2026 Asia Conference on Computer Vision, Image Processing and Pattern Recognition (CVIPPR 2026)

Shanghai, China
22-24 May 2026

IEEE Catalog Number: CFP265BJ-POD
ISBN: 979-8-3315-5229-9

IEEE Advancing Technology for Humanity

Copyright © 2026, IEEE

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:	CFP265BJ-POD
ISBN (Print-On-Demand):	979-8-3315-5229-9
ISBN (Online):	979-8-3315-5228-2

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



TABLE OF CONTENTS

Sequential Curriculum Training with Elite Hard Negative Mining for Robust Face Anti-Spoofing 1
Tanaka Michelle Sandati

Dual-Branch Low-Light Image Enhancement Method Guided by Pixel-Level Estimated Illumination Maps 9
Wanwei Wang, Bingqian Zhou, Siyuan Li

MDNUNet: A U-Net-Based Multi-Scale Feature Fusion and Dynamic Weight Generation Network for Liver Tumor Segmentation 15
Yuhang Feng, Xiaosheng Wu, Yuwen Wang

Lightweight Binocular Obstacle Detection Network Based on Multi-Scale Convolutional Attention 23
Yuhao Tang, Mingbo Zhao, Hongtao Zhang

HMSFA-YOLO: Hierarchical Multi-Scale Feature Aggregation Model for Smoke and Flame Detection 29
Zheng Fang, Bo Jia, Xuesheng Li, Tingting Zhou, Bifeng Xie, Wenjing Li

Uncertainty- and Consistency-aware Stabilized Residual Fusion for Multimodal Action Recognition with TSM 35
Yuntian Wang

MA³N: Modality-Aware Adaptive Attention Network for Multi-modal Recommendation 40
Jie Li, Xiang Li, Ningzhuo Wang, Dongbing Pu, Bangzuo Zhang

Comparative Analysis of Multi-Scale Multimodal Language Models for Facial Expression Recognition 46
Meet Raval, Tejul Pandit, Dhvani Upadhyay

DIFNet: Dynamic Interactive State Space Model with Frequency Adaptive Injection for Infrared and Visible Image Fusion 53
Haozheng Zhang, Yanhong Yang, Fei Wang, Chengkun Li

Quality-Aware Two-Stage Normal Estimation for Uncalibrated Photometric Stereo 59
Zhongyi Xiong, Xu Zhang, Dawei Tu

Progressive Pansharpening Network via Wavelet-Guided Frequency Modeling and Sparse Attention 67
Wenqing Wang, Wenchao Han, Qi Lei, Liu Han

Low-Cost Weakly-Supervised Box-Prompt Segmentation for Traditional Chinese Medicine Decoction Pieces via LoRA-Tuned SAM2 75
Yiqian Yang, Jianpeng Chen, Jiaheng Li, Yujuan Xu, Yihang Mao, Shirong Guo

Image Analysis and Segmentation for Corrosion Detection and Leakage Risk Assessment in Oil and Gas Pipeline Infrastructure 81
Ning MeiJun, Lv Hui

Low-Light Infrared and Visible Image Fusion via Restoration-Guided Diffusion Prior Fusion 87
Canfeng Chen, Jing Yang

Dim and Small Target Detection Method Based on Multi-Directional Compensation and Enhancement 96
Haiqing Wang, Yiran Zhou, Weixian Qian, Minghan Wu

Detection of dim and small targets based on large-field-of-view distorted images 104
XiaoXia Du, RunFei Zhang, WeiXian Qian

Moving Platform Target Detection Based on Multi-Sensor Attitude Fusion and World Coordinate System Projection 112
Minghan Wu, Xiaogang Huang, Weixian Qian, Haiqing Wang

High-Precision Infrared Small Target Tracking via Frequency-Domain Plane Fitting on Mobile Platforms 118
Runmeng Tang, Weixian Qian

Adaptive Spectral-Spatial Knowledge Transfer for Few-shot Hyperspectral Image Classification 124
Mingshuo Ma, Huili Gong

A Comparative Study of FFT, RFFT, and DCT Convolution 130
Libin Yang, Xianglei Yu, Zhuo Qian, Haoyu Tan, Jiantao Yin, Guoyou Gan

GSPA-3D: Geometric-Semantic Prototype Augmentation for 3D FSCIL 135
Yuhang Ren, Shengwei Qin, Zhong Li

TTSP: Test-Time Subject Probing for Cross-Subject Sitting-Posture Classification from a Single Front Webcam 141
Worakan Lasudee, Taspol Sawasdee, Thepbordin Jaiinsom, Nonthapan Wongkanha, Ratchanon Mookkaew Supachok Butdeekhan, Tanakrit Saiphan, Narongdech Keeratipranon, Somnuek Songvanich

Detecting Small Traffic Lights via Mask-Guided Feature Denoising of Pyramid Representations 151
Yaqi Hu, Lan Yang, Kunfeng Wang

SFMamba3D: A Spatial-Frequency Dual-Branch Mamba Network for 3D Medical Image Segmentation 158
Chenyu Li, Xingchun Ru

Relation-Aware Representation Learning for Label-Free Unified Multi-Class Anomaly Detection 166
Hanbi Kim, Jeongwoo Heo, Jongmin Kim, Jaewon Lee, Wongyeom Seo

Semantic-Conditioned Diffusion for Task-Driven Underwater Image Enhancement 173
Yingda Yu, Jiaqi Xuan

DWM-YOLOv11: Lightweight RGB-T Pedestrian Detection with Selective Wavelet Enhancement 178
Haokun Li, Haodong Xu, Daheng Chen

Semantic-SVG: Semantic-Aware and Progressive SVG Generation 185
Zheng Shi, Yi Zhao, Wenbo Chen, Zhewen Wang, Shixing Zhang

Enhanced Graph Pattern Classification Based on Data Augmentation and Contrastive Learning 191
Zhengyuan Li

Attention-Driven Seam Optimization for Parallax-Tolerant Image Stitching 197
XinYuan Wang, Huanhuan Li

DCP-Net: A Dual-Constraint Predictive Network for Frame-Level Video Anomaly Detection 202
Guanghan Sun, Danning Wang, Heng Liu, Keshuai Yang, Wenhang Zhu

An Extension of Group Convolution in Multiple Classic Neural Networks 208
Ying Wang, Yanbo Wang

A CycleGAN-Based Method for Generating Images of Snow Accretion on Transmission Lines 214
Leshui Xiao, Yanfeng Lu, Ligang Li, Yi Li

QSM-GCN: A Multi-Scale Graph Convolutional Network for Link Prediction 219
Zhengyuan Li

HPBA: A Dynamic Hybrid Algorithm for Enhanced Exploration and Exploitation in Optimization 225
Leyi Hu, Yinghang Li, Bingmeng Hu, Jieyao Wei, Xuan Liu

AF-FLA Net: Adaptive Feature Pyramid and Focused Linear Attention for Efficient BEV Perception 232
Zeen Pan, Yuhan Dong

Video snow removal based on low-rank tensor decomposition and phase congruency 238
Pucheng Zhou, Yunpeng Zhang, Ling Ma, Jiong Zhao

Automated Diabetic Retinopathy Severity Assessment and Narrative Reporting via CNN Classification and Retrieval-Augmented Generation (RAG) 244
Afrah Alshehri, Arwa Basbrain

Multi-Class Object Detection for Safety-Critical Application Based on Camera and LiDAR Fusion 254
Tianyi Zhu, Hong Liu, Haoming Duan, Yiyang Liu, Jinjun Rao

Enhanced Local Watermarking for Multi-Watermark Coexistence via Discriminative Decoding 260
Lebin Zhou