated Learning and Fusion Techniques	654
<i>Mehdi Setayesh (The University of British Columbia, Canada) and Vincent W.S. Wong (The University of British Columbia, Canada)</i>	
Learning Cautiously in Federated Learning with Noisy and Heterogeneous Clients	660
<i>Chenrui Wu (The Chinese University of Hong Kong, China), Zexi Li (Zhejiang University, China), Fangxin Wang (The Chinese University of Hong Kong; The Guangdong Provincial Key Laboratory of Future Networks of Intelligence; Peng Cheng Laboratory, China), and Chao Wu (Zhejiang University, China)</i>	
Multi-tier Client Selection for Mobile Federated Learning Networks	666
<i>Yulan Gao (Nanyang Technological University, Singapore), Yansong Zhao (Nanyang Technological University, Singapore), and Han Yu (Nanyang Technological University, Singapore)</i>	
SWATM: Contribution-Aware Adaptive Federated Learning Framework Based on Augmented Shapley Values	672
<i>Chengyi Yang (ENN Group, China), Zhaoxiang Hou (ENN Group, China), Sheng Guo (ENN Group, China), Hui Chen (ENN Group, China), and Zengxiang Li (ENN Group, China)</i>	
FedDBM: Federated Digital Biomarker for Detecting Parkinson's Disease Progress	678
<i>Yiqiang Chen (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Xiaodong Yang (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Yuting He (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Chunyan Miao (Nanyang Technologies University, Singapore), and Piu Chan (Capital Medical University, China)</i>	

Media Coding

MIXLIC: Mixing Global and Local Context Model for Learned Image Compression	684
<i>Haihang Ruan (Tsinghua University; Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Feng Wang (Peking University, China), Tongda Xu (Tsinghua University, China), Zhiyong Tan (Chinese Academy of Sciences, China), and Yan Wang (Tsinghua University, China)</i>	
Model-Driven Compression for Digital Human using Multi-granularity Representations	690
<i>Ruoke Yan (Peking University, China), Qian Yin (Peking University, China), Xinfeng Zhang (University of Chinese Academy of Sciences, China), and Siwe Ma (Peking University, China)</i>	

Meta-ILF: In-Loop Filter with Customized Weights for VVC Intra Coding	696
<i>Hengyu Man (Harbin Institute of Technology, China), Xingtao Wang (Harbin Institute of Technology, China), Riyu Lu (Harbin Institute of Technology, China), and Xiaopeng Fan (Harbin Institute of Technology, China)</i>	
Variable-Rate Neural Image Compression with Joint Content-Channel Features and Accurate R- λ Model	702
<i>Yunhui Shi (Beijing University of Technology, China), Pengquan Wang (Beijing University of Technology, China), Jin Wang (Beijing University of Technology, China), Baocai Yin (Beijing University of Technology, China), and Nam Ling (Santa Clara University, USA)</i>	
Peer Upsampled Transform Domain Prediction for G-PCC	708
<i>Wenyi Wang (ByteDance Inc., China), Yingzhan Xu (ByteDance Inc., China), Kai Zhang (ByteDance Inc., US), and Li Zhang (ByteDance Inc., US)</i>	
Optimizing DNN Based Quality Assessment Metric for Image Compression: A Novel Rate Control Method	714
<i>Qiuyue Fang (Beihang University, China), Tao Xu (Beihang University, China), Lai Jiang (University of British Columbia, Canada), Shengxi Li (Beihang University, China), Mai Xu (Beihang University, China), Yunjin Chen (n/a), and Leonid Sigal (University of British Columbia, Canada)</i>	

Few-Shot Learning

Dual-Expert Distillation Network for Few-Shot Segmentation	720
<i>Junhang Zhang (East China Normal University, China), Zisong Zhuang (East China Normal University, China), Luwei Xiao (East China Normal University, China), Xingjiao Wu (Fudan University, China), Tianlong Ma (East China Normal University, China), and Liang He (East China Normal University, China)</i>	
Rethinking Self-Supervision for Few-Shot Class-Incremental Learning	726
<i>Linglan Zhao (Shanghai Jiao Tong University, China), Jing Lu (University of Science and Technology of China, China), Zhazhan Cheng (Zhejiang University & Hikvision Research Institute, China), Duo Liu (Shanghai Jiao Tong University, China), and Xiangzhong Fang (Shanghai Jiao Tong University, China)</i>	
Two-Level Graph Network for Few-Shot Class-Incremental Learning	732
<i>Hao Chen (Suzhou University of Science and Technology), Linyan Li (Suzhou Institute of Trade & Commerce), Fan Lyu (Tianjin University), Fuyuan Hu (Suzhou University of Science and Technology; Jiangsu Industrial Intelligent and Low-carbon Technology Engineering Center; Suzhou Key Laboratory of Intelligent Low-carbon Technology Application), Zhenping Xia (Suzhou University of Science and Technology), and Fenglei Xu (Suzhou University of Science and Technology)</i>	

Learning Component-Level and Inter-Class Glyph Representation for Few-Shot Font Generation... 738

Yongliang Su (Shandong University, China), Xu Chen (Shandong University, China), Lei Wu (Shandong University, China), and Xiangxu Meng (Shandong University, China)

Masked Cross-Image Encoding for Few-Shot Segmentation 744

Wenbo Xu (University of Technology Sydney, Australia, Australia), Huaxi Huang (Commonwealth Scientific and Industrial Research Organisation, Australia), Ming Cheng (University of Technology Sydney, Australia, Australia), Litao Yu (University of Technology Sydney, Australia, Australia), Qiang Wu (University of Technology Sydney, Australia, Australia), and Jian Zhang (University of Technology Sydney, Australia, Australia)

Frame-Level Embedding Learning for Few-Shot Bioacoustic Event Detection 750

Xueyang Zhang (iFlytek Research, China), Shuxian Wang (University of Science and Technology of China, China), Jun Du (University of Science and Technology of China, China), Genwei Yan (China University of Mining and Technology, China), Jigang Tang (iFlytek Research, China), Tian Gao (iFlytek Research, China), Xin Fang (iFlytek Research, China), Jia Pan (iFlytek Research, China), and Jianqing Gao (iFlytek Research, China)

Action Recognition

End-to-End Part-Level Action Parsing with Transformer 756

Xiaojia Chen (University of Electronic Science and Technology of China, China), Xuanhan Wang (University of Electronic Science and Technology of China, China), Beitao Chen (University of Electronic Science and Technology of China, China), and Lianli Gao (University of Electronic Science and Technology of China, China)

Leveraging Attribute Knowledge for Open-set Action Recognition 762

Kaixiang Yang (Tianjin University of Technology, China), Junyu Gao (Chinese Academy of Sciences, China), Yangbo Feng (Tianjin University of Technology, China), and Changsheng Xu (Chinese Academy of Sciences, China)

ConCAP: Contrastive Context-Aware Prompt for Resource-Hungry Action Recognition 768

Hailun Zhang (Sichuan University), Ziyun Zeng (Tsinghua University), Qijun Zhao (Sichuan University), and Zhen Zhai (Sichuan University)

Bullet-Screen-Emoj Attack with Temporal Difference Noise for Video Action Recognition 774

Yongkang Zhang (Capital Normal University, China), Jun Li (Capital Normal University, China), Zhiping Shi (Capital Normal University, China), Na Jiang (Capital Normal University, China), Qichuan Geng (Capital Normal University, China), and Yifei Pei (Capital Normal University, China)

Is Really Correlation Information Represented Well in Self-Attention for Skeleton-Based Action Recognition?	780
<i>Wentian Xin (Xidian University, China), Hongkai Lin (Xidian University, China), Ruyi Liu (Xidian University, China), Yi Liu (Xidian University, China), and Qiguang Miao (Xidian University, China)</i>	
DD-GCN: Directed Diffusion Graph Convolutional Network for Skeleton-Based Human Action Recognition	786
<i>Chang Li (Key Laboratory of Water Big Data Technology of Ministry of Water Resources, China; Hohai University, China;), Qian Huang (Key Laboratory of Water Big Data Technology of Ministry of Water Resources, China; Hohai University, China), and Yingchi Mao (Key Laboratory of Water Big Data Technology of Ministry of Water Resources, China; Hohai University, China)</i>	

Knowledge Distillation I

Improving CTC-Based Handwritten Chinese Text Recognition with Cross-Modality Knowledge Distillation and Feature Aggregation	792
<i>Shilian Wu (University of Science and Technology of China, China), Yongrui Li (University of Science and Technology of China, China), and Zengfu Wang (University of Science and Technology of China, China)</i>	
Decoupled Mutual Distillation for Incremental Object Detection	798
<i>Gao-Dong Liu (Xiamen University, China), Wan-Lei Zhao (Xiamen University, China), and Jie Zhao (BodenAI, China)</i>	
Robust Cross-Modal Knowledge Distillation for Unconstrained Videos	804
<i>Wenke Xia (Renmin University of China; Baidu Research), Xingjian Li (Baidu Research; University of Macau), Andong Deng (Baidu Research; University of Central Florida), Haoyi Xiong (Baidu Research), Dejing Dou (BCG X), and Di Hu (Renmin University of China)</i>	
Accelerating Diffusion Sampling with Classifier-Based Feature Distillation	810
<i>Wujie Sun (Zhejiang University, China), Defang Chen (Zhejiang University, China), Can Wang (Zhejiang University, China), Deshi Ye (Zhejiang University, China), Yan Feng (Zhejiang University, China), and Chun Chen (Zhejiang University, China)</i>	
Semantic Stage-Wise Learning for Knowledge Distillation	816
<i>Dongqin Liu (Institute of Information Engineering, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Wentao Li (Institute of Information Engineering, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Wei Zhou (Chinese Academy of Sciences of Sciences, China), Zhaoxing Li (Institute of Information Engineering, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Jiao Dai (Institute of Information Engineering, Chinese Academy of Sciences of Sciences, China), Jizhong Han (Institute of Information Engineering, Chinese Academy of Sciences of Sciences, China), Ruixuan Li (Huazhong University of Science and Technology, China), and Songlin Hu (Institute of Information Engineering, Chinese Academy of Sciences of Sciences, China)</i>	

Discriminative Gradient Adjustment with Coupled Knowledge Distillation for Class Incremental Learning	822
<i>Hao Zhang (Sun Yat-sen University, China), Yanxu Hu (Sun Yat-sen University, China), Jiawen Peng (Sun Yat-sen University, China), and Andy J Ma (Sun Yat-sen University, China; Guangdong Province Key Laboratory of Information Security Technology, China; Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, China)</i>	

Semantic Processing II

SACANet: Scene-Aware Class Attention Network for Semantic Segmentation of Remote Sensing Images	828
<i>Xiaowen Ma (Zhejiang University), Rui Che (Zhejiang University), Tingfeng Hong (Zhejiang University), Mengting Ma (Zhejiang University), Ziyang Zhao (Zhejiang University), Tian Feng (Zhejiang University; Alibaba-Zhejiang University Joint Research Institute of Frontier Technologies), and Wei Zhang (Zhejiang University)</i>	
Few-Shot Semantic Segmentation by Exploiting Dynamic and Regional Contexts	834
<i>Hongyu Gu (Dalian University of Technology, China), Yunzhi Zhuge (The University of Adelaide, Australia), Lu Zhang (Dalian University of Technology, China), Jinqing Qi (Dalian University of Technology, China), and Huchuan Lu (Dalian University of Technology, China)</i>	
DiST-GAN: Distillation-Based Semantic Transfer for Text-Guided Face Generation	840
<i>Guoxing Yang (Renmin University of China, China), Feifei Fu (Renmin University of China, China), Nanyi Fei (Renmin University of China, China), Haoran Wu (China Unicom Research Institute, China), Ruitao Ma (China Unicom Research Institute, China), and Zhiwu Lu (Renmin University of China, China)</i>	
Self-Attention Prediction Correction with Channel Suppression for Weakly-Supervised Semantic Segmentation	846
<i>Guoying Sun (Sun Yat-sen University; Xidian University, China) and Meng Yang (Sun Yat-sen University; Xidian University, China)</i>	
Semi-Supervised Semantic Segmentation With Region Relevance	852
<i>Rui Chen (Nanjing University of Science and Technology, China), Tao Chen (Nanjing University of Science and Technology, China), Qiong Wang (Nanjing University of Science and Technology, China), and Yazhou Yao (Nanjing University of Science and Technology, China)</i>	
Who, What and Where: Composite-Semantic Instance Search for Story Videos	858
<i>Jiahao Guo (Wuhan University), Chao Liang (Wuhan University), and Zhongyuan Wang (Wuhan University)</i>	
Semantic Embedding Uncertainty Learning for Image and Text Matching	864
<i>Yan Wang (Tianjin University; Hefei Comprehensive National Science Center, China), Yu-ting Su (Tianjin University, China), Wenhui Li (Tianjin University, China), Chenggang Yan (Hangzhou Dianzi University, China), Bolun Zheng (Hangzhou Dianzi University, China), Xuanya Li (Baidu Inc, China), and An-An Liu (Tianjin University, China; Hefei Comprehensive National Science Center, China)</i>	

Multi-view Network Embedding with Structure and Semantic Contrastive Learning	870
<i>Yifan Shang (University of Tsukuba, Japan), Xiucai Ye (University of Tsukuba, Japan), and Tetsuya Sakurai (University of Tsukuba, Japan)</i>	
A Self-Training Framework Based on Multi-scale Attention Fusion for Weakly Supervised Semantic Segmentation	876
<i>Guoqing Yang (Beijing University of Posts and Telecommunications, China), Chuang Zhu (Beijing University of Posts and Telecommunications, China), and Yu Zhang (Beijing University of Posts and Telecommunications, China)</i>	

3D & Depth II

Depth and DOF Cues Make a Better Defocus Blur Detector	882
<i>Yuxin Jin (Wuhan University, China), Ming Qian (Wuhan University, China), Jincheng Xiong (Wuhan University, China), Nan Xue (Wuhan University, China), and Gui-Song Xia (Wuhan University, China)</i>	
Explainable Unfolding Network for Joint Edge-Preserving Depth Map Super-Resolution	888
<i>Jialong Zhang (Taiyuan University of Science and Technology, China), Lijun Zhao (Taiyuan University of Science and Technology, China), Jinjing Zhang (North University of China, China), Ke Wang (Taiyuan University of Science and Technology, China), and Anhong Wang (Taiyuan University of Science and Technology, China)</i>	
DFR: Depth from Rotation by Uncalibrated Image Rectification with Latitudinal Motion Assumption	894
<i>Yongcong Zhang (Hunan University), Yifei Xue (Jiangxi Provincial Natural Resources Cause Development Center), Ming Liao (Jiangxi Provincial Natural Resources Cause Development Center), Huiqing Zhang (Hunan University), and Yizhen Lao (Hunan University)</i>	
Robust 3D Craniofacial Landmarks Localization by an End-to-End Regression Network	900
<i>Xianhe Jiao (Qingdao University, China), Junli Zhao (Qingdao University, China), Chenlei Lv (Shenzhen University, China), Fuqing Duan (Beijing Normal University, China), Zhenkuan Pan (Qingdao University, China), and Xin Li (Texas A&M University)</i>	
Twins-Mix: Self Mixing in Latent Space for Reasonable Data Augmentation of 3D Computer-Aided Design Generative Modeling	906
<i>Xueyang Li (Fudan University, China), Minyang Xu (Fudan University, China), and Xiangdong Zhou (Fudan University, China)</i>	
Rendering and Reconstruction Based 3D Portrait Stylization	912
<i>Shaoxu Li (Shanghai Jiao Tong University, China) and Ye Pan (Shanghai Jiao Tong University, China)</i>	
GT-Net: Variational Autoencoder Networks Based on Graph Transformer for 3D Shape Learning ..	918
<i>Zhenjiang Du (University of Electronic Science and Technology of China, China), Yi Lu (University of Electronic Science and Technology of China, China), Guan Wang (University of Electronic Science and Technology of China, China), Ning Xie (University of Electronic Science and Technology of China, China), and Yang Yang (University of Electronic Science and Technology of China, China)</i>	

M-GCN: Multi-scale Graph Convolutional Network for 3D Point Cloud Classification	924
<i>Jing Hu (University of Electronic Science and Technology of China, China), Xincheng Wang (University of Melbourne, Australia), Ziheng Liao (University of Electronic Science and Technology of China, China), and Tingsong Xiao (University of Florida, USA)</i>	
A Lightweight Grouped Low-Rank Tensor Approximation Network for 3D Mesh Reconstruction from Videos	930
<i>Boyang Zhang (Ningxia University, China), Suping Wu (Ningxia University, China), Leyang Yang (Ningxia University, China), Bin Wang (Ningxia University, China), and Wenlong Lu (Ningxia University, China)</i>	

Multimedia-based Health Computing

Hierarchical Attention Learning for Multimodal Classification	936
<i>Xin Zou (China University of Geosciences, China), Chang Tang (China University of Geosciences, China), Wei Zhang (Qilu University of Technology (Shandong Academy of Sciences), China), Kun Sun (China University of Geosciences, China), and LiangXiao Jiang (China University of Geosciences, China)</i>	
An End-to-End Food Portion Estimation Framework Based on Shape Reconstruction from Monocular Image	942
<i>Zeman Shao (Purdue University, USA), Gautham Vinod (Purdue University, USA), Jiangpeng He (Purdue University, USA), and Fengqing Zhu (Purdue University, USA)</i>	
Unsupervised Domain Adaptation for Neuron Membrane Segmentation Based on Structural Features	948
<i>Yuxiang An (The University of Sydney, Australia), Dongnan Liu (The University of Sydney, Australia), and Weidong Cai (The University of Sydney, Australia)</i>	
Latent Feature Regularization Based Adversarial Network for Brain Tumor Anomaly Detection	954
<i>Nan Wang (East China Normal University, China), Chengwei Chen (East China Normal University, China), Shaohui Lin (East China Normal University, China), and Lizhuang Ma (East China Normal University, China)</i>	
Development of Deep Learning Algorithms for Automated Scoliosis and Abnormal Posture Screening using 2D Back Image	960
<i>Zhenda Xu (Hong Kong Polytechnic University, Hong Kong), Jiaohao Hu (Northwestern Polytechnical University, China), Qiang Gao (Sichuan University, China), Donghua Hang (Shanghai Jiao Tong University, China), Qihua Zhou (Hong Kong Polytechnic University, Hong Kong), Song Guo (Hong Kong Polytechnic University, Hong Kong), and Aiqian Gan (The University of Sydney, Australia)</i>	
LACL: Lesion-Aware Contrastive Learning Framework for Medical Image Classification	966
<i>Yu Tang (Renmin University of China, China), Gang Yang (Renmin University of China, China), Jianchun Zhao (Vistel Inc., China), Dayong Ding (Vistel Inc., China), and Jun Wu (Northwestern Polytechnical University, China)</i>	

Storage, Transmission & Communication

- Collaborative Edge Caching: A Meta Reinforcement Learning Approach with Edge Sampling ... 972
Yinan Mao (Tsinghua University, China), Bowei He (City University of Hong Kong, China), Shiji Zhou (Tsinghua University, China), Chen Ma (City University of Hong Kong, China), and Zhi Wang (Tsinghua University, China)
- PACC: Perception Aware Congestion Control for Real-Time Communication 978
Feng Peng (Shanghai Jiao Tong University, China), Bingcong Lu (Shanghai Jiao Tong University, China), Li Song (Shanghai Jiao Tong University, China), Rong Xie (Shanghai Jiao Tong University, China), Yanmei Liu (Alibaba Group, China), and Ying Chen (Alibaba Group, China)
- Label-Semantic-Enhanced Online Hashing for Efficient Cross-Modal Retrieval 984
Xueting Jiang (Huaqiao University, China), Xin Liu (Huaqiao University, China; Zhejiang Lab, China; Xiamen Key Lab. of Comput. Vis. Pattern Recognit. & Fujian Key Lab. of Big Data Intell. and Security, China), Yiu-ming Cheung (Hong Kong Baptist University, Hong Kong SAR, China), Xing Xu (University of Electronic Science and Technology of China, China), Shukai Zheng (Zhejiang Lab, China), and Taihao Li (Zhejiang Lab, China)
- QoE Maximization for Aerial Video Streaming with Multiple Cellular Connected UAVs 990
Cheng Zhan (Southwest University, China), Huan Yan (Southwest University, China), Han Hu (Beijing Institute of Technology, China), Liyue Zhu (Academy of Broadcasting Science, SAPPRFT, China), and Shubin Xu (CETC Advanced Mobile Communication Innovation Center, China)
- Multi-stream Adaptive Offloading of Joint Compressed Video Streams, Feature Streams, and Semantic Streams in Edge Computing Systems 996
Dieli Hu (Institute of Computing Technology, Chinese Academy of Sciences; Peng Cheng Laboratory; University of Chinese Academy of Sciences, China), Wen Ji (Institute of Computing Technology, Chinese Academy of Sciences; Peng Cheng Laboratory, China), and Zhi Wang (Tsinghua University; Peng Cheng Laboratory, China)
- L4S Congestion Control Algorithm for Interactive Low Latency Applications over 5G 1002
Jangwoo Son (Fraunhofer HHI, Germany), Yago Sanchez (Fraunhofer HHI, Germany), Christian Hampe (Deutsche Telekom, Germany), Dominik Schnieders (Deutsche Telekom, Germany), Thomas Schierl (Fraunhofer HHI, Germany), and Cornelius Hellge (Fraunhofer HHI, Germany)

Action Detection & Localization

- Weakly-Supervised Temporal Action Localization with Adaptive Clustering and Refining Network 1008
Hao Ren (Fudan University, China), Wu Ran (Fudan University, China), Xingson Liu (Fudan University, China), Haoran Ren (Fudan University, China), Hong Lu (Fudan University, China), Rui Zhang (Fudan University, China), and Cheng Jin (Fudan University, China; Peng Cheng Laboratory, China)

Do we Really Need Temporal Convolutions in Action Segmentation?	1014
<i>Dazhao Du (University of Chinese Academy of Sciences), Bing Su (Renmin University of China), Yu Li (International Digital Economy Academy), Zhongang Qi (Tencent PCG), Lingyu Si (University of Chinese Academy of Sciences), and Ying Shan (Tencent PCG)</i>	
ELAN: Enhancing Temporal Action Detection with Location Awareness	1020
<i>Guo Chen (Nanjing University), Yin-Dong Zheng (Nanjing University), Zhe Chen (Nanjing University), Jiahao Wang (Nanjing University), and Tong Lu (Nanjing University)</i>	
MRSN: Multi-relation Support Network for Video Action Detection	1026
<i>Yin-Dong Zheng (Nanjing University), Guo Chen (Nanjing University), Minglei Yuan (Nanjing University), and Tong Lu (Nanjing University)</i>	
Unleashing the Potential of Adjacent Snippets for Weakly-Supervised Temporal Action Localization	1032
<i>Qinying Liu (University of Science and Technology of China, China), Zilei Wang (University of Science and Technology of China, China), Ruoxi Chen (University of Science and Technology of China, China), and Zhilin Li (University of Science and Technology of China, China)</i>	
Compositional Learning in Transformer-Based Human-Object Interaction Detection	1038
<i>Zikun Zhuang (Tongji University, China), Ruihao Qian (Tongji University, China), Chi Xie (Tongji University, China), and Shuang Liang (Tongji University, China)</i>	

Contrastive Learning

Self-Supervised Cross-Stage Regional Contrastive Learning for Object Detection	1044
<i>Junkai Yan (Sun Yat-sen University; Ministry of Education, China), Lingxiao Yang (Sun Yat-sen University, China), Yipeng Gao (Sun Yat-sen University; Ministry of Education, China), and Wei-Shi Zheng (Sun Yat-sen University; Ministry of Education, China)</i>	
Hierarchical and Contrastive Representation Learning for Knowledge-Aware Recommendation ...	1050
<i>Bingchao Wu (Institute of Software, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Yangyuxuan Kang (Intel Labs China, China), Daoguang Zan (Institute of Software, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Bei Guan (Institute of Software, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), and Yongji Wang (State Key Laboratory of Computer Science, Institute of Software, China; Institute of Software, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China)</i>	
Colo-SCRL: Self-Supervised Contrastive Representation Learning for Colonoscopic Video Retrieval	1056
<i>Qingzhong Chen (Shanghai Jiao Tong University, China), Shilun Cai (Zhongshan Hospital of Fudan University, China), Crystal Cai (Shanghai Jiao Tong University, China), Zefang Yu (Shanghai Jiao Tong University, China), Dahong Qian (Shanghai Jiao Tong University, China), and Suncheng Xiang (Shanghai Jiao Tong University, China)</i>	

Establishing a Stronger Baseline for Lightweight Contrastive Models	1062
<i>Wenye Lin (Tsinghua University, China), Yifeng Ding (Tsinghua University, China), Zhixiong Cao (Tsinghua University, China), and Hai-Tao Zheng (Tsinghua University, China; Peng Cheng Laboratory)</i>	
Graph Information Interaction on Feature and Structure via Cross-Modal Contrastive Learning	1068
<i>Jinyong Wen (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Yuhu Wang (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Chunxia Zhang (Beijing Institute of Technology, China), Shiming Xiang (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), and Chunhong Pan (Institute of Automation, Chinese Academy of Sciences, China)</i>	
Discriminative and Contrastive Consistency for Semi-Supervised Domain Adaptive Image Classification	1074
<i>Yidan Fan (Tianjin University, China), Wenhuan Lu (Tianjin University, China), and Yahong Han (Tianjin University; Peng Cheng Laboratory, China)</i>	

Sound Processing

CoverHunter: Cover Song Identification with Refined Attention and Alignments	1080
<i>Feng Liu (Huya Inc, China), Deyi Tuo (Huya Inc, China), Yanan Xu (Huya Inc, China), and Xintong Han (Huya Inc, China)</i>	
Exploring Pre-Trained Neural Audio Representations for Audio Topic Segmentation	1086
<i>Iacopo Ghinassi (Queen Mary University of London, UK; Centre for Digital Music), Matthew Purver (Queen Mary University of London, UK; Jožef Stefan Institut, Slovenia), Huy Phan (Amazon Alexa, USA), and Chris Newell (BBC R&D, UK)</i>	
A High-Quality Melody-Aware Peking Opera Synthesizer using Data Augmentation	1092
<i>Xun Zhou (Xiamen University, China), Wujin Sun (Xiamen University, China), and Xiaodong Shi (Xiamen University, China)</i>	
LC-Beating: An Online System for Beat and Downbeat Tracking using Latency-Controlled Mechanism	1098
<i>Xinlu Liu (Fudan University, China), Jiale Qian (Fudan University, China), Qiqi He (Fudan University, China), Yi Yu (National Institute of Informatics (NII), Japan), and Wei Li (Fudan University, China)</i>	
Improving Domain Generalization for Sound Classification with Sparse Frequency-Regularized Transformer	1104
<i>Honglin Mu (Harbin Institute of Technology, China), Wentian Xia (Harbin Institute of Technology, China), and Wanxiang Che (Harbin Institute of Technology, China)</i>	
MFAE: Masked Frame-Level Autoencoder with Hybrid-Supervision for Low-Resource Music Transcription	1109
<i>Yulun Wu (Fudan University, China), Jiahao Zhao (Fudan University, China), Yi Yu (n/a), and Wei Li (Fudan University, China)</i>	

3D & Depth I

Self-Supervised Implicit 3D Reconstruction via RGB-D Scans	1115
<i>Hongji Yang (Nankai University, China), Jiao Liu (Nankai University, China), Shaoping Lu (Nankai University, China), and Bo Ren (Nankai University, China)</i>	
Object-Aware Calibrated Depth-Guided Transformer for RGB-D Co-Salient Object Detection .	1121
<i>Yang Wu (Inspur Suzhou Intelligent Technology Co., Ltd; State Key Laboratory of High-End Server & Storage Technology (Inspur Group Company Limited; Nanjing University of Information Science and Technology, China), Lingyan Liang (Inspur Suzhou Intelligent Technology Co., Ltd; 2State Key Laboratory of High-End Server & Storage Technology (Inspur Group Company Limited), China), Yaqian Zhao (Inspur Suzhou Intelligent Technology Co., Ltd; 2State Key Laboratory of High-End Server & Storage Technology (Inspur Group Company Limited), China), and Kaihua Zhang (Inspur Suzhou Intelligent Technology Co., Ltd; State Key Laboratory of High-End Server & Storage Technology (Inspur Group Company Limited; Nanjing University of Information Science and Technology, China)</i>	
A Two-Stage Hybrid CNN-Transformer Network for RGB Guided Indoor Depth Completion ...	1127
<i>Yufan Deng (Beihang University, China), Xin Deng (Beihang University, China), and Mai Xu (Beihang University, China)</i>	
Feature Decoupling and Uncertainty Estimation for 3D Object Detection	1133
<i>Peiyuan Zhi (Tsinghua University, China), Kaiyue Zhou (Tsinghua University, China), Yali Li (Tsinghua University, China), and Shengjin Wang (Tsinghua University, China)</i>	
Scene Graph Generation using Depth-Based Multimodal Network	1139
<i>Lianggangxu Chen (East China Normal University, China), Jiale Lu (East China Normal University, China), Changbo Wang (East China Normal University, China), and Gaoqi He (East China Normal University, China)</i>	
Multi-view Token Clustering and Fusion for 3D Object Recognition and Retrieval	1145
<i>Linlong Fan (University of Electronic Science and Technology of China, China), Yanqi Ge (University of Electronic Science and Technology of China, China), Wen Li (University of Electronic Science and Technology of China, China), and Lixin Duan (University of Electronic Science and Technology of China, China)</i>	

Transformer I

Local Consensus Transformer for Correspondence Learning	1151
<i>Gang Wang (Shanghai University of Finance and Economics, China) and Yufei Chen (Tongji University, China)</i>	
Preserving Locality in Vision Transformers for Class Incremental Learning	1157
<i>Bowen Zheng (Nanjing University, China), Da-Wei Zhou (Nanjing University, China), Han-Jia Ye (Nanjing University, China), and De-Chuan Zhan (Nanjing University, China)</i>	

MTNet: Learning Modality-Aware Representation with Transformer for RGBT Tracking	1163
<i>Ruichao Hou (Nanjing University, China), Boyue Xu (Nanjing University, China), Tongwei Ren (Nanjing University, China), and Gangshan Wu (Nanjing University, China)</i>	
Adaptive Split-Fusion Transformer	1169
<i>Zixuan Su (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Jingjing Chen (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Lei Pang (City University of Hong Kong), Chong-Wah Ngo (Singapore Management University), and Yu-Gang Jiang (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing)</i>	
Gsformer: Geometric-Spatial Transformer on Point Cloud Completion	1175
<i>Yijun Long (Fudan University, China), Zhaoyu Chen (Fudan University, China), Hong Lu (Fudan University, China), and Wenqiang Zhang (Fudan University, China)</i>	
SDGFormer: An Efficient Convolution Network Structurally Similar to Transformer	1181
<i>Chaohao Wen (SWJTU, China) and Xun Gong (SWJTU; Ministry of Education, China)</i>	

Security, Privacy & Forensics I

General GAN-Generated Image Detection by Data Augmentation in Fingerprint Domain	1187
<i>Huaming Wang (Jinan University, China), Jianwei Fei (Nanjing University of Information Science and Technology, China), Yunshu Dai (Nanjing University of Information Science and Technology, China), Lingyun Leng (Jinan University, China), and Zhihua Xia (Jinan University, China)</i>	
Image Protection for Robust Cropping Localization and Recovery	1193
<i>Qichao Ying (Fudan University, China), Hang Zhou (Simon Fraser University, Canada), Xiaoxiao Hu (Fudan University, China), Zhenxing Qian (Fudan University, China), Sheng Li (Fudan University, China), and Xinpeng Zhang (Fudan University, China)</i>	
Towards Diverse Liveness Feature Representation and Domain Expansion for Cross-Domain Face Anti-Spoofing	1199
<i>Pei-Kai Huang (National Tsing Hua University, Taiwan), Jun-Xiong Chong (National Tsing Hua University, Taiwan), Hui-Yu Ni (National Tsing Hua University, Taiwan), Tzu-Hsien Chen (National Tsing Hua University, Taiwan), and Chiou-Ting Hsu (National Tsing Hua University, Taiwan)</i>	

Joint Statistical and Causal Feature Modulated Face Anti-Spoofing	1205
<i>Xin Dong (Ningxia University, China), Tao Wang (Ningxia University, China), Zhendong Li (Ningxia University; Collaborative Innovation Center for Ningxia Big Data and Artificial Intelligence Co-founded by Ningxia Municipality and Ministry of Education; Key Laboratory of the Internet of Water and Digital Water Governance of the Yellow River in Ningxia, China), and Hao Liu (Ningxia University; Collaborative Innovation Center for Ningxia Big Data and Artificial Intelligence Co-founded by Ningxia Municipality and Ministry of Education; Key Laboratory of the Internet of Water and Digital Water Governance of the Yellow River in Ningxia, China)</i>	
Watermarks for Generative Adversarial Network Based on Steganographic Invisible Backdoor	1211
<i>Yuwei Zeng (Fudan University, China), Jingxuan Tan (Fudan University, China), Zhengxin You (Fudan University, China), Zhenxing Qian (Fudan University, China), and Xinpeng Zhang (Fudan University, China)</i>	
Promoting Adversarial Transferability with Enhanced Loss Flatness	1217
<i>Yan Fang (Wuhan University, China), Zhongyuan Wang (Wuhan University, China), Jikang Cheng (Wuhan University, China), Ruoxi Wang (Jiangnan University, China), and Chao Liang (Wuhan University, China)</i>	
Face Poison: Obstructing DeepFakes by Disrupting Face Detection	1223
<i>Yuezun Li (Ocean University of China, China), Jiaran Zhou (Ocean University of China, China), and Siwei Lyu (University at Buffalo, State University of New York, USA)</i>	
ABTD-Net: Autonomous Baggage Threat Detection Networks for X-ray Images	1229
<i>Wen Liu (Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Degang Sun (University of Chinese Academy of Sciences, China), Yan Wang (Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Zhongyuan Chen (Beijing University Of Technology, China), Xinbo Han (Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), and Haitian Yang (Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China)</i>	

An Explainable Multi-view Semantic Fusion Model for Multimodal Fake News Detection	1235
<i>Zhi Zeng (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), Mingmin Wu (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), Guodong Li (Xinjiang Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, China), Xiang Li (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), Zhongqiang Huang (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), and Ying Sha (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China)</i>	
Improving CoatNet for Spatial and JPEG Domain Steganalysis	1241
<i>Hao Li (State Key Laboratory of Mathematical Engineering and Advanced Computing, China; Advanced Computing and Key Laboratory of Cyberspace Situation Awareness of Henan Provinc, China), Xiangyang Luo (State Key Laboratory of Mathematical Engineering and Advanced Computing, China), and Yi Zhang (State Key Laboratory of Mathematical Engineering and Advanced Computing, China)</i>	

Super Resolution & Inpainting II

Image Super-Resolution with Implicit Texture Pattern Modulation	1247
<i>Shuai Hao (Dalian University of Technology, China), Jialin Yang (Dalian University of Technology, China), Xu Jia (Dalian University of Technology, China), You He (Naval Aviation University, China), and Huchuan Lu (Dalian University of Technology, China)</i>	
Towards Efficient Large Mask Inpainting via Knowledge Transfer	1253
<i>Feihong Qin (Nanjing University of Aeronautics and Astronautics, China) and Liyan Zhang (Nanjing University of Aeronautics and Astronautics, China)</i>	
GRNN: Recurrent Neural Network Based on Ghost Features for Video Super-Resolution	1259
<i>Yutong Guo (East China University of Science and Technology, China)</i>	

Structure First Detail Next: Image Inpainting with Pyramid Generator	1265
<i>Shuyi Qu (Northwest University; Xi'an Jiaotong-Liverpool University, China), Zhenxing Niu (Xidian University, China), Jianke Zhu (Zhejiang University, China), Bin Dong (Ricoh Co., Ltd, China), and Kaizhu Huang (Duke Kunshan University, China)</i>	
Learning a Multilevel Cooperative View Reconstruction Network for Light Field Angular Super-Resolution	1271
<i>Deyang Liu (Jiangxi University of Finance and Economics), Yifan Mao (Anqing Normal University, China), Xiaofei Zhou (Hangzhou Dianzi University, China), Ping An (Shanghai University, China), and Yuming Fang (Jiangxi University of Finance and Economics, China)</i>	
NLCUnet: Single-Image Super-Resolution Network with Hairline Details	1277
<i>Jiancong Feng (Guangzhou University, China), Yuan-Gen Wang (Guangzhou University, China), and Fengchuang Xing (Guangzhou University, China)</i>	
Progressive Generative Adversarial Network for High-Resolution Image Inpainting	1283
<i>Muzi Cui (Jinan University, China), Chaozhuo Li (Microsoft Research Asia, China), Zhiying Li (Jinan University, China), Yingze Xie (Beijing Foreign Studies University, China), Xing Xie (Microsoft Research Asia, China), Weixin Wu (Jinan University, China), and Feiran Huang (Jinan University, China)</i>	

Quality Enhancement And Assessment For Low-Quality Multimedia Data Understanding

An Order-Complexity Model for Aesthetic Quality Assessment of Symbolic Homophony Music Scores	1289
<i>Xin Jin (Beijing Electronic Science and Technology Institute, China; Beijing Institute for General Artificial Intelligence, China), Wu Zhou (Beijing Electronic Science and Technology Institute, China), Jinyu Wang (Beijing Electronic Science and Technology Institute, China), Duo Xu (Tianjin Conservatory of Music, China; Tongji University, China), Yiqing Rong (Beijing Electronic Science and Technology Institute, China), and Shuai Cui (University of California, USA)</i>	
Collaborative Auto-Encoding for Blind Image Quality Assessment	1295
<i>Zehong Zhou (Shenzhen University, China; Guangdong Key Laboratory of Intelligent Information Processing, China; Shenzhen Institute for Artificial Intelligence and Robotics for Society, China, China), Fei Zhou (Shenzhen University, China; Peng Cheng Laboratory, China; Guangdong Key Laboratory of Intelligent Information Processing, China; Guangdong-Hong Kong Joint Laboratory for Big Data Image and Communication, China), and Guoping Qiu (Shenzhen University, China; Peng Cheng Laboratory, China; Guangdong Key Laboratory of Intelligent Information Processing, China; Shenzhen Institute for Artificial Intelligence and Robotics for Society, China, China; University of Nottingham, U.K; Guangdong-Hong Kong Joint Laboratory for Big Data Image and Communication, China)</i>	

No Reference Image Quality Assessment via Quality Difference Learning	1301
<i>Jiaming Xie (Guangdong University of Technology, China), Yu Luo (Guangdong University of Technology, China), Jie Ling (Guangdong University of Technology, China), and Guanghui Yue (Shenzhen University, China)</i>	
Low-Light Image Enhancement by Learning Contrastive Representations in Spatial and Frequency Domains	1307
<i>Yi Huang (Southwest University of Science and Technology, China), Xiaoguang Tu (Civil Aviation Flight University of China; Sichuan University, China), Gui Fu (Southwest University of Science and Technology; Civil Aviation Flight University of China, China), Tingting Liu (Southwest University of Science and Technology, China), Bokai Liu (Civil Aviation Flight University of China, China), Ming Yang (Civil Aviation Flight University of China, China), and Ziliang Feng (Sichuan University, China)</i>	
Noise Adaptive Speech Intelligibility Enhancement Based on Improved StarGAN	1313
<i>Lanxin Zhao (Jiangnan University, China), Dengshi Li (Jiangnan University, China), Jing Xiao (Wuhan University, China), and Chenyi Zhu (Jiangnan University, China)</i>	
Image Template Matching via Dense and Consistent Contrastive Learning	1319
<i>Bo Li (Northwestern Polytechnical University, China), Lin Yuanbo Wu (Swansea University, United Kingdom), Deyin Liu (Anhui University, China), Hongyang Chen (Zhejiang Lab, China), Yuanxin Ye (Southwest Jiaotong University, China), and Xianghua Xie (Swansea University, United Kingdom)</i>	

UAV & Underwater Media Processing

Deep Reinforcement Learning with Semi-Expert Distillation for Autonomous UAV Cinematography	1325
<i>Andreas Sochopoulos (Aristotle University of Thessaloniki, Greece), Ioannis Mademlis (Aristotle University of Thessaloniki, Greece), Evangelos Charalampakis (Aristotle University of Thessaloniki, Greece), Sotirios Papadopoulos (Aristotle University of Thessaloniki, Greece), and Ioannis Pitas (Aristotle University of Thessaloniki, Greece)</i>	
Learning Disentangled Representation with Mutual Information Maximization for Real-Time UAV Tracking	1331
<i>Xucheng Wang (Guilin University of Technology, China), Xiangyang Yang (Guilin University of Technology, China), Hengzhou Ye (Guilin University of Technology, China), and Shuiwang Li (Guilin University of Technology, China)</i>	
Transmission and Color-Guided Network for Underwater Image Enhancement	1337
<i>Pan Mu (Zhejiang University of Technology, China), Jing Fang (Zhejiang University of Technology, China), Haotian Qian (Zhejiang University of Technology, China), and Cong Bai (Zhejiang University of Technology, China)</i>	

Invertible Underwater Image Enhancement Network	1343
<i>Fei Li (China Agricultural University, China), Zhenbo Li (China Agricultural University, China), Xinxin Zhang (China Agricultural University, China), Meng Ding (The State University of New York, USA), and Zikun Liu (Samsung Research China (SRC-B), China)</i>	
Towards Discriminative Representations with Contrastive Instances for Real-Time UAV Tracking	1349
<i>Dan Zeng (Southern University of Science and Technology, China), Mingliang Zou (Guilin University of Technology, China), Xucheng Wang (Guilin University of Technology, China), and Shuiwang Li (Guilin University of Technology, China)</i>	
Underwater Image Enhancement with an Adaptive Self Supervised Network	1355
<i>Rizwan Khan (Zhejiang Normal University (ZJNU); Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, ZJNU), Atif Mehmood (Zhejiang Normal University (ZJNU)), Saeed Akbar (Zhejiang Normal University (ZJNU)), and Zhonglong Zheng (Zhejiang Normal University (ZJNU); Key Laboratory of Intelligent Education Technology and Application of Zhejiang Province, ZJNU)</i>	

Sentiment, Expression & Emotion

Privacy-Protected Facial Expression Recognition Augmented by High-Resolution Facial Images.....	1361
<i>Cong Liang (University of Science and Technology of China, China), Shangfei Wang (University of Science and Technology of China, China), and Xiaoping Chen (University of Science and Technology of China, China)</i>	
Multimodal Sentiment Analysis with Preferential Fusion and Distance-Aware Contrastive Learning	1367
<i>Feipeng Ma (University of Science and Technology of China, China), Yueyi Zhang (University of Science and Technology of China, China), and Xiaoyan Sun (University of Science and Technology of China, China)</i>	
A Multi-view Co-learning Method for Multimodal Sentiment Analysis	1373
<i>Wenxiu Geng (Shandong University, China), Yulong Bian (Shandong University, China; Key Laboratory of Shandong Province for Software Engineering, China), and Xiangxian Li (Shandong University, China)</i>	
Multimodal Aspect-Based Sentiment Classification with Knowledge-Injected Transformer	1379
<i>Zenan Xu (Sun Yat-sen University, China), Qinliang Su (Sun Yat-sen University; Guangdong Key Laboratory of Big Data Analysis and Processing, China), and Junxi Xiao (Sun Yat-sen University, China)</i>	
STA-GCN:Spatial Temporal Adaptive Graph Convolutional Network for Gait Emotion Recognition..	1385
<i>Chuang Chen (Anhui University, China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, China) and Xiao Sun (Hefei University of Technology, China; Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, China)</i>	

Adaptive Graph Attention Network with Temporal Fusion for Micro-Expressions Recognition 1391
Yiming Zhang (University of Science and Technology of China, China), Hao Wang (University of Science and Technology of China, China), Yifan Xu (University of Science and Technology of China, China), Xinglong Mao (University of Science and Technology of China, China), Tong Xu (University of Science and Technology of China, China), Sirui Zhao (University of Science and Technology of China, China), and Enhong Chen (University of Science and Technology of China, China)

Transformer II

Deep Homography Estimation with Feature Correlation Transformer 1397
Haoyu Zhou (Wuhan University, China), Wei Hu (Wuhan University, China), Ying Li (Wuhan University, China), Chu He (Wuhan University, China), and Xi Chen (Wuhan University, China)

ADATS: Adaptive RoI-Align Based Transformer for End-to-End Text Spotting 1403
Zepeng Huang (Shenzhen University, China), Qi Wan (Shenzhen University, China), Junliang Chen (Shenzhen University, China), Xiaodong Zhao (Shenzhen University, China), Kai Ye (Shenzhen University, China), and Linlin Shen (Shenzhen University, China)

Trajectory Alignment Based Multi-scaled Temporal Attention for Efficient Video Transformer 1409
Zao Zhang (The University of Sydney, Australia), Dong Yuan (The University of Sydney, Australia), Yu Zhang (The University of Sydney, Australia), and Wei Bao (The University of Sydney, Australia)

Swin-ASNet: An Adaptive RGB-Selection Network with Swin Transformer for Retinal Vessel Segmentation 1415
Qunchao Jin (East China Normal University, China), Hongyu Hou (East China Normal University, China), Guixu Zhang (East China Normal University, China), Haoan Wang (East China Normal University, China), and Zhi Li (East China Normal University, China)

OAFFormer: Occlusion Aware Transformer for Camouflaged Object Detection 1421
Xin Yang (Fujian University of Technology, China), Hengliang Zhu (Fujian University of Technology, China), Guojun Mao (Fujian University of Technology, China), and Shuli Xing (Fujian University of Technology, China)

Know Who You Are: Learning Target-Aware Transformer for Object Tracking 1427
Zhuojun Zou (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Xuexin Liu (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Yuanpei Zhang (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Lin Shu (Chinese Academy of Sciences; Guangdong Institute of Artificial Intelligence and Advanced Computing, China), and Jie Hao (Chinese Academy of Sciences; Guangdong Institute of Artificial Intelligence and Advanced Computing, China)

Model Simplification

A Novel Channel Pruning Approach Based on Local Attention and Global Ranking for CNN Model Compression	1433
<i>Wei Lu (Tianjin University, China), Yang Jiang (Tianjin University, China), Peiguang Jing (Tianjin University, China), Jinghui Chu (Tianjin University, China), and Fugui Fan (Tianjin University, China)</i>	
Splittable Pattern-Specific Weight Pruning for Deep Neural Networks	1439
<i>Yiding Liu (Beijing University of Posts and Telecommunications, China), Yinglei Teng (Beijing University of Posts and Telecommunications, China), and Tao Niu (Beijing University of Posts and Telecommunications, China)</i>	
Dynamic Dense-Sparse Representations for Real-Time Question Answering	1445
<i>Minyu Sun (Hunan University, China), Bin Jiang (Hunan University, China), and Chao Yang (Hunan University, China)</i>	
DynaSlim: Dynamic Slimming for Vision Transformers	1451
<i>Da Shi (Shanghai Jiao Tong University, China), Jingsheng Gao (Shanghai Jiao Tong University, China), Ting Liu (Shanghai Jiao Tong University, China), and Yuzhuo Fu (Shanghai Jiao Tong University, China)</i>	
Post-Training Quantization for Vision Transformer in Transformed Domain	1457
<i>Kai Feng (Peng Cheng Laboratory; China-Singapore International Joint Research Institute (CSIJRI), China), Zhuo Chen (Peng Cheng Laboratory, China), Fei Gao (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Zhe Wang (Nanyang Technological University, Singapore), Long Xu (Chinese Academy of Sciences, China), and Weisi Lin (China-Singapore International Joint Research Institute (CSIJRI); Nanyang Technological University, Singapore)</i>	
Residual Based Hierarchical Feature Compression for Multi-task Machine Vision	1463
<i>Chaoran Chen (Beihang University, China), Mai Xu (Beihang University, China), Shengxi Li (Beihang University, China), Tie Liu (Beihang University, China), Minglang Qiao (Beihang University, China), and Zhuoyi Lv (VIVO Mobile Communication Company Limited, China)</i>	

Cross-Modality & Cross-Domain

Cross-Domain Federated Object Detection	1469
<i>Shangchao Su (Fudan University), Bin Li (Fudan University), Chengzhi Zhang (Fudan University), Mingzhao Yang (Fudan University), and Xiangyang Xue (Fudan University)</i>	

Cross-Modality Fourier Feature for Medical Image Synthesis	1475
<i>Mei Ma (NingXia University, China), Ling Lin (NingXia University, China), Heng Wang (Ningxia University, China; Collaborative Innovation Center for Ningxia Big Data and Artificial Intelligence Co-founded by Ningxia Municipality and Ministry of Education, China), Zhendong Li (Ningxia University, China; Collaborative Innovation Center for Ningxia Big Data and Artificial Intelligence Co-founded by Ningxia Municipality and Ministry of Education, China; Key Laboratory of the Internet of Water and Digital Water Governance of the Yellow River in Ningxia, Yinchuan, China), and Hao Liu (Ningxia University, China; Collaborative Innovation Center for Ningxia Big Data and Artificial Intelligence Co-founded by Ningxia Municipality and Ministry of Education, China; Key Laboratory of the Internet of Water and Digital Water Governance of the Yellow River in Ningxia, Yinchuan, China)</i>	
Point-Syn2Real: Semi-Supervised Synthetic-to-Real Cross-Domain Learning for Object Classification in 3D Point Clouds	1481
<i>Ziwei Wang (CSIRO Data61), Reza Arablouei (CSIRO Data61), Jiajun Liu (CSIRO Data61), Paulo Borges (CSIRO Data61), Greg Bishop-Hurley (CSIRO Agriculture & Food), and Nicholas Heaney (Evolve Group)</i>	
Temporal-Enhanced Cross-Modality Fusion Network for Video Sentence Grounding	1487
<i>Ze Zhong Lv (Renmin University of China; Beijing Key Laboratory of Big Data Management and Analysis Methods, China) and Bing Su (Renmin University of China; Beijing Key Laboratory of Big Data Management and Analysis Methods, China)</i>	
A Cross-direction Task Decoupling Network for Small Logo Detection	1493
<i>Sujuan Hou (Shandong Normal University, China), Xingzhuo Li (Shandong Normal University, China), Weiqing Min (Institute of Computing Technology, Chinese Academy of Sciences, China), Jiacheng Li (Shandong Normal University, China), Jing Wang (Shandong Normal University, China), Yuanjie Zheng (Shandong Normal University, China), and Shuqiang Jiang (Institute of Computing Technology, Chinese Academy of Sciences, China)</i>	
CHAN: Cross-Modal Hybrid Attention Network for Temporal Language Grounding in Videos .	1499
<i>Wen Wang (Zhejiang Lab, China), Ling Zhong (Zhejiang Lab, China), Guang Gao (Zhejiang Lab, China), Minhong Wan (Zhejiang Lab, China), and Jason Gu (Dalhousie University, Canada)</i>	

Multimodal

DMRL-Net: Differentiable Multi-view Representation Learning Network	1505
<i>Zihan Fang (Fuzhou University, China), Shide Du (Fuzhou University, China), Yaqing Chen (Fuzhou University, China), and Shiping Wang (Fuzhou University, China)</i>	
Conditional Video-Text Reconstruction Network with Cauchy Mask for Weakly Supervised Temporal Sentence Grounding	1511
<i>Jueqi Wei (Fudan University, China), Yuanwu Xu (Fudan University, China), Mohan Chen (Fudan University, China), Yuejie Zhang (Fudan University, China), Rui Feng (Fudan University, China), and Shang Gao (Deakin University)</i>	

FlowText: Synthesizing Realistic Scene Text Video with Optical Flow Estimation	1517
<i>Yuzhong Zhao (University of Chinese Academy of Sciences, China), Weijia Wu (Zhejiang University, China), Zhuang Li (MMU Kuaishou Technology, China), Jiahong Li (MMU Kuaishou Technology, China), and Weiqiang Wang (University of Chinese Academy of Sciences, China)</i>	
Atomic-Action-Based Contrastive Network for Weakly Supervised Temporal Language Grounding	1523
<i>Hongzhou Wu (National University of Defense Technology, China), Yifan Lyu (University of Chinese Academy of Sciences, China), Xingyu Shen (National University of Defense Technology, China), Xuechen Zhao (National University of Defense Technology, China), Mengzhu Wang (National University of Defense Technology, China), Xiang Zhang (National University of Defense Technology, China), and Zhigang Luo (National University of Defense Technology, China)</i>	
Mandari: Multi-modal Temporal Knowledge Graph-Aware Sub-Graph Embedding for Next-POI Recommendation	1529
<i>Xiaoqian Liu (Beijing University of Posts and Telecommunications, China), Xiuyun Li (The Technology Innovation Center of Cultural Tourism Big Data of Hebei Province, China; Hebei Normal University for Nationalities, China), Yuan Cao (Beijing University of Posts and Telecommunications, China), Fan Zhang (Beijing University of Posts and Telecommunications, China), Xiongnan Jin (Knowledge Discovery and Data Mining Research Center, Zhejiang Lab, China), and Jinpeng Chen (Beijing University of Posts and Telecommunications, China)</i>	
Movie Box Office Prediction With Self-Supervised and Visually Grounded Pretraining	1535
<i>Qin Chao (Alibaba Group and the Alibaba-NTU Joint Research Institute; School of Computer Science and Engineering, NTU, Singapore), Eunsoo Kim (Nanyang Technological University (NTU), Singapore), and Boyang Li (School of Computer Science and Engineering, NTU, Singapore)</i>	

Data & Labelling for Machine Learning II

Need a Dog for Seeing eye? A Walk Viewpoint Dataset for Freespace Detection in Unstructured Environments	1541
<i>Wenbin Zou (Shenzhen University, China), Guoguang Hua (Shenzhen University, China), Guangxu Chen (Shenzhen University, China), Zaiyue He (Shenzhen University, China), Guangli Liu (Shenzhen University, China), Pengfei Chen (Shenzhen University, China), Yuyang Li (Shenzhen University, China), Huakun Li (Shenzhen University, China), Lei Zheng (Shenzhen University, China), and Shishun Tian (Shenzhen University, China)</i>	
Intensifying the Consistency of Pseudo Label Refinement for Unsupervised Domain Adaptation Person Re-Identification	1547
<i>Linfan Zha (Anhui University, China), Yanming Chen (Anhui University, China), Peng Zhou (Anhui University, China), and Yiwen Zhang (Anhui University, China)</i>	

Noisy-to-Clean Label Learning for Medical Image Segmentation	1553
<i>Zihao Bu (Jiangsu University, China), Xiaoxiao Wang (Zhenjiang First People's Hospital, China), Chengjian Qiu (Jiangsu University, China), Zhixuan Wang (Zhenjiang First People's Hospital, China), Kai Han (Jiangsu University, China), Xiuhong Shan (Zhenjiang First People's Hospital, China), and Zhe Liu (Jiangsu University, China)</i>	
Learning Discrimination from Contaminated Data: Multi-instance Learning for Unsupervised Anomaly Detection	1559
<i>Wenhao Hu (Zhejiang University, China), Yingying Liu (Zhejiang University, China), Jiazhen Xu (Zhejiang University, China), Xuanyu Chen (Zhejiang University, China), and Gaoang Wang (Zhejiang University, China)</i>	
Rethinking Video Error Concealment: A Benchmark Dataset	1565
<i>Bin Zheng (Shenzhen University, China) and Miaohui Wang (Shenzhen University, China)</i>	
Visual Place Recognition Datasets for Indoor Spaces	1571
<i>Zemian Guo (Shenzhen University, China) and Yingying Zhu (Shenzhen University, China)</i>	
AutoKary2022: A Large-Scale Densely Annotated Dataset for Chromosome Instance Segmentation.....	1577
<i>Dan You (Hangzhou City University, China), Pengcheng Xia (Hangzhou City University, China), Qiuzhu Chen (Hangzhou City University, China), Minghui Wu (Hangzhou City University, China), Suncheng Xiang (Shanghai Jiao Tong University, China), and Jun Wang (Hangzhou City University, China)</i>	

Visual Information Processing II

Designing Optics and Algorithm for Ultra-Thin, High-Speed Lensless Cameras	1583
<i>Salman S. Khan (IIT Madras, India), Vivek Boominathan (Rice University, USA), Ashok Veeraraghavan (Rice University, USA), and Kaushik Mitra (IIT Madras, India)</i>	
Dual-Domain Feature Learning and Memory-Enhanced Unfolding Network for Spectral Compressive Imaging	1589
<i>Yangke Ying (Beijing University of Technology, China), Jin Wang (Beijing University of Technology, China), Yunhui Shi (Beijing University of Technology, China), and Baocai Yin (Beijing University of Technology, China)</i>	
Image Compressed Sensing using Multi-scale Characteristic Residual Learning	1595
<i>Shumian Yang (Qilu University of Technology (Shandong Academy of Sciences), China), Xinxin Xiang (Qilu University of Technology (Shandong Academy of Sciences), China), Fenghua Tong (Qilu University of Technology (Shandong Academy of Sciences), China), Dawei Zhao (Qilu University of Technology (Shandong Academy of Sciences), China), and Xin Li (Qilu University of Technology (Shandong Academy of Sciences), China)</i>	

LKD-Net: Large Kernel Convolution Network for Single Image Dehazing	1601
<i>Pinjun Luo (Southwest University, China), Guoqiang Xiao (Southwest University, China), Xinbo Gao (Chongqing University of Posts and Telecommunications, China), and Song Wu (Southwest University, China)</i>	
Video Noise Removal using Progressive Decomposition with Conditional Invertibility	1607
<i>Haoran Huang (South China University of Technology, China), Yuhui Quan (South China University of Technology; Pazhou Lab, China), Zhenghua Lei (South China University of Technology, China), Jinlong Hu (South China University of Technology, China), and Yan Huang (South China University of Technology; Pazhou Lab, China)</i>	
DocMAE: Document Image Rectification via Self-Supervised Representation Learning	1613
<i>Shaokai Liu (University of Science and Technology of China, China), Hao Feng (University of Science and Technology of China, China), Wengang Zhou (University of Science and Technology of China, China), Houqiang Li (University of Science and Technology of China, China), Cong Liu (iFLYTEK Research, China), and Feng Wu (University of Science and Technology of China, China)</i>	
Information-Density Masking Strategy for Masked Image Modeling	1619
<i>He Zhu (Chinese Academy of Sciences(CASIA); University of Chinese Academy of Sciences(UCAS)), Yang Chen (Chinese Academy of Sciences(CASIA)), Guyue Hu (Nanyang Technological University), and Shan Yu (Chinese Academy of Sciences(CASIA); University of Chinese Academy of Sciences(UCAS))</i>	
Histogram-Guided Video Colorization Structure with Spatial-Temporal Connection	1625
<i>Zheyuan Liu (Zhejiang University of Technology, China), Pan Mu (Zhejiang University of Technology, China), Hanning Xu (Zhejiang University of Technology, China), and Cong Bai (Zhejiang University of Technology, China)</i>	
Mask-Guided Stamp Erasure for Real Document Image	1631
<i>Xinye Yang (University of Chinese Academy of Sciences, China), Dongbao Yang (University of Chinese Academy of Sciences, China), Yu Zhou (University of Chinese Academy of Sciences, China), Youhui Guo (University of Chinese Academy of Sciences, China), and Weiping Wang (Institute of Information Engineering, Chinese Academy of Sciences, China)</i>	
Attention-Aware Anime Line Drawing Colorization	1637
<i>Yu Cao (The Hong Kong Polytechnic University, Hong Kong), Hao Tian (The Hong Kong Polytechnic University, Hong Kong), and P. Y. Mok (The Hong Kong Polytechnic University, Hong Kong)</i>	
Edge-Aware Neural Implicit Surface Reconstruction	1643
<i>Xinghui Li (Tsinghua University, China), Yikang Ding (Tsinghua University, China), Jia Guo (Huawei Technologies, China), Xiansong Lai (Tsinghua University, China), Shihao Ren (Tsinghua University, China), Wensen Feng (Huawei Technologies, China), and Long Zeng (Tsinghua University, China)</i>	
Handwriting Curve Interpolation using Gradient Graph Laplacian Regularizer	1649
<i>Yinhe Lin (Fuzhou University, China), Fei Chen (Fuzhou University, China), Hang Cheng (Fuzhou University, China), and Meiqing Wang (Fuzhou University, China)</i>	

Optimized Media Delivery

- Comparison of HDR Quality Metrics in Per-Clip Lagrangian Multiplier Optimisation with AV1 . 1655
Vibhoothi Vibhoothi (Trinity College Dublin, Ireland), François Pitié (Trinity College Dublin, Ireland), Angeliki Katsenou (Trinity College Dublin, Ireland), Yeping Su (Google Inc, USA), Balu Adsumilli (Google Inc, USA), and Anil Kokaram (Trinity College Dublin, Ireland)
- A Real-Time Blind Quality-of-Experience Assessment Metric for HTTP Adaptive Streaming 1661
Chunyi Li (Shanghai Jiao Tong University, China), May Lim (National University of Singapore, Singapore), Abdelhak Bentaleb (Concordia University, Canada), and Roger Zimmermann (National University of Singapore, Singapore)
- Towards Guidelines for Subjective Haptic Quality Assessment: A Case Study on Quality Assessment of Compressed Haptic Signals 1667
Andréas Pastor (Nantes Université, France) and Patrick Le Callet (Nantes Université, France)
- Just Noticeable Difference-Aware Per-Scene Bitrate-Laddering for Adaptive Video Streaming 1673
Vignesh V Menon (Alpen-Adria-Universität Klagenfurt, Austria), Jingwen Zhu (Nantes Université, France), Prajit T Rajendran (Université Paris-Saclay, France), Hadi Amirpour (Alpen-Adria-Universität Klagenfurt, Austria), Patrick Le Callet (Nantes Université, France), and Christian Timmerer (Alpen-Adria-Universität Klagenfurt, Austria)
- Optimizing Video Streaming for Sustainability and Quality: The Role of Preset Selection in Per-Title Encoding 1679
Hadi Amirpour (Alpen-Adria-Universität, Austria), Vignesh V Menon (Alpen-Adria-Universität, Austria), Samira Afzal (Alpen-Adria-Universität, Austria), Radu Prodan (Alpen-Adria-Universität, Austria), and Christian Timmerer (Alpen-Adria-Universität, Austria)
- Anableps: Adapting Bitrate for Real-Time Communication using VBR-Encoded Video 1685
Zicheng Zhang (Nanjing University, China), Hao Chen (Nanjing University, China), Xun Cao (Nanjing University, China), and Zhan Ma (Nanjing University, China)

Speech Processing

- Adversarial Speaker Disentanglement using Unannotated External Data for Self-Supervised Representation-Based Voice Conversion 1691
Xintao Zhao (Tsinghua University, China), Shuai Wang (Tencent Inc, China), Yang Chao (Tencent Inc, China), Zhiyong Wu (Tsinghua University, China), and Helen Meng (The Chinese University of Hong Kong, China)
- A Disentangled Recurrent Variational Autoencoder for Speech Enhancement 1697
Hegen Yan (Ningbo University, China) and Zhihua Lu (Ningbo University, China)

SnakeGAN: A Universal Vocoder Leveraging DDSP Prior Knowledge and Periodic Inductive Bias	1703
<i>Sipan Li (Tsinghua University, China), Songxiang Liu (Tencent Inc, China), Luwen Zhang (Tsinghua University, China), Xiang Li (Tsinghua University, China), Yanyao Bian (Tencent Inc, China), Chao Weng (Tencent Inc, China), Zhiyong Wu (Tsinghua University, China), and Helen Meng (The Chinese University of Hong Kong, China)</i>	
Cra-Diffuse: Improved Cross-Domain Speech Enhancement Based on Diffusion Model with T-F Domain Pre-Denoising	1709
<i>Zhibin Qiu (Xinjiang University, China), Yachao Guo (Xinjiang University, China), Mengfan Fu (Xinjiang University, China), Hao Huang (Xinjiang University; Xinjiang Key Laboratory of Multi-lingual Information Technology, China), Ying Hu (Xinjiang University, China), Liang He (Xinjiang University; Tsinghua University, China), and Fuchun Sun (Tsinghua University, China)</i>	
A Joint Network Based on Interactive Attention for Speech Emotion Recognition	1715
<i>Ying Hu (Xinjiang University, China), Shijing Hou (Xinjiang University, China), Huamin Yang (Xinjiang University, China), Huang Hao (Xinjiang University, China), and Liang He (Tsinghua University, China)</i>	
Speech Topic Classification Based on Pre-Trained and Graph Networks	1721
<i>Fangjing Niu (Xinjiang University, China), Tengfei Cao (Xinjiang University, China), Ying Hu (Xinjiang University, China), Hao Huang (Xinjiang University, China), and Liang He (Xinjiang University; Tsinghua University, China)</i>	

Face Computing

Unsupervised 3D Face Reconstruction with Reprogramming Skip Connections	1727
<i>Zhuoming Dong (Sun Yat-sen University, China), Huajun Zhou (Sun Yat-sen University, China), and Jianhuang Lai (Sun Yat-sen University, China; Guangdong Province Key Laboratory of Information Security Technology, China; Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, China)</i>	
EvenFace: Deep Face Recognition with Uniform Distribution of Identities	1733
<i>Pengfei Hu (Tsinghua University, China), Yingfan Tao (Tsinghua University, China), Qiqi Bao (Tsinghua University, China), Guijin Wang (Tsinghua University, China), and Wenming Yang (Tsinghua University, China)</i>	
Large Pose Friendly Face Reenactment using Subtle Motions	1739
<i>Xiaomeng Fu (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Xi Wang (Chinese Academy of Sciences, China), Jin Liu (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Jiao Dai (Chinese Academy of Sciences, China), and Jizhong Han (Chinese Academy of Sciences, China)</i>	

MSABox: A Spatially Stable Face Detector	1745
<i>Wei Xu (Beijing University of Posts and Telecommunications, China), Kangkang Wang (Baidu Inc., China), Ziliang Chen (Baidu Inc., China), Bin He (Baidu Inc., China), Bi Li (Baidu Inc., China), Haocheng Feng (Baidu Inc., China), Gang Zhang (Baidu Inc., China), Jingtuo Liu (Baidu Inc., China), Junyu Han (Baidu Inc., China), and Errui Ding (Baidu Inc., China)</i>	
DR-Net: A Multi-view Face Synthesis Network Driven by Dual Representation	1751
<i>Xianliang Huang (Fudan University), Yining Lang (Alibaba Group), Ying Guo (Alibaba Group), Yuan He (Alibaba Group), Hui Xue (Alibaba Group), Li Zhao (Zhejiang University), and Shuigeng Zhou (Fudan University)</i>	
MA-NeRF: Motion-Assisted Neural Radiance Fields for Face Synthesis from Sparse Images ...	1757
<i>Weichen Zhang (Tsinghua University, China), Xiang Zhou (Tsinghua University, China), Yukang Cao (The University of Hong Kong, China), WenSen Feng (Huawei Technologies, China), and Chun Yuan (Tsinghua University, China)</i>	

Robustness

Enhancing Robustness of Deep Networks Against Noisy Labels Based on a Two-Phase Formulation of Their Learning Behavior	1763
<i>Yaoru Luo (Institute of Automation, Chinese Academy of Sciences) and Ge Yang (University of Chinese Academy of Sciences)</i>	
Robust and Efficient Memory Network for Video Object Segmentation	1769
<i>Yadang Chen (Nanjing University of Information Science and Technology, China), Dingwei Zhang (Nanjing University of Information Science and Technology, China), Zhi-Xin Yang (University of Macau, China), and Enhua Wu (University of Chinese Academy of Sciences, China)</i>	
Weight-Based Regularization for Improving Robustness in Image Classification	1775
<i>Hao Yang (National University of Defense Technology, China), Min Wang (National University of Defense Technology, China), Zhengfei Yu (National University of Defense Technology, China), and Yun Zhou (National University of Defense Technology, China)</i>	
Robust Structured Sparse Subspace Clustering with Neighborhood Preserving Projection	1781
<i>Wenyi Feng (East China University of Science and Technology; Qinghai University, P. R. China), Wei Guo (East China University of Science and Technology, P. R. China), Ting Xiao (East China University of Science and Technology, P. R. China), and Zhe Wang (East China University of Science and Technology, P. R. China)</i>	
Improving Robustness of Learning-Based Adaptive Video Streaming in Wildly Fluctuating Networks	1787
<i>Jiawei Lin (Shenzhen University, China) and Shuoyao Wang (Shenzhen University, China)</i>	
Robust Person Re-identification with Wireless Signals	1793
<i>Dong Xi (University of Science and Technology of China, China), Wengang Zhou (University of Science and Technology of China, China), and Houqiang Li (University of Science and Technology of China, China)</i>	

Data & Labelling for Machine Learning I

GradSalMix: Gradient Saliency-Based Mix for Image Data Augmentation	1799
<i>Tao Hong (Peking University, China), Ya Wang (Tencent Inc., China), Xingwu Sun (Tencent Inc.; University of Macau, China), Fengzong Lian (Tencent Inc., China), Zhanhui Kang (Tencent Inc., China), and Jinwen Ma (Peking University, China)</i>	
Get a Head Start: Targeted Labeling at Source with Limited Annotation Overhead for Semi-Supervised Learning	1805
<i>Hui Zhu (Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences), Yongchun Lu (Mashang Consumer Finance Co., Ltd.), Qin Ma (China Agricultural University), Xunyi Zhou (Mashang Consumer Finance Co., Ltd.), Fen Xia (Mashang Consumer Finance Co., Ltd.), Guoqing Zhao (Mashang Consumer Finance Co., Ltd.), Ning Jiang (Mashang Consumer Finance Co., Ltd.), and Xiaofang Zhao (Institute of Computing Technology, Chinese Academy of Sciences; Institute of Intelligent Computing Technology; University of Chinese Academy of Sciences)</i>	
Partial Multi-label Learning: Exploration of Binary Ground-Truth Labels	1811
<i>Yan Hu (Guangdong University of Technology, China), Xiaozhao Fang (Guangdong University of Technology, China), Weijun Lv (Guangdong University of Technology, China), and Peipei Kang (Guangdong University of Technology, China)</i>	
Customizing Synthetic Data for Data-Free Student Learning	1817
<i>Shiya Luo (Zhejiang University, China), Defang Chen (Zhejiang University, China), and Can Wang (Zhejiang University, China)</i>	
A Geometrical Characterization on Feature Density of Image Datasets	1823
<i>Zhen Liang (National University of Defense Technology, China), Changyuan Zhao (Chinese Academy of Sciences, China), Wanwei Liu (National University of Defense Technology, China), Bai Xue (Chinese Academy of Sciences, China), and Wenjing Yang (National University of Defense Technology, China)</i>	
Federated Domain Adaptation via Pseudo-Label Refinement	1829
<i>Gang Li (Zhejiang University, China), Qifei Zhang (Zhejiang University, China), Peizheng Wang (Zhejiang University, China), Jie Zhang (Zhejiang University, China), and Chao Wu (Zhejiang University, China)</i>	

Learning Techniques I

Learning Continuous Piecewise non-Linear Activation Functions for Deep Neural Networks ..	1835
<i>Xinchen Gao (UESTC, China), Yawei Li (ETH Zurich, Switzerland), Wen Li (UESTC, China), Lixin Duan (UESTC, China), Luc Van Gool (ETH Zurich, Switzerland), Luca Benini (ETH Zurich, Switzerland), and Michele Magno (ETH Zurich, Switzerland)</i>	

Discriminative Spatiotemporal Alignment for Self-Supervised Video Correspondence Learning	1841
<i>Qiaoqiao Wei (Tsinghua University, China), Hui Zhang (Tsinghua University, China), and Jun-Hai Yong (Tsinghua University, China)</i>	
Unsupervised Fashion Style Learning by Solving Fashion Jigsaw Puzzles	1847
<i>Jia Chen (Wuhan Textile University; Engineering Research Center of Hubei Province for Clothing Information, China), Haidongqing Yuan (Wuhan Textile University), Fei Fang (Wuhan Textile University), Tao Peng (Wuhan Textile University), and Xinrong Hu (Wuhan Textile University)</i>	
Anchor-Free Action Proposal Network with Uncertainty Estimation	1853
<i>Selen Pehlivan (Aalto University, Finland) and Jorma Laaksonen (Aalto University, Finland)</i>	
Scale-Aware Task Message Transferring for Multi-task Learning	1859
<i>Shalayiding Sirejiding (Shanghai Jiao Tong University, China), Yuxiang Lu (Shanghai Jiao Tong University, China), Hongtao Lu (Shanghai Jiao Tong University, China), and Yue Ding (Shanghai Jiao Tong University, China)</i>	
Improving the Homophily of Heterophilic Graphs for Semi-Supervised Node Classification ...	1865
<i>Yuhu Wang (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Shiming Xiang (Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), and Chunhong Pan (Institute of Automation, Chinese Academy of Sciences, China)</i>	

Transformer III

Sitpose: A Siamese Convolutional Transformer for Relative Camera Pose Estimation	1871
<i>Kai Leng (Harbin Institute of Technology), Cong Yang (Soochow University), Wei Sui (Horizon Robotics), Jie Liu (Harbin Institute of Technology), and Zhijun Li (Harbin Institute of Technology; Soochow University)</i>	
TextFormer: Component-Aware Text Segmentation with Transformer	1877
<i>Xiaocong Wang (Fudan University, China), Chaoyue Wu (Fudan University, China), Haiyang Yu (Fudan University, China), Bin Li (Fudan University, China), and Xiangyang Xue (Fudan University, China)</i>	
SCFormer: Integrating Hybrid Features in Vision Transformers	1883
<i>Hui Lu (Utrecht University), Ronald Poppe (Utrecht University), and Albert Ali Salah (Bogazici University, Turkey)</i>	
Image Deraining Transformer with Sparsity and Frequency Guidance	1889
<i>Tianyu Song (Dalian Polytechnic University, China), Pengpeng Li (Dalian Polytechnic University, China), Guiyue Jin (Dalian Polytechnic University, China), Jiyu Jin (Dalian Polytechnic University, China), Shumin Fan (Dalian Polytechnic University, China), and Xiang Chen (Nanjing University of Science and Technology, China)</i>	

ShiftFormer: Spatial-Temporal Shift Operation in Video Transformer	1895
<i>Beiyang Yang (University of Chinese Academy of Sciences; Chinese Academy of Sciences), Guibo Zhu (University Chinese Academy of Sciences; Chinese Academy of Sciences; Wuhan AI Research), Guojing Ge (Chinese Academy of Sciences), Jinzhao Luo (University of Chinese Academy of Sciences; Chinese Academy of Sciences), and Jinqiao Wang (University Chinese Academy of Sciences; Chinese Academy of Sciences; Wuhan AI Research; Peng Cheng Laboratory)</i>	
ABMNet: Coupling Transformer with CNN Based on Adams-Bashforth-Moulton Method for Infrared Small Target Detection	1901
<i>Tianxiang Chen (University of Science and Technology of China; Chinese Academy of Sciences, China), Qi Chu (University of Science and Technology of China; Chinese Academy of Sciences, China), Zhentao Tan (University of Science and Technology of China; Alibaba DAMO Academy; Chinese Academy of Sciences, China), Bin Liu (University of Science and Technology of China; Chinese Academy of Sciences, China), and Nenghai Yu (University of Science and Technology of China; Chinese Academy of Sciences, China)</i>	
ART: An Efficient Transformer with Atrous Residual Learning for Medical Images	1907
<i>Yue He (Northeastern University, China), Yufan Wang (Northeastern University, China), Linlong He (Northeastern University, China), Guangyao Pan (National University of Singapore (Suzhou) Research Institute, China), and He Ma (Northeastern University, China)</i>	
MedFCT: A Frequency Domain Joint CNN-Transformer Network for Semi-Supervised Medical Image Segmentation	1913
<i>Shiao Xie (Zhejiang University, China), Huimin Huang (Zhejiang University, China), Ziwei Niu (Zhejiang University, China), Lanfen Lin (Zhejiang University, China), and Yen-Wei Chen (Ritsumeikan University, Japan)</i>	
Cross-Cycle Transformer-Based Stitching Method for Low-Resolution Borehole Images	1919
<i>Jia Chen (Wuhan Textile University; Engineering Research Center of Hubei Province for Clothing Information, China), Zhenpeng Fu (WuHan textile university), Fang Fei (Wuhan Textile University), Mingfu Xiong (WuHan Textile University), Xinrong Hu (Wuhan Textile University), and Tao Peng (Wuhan Textile University)</i>	
Improving Vision Transformers with Nested Multi-head Attentions	1925
<i>Jiquan Peng (Yanshan University, China; The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province), Chaozhuo Li (Microsoft Research Asia, China), Yi Zhao (Yanshan University, China; The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province), Yuting Lin (Yanshan University, China; The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province), Xiaohan Fang (Yanshan University, China; The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province), and Jibing Gong (Yanshan University, China; The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province)</i>	

Knowledge Distillation II

KnowledgeIE: Unifying Online-Offline Distillation Based on Knowledge Inheritance and Evolution	1931
<i>Yiqing Shen (Johns Hopkins University) and Jing Ke (Shanghai Jiao Tong University)</i>	
Collaborative Spatial-Temporal Distillation for Efficient Video Deraining	1937
<i>Yuzhang Hu (Peking University, China), Minghao Liu (Peking University, China), Wenhan Yang (Peng Cheng Laboratory, China), Jiaying Liu (Peking University, China), and Zongming Guo (Peking University, China)</i>	
Adaptive Multi-teacher Knowledge Distillation with Meta-Learning	1943
<i>Hailin Zhang (Zhejiang University, China), Defang Chen (Zhejiang University, China), and Can Wang (Zhejiang University, China)</i>	
Towards General and Fast Video Derain via Knowledge Distillation	1949
<i>Defang Cai (Zhejiang University of Technology, China), Pan Mu (Zhejiang University of Technology, China), Sixian Chan (Zhejiang University of Technology, China), Zhanpeng Shao (Hunan Normal University, China), and Cong Bai (Zhejiang University of Technology, China)</i>	

Media Retrieval II

Deep Metric Multi-view Hashing for Multimedia Retrieval	1955
<i>Jian Zhu (Zhejiang Lab, China), Xiaohu Ruan (vivo AI lab, China), Yongli Cheng (Fuzhou University, China), Zhangmin Huang (Zhejiang Lab, China), Yu Cui (Zhejiang Lab, China), and Lingfang Zeng (Zhejiang Lab, China)</i>	
Mim: Lightweight Multi-Modal Interaction Model for Joint Video Moment Retrieval and Highlight Detection	1961
<i>Jinyu Li (National Engineering Research Center of Digital Life), Fuwei Zhang (National Engineering Research Center of Digital Life), Shujin Lin (Guangdong Key Laboratory for Big Data Analysis and Simulation of Public Opinion; Sun Yat-sen University, China), Fan Zhou (National Engineering Research Center of Digital Life), and Ruomei Wang (National Engineering Research Center of Digital Life)</i>	
Image-Text Retrieval via Preserving Main Semantics of Vision	1967
<i>Xu Zhang (University of Electronic Science and Technology of China, China), Xinzheng Niu (University of Electronic Science and Technology of China, China), Philippe Fournier-Viger (Shenzhen University, China), and Xudong Dai (University of Electronic Science and Technology of China, China)</i>	
Progressive Event Alignment Network for Partial Relevant Video Retrieval	1973
<i>Xun Jiang (University of Electronic Science and Technology of China, China), Zhiguo Chen (University of Electronic Science and Technology of China, China), Xing Xu (University of Electronic Science and Technology of China, China), Fumin Shen (University of Electronic Science and Technology of China, China), Zuo Cao (Meituan, China), and Xunliang Cai (Meituan, China)</i>	

Immersive Video Technologies for the Metaverse

- Understanding and Improving Perceptual Quality of Volumetric Video Streaming 1979
Mengyu Yang (Sun Yat-sen University, China), Di Wu (Sun Yat-sen University, China), Zelong Wang (Sun Yat-sen University, China), Miao Hu (Sun Yat-sen University, China), and Yipeng Zhou (Macquarie University, Australia)
- Adaptive Geometry Reconstruction for Geometry-Based Point Cloud Compression 1985
Lei Wei (Northwestern Polytechnical University, China), Shuai Wan (Northwestern Polytechnical University, China), Xiaobin Ding (Northwestern Polytechnical University, China), FuZheng Yang (Xidian University, China), and Zhecheng Wang (Northwestern Polytechnical University, China)
- CAS-Net Cascade Attention-Based Sampling Neural Network for Point Cloud Simplification .. 1991
Chen Chen (Shandong University, China), Hui Yuan (Shandong University, China), Hao Liu (Yantai University, China), Junhui Hou (City University of Hong Kong, China), and Raouf Hamzaoui (De Montfort University, UK)
- PCHM-Net: A New Point Cloud Compression Framework for Both Human Vision and Machine Vision..... 1997
Lei Liu (Beihang University, China), Zhihao Hu (Beihang University, China), and Jing Zhang (Beihang University, China)
- Large-Scale Spatio-Temporal Attention Based Entropy Model for Point Cloud Compression .. 2003
Rui Song (Peking University, China; Peng Cheng Laboratory, China), Chunyang Fu (Peking University, China), Shan Liu (Tencent America, United States of America), and Ge Li (Peking University, China)
- Edge-FVV: Free Viewpoint Video Streaming by Learning at the Edge 2009
Haipeng Zhang (National Key Laboratory for Multimedia Information Processing; Peking University, China), Jie Zhang (University of Bath, UK), Weimiao Feng (Institute of Information Engineering of CAS, China), Kaigui Bian (National Key Laboratory for Multimedia Information Processing; Peking University, China), and Hu Tuo (IQIYI Science & Technology Co., Ltd., China)

3D & Point Cloud

- Weighted Point Cloud Normal Estimation 2015
Weijia Wang (Deakin University, Australia), Xuequan Lu (Deakin University, Australia), Di Shao (Deakin University, Australia), Xiao Liu (Deakin University, Australia), Richard Dazeley (Deakin University, Australia), Antonio Robles-Kelly (Deakin University, Australia), and Wei Pan (OPT Machine Vision Tech Co., Ltd, Japan)
- HybridPoint: Point Cloud Registration Based on Hybrid Point Sampling and Matching 2021
Yiheng Li (Xi'an Jiaotong University, China), Canhui Tang (Xi'an Jiaotong University, China), Runzhao Yao (Xi'an Jiaotong University, China), Aixue Ye (Huawei Noah's Ark Lab, China), Feng Wen (Huawei Noah's Ark Lab, China), and Shaoyi Du (Xi'an Jiaotong University, China)

Learning Deep Photometric Stereo Network with Reflectance Priors	2027
<i>Yakun Ju (The Hong Kong Polytechnic University, Hong Kong SAR), Cong Zhang (The Hong Kong Polytechnic University, Hong Kong SAR), Kin-Man Lam (The Hong Kong Polytechnic University, Hong Kong SAR), Songsong Huang (Ocean University of China, China), and Yuan Rao (Ocean University of China, China)</i>	
SST: Real-Time End-to-end Monocular 3D Reconstruction via Sparse Spatial-Temporal Guidance... 2033	
<i>Chenyangguang Zhang (Tsinghua University), Zhiqiang Lou (Tsinghua University), Yan Di (Technical University of Munich), Federico Tombari (Technical University of Munich; Google), and Xiangyang Ji (Tsinghua University, China)</i>	
Implicit Neural Distance Optimization for Mesh Neural Subdivision	2039
<i>Ke Liu (Zhejiang University, China), Ning Ma (Zhejiang University, China), Zhihua Wang (Zhejiang University, China), Jingjun Gu (Zhejiang University, China), Jiajun Bu (Zhejiang University, China), and Haishuai Wang (Zhejiang University, China)</i>	
MRRA-GAN: Multi-resolution Relation-Aware GAN for Point Cloud Completion	2045
<i>Ke Ren (University of Electronic Science and Technology of China, China), Zhenjiang Du (University of Electronic Science and Technology of China, China), Qifeng He (University of Electronic Science and Technology of China, China), Ning Xie (University of Electronic Science and Technology of China, China), and Guan Wang (University of Electronic Science and Technology of China, China)</i>	

Attention based Machine Learning

LSAS: Lightweight Sub-Attention Strategy for Alleviating Attention Bias Problem	2051
<i>Shanshan Zhong (Sun Yat-Sen University, China), Wushao Wen (Sun Yat-Sen University, China), Jinghui Qin (Guangdong University of Technology, China), Qiangpu Chen (Sun Yat-sen University, China), and Zhongzhan Huang (Sun Yat-Sen University, China)</i>	
LA-Layer: General Local Attention Layer for full Attention Networks	2057
<i>Hui Lu (Utrecht University), Ronald Poppe (Utrecht University), and Albert Ali Salah (Bogazici University, Turkey)</i>	
A Progressive Gated Attention Model for Fine-Grained Visual Classification	2063
<i>Qiangxi Zhu (Guangxi Normal University, China) and Zhixin Li (Guangxi Normal University, China)</i>	
Flow-Guided Attention Deformation for Person Image Generation	2069
<i>Yubo Wu (Peking University, China), Yurui Ren (Bytedance Inc.), Yuanqi Chen (Peking University, China), and Ge Li (Peking University, China)</i>	

Explicit Attention Modeling for Pedestrian Attribute Recognition	2075
<i>Jinyi Fang (Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Bingke Zhu (Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Yingying Chen (Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Jinqiao Wang (Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), and Ming Tang (Institute of Automation, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China)</i>	
Hidden Follower Detection via Refined Gaze and Walking State Estimation	2081
<i>Yaxi Chen (Wuhan University, China), Ruimin Hu (Wuhan University, China), Danni Xu (the National University of Singapore, Singapore), Zheng Wang (Wuhan University, China), Linbo Luo (Xidian University, China), and Dengshi Li (Wuhan University; Jiangnan University, China)</i>	

Generation

SketchScene: Scene Sketch to Image Generation with Diffusion Models	2087
<i>Zhenbei Wu (Beijing University of Posts and Telecommunications, China), Haoge Deng (Beijing University of Posts and Telecommunications, China), Qiang Wang (Beijing University of Posts and Telecommunications, China), Di Kong (Beijing University of Posts and Telecommunications, China), Jie Yang (Beijing University of Posts and Telecommunications, China), and Yonggang Qi (Beijing University of Posts and Telecommunications, China)</i>	
DanceU: Motion-and-Music-Based Automatic Effect Generation for Dance Videos	2093
<i>Yanjie Pan (Intel, China), Yaru Du (Intel, China), Shandong Wang (Intel, China), Yun Ye (Intel, China), Yong Jiang (Intel, China), Zhen Zhou (Intel, China), Li Xu (Intel, China), Ming Lu (Intel, China), Yunbiao Lin (Intel, China), and Jiehui Lu (Intel, China)</i>	
FONT: Flow-guided One-shot Talking Head Generation with Natural Head Motions	2099
<i>Jin Liu (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Xi Wang (Chinese Academy of Sciences, China), Xiaomeng Fu (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Yesheng Chai (Chinese Academy of Sciences, China), Cai Yu (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Jiao Dai (Chinese Academy of Sciences, China), and Jizhong Han (Chinese Academy of Sciences, China)</i>	
UFS-Net: Unsupervised Network For Fashion Style Editing and Generation	2105
<i>Wanqing Wu (South China University of Technology, China), Aihua Mao (South China University of Technology, China), Wenwei Yan (South China University of Technology, China), and Qing Liu (South China University of Technology, China)</i>	
Graph Convolutional GRU for Music-Oriented Dance Choreography Generation	2111
<i>Yuxin Hou (Harbin Institute of Technology, China), Hongxun Yao (Harbin Institute of Technology, China), and Haoran Li (Harbin Institute of Technology, China)</i>	

Cclap: Controllable Chinese Landscape Painting Generation Via Latent Diffusion Model	2117
<i>Zhongqi Wang (Beijing Institute of Technology, China), Jie Zhang (Chinese Academy of Sciences, China), Zhilong Ji (Tomorrow Advancing Life, China), Jinfeng Bai (Tomorrow Advancing Life, China), and Shiguang Shan (Chinese Academy of Sciences, China)</i>	

Understanding

Prototype Calibration for Long Tailed Recognition	2123
<i>Zhongan Wang (ShanghaiTech University, China), Shuai Shi (ShanghaiTech University, China, China), Yingna Wu (ShanghaiTech University, China), and Rui Yang (ShanghaiTech University, China)</i>	
ReadLM: Understanding Structured Document using Reading Order and Relationship of Adjacent Texts	2129
<i>Songjian Chen (Fudan University, China), Yuejie Zhang (Fudan University, China), and Rui Feng (Fudan University, China)</i>	
Open-Vocabulary Multi-label Image Classification with Pretrained Vision-Language Model	2135
<i>Son D.Dao (Monash University, Australia), Dat Huynh (Northeastern University, United States), He Zhao (CSIRO's Data61, Australia), Dinh Phung (Monash University, Australia), and Jianfei Cai (Monash University, Australia)</i>	
RASNet: A Reinforcement Assistant Network for Frame Selection in Video-Based Posture Recognition	2141
<i>Ruotong Hu (Ocean University of China, China), Xianzhi Wang (University of Technology Sydney, Australia), Xiaojun Chang (University of Technology Sydney, Australia), Yeqi Hu (Ocean University of China, China), Xiaowei Xin (Ocean University of China, China), Xiangqian Ding (Ocean University of China, China), and Baoqi Guo (NEWSTAR Software & Consulting CO.LTD, China)</i>	
Dynamic Spatial-Temporal Hypergraph Convolutional Network for Skeleton-Based Action Recognition	2147
<i>Shengqin Wang (Jilin University, China), Yongji Zhang (Jilin University, China), Hong Qi (Jilin University, China), Minghao Zhao (Jilin University, China), and Yu Jiang (Jilin University, China)</i>	
A Token-Wise Graph-Based Framework for Multimodal Named Entity Recognition	2153
<i>Zhengxuan Zhang (South China Normal University, China), Weixing Mai (South China Normal University, China), Haoliang Xiong (South China Normal University, China), Chuhan Wu (Tsinghua University, China), and Yun Xue (South China Normal University, China)</i>	

Visual Information Processing I

Multi-Focus Image Fusion via Gradient Guidance Progressive Network	2159
<i>Zhao Duan (Chongqing University, China), Xiaoliu Luo (Chongqing University, China), and Taiping Zhang (Chongqing University, China)</i>	

Efficient Video Matting on Human Video Clips for Real-Time Application	2165
<i>Chao-Liang Yu (National Yang Ming Chiao Tung University, Taiwan) and I-Chen Lin (National Yang Ming Chiao Tung University, Taiwan)</i>	
Render-and-Compare: Cross-View 6-DoF Localization from Noisy Prior	2171
<i>Shen Yan (National University of Defense Technology, China), Xiaoya Cheng (National University of Defense Technology, China), Yuxiang Liu (National University of Defense Technology, China), Juelin Zhu (National University of Defense Technology, China), Rouwan Wu (National University of Defense Technology, China), Yu Liu (National University of Defense Technology, China), and Maojun Zhang (National University of Defense Technology, China)</i>	
Video Snapshot Compressive Imaging via Optical Flow	2177
<i>Zan Chen (Zhejiang University of Technology, China), Ran Li (Zhejiang University of Technology, China), Yongqiang Li (Zhejiang University of Technology, China), and Yuanjing Feng (Zhejiang University of Technology, China)</i>	
Castensorf: Cascaded Tensorial Radiance Fields for Novel View Synthesis	2183
<i>Wenpeng Xing (Hong Kong Baptist University, China) and Jie Chen (Hong Kong Baptist University, China)</i>	
Compact Real-Time Radiance Fields with Neural Codebook	2189
<i>Lingzhi Li (Alibaba DAMO Academy, China), Zhongshu Wang (Alibaba DAMO Academy, China), Zhen Shen (Alibaba DAMO Academy, China), Li Shen (Alibaba DAMO Academy, China), and Ping Tan (Alibaba DAMO Academy, China)</i>	

Detection II

STNet: Spatial and Temporal Feature Fusion Network for Change Detection in Remote Sensing Images	2195
<i>Xiaowen Ma (Zhejiang University), Jiawei Yang (Zhejiang University), Tingfeng Hong (Zhejiang University), Mengting Ma (Zhejiang University), Ziyang Zhao (Zhejiang University), Tian Feng (Zhejiang University; Alibaba-Zhejiang University Joint Research Institute of Frontier Technologies), and Wei Zhang (Zhejiang University)</i>	
Social Bot Detection Based on Window Strategy	2201
<i>Boyu Qiao (UCAS, China), Kun Li (UCAS, China), Wei Zhou (UCAS, China), Zhou Yan (People's Daily Online, China), Shilong Li (UCAS, China), and Songlin Hu (UCAS, China)</i>	
A Semantics-Aware Normalizing Flow Model for Anomaly Detection	2207
<i>Wei Ma (Sichuan University, China), Shiyong Lan (Sichuan University, China), Weikang Huang (National Key Laboratory of Fundamental Science on Synthetic Vision, China), Wenwu Wang (Center for Vision Speech and Signal Processing, University of Surrey, UK), Hongyu Yang (Sichuan University, China), Yitong Ma (Sichuan University, China), and Yongjie Ma (Sichuan University, China)</i>	

Online Action Detection with Learning Future Representations by Contrastive Learning	2213
<i>Haitao Leng (Alibaba Group, China), Xiaoming Shi (Shanghai Artificial Intelligence Laboratory, China), Wei Zhou (Alibaba Group, China), Kuncai Zhang (Alibaba Group, China), Qiankun Shi (Alibaba Group, China), and Pengcheng Zhu (Alibaba Group, China)</i>	
HOD: Human-Object Decoupling Network for HOI Detection	2219
<i>Hantao Zhang (University of Science and Technology of China, China), Shouhong Wan (University of Science and Technology of China, China), Weidong Guo (University of Science and Technology of China, China), Peiquan Jin (University of Science and Technology of China, China), and Mingguang Zheng (University of Science and Technology of China, China)</i>	
Fixing Domain Bias for Generalized Deepfake Detection	2225
<i>Yuzhe Mao (Beijing University of Posts and Telecommunications, China), Weike You (Beijing University of Posts and Telecommunications, China), Linna Zhou (Beijing University of Posts and Telecommunications, China), and Zhigao Lu (University of International Relations, China)</i>	
Variational Information Bottleneck for Cross Domain Object Detection	2231
<i>Jiangming Chen (National University of Defense Technology, China), Wanxia Deng (National University of Defense Technology, China), Bo Peng (National University of Defense Technology, China), Tianpeng Liu (National University of Defense Technology, China), Yingmei Wei (National University of Defense Technology, China), and Li Liu (National University of Defense Technology, China)</i>	
Attention Based Network with DA-Loss for X-ray Contraband Automatic Detection	2237
<i>Peiwen Li (University of Chinese Academy of Sciences, China), Lijun Zhang (University of Chinese Academy of Sciences, China), Xiang-Dong Zhou (University of Chinese Academy of Sciences, China), Yu Shi (University of Chinese Academy of Sciences, China), and Xiaohu Shao (University of Chinese Academy of Sciences, China)</i>	
Cross-Level Attention Based Adaptive Feature Alignment Network for Arbitrary-Shaped Text Detection	2243
<i>Haiyan Zhang (Tianjin University, China) and Sumei Li (Tianjin University, China)</i>	

Generation II

Multi-object Video Generation from Single Frame Layouts	2249
<i>Yang Wu (Sun Yat-sen University), Zhibin Liu (Sun Yat-sen University), Hefeng Wu (Sun Yat-sen University), and Liang Lin (Sun Yat-sen University)</i>	
Towards Confidence-Aware Commonsense Knowledge Integration for Scene Graph Generation ...	2255
<i>Hongshuo Tian (Tianjin University; Hefei Comprehensive National Science Center, China), Ning Xu (Tianjin University, China), Yanhui Wang (Tianjin University, China), Chenggang Yan (Hangzhou Dianzi University, China), Bolun Zheng (Hangzhou Dianzi University, China), Xuanya Li (Baidu Inc., China), and An-An Liu (Tianjin University; Hefei Comprehensive National Science Center, China)</i>	

Image Layer Modeling for Complex Document Layout Generation	2261
<i>Tianlong Ma (Fudan University, China), Xingjiao Wu (Fudan University, China), Xiangcheng Du (Fudan University, China), Yanlong Wang (Shanxi University, China), and Cheng Jin (Fudan University, China)</i>	
Knowledge Enhanced Model for Live Video Comment Generation	2267
<i>Jieting Chen (Renmin University of China, China), Junkai Ding (Columbia University, United States), Wenping Chen (Renmin University of China, China), and Qin Jin (Renmin University of China, China)</i>	

Segmentation II

Mutual Query Network for Multi-modal Product Image Segmentation	2273
<i>Yun Guo (Retail Platform Operation and Marketing Center, China), Wei Feng (Retail Platform Operation and Marketing Center, China), Zheng Zhang (Retail Platform Operation and Marketing Center, China), Xiancong Ren (Retail Platform Operation and Marketing Center, China), Yaoyu Li (Retail Platform Operation and Marketing Center, China), Jingjing Lv (Retail Platform Operation and Marketing Center, China), Xin Zhu (Retail Platform Operation and Marketing Center, China), Zhangang Lin (Retail Platform Operation and Marketing Center, China), and Jingping Shao (Retail Platform Operation and Marketing Center, China)</i>	
ATENet: Adaptive Tiny-Object Enhanced Network for Polyp Segmentation	2279
<i>Xiaogang Du (Shaanxi University of Science and Technology, China), Yinghao Wu (Shaanxi University of Science and Technology, China), Tao Lei (Shaanxi University of Science and Technology, China), Dongxin Gu (Shaanxi University of Science and Technology, China), Yinyin Nie (Shaanxi University of Science and Technology, China), and Asoke K. Nandi (Brunel University London, United Kingdom)</i>	
Adaptive-Masking Policy with Deep Reinforcement Learning for Self-Supervised Medical Image Segmentation	2285
<i>Gang Xu (Hebei University of Technology, China), Shengxin Wang (Hebei University of Technology, China), Thomas Lukasiewicz (Institute of Logic and Computation, Austria; University of Oxford, UK), and Zhenghua Xu (Hebei University of Technology, China)</i>	
MSAANet : Multi-scale Axial Attention Network for Medical Image Segmentation	2291
<i>Hao Zeng (East China Normal University), Xinxin Shan (EastChinaNormalUniversity), Yu Feng (East China Normal University), and Ying Wen (East China Normal University)</i>	

DeepFake & Adversarial Robustness

A Simple Stochastic Neural Network for Improving Adversarial Robustness	2297
<i>Hao Yang (National University of Defense Technology, China), Min Wang (National University of Defense Technology, China), Zhengfei Yu (National University of Defense Technology, China), and Yun Zhou (National University of Defense Technology, China)</i>	

DFCP: Few-Shot DeepFake Detection via Contrastive Pretraining	2303
<i>Bo Zou (Tsinghua University, China), Chao Yang (Shanghai AI Laboratory, China), Jiazhi Guan (Tsinghua University, China), Chengbin Quan (Tsinghua University, China), and Youjian Zhao (Tsinghua University; Zhongguancun Laboratory, China)</i>	
Forensics Forest: Multi-scale Hierarchical Cascade Forest for Detecting GAN-Generated Faces	2309
<i>Jiucui Lu (Ocean University of China, China), Yuezun Li (Ocean University of China, China), Jiaran Zhou (Ocean University of China, China), Bin Li (Shenzhen University, China), and Siwei Lyu (University at Buffalo, State University of New York, USA)</i>	
Transferable Waveform-Level Adversarial Attack Against Speech Anti-Spoofing Models	2315
<i>Bingyuan Huang (Sun Yat-Sen University, China), Sanshuai Cui (Sun Yat-Sen University, China), Xiangui Kang (Sun Yat-Sen University, China), and Enping Li (Bridgewater State University, USA)</i>	
Domain-Invariant Feature Learning for General Face Forgery Detection	2321
<i>Jian Zhang (Sun Yat-sen University; Zhengzhou Xinda Institute of Advanced Technology, China) and Jiangqun Ni (Sun Yat-sen University; Cyberspace Security Research Center, Peng Cheng Laboratory, China)</i>	
Image Copy-Move Forgery Detection via Deep Cross-Scale PatchMatch	2327
<i>Yingjie He (Shenzhen University, China), Yuanman Li (Shenzhen University, China), Changsheng Chen (Shenzhen University, China), and Xia Li (Shenzhen University, China)</i>	

Segmentation I

BAProto: Boundary-Aware Prototype for High-Quality Instance Segmentation	2333
<i>Yuxuan Zhang (University of Science and Technology of China, China), Wei Yang (University of Science and Technology of China; Hefei National Laboratory, China), and Rong Hu (University of Science and Technology of China, China)</i>	
PMDA: Domain Alignment with Prototype Matching for Cross-Domain Adaptive Segmentation	2339
<i>Weiwei Li (University of Electronic Science and Technology of China, China), Yuanyuan Ren (Shihezi University, China), Junzhuo Liu (University of Electronic Science and Technology of China, China), Chenyang Wang (Chinese Academy of Sciences, China), and Yuchen Zheng (Shihezi University, China)</i>	
Cross-Slice Context Consistency for Semi-Supervised 3D Left Atrium Segmentation	2345
<i>Yongchao Wang (Chongqing University of Posts and Telecommunications, China), Bin Xiao (Chongqing University of Posts and Telecommunications, China), Xiuli Bi (Chongqing University of Posts and Telecommunications, China), Weisheng Li (Chongqing University of Posts and Telecommunications, China), and Xinbo Gao (Chongqing University of Posts and Telecommunications, China)</i>	

CTSSeg: Consistent Teacher-Student Model for Magnetic Resonance Image Segmentation ...	2351
<i>Chenbin Zhang (Peking University, China), Qingyuan He (Peking University Third Hospital, China), Kun Yan (Peking University, China), Meng Ma (Peking University, China), Defeng Liu (Peking University Third Hospital, China), and Ping Wang (Peking University, China)</i>	
Adaptive Non-Local Affinity Graph for Unsupervised Image Segmentation	2357
<i>Xin Lv (Lanzhou University, China), Zhenming Su (Lanzhou University, China), Taiyi Zhang (Lanzhou University, China), Wenxiang Cheng (Lanzhou University, China), and Xiaoqiong Qi (The University of Queensland, Australia)</i>	
Fine-Grained Domain Adaptive Crowd Counting via Point-Derived Segmentation	2363
<i>Yongtuo Liu (South China University of Technology; University of Amsterdam), Dan Xu (Hong Kong University of Science and Technology), Sucheng Ren (South China University of Technology), Hanjie Wu (South China University of Technology), Hongmin Cai (South China University of Technology), and Shengfeng He (Singapore Management University)</i>	

Attention & Saliency

Scribble-Supervised RGB-T Salient Object Detection	2369
<i>Zhengyi Liu (Anhui University, China), Xiaoshen Huang (Anhui University, China), Guanghui Zhang (Anhui University, China), Xianyong Fang (Anhui University, China), Linbo Wang (Anhui University, China), and Bin Tang (Hefei University, China)</i>	
CA-GAN: Object Placement via Coalescing Attention Based Generative Adversarial Network .	2375
<i>Yibin Wang (Zhejiang University of Technology, China), Yuchao Feng (Zhejiang University of Technology, China), Jie Wu (Zhejiang University of Technology, China), Honghui Xu (Zhejiang University of Technology, China), and Jianwei Zheng (Zhejiang University of Technology, China)</i>	
ABC: Attention with Bilinear Correlation for Infrared Small Target Detection	2381
<i>Peiwen Pan (Nanjing University of Science and Technology, China), Huan Wang (Nanjing University of Science and Technology, China), Chenyi Wang (Nanjing University of Science and Technology, China), and Chang Nie (Nanjing University of Science and Technology, China)</i>	
Guided Focal Stack Refinement Network for Light Field Salient Object Detection	2387
<i>Bo Yuan (Sichuan University, China), Yao Jiang (Sichuan University, China), Keren Fu (Sichuan University, China), and Qijun Zhao (Sichuan University, China)</i>	
Triplet Spatiotemporal Aggregation Network for Video Saliency Detection	2393
<i>Zhenshan Tan (Fudan University, China), Cheng Chen (Fudan University, China), and Xiaodong Gu (Fudan University, China)</i>	
GFNet: Gaze Focus Network using Attention for Gaze Estimation	2399
<i>Daosong Hu (Sun Yat-sen University, China) and Kai Huang (Sun Yat-sen University, China)</i>	

Person Re-identification

- Feature Mixing and Disentangling for Occluded Person Re-identification 2405
Zepeng Wang (Shanghai Jiao Tong University, China), Ke Xu (Shanghai Jiao Tong University, China), Yuting Mou (Shanghai Jiao Tong University, China), and Xinghao Jiang (Shanghai Jiao Tong University, China)
- Multi-scale Query-Adaptive Convolution for Generalizable Person Re-identification 2411
Kaixiang Chen (Nanjing University of Aeronautics and Astronautics, China), Tiantian Gong (Nanjing University of Aeronautics and Astronautics, China), and Liyan Zhang (Nanjing University of Aeronautics and Astronautics, China)
- Fine-Grained Learning for Visible-Infrared Person Re-identification 2417
Mengzan Qi (East China Normal University, China), Sixian Chan (Zhejiang University of Technology; Xsuan Technology Co., Ltd., China), Chen Hang (East China Normal University, China), Guixu Zhang (East China Normal University, China), and Zhi Li (East China Normal University, China)
- Camera Proxy Based Contrastive Learning with Hard Sampling for Unsupervised Person Re-identification 2423
Yimin Liu (Hefei University of Technology, China), Meibin Qi (Hefei University of Technology, China), Qiang Wu (University of Technology Sydney, Australia), Yanfang Yang (Hefei University of Technology, China), Xiaohong Li (Hefei University of Technology, China), and Jian Zhang (University of Technology Sydney, Australia)
- Inter-Intra Camera Identity Learning for Person Re-Identification with Training in Single Camera 2429
Guoqing Zhang (Nanyang Technological University, Singapore), Zhiyuan Luo (Nanjing University of Information Science and Technology, China), Weisi Lin (Nanyang Technological University, Singapore), and Jing Xuan (NCS Pte Ltd, Singapore)
- Dynamically Adaptive Instance Normalization and Attention-Aware Incremental Meta-Learning for Generalizable Person Re-identification 2435
Tiantian Gong (Nanjing University of Aeronautics and Astronautics, China), Kaixiang Chen (Nanjing University of Aeronautics and Astronautics, China), Junsheng Wang (Nanjing University of Science And Technology, China), and Liyan Zhang (Nanjing University of Aeronautics and Astronautics, China)

Object Detection

- CFANet: A Cross-Layer Feature Aggregation Network for Camouflaged Object Detection 2441
Qing Zhang (Shanghai Institute of Technology, China) and Weiqi Yan (Shanghai Institute of Technology, China)
- Multibox Sample Selection for Active Object Detection 2447
Jiaxiang Dong (n/a) and Li Zhang (n/a)

Run and Chase: Towards Accurate Source-Free Domain Adaptive Object Detection	2453
<i>Luojun Lin (Fuzhou University, China), Zhifeng Yang (FuZhou University, China), Qipeng Liu (FuZhou University, China), Yuanlong Yu (Fuzhou University, China), and Qifeng Lin (Fuzhou University, China)</i>	
Camouflaged Object Detection with Feature Grafting and Distractor Aware	2459
<i>Yuxuan Song (Ocean University of China, China), Xinyue Li (Ocean University of China, China), and Lin Qi (Ocean University of China, China)</i>	
Edge-Aware Mirror Network for Camouflaged Object Detection	2465
<i>Dongyue Sun (Ocean University of China, China), Shiyao Jiang (Ocean University of China, China), and Lin Qi (Ocean University of China, China)</i>	
'Skimming-Perusal' Detection: A Simple Object Detection Baseline in GigaPixel-Level Images.	2471
<i>Zhibin Zhang (Tianjin University of Technology, China), Wanli Xue (Tianjin University of Technology, China), Kaihua Zhang (Nanjing University of Information Science and Technology, China), and Shengyong Chen (Tianjin University of Technology, China)</i>	

Quality Assessment II

Attribute-Assisted Multimodal Network for Image Aesthetics Assessment	2477
<i>Tong Zhu (Xidian University, China), Leida Li (Xidian University, China), Pengfei Chen (Xidian University, China), Jinjian Wu (Xidian University, China), Yuzhe Yang (OPPO Research Institute, China), Yaqian Li (OPPO Research Institute, China), and Yandong Guo (OPPO Research Institute, China)</i>	
EEP-3DQA: Efficient and Effective Projection-Based 3D Model Quality Assessment	2483
<i>Zicheng Zhang (Shanghai Jiao Tong University), Wei Sun (Shanghai Jiao Tong University), Yingjie Zhou (Shanghai Jiao Tong University), Wei Lu (Shanghai Jiao Tong University), Yucheng Zhu (Shanghai Jiao Tong University), Xionghuo Min (Shanghai Jiao Tong University), and Guangtao Zhai (Shanghai Jiao Tong University)</i>	
Exploring the Influence of View and Camera Path Selection for Dynamic Mesh Quality Assessment	2489
<i>Kaifa Yang (Shanghai Jiao Tong University, China), Qi Yang (Tencent, China), Joel Jung (Tencent, United States), Yiling Xu (Shanghai Jiao Tong University, China), Xiaozhong Xu (Tencent, United States), and Shan Liu (Tencent, United States)</i>	
HandGCAT: Occlusion-Robust 3D Hand Mesh Reconstruction from Monocular Images	2495
<i>Shuaibing Wang (Fudan University; Institute of Meta-Medical), Shunli Wang (Fudan University; Institute of Meta-Medical), Dingkan Yang (Fudan University; Institute of Meta-Medical), Mingcheng Li (Fudan University; Institute of Meta-Medical), Ziyun Qian (Fudan University; Institute of Meta-Medical), Liuzhen Su (Fudan University; Institute of Meta-Medical), and Lihua Zhang (Fudan University; Institute of Meta-Medical; Jilin Provincial Key Laboratory of Intelligence Science and Engineering; AI & Unmanned Systems Engineering Research Center of Jilin Province)</i>	

BH-VQA: Blind High Frame Rate Video Quality Assessment	2501
<i>Wei Lu (Shanghai Jiao Tong University), Wei Sun (Shanghai Jiao Tong University), Zicheng Zhang (Shanghai Jiao Tong University), Danyang Tu (Shanghai Jiao Tong University), Xionghuo Min (Shanghai Jiao Tong University), and Guangtao Zhai (Shanghai Jiao Tong University)</i>	
Coordinate Channel Attention and a High-Dimensional Fusion Network for Speech Depression Level Estimation	2507
<i>Shizhao Liu (Tianjin Normal University, China), Mingyue Niu (Tianjin Normal University, China), and Ziping Zhao (Tianjin Normal University, China)</i>	
Multi-level Feature-Guided Stereoscopic Video Quality Assessment Based on Transformer and Convolutional Neural Network	2513
<i>Yuan Chen (Tianjin University, China) and Sumei Li (Tianjin University, China)</i>	
DDH-QA: A Dynamic Digital Humans Quality Assessment Database	2519
<i>Zicheng Zhang (Shanghai Jiao Tong university; Peng Cheng Laboratory, China), Yingjie Zhou (Shanghai Jiao Tong University; Peng Cheng Laboratory, China), Wei Sun (Shanghai Jiao Tong University; Peng Cheng Laboratory, China), Wei Lu (Shanghai Jiao Tong University; Peng Cheng Laboratory, China), Xionghuo Min (Shanghai Jiao Tong University; Peng Cheng Laboratory, China), Yu Wang (Shanghai Jiaotong University, China), and Guangtao Zhai (Shanghai Jiao Tong University; Peng Cheng Laboratory, China)</i>	

Media Storage & Coding

Improving Multi-generation Robustness of Learned Image Compression	2525
<i>Litian Li (Peking University, China), Zheng Yang (Peking University, China), Yongqi Zhai (Peking University; Peng Cheng Laboratory, China), Jiayu Yang (Peking University; Peng Cheng Laboratory, China), and Ronggang Wang (Peking University; Peng Cheng Laboratory, China)</i>	
Code Verification Hashing for Image Retrieval	2531
<i>Yinqi Chen (Jihua Laboratory, China), Zhiyi Lu (Jihua Laboratory, China), Ya Lu (Jihua Laboratory, China), Yangting Zheng (JiHua Laboratory, China), Peiwen Li (JiHua Laboratory, China), and Shuo Kang (Jihua Laboratory, China)</i>	
Low-Complexity Deep Video Compression with a Distributed Coding Architecture	2537
<i>Xinjie Zhang (The Hong Kong University of Science and Technology, China), Jiawei Shao (The Hong Kong University of Science and Technology, China), and Jun Zhang (The Hong Kong University of Science and Technology, China)</i>	
Perceptual Audio Object Coding using Adaptive Subband Grouping with CNN and Residual Block..	2543
<i>Yulin Wu (Wuhan University, China), Ruimin Hu (Wuhan University, China), and Xiaochen Wang (Wuhan University, China)</i>	

Learning Lossless Compression for High Bit-Depth Medical Imaging	2549
<i>Kai Wang (Harbin Institute of Technology, China; Peng Cheng Laboratory, China), Yuanchao Bai (Harbin Institute of Technology, China), Deming Zhai (Harbin Institute of Technology, China), Daxin Li (Harbin Institute of Technology, China; Peng Cheng Laboratory, China), Junjun Jiang (Harbin Institute of Technology, China; Peng Cheng Laboratory, China), and Xianming Liu (Harbin Institute of Technology, China; Peng Cheng Laboratory, China)</i>	
Sparse Representation Based Deep Residual Geometry Compression Network for Large-Scale Point Clouds	2555
<i>Pengpeng Yu (Sun Yat-sen University, China), Dian Zuo (Sun Yat-sen University, China), Yueer Huang (Sun Yat-sen University, China), Ruishan Huang (Sun Yat-sen University, China), Hanyun Wang (Information Engineering University, China), Yulan Guo (Sun Yat-sen University, China), and Fan Liang (Sun Yat-sen University; Peng Cheng Laboratory, China)</i>	
An Efficient Real-Time Hardware Architecture for Deblocking Filter in AVS3	2561
<i>Shaokang Wang (Peking University, China), Xiaofeng Huang (Hangzhou Dianzi University, China), Guoqing Xiang (Peking university, China), Xizhong Zhu (Peking University, China), Jiaojiao Yang (Peking University, China), Peng Zhang (Peking University, China), Huizhu Jia (Peking University, China), and Xiaodong Xie (Peking University, China)</i>	
Microimage-Based Two-Step Search For Plenoptic 2.0 Video Coding	2567
<i>Yuqing Yang (Tsinghua University, China), Xin Jin (Tsinghua University, China), Kedeng Tong (Tsinghua University, China), Chen Wang (Tsinghua University, China), and Haitian Huang (Tsinghua University, China)</i>	
Low Complexity Transcoding from HEVC to VVC	2573
<i>Xi Xie (City University of Hong Kong, CN), Kai Zhang (Bytedance Inc., USA), Li Zhang (Bytedance Inc., USA), Meng Wang (City University of Hong Kong, CN), Junru Li (Douyin Group (HK) Limited, CN), and Shiqi Wang (City University of Hong Kong, CN)</i>	

Re-ID

Visible-Xray Cross-Modality Package Re-Identification	2579
<i>Sixian Chan (Zhejiang University of Technology; Chinese Academy of Science; Anhui QixinMingzhi Technology Co., Ltd., China), Jiaao Cui (Zhejiang University of Technology, China), Yonggan Wu (Chinese Academy of Science; Anhui QixinMingzhi Technology Co., Ltd., China), Hongqiang Wang (Chinese Academy of Science, China), and Cong Bai (Zhejiang University of Technology, China)</i>	
Aerial-Ground Person Re-ID	2585
<i>Huy Nguyen (Queensland University of Technology, Australia), Kien Nguyen (Queensland University of Technology, Australia), Sridha Sridharan (Queensland University of Technology, Australia), and Clinton Fookes (Queensland University of Technology, Australia)</i>	

A Knowledge-Driven Cross-Period Network for Clothes Changing RE-ID	2591
<i>Keyang Cheng (Jiangsu University, China), Sai Liang (Jiangsu University, China), Wenxuan Zou (Jiangsu University, China), and Hao Wan (Jiangsu University, China)</i>	
Meta Perturbed Re-Id Defense	2597
<i>Astha Verma (IIIT Delhi, India), A V Subramanyam (IIIT Delhi, India), Mohammad Ali Jauhar (IIIT Delhi, India), Divij Gera (IIIT Delhi, India), and Rajiv Ratn Shah (IIIT Delhi, India)</i>	

Multimodal Processing I

EFT: Expert Fusion Transformer for Voice-Face Association Learning	2603
<i>Guangyu Chen (Renmin University of China, China), Deyuan Zhang (Shenyang Aerospace University, China), Tao Liu (Renmin University of China, China), and Xiaoyong Du (Renmin University of China, China)</i>	
Social Context-Aware GCN for Video Character Search via Scene-Prior Enhancement	2609
<i>Wenjun Peng (University of Science and Technology of China; State Key Laboratory of Cognitive Intelligence, China), Weidong He (University of Science and Technology of China; State Key Laboratory of Cognitive Intelligence, China), Derong Xu (University of Science and Technology of China; State Key Laboratory of Cognitive Intelligence, China), Tong Xu (University of Science and Technology of China; State Key Laboratory of Cognitive Intelligence, China), Chen Zhu (University of Science and Technology of China; Alibaba, China), and Enhong Chen (University of Science and Technology of China; State Key Laboratory of Cognitive Intelligence, China)</i>	
MRCap: Multi-Modal and Multi-Level Relationship-Based Dense Video Captioning	2615
<i>Wei Chen (Beihang University, China), Jianwei Niu (Beihang University; Zhongguancun Laboratory; Zhengzhou University, China), and Xuefeng Liu (Beihang University; Zhongguancun Laboratory, China)</i>	
Auxiliary Fine-Grained Alignment Constraints for Vision-and-Language Navigation	2621
<i>Yibo Cui (Academy of Military Sciences China; Tianjin Artificial Intelligence Innovation Center (TAIIC), China), Ruqiang Huang (Tianjin Artificial Intelligence Innovation Center (TAIIC), China), Yakun Zhang (Academy of Military Sciences China; Tianjin Artificial Intelligence Innovation Center (TAIIC), China), Yingjie Cen (Tianjin Artificial Intelligence Innovation Center (TAIIC); East China University of Science and Technology), Liang Xie (Academy of Military Sciences China; Tianjin Artificial Intelligence Innovation Center (TAIIC), China), Ye Yan (Academy of Military Sciences China; Tianjin Artificial Intelligence Innovation Center (TAIIC), China), and Erwei Yin (Academy of Military Sciences China; Tianjin Artificial Intelligence Innovation Center (TAIIC), China)</i>	

Improving Audio-Visual Speech Recognition by Lip-Subword Correlation Based Visual Pre-Training and Cross-Modal Fusion Encoder	2627
<i>Yusheng Dai (University of Science and Technology, China), Hang Chen (University of Science and Technology, China), Jun Du (University of Science and Technology, China), Xiaofei Ding (Alibaba Group, China), Ning Ding (Alibaba Group, China), Feijun Jiang (Alibaba Group, China), and Chin-Hui Lee (Georgia Institute of Technology, USA)</i>	
Detection of Public Speaking Anxiety: A New Dataset and Algorithm	2633
<i>Wei Song (Beijing University of Posts and Telecommunications, China), Bin Wu (Beijing University of Posts and Telecommunications, China), Chunping Zheng (Beijing University of Posts and Telecommunications, China), and Huayang Zhang (Beijing Normal University, China)</i>	

Pose Estimation

EANet: Towards Lightweight Human Pose Estimation with Effective Aggregation Network	2639
<i>Beitao Chen (University of Electronic Science and Technology of China, China), Xuanhan Wang (University of Electronic Science and Technology of China, China), Xiaojia Chen (University of Electronic Science and Technology of China, China), Yulan He (the University of Warwick, UK), and Jingkuan Song (University of Electronic Science and Technology of China, China)</i>	
Effective Occlusion Suppression Network via Grouped Pose Estimation for Occluded Person Re-Identification	2645
<i>Zhihao Li (Shandong Normal University, China), Huaxiang Zhang (Shandong Normal University, China), Lei Zhu (Shandong Normal University, China), Jiande Sun (Shandong Normal University, China), and Li Liu (Shandong Normal University, China)</i>	
Structural Equivariance Self-Supervised Learning for Facial Pose Estimation	2651
<i>Yaoxing Wang (Ningxia University, China), Heng Zhou (Ningxia University, China), Zhendong Li (Ningxia University, China; Collaborative Innovation Center for Ningxia Big Data and Artificial Intelligence Co-founded by Ningxia Municipality and Ministry of Education, China; Key Laboratory of the Internet of Water and Digital Water Governance of the Yellow River in Ningxia, China), Xian Mo (Ningxia University, China; Collaborative Innovation Center for Ningxia Big Data and Artificial Intelligence Co-founded by Ningxia Municipality and Ministry of Education, China; Key Laboratory of the Internet of Water and Digital Water Governance of the Yellow River in Ningxia, China), and Hao Liu (Ningxia University, China; Collaborative Innovation Center for Ningxia Big Data and Artificial Intelligence Co-founded by Ningxia Municipality and Ministry of Education, China; Key Laboratory of the Internet of Water and Digital Water Governance of the Yellow River in Ningxia, China)</i>	

ActionPrompt: Action-Guided 3D Human Pose Estimation With Text and Pose Prompting	2657
<i>Hongwei Zheng (Shanghai Jiao Tong University, China), Han Li (Shanghai Jiao Tong University, China), Bowen Shi (Shanghai Jiao Tong University, China), Wenrui Dai (Shanghai Jiao Tong University, China), Botao Wang (Qualcomm AI Research, China), Yu Sun (Qualcomm AI Research, China), Min Guo (Qualcomm AI Research, China), and Hongkai Xiong (Shanghai Jiao Tong University, China)</i>	
Hierarchical Graph Neural Network for Human Pose Estimation	2663
<i>Guanghua Zheng (Hefei University of Technology, China), Zhongqiu Zhao (Hefei University of Technology, China), Zhao Zhang (Hefei University of Technology, China), and Yi Yang (Hefei University of Technology, China)</i>	
RF-Based Multi-view Pose Machine for Multi-person 3D Pose Estimation	2669
<i>Chunyang Xie (University of Electronic Science and Technology of China; University of Science and Technology of China, China), Dongheng Zhang (University of Science and Technology of China, China), Zhi Wu (University of Science and Technology of China, China), Cong Yu (University of Electronic Science and Technology of China; University of Science and Technology of China, China), Yang Hu (University of Science and Technology of China, China), Qibin Sun (University of Science and Technology of China, China), and Yan Chen (University of Science and Technology of China, China)</i>	

Domain Adaptation & Style Transfer

Coarse Helps Fine: A Multi-Granularity Discriminative Adversarial Network for Fine-Grained Open-Set Domain Adaptation	2675
<i>Jing Li (Tianjin University, China), Liu Yang (Tianjin University, China), Qilong Wang (Tianjin University, China), and Qinghua Hu (Tianjin University, China)</i>	
Adversarially Robust Source-free Domain Adaptation with Relaxed Adversarial Training	2681
<i>Yao Xiao (Sun Yat-sen University, China), Pengxu Wei (Sun Yat-sen University, China), Cong Liu (Sun Yat-sen University, China), and Liang Lin (Sun Yat-sen University, China)</i>	
A Compact Transformer for Adaptive Style Transfer	2687
<i>Yi Li (Dalian University of Technology, China), Xin Xie (Dalian University of Technology, China), Haiyan Fu (Dalian University of Technology, China), Xiangyang Luo (State Key Laboratory of Mathematical Engineering and Advanced Computing, China), and Yanqing Guo (Dalian University of Technology, China)</i>	
Fine-Grained Alignment for Boundary Samples Under Open Set Domain Adaptation	2693
<i>Jianglin Wei (Hunan University, China), Guangyi Xiao (Hunan University, China), Shun Peng (Hunan University, China), Hao Chen (Hunan University, China), Jingzhi Guo (University of Macau, China), and Zhiguo Gong (University of Macau, China)</i>	

Information Selection-Based Domain Adaptation from Black-box Predictors	2699
<i>Kai Wang (University of Electronic Science and Technology of China, China), Xing Xu (University of Electronic Science and Technology of China, China), Jialin Tian (University of Electronic Science and Technology of China, China), Zuo Cao (Meituan, China), and Gong Zhang (Meituan, China)</i>	
E ² : Entropy Discrimination and Energy Optimization for Source-free Universal Domain Adaptation	2705
<i>Meng Shen (Sun Yat-sen University, China), Andy J. Ma (Sun Yat-sen University, China; Guangdong Province Key Laboratory of Information Security Technology, China; Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, China), and Pong C. Yuen (Department of Computer Science, Hong Kong Baptist University, Hong Kong)</i>	

Anomaly Detection

Long-Short Temporal Co-Teaching for Weakly Supervised Video Anomaly Detection	2711
<i>Shengyang Sun (Zhejiang University, China) and Xiaojin Gong (Zhejiang University, China)</i>	
Multi-level Memory-Augmented Appearance-Motion Correspondence Framework for Video Anomaly Detection	2717
<i>Xiangyu Huang (Xiamen University, China), Caidan Zhao (Xiamen University, China), Jinhui Yu (Xiamen University, China), Chenxing Gao (Xiamen University, China), and Zhiqiang Wu (PKU-Wuhan Institute for Artificial Intelligence, China)</i>	
Weakly Supervised Video Anomaly Detection Based on Cross-Batch Clustering Guidance	2723
<i>Congqi Cao (Northwestern Polytechnical University, China), Xin Zhang (Northwestern Polytechnical University, China), Shizhou Zhang (Northwestern Polytechnical University, China), Peng Wang (Northwestern Polytechnical University, China), and Yanning Zhang (Northwestern Polytechnical University, China)</i>	
Pose-Motion Video Anomaly Detection via Memory-Augmented Reconstruction and Conditional Variational Prediction	2729
<i>Weilin Wan (Fudan University, China), Weizhong Zhang (Fudan University, China), and Cheng Jin (Fudan University, China)</i>	
Rethinking Graph Anomaly Detection: A Self-Supervised Group Discrimination Paradigm with Structure-Aware	2735
<i>Junyi Yan (Xinjiang University, China), Enguang Zuo (Xinjiang University, China), Chen Chen (Xinjiang University, China), Cheng Chen (Xinjiang University, China), Jie Zhong (Xinjiang University, China), Tianle Li (Xinjiang University, China), and Xiaoyi Lv (Xinjiang University, China)</i>	

A Masked Attention Network with Query Sparsity Measurement for Time Series Anomaly Detection	2741
<i>Jie Zhong (Xinjiang University, China), Enguang Zuo (Xinjiang University, China), Chen Chen (Xinjiang University, China), Cheng Chen (Xinjiang University, China), Junyi Yan (Xinjiang University, China), Tianle Li (Xinjiang University, China), and Xiaoyi Lv (Xinjiang University, China)</i>	

Post Processing

HPCNet: A Hybrid Progressive Coupled Network for Image Deraining	2747
<i>Qiong Wang (Institute of Artificial Intelligence; Wuhan University; Hubei Key Laboratory of Multimedia and Network Communication Engineering), Kui Jiang (Huawei Technologies), Jinyi Lai (Institute of Artificial Intelligence; Wuhan University; Hubei Key Laboratory of Multimedia and Network Communication Engineering), Zheng Wang (Institute of Artificial Intelligence; Wuhan University; Hubei Key Laboratory of Multimedia and Network Communication Engineering), and Jianhui Zhang (Hangzhou Dianzi University)</i>	
Iterative Refinement Network for Hyperspectral Image Denoising	2753
<i>Fengchao Xiong (Nanjing University of Science and Technology, China), Jun Zhou (Griffith University, Australia), Zhuang Zhao (Nanjing University of Science and Technology, China), and Yuntao Qian (Zhejiang University, China)</i>	
CS-PCN: Context-Space Progressive Collaborative Network for Image Denoising	2759
<i>Yuqi Jiang (Beijing Jiaotong University, China), Chune Zhang (Beijing Jiaotong University, China), and Jiao Liu (Nankai University, China)</i>	
DDT: Dual-Branch Deformable Transformer for Image Denoising	2765
<i>Kangliang Liu (Fudan University, China), Xiangcheng Du (Fudan University; Videt Technology, China), Sijie Liu (Fudan University, China), Yingbin Zheng (Videt Technology, China), Xingjiao Wu (Fudan University, China), and Cheng Jin (Fudan University, China)</i>	
Adaptively Hashing 3DLUTs for Lightweight Real-Time Image Enhancement	2771
<i>Fengyi Zhang (Tongji University, China), Lin Zhang (Tongji University, China), Tianjun Zhang (Tongji University, China), and Dongqing Wang (Tongji University, China)</i>	
Frequency-Assisted Adaptive Sharpening Scheme Considering Bitrate and Quality Tradeoff .	2777
<i>Yingxue Pang (Bytedance Inc., China), Shijie Zhao (Bytedance Inc., China), Haiqiang Wang (Bytedance Inc., China), Gen Zhan (Bytedance Inc., China), Junlin Li (Bytedance Inc., China), and Li Zhang (Bytedance Inc., China)</i>	

Prompt & Query based Computing

Text-Guided Mask-Free Local Image Retouching	2783
<i>Zerun Liu (Zhejiang University of Technology & Zhejiang Lab, China), Fan Zhang (Zhejiang University of Technology, China), Jingxuan He (Zhejiang University, China), Jin Wang (Zhejiang Lab, China), Zhangye Wang (Zhejiang University, China), and Lechao Cheng (Zhejiang Lab, China)</i>	

P3O: Transferring Visual Representations for Reinforcement Learning via Prompting	2789
<i>Guoliang You (University of Science and Technology of China, China), Xiaomeng Chu (University of Science and Technology of China, China), Yifan Duan (University of Science and Technology of China, China), Jie Peng (University of Science and Technology of China, China), Jianmin Ji (University of Science and Technology of China, China), Yu Zhang (University of Science and Technology of China, China), and Yanyong Zhang (University of Science and Technology of China, China)</i>	
Accurate and Complete Captions for Question-Controlled Text-Aware Image Captioning	2795
<i>Yehuan Wang (Nanjing University, China), Jian Hu (Nanjing University of Science and Technology, China), and Lin Shang (Nanjing University, China)</i>	
Multi-Level Part-Aware Feature Disentangling for Text-Based Person Search	2801
<i>Yuhao Chen (Nanjing University of Information Science and Technology, China), Guoqing Zhang (Nanjing University of Information Science and Technology, China), Hongwei Zhang (Nanjing University of Information Science and Technology, China), Yuhui Zheng (Nanjing University of Information Science and Technology, China), and Weisi Lin (Nanyang Technological University, Singapore)</i>	
SPTNET: Span-Based Prompt Tuning for Video Grounding	2807
<i>Yiren Zhang (Fudan University), Yuanwu Xu (Fudan University), Mohan Chen (Fudan University), Yuejie Zhang (Fudan University), Rui Feng (Fudan University), and Shang Gao (Deakin University)</i>	
ERPG: Enhancing Entity Representations with Prompt Guidance for Complex Named Entity Recognition	2813
<i>Xingyu Zhu (Institute of Information Engineering, Chinese Academy of Science; University of Chinese Academy of Sciences, China), Feifei Dai (Institute of Information Engineering, Chinese Academy of Science, China), Xiaoyan Gu (Institute of Information Engineering, Chinese Academy of Science, China), Haihui Fan (Institute of Information Engineering, Chinese Academy of Science, China), Bo Li (Institute of Information Engineering, Chinese Academy of Science, China), and Weiping Wang (Institute of Information Engineering, Chinese Academy of Science, China)</i>	

Security, Privacy & Forensics II

Dbia: Data-Free Backdoor Attack Against Transformer Networks	2819
<i>Peizhuo Lv (SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Hualong Ma (SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Jiachen Zhou (SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Ruigang Liang (SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Kai Chen (SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Shengzhi Zhang (Boston University, USA), and Yunfei Yang (SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China)</i>	

Multimodal Fake News Detection via CLIP-Guided Learning	2825
<i>Yangming Zhou (Fudan University, China), Yuzhou Yang (Fudan University, China), Qichao Ying (Fudan University, China), Zhenxing Qian (Fudan University, China), and Xinpeng Zhang (Fudan University, China)</i>	
Downstream Task-Agnostic Transferable Attacks on Language-Image Pre-Training Models ...	2831
<i>Yiqiang Lv (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Jingjing Chen (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Zhipeng Wei (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Kai Chen (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), Zuxuan Wu (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing), and Yu-Gang Jiang (Fudan University; Shanghai Collaborative Innovation Center of Intelligent Visual Computing)</i>	
Multimodal Stacked Cross Attention Network for Fine-Grained Fake News Detection	2837
<i>Zhongqiang Huang (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), Yuxue Hu (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), Zhi Zeng (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), Xiang Li (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), and Ying Sha (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China)</i>	
Adaptive and Robust Fourier-Mellin-Based Image Watermarking for Social Networking Platforms	2843
<i>Jinghong Xia (Sichuan University, China), Hongxia Wang (Sichuan University, China), Sani M. Abdullahi (Three Gorges University, China), Heng Wang (Sichuan University, China), Fei Zhang (Sichuan University, China), and Bingling Luo (Sichuan University, China)</i>	

Adversarial Attacks on Generated Text Detectors	2849
<i>Pengcheng Su (Sun Yat-Sen University, China), Rongxin Tu (Sun Yat-Sen University, China), Hongmei Liu (Sun Yat-Sen University, China), Yue Qing (Sun Yat-Sen University, China), and Xiangui Kang (Sun Yat-Sen University, China)</i>	
Automated Software Vulnerability Detection via Curriculum Learning	2855
<i>Qianjin Du (Tsinghua University, China), Wei Kong (National Key Laboratory of Science and Technology on Information System Security, China; Tsinghua University, China), Xiaohui Kuang (National Key Laboratory of Science and Technology on Information System Security, China), Xiang Li (National Key Laboratory of Science and Technology on Information System Security, China), and Gang Zhao (National Key Laboratory of Science and Technology on Information System Security, China)</i>	
Correcting the Bias: Mitigating Multimodal Inconsistency Contrastive Learning for Multimodal Fake News Detection	2861
<i>Zhi Zeng (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), Mingmin Wu (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), Guodong Li (Xinjiang Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, China), Xiang Li (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), Zhongqiang Huang (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China), and Ying Sha (Huazhong Agricultural University, China; Key Laboratory of Smart Farming for Agricultural Animals, China; Hubei Engineering Technology Research Center of Agricultural Big Data, China; Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education, China)</i>	
Meaningful Ciphertext Image Encryption Based on Histogram Shift and ND-ICM Hyperchaos	2867
<i>Shiwei Jing (Hangzhou Dianzi University, China), Jianjun Li (Hangzhou Dianzi University, China), and Wanyong Tian (The Key laboratory of Data Link of China Electronics Technology Group Corporation, China)</i>	

Zero/Few-Shot Processing

- Hierarchical Class Level Attribute Guided Generative Meta Learning for Pest Image
Zero-Shot Learning 2873
*Shansong Wang (Shandong University of Science and Technology, China),
Qingtian Zeng (Shandong University of Science and Technology, China),
Weijian Ni (Shandong University of Science and Technology, China), Xue
Zhang (Shandong University of Science and Technology, China), and
Cheng Cheng (Shandong University of Science and Technology, China)*
- Knowledge Prompt Makes Composed Pre-Trained Models Zero-Shot News Captioner 2879
*Yanhui Wang (Tianjin University; Institute of Artificial Intelligence
Science Center, China), Ning Xu (Tianjin University, China), Hongshuo
Tian (Tianjin University, China), Bo Lv (China Electronics Technology
Corporation, China), YuLong Duan (China Electronics Technology
Corporation, China), Xuanya Li (Baidu Inc., China), and An-An Liu
(Tianjin University; Institute of Artificial Intelligence Science
Center, China)*
- Semantic-Visual Guided Transformer for Few-Shot Class-Incremental Learning 2885
*Wenhao Qiu (Huazhong University of Science and Technology, China),
Sichao Fu (Huazhong University of Science and Technology, China),
Jingyi Zhang (Huazhong University of Science and Technology, China),
Chengxiang Lei (Huazhong University of Science and Technology, China),
and Qinmu Peng (Huazhong University of Science and Technology, China)*
- Dual Episodic Sampling and Momentum Consistency Regularization for Unsupervised Few-Shot
Learning 2891
*Jiaxin Chen (Sun Yat-sen University, China), Yanxu Hu (Sun Yat-sen
University, China), Meng Shen (Sun Yat-sen University, China), and
Andy J. Ma (Sun Yat-sen University, China; Guangdong Province Key
Laboratory of Information Security Technology, China; Key Laboratory
of Machine Intelligence and Advanced Computing, Ministry of Education,
China)*
- Cross-Domain Prototype Contrastive Loss for Few-Shot 2D Image-Based 3D Model Retrieval
2897
*Yaqian Zhou (Tianjin University; Hefei Comprehensive National Science
Center, China), Yu Liu (Tianjin University, China), Dan Song (Tianjin
University, China), Jiayu Li (Tianjin University, China), Xuanya Li
(Baidu Inc., China), and An-An Liu (Tianjin University; Hefei
Comprehensive National Science Center, China)*
- Few-Shot Object Detection via Back Propagation and Dynamic Learning 2903
*Dianlong You (Yanshan University, China), Peng Wang (Yanshan
University, China), Yi Zhang (Yanshan University, China), Ling Wang
(Chongqing University of Posts and Telecommunications, China), and
Shunfu Jin (Yanshan University, China)*
- Multi-Level Correlation Network for Few-Shot Image Classification 2909
*Yunkai Dang (Tianjin University, P.R. China), Meijun Sun (Tianjin
University, P.R. China), Min Zhang (Westlake University, P.R. China),
Zhengyu Chen (Westlake University, P.R. China), Xinliang Zhang
(Tianjin University, P.R. China), Zheng Wang (Tianjin University, P.R.
China), and Donglin Wang (Westlake University, P.R. China)*

Irecut+MM: Data Generalization and Metric Improvement for Few-Shot Learning	2915
<i>Xixiang Lin (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Zhenghao Li (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), Liangchen Liu (the University of Melbourne, Australia), Jun Wu (Chongqing Jiao Tong University, China), Lijun Zhang (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China), and Xiang-Dong Zhou (Chinese Academy of Sciences; University of Chinese Academy of Sciences, China)</i>	
Counting and Locating Anything: Class-Agnostic Few-Shot Object Counting and Localization .	2921
<i>Yiwen Zhang (Sichuan University, China), Hailun Zhang (Sichuan University, China), and Qijun Zhao (Sichuan University, China)</i>	

Author Index