2026 7th International Conference on Computer Vision, Image and Deep Learning (CVIDL 2026)

Changsha, China
22-24 May 2026

Pages 1–742

IEEE Catalog Number: CFP26Y58-POD
ISBN: 979-8-3195-0693-1

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:	CFP26Y58-POD
ISBN (Print-On-Demand):	979-8-3195-0693-1
ISBN (Online):	979-8-3195-4465-0

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

TABLE OF CONTENTS

Semantic-Conditioned Prior Learning for Action Anticipation 1
Peng Shen, Shenghao Li, Jinkang Luo

Pedestrian Detection in Complex Scenes Using an Improved YOLOV11 Model 5
QiChao Hong, YinXia Xu

Power Spectrum Adaptive Line Enhancement Algorithm Based on Weighted Signal-to-Noise Ratio Criterion 11
Zongtang Zhang, Xinyuan Han, Yapin Yin

Design and Analysis of Multi-Source FusionSystem for Helicopter Obstacle Detection and Identification 17
Xiaolong Pei, Ling Li, Lingxiao Wang, Jing Yang, Hongyu Xiao

Construction of a Multimodal Tree-Drawing Dataset with Multi-Dimensional Annotation for Mental Health Assessment 23
Yuze She, Hefeng Wang

Cross-View Token-Enhanced Neural Radiance Fields for Sparse-View Novel View Synthesis 29
Kehan Zhou

NUR: 3D Geometric Priors for Rethinking 2D Annotation in Multimodal Detection 35
Kangjian Rao, Liu Yu, Xiaolei Wang

Research on Basketball Recognition of Humanoid Shooting Robot Based on Improved YOLOv12 41
Shiji Yang, Hui Zhang, Fan Guo, Pengfei Ma, Zheng Yuan, Qian Li

TSF-YOLO for Tiny Fallen-Person Detection in Cluttered Water-Surface Scenes 47
Hao Sun, Haiyan Zeng, Xiansong Zeng, Wanyu Liu, Ping Zhang, Xinguo Zhou, Min Yang, Tao Wang

TWD-KD: Temperature-Guided Wavelet Decomposition Knowledge Distillation 53
Zijun Bai, Xuemei Lei

Multi-Scale Feature Fusion and Attention-Enhanced Image Deraining Network for Computer Vision 59
Xinyi Zhu, Jun Wu, Han Wang

S2MNET: a Contrastive Learning Framework with Geometric Semantics for 2D Map Geo-Localization 64
Yongjian Tan, Zhinong Zhong, Qingren Jia, Xiaohan Gao

Learnable Trusted Sparse Attention (LTSA) for UHD Image Restoration 70
Yingda Yu, Jiaqi Xuan, Haodong Zhang, Shuhui Shi, Xuanyu Teng, Yulin Tang

Incremental Learning Method for Intelligent Identification of Asphalt Pavement Distress Based on YOLOv11 76
Li Wang

A Frequency-Guided Edge-Texture Dual-Assisted Camouflaged Object Detection Method 81
Tian Tian, Liping Liu, Xuehong Sun, Zhenkun Ma

A Two-Dimensional Budget Learning and Fast Deployment Method for Resource-Constrained Visual Inference 85
Kaishen Wu, Mohan Li, Heng Zhao

An Improved YOLOv8n-Based Method for UAV Small Object Detection 90
Ruochen Cheng

SAS-YOLO: A Scale-Aware and Selective Object Detector for Complex Autonomous Driving Scenes 94
Xin Wang, Jiarui Hu, Pengpeng Sun

YOLOv8n-Based Lightweight Target Detection for Silkworm Age Recognition in Complex Sericulture Scenarios 98
Dexin Meng, Xin Dong, Shicheng Mo, Chunming Wen

Robust Object Detection via GAN-Based Data Augmentation 103
Enze Zhao

An Improved Method for Target Detection in Coal Mine Drilling Sites Under Complex Environments 108
Qiang Li, Yanyi Zhou, Boxuan Zhang

An Improved YOLO Method for Detecting Tiny Defects in Strip Steel Based on High-Resolution Feature Fusion 113
Jianke Du, Junkai Ma, Yan Dong, Jun Ma

Research on Plastic Particle Defect Detection Method Based on Improved RT-DETR 117
Meiling Ding, Wei Li

Multi-Scale Attention Enhancement for Tiny Defect Generation in Industrial Anomaly Detection 121
Xiaoling Bai, Kim Young Chul

Human-Machine Collaborative Adaptive Labeling for Concrete Cracks in Zero-Shot Scenarios 126
Jiewen Yang, Zhankui Huo, Haibo Wu, Jianming Wang, Man Ni, Yisheng Cui

LTS-YOLO: A YOLOv11-Based Method for Visual Scale-Based Lunar Crater Detection in Earth-Moon Transportation 131
Chengyu Zhang, Yilin Wang, Qiwen Xue

YOLO-KFI: A Lightweight Model for Sunburn Detection in Densely Distributed Kiwifruit 135
Baohua Li, Zhen Niu, Yunwang Su, Suguang Xu, Geni Cao, Qiheng Ren, Dong Han, Dongyan Zhang

Improved Photovoltaic Panel Defect Detection Algorithm Based on YOLOv11 143
Junjie Feng, Zuxing Chen

NoduleNet: A High-Precision Lung Nodule Detection Algorithm Based on Deep Learning 149
Xinyuan Liu

Foreground-Background Contrast-Driven Infrared Small Target Detection Network 153
Liubing Yang, Xiaoyu Li, Mengen Shao

Intelligent Detection of Moldy Tobacco Based on Multi-Color Illumination and Image Recognition 158
Ting Fang, Yuejun Ma, Wen Tan, Cunrui Pang, ZhiChao Tang

VISION-I: Real-Time Object Detection and Distance Estimation With Adaptive Multimodal Feedback for Visually Impaired Navigation 163
Xin Yan

An Improved YOLOv8-Based Method for Ornamental Fish Detection 170
Ziwen Zhang, Changming Zhu

Tiny Impurity Detection in Baijiu Bottles using YOLOv8s-P2 and Dark-Field Imaging 176
Chunyu Rong, Huali Dong, Qiangjun Tan

An Infrared Small Target Detection Method for Remote Sensing Images Based on Improved RT-DETR 181
Liangji Wang, Hui Zhang, Kehao Wang, Chenxiao Li, Hanshuo Huo, Xiaofeng Zhao

Automatic Reading System for Industrial Environment Pressure Gauge Using Deep Learning and Geometric Transformation 185
Qihang Su

Research on Target Detection Algorithms for Anti-Drone Applications 189
Yuhang Wu, Zhiduan Cai

Pseudo-Label Enhanced Semi-Supervised Detection Method for Textured Surface Defects Based on Improved Soft Teacher 194
Kedi Liu

Lightweight Improved YOLO26 for UAV-Based Face Detection and Recognition 200
Xueda Jia, Haiyan Cao, Haotian Zhu, Jun Ma

Quantitative Detection of Overhead Contact System Insulator Damage Defects Based on YOLOv8-OVC 205
Donghu Peng, Weiguo Li, Rui Shi, Wenzheng Yang, Mingyang Jiang, Wenming Hu

Improved YOLOv8-Based Small Object Intelligent Detection Method for Sample Morphology 211
Ningning Zhu, Yongjiang Zhang, Jian Cui, Chunxia Li, Qi Zhao, Pingli Mu, Yinlong Feng

YOLO-SCCD: An Efficient YOLO11-Based Method for Small Object Detection in UAV Aerial Imagery 215
Linhao Liu, Xiaoming Su, Yongwang Jia

Research Review on Deep Learning Based Target Detection Methods 221
Kangyue Hu

Topic-Aware Action Tube Detection 226
LiJun Ren, Yihua Tan

Research on Surface Defect Detection of Wind Turbine Blades Based on Deep Learning 231
Yingli Wang, Peng Jin, Sike Tian, Zeyu Cao

MeRD: Memory-Enhanced Reverse Distillation for Industrial Image Anomaly Detection 236
Junrui Zhang, Shifeng Li

Research on Automatic Bridge Crack Detection Method Based on Machine Vision 240
Bo Tang

Research on Intelligent Defect Detection of Wind Turbine Blades Based on Computer Vision and Improved RT-DETR 246
Weicheng Wang, Lixin He, Jie Li, Hechun Guo

Investigation of Deep Learning Image Processing Algorithms for Real-Time Video Analytics 251
Zihua Liu

Research on Improved Dual Channel Generative Anomaly Detection Based on Graph Neural Network 258
Siyuan Ma, Fang Wang, Runli Qu, Deyin Kong

Lightweight Target Detection Method for UAVs Adapted to Edge Computing Deployment 264
Junjie Li, Yisheng Wang, Jie Xiao, Bo Zhang

An Improved YOLO26 Algorithm for Small-Target Detection Based on Image Enhancement 268
Xinyu Tan, JingXuan Huang, Haoran Lin, Hongxue WU

Improved Lightweight Routed Attention and Occlusion-Aware Training for X-Ray Contraband Detection 274
Yixuan Zheng, JingAo Shen, Lv Zhou, Tingting Li, Jiandong Huang, Jiaxin Ye

Small Target Detection Algorithm for Aerial Lmages based on YOLOv8n 278
Youjuan Lv, Weiguang Dong

Spatial Semantic Enhancement Guided Learning and Consistency Constraints for Object Detection 283
Wangwang Cao, Xin Liu, Yibulayi Kuwan, Gulinazi Ailimujiang

Research on Image Object Detection and Intelligent Visual Recognition Based on Deep Learning 290
Getong Shi

Research on Industrial Heat Source Recognition Methods Integrating Multi-Scale Spatial Analysis and Deep Learning 295
Yu Liu, Ping Ai, Yongxin Wang, Lijing Xu

SA-YOLOv11: Scale-Aware Enhanced YOLOv11 for Tiny Object Detection in VisDrone Dataset 300
Xinyuan Yan

YOLOV11-HCP: a Lightweight Plant Leaf Disease Detection Method Based on YOLOv11 305
Junwei Zhang

Research on Object Recognition in Complex Backgrounds Based on the Improved YOLOv8n 310
Hu Da

Spatial-Frequency Wavelet Analysis for Synthetic Face Image Detection with Controlled Ablations 318
Haoxing Zhou

An Empirical Study of a Vision Foundation Model for Remote Sensing Wind Turbine Detection 322
Wei Zhang, Bin Li, Guoqing Yao, Lijuan Zheng, Tao Zhang, Guanghui Wang

YOLO-EA: A Railway Clearance Foreign Object Intrusion Detection Method Based on YOLO11n 327
Zhibin Zhang, Wei Wu, Faxu Deng, Qingfeng Gao, Zihan Wang

Design and Implementation of an Embedded Vision-Based Liner Board Detection System for Blind Areas in the Tobacco Bale Unpacking Process 331
Wentao Xu, Xiang Wang, Xiaoshan Si, Zhixin Zhao, Hailong Liu

MesoFuseNet: A Two-Stream Fusion Network for Multi-Granularity Forgery Perception in Images 339
Zihan Ye, Ruifang Zeng, Ying Xi, Guiru Li, Zikang Wang, Mengqin Shi, Cong Li

Lightweight Small Object Detection Algorithm for UAV Aerial Photography Based on RT-DETR 347
Liyuan Qi

Edge Deployment-Friendly Real-Time Detection Algorithm for UAV Aerial Photography 351
Liyuan Qi

RailSC-YOLO: Lightweight Surface Defect Detection for Railway Metal Components 355
Demin Fu, Jianhui Yi, DeShui Han, Xuemei Jia, ChunSheng Yuan

A Generated Image Detection Method Based on Data Augmentation and a Global–Local Dual Branch Architecture 363
Xiao Zeng, Xinyu Wang, XiaoMing Guo

Benchmarking Consecutive YOLO Generations for Photovoltaic Electroluminescence Defect Detection on the PVEL-AD Dataset 367
Zihao Hu, Shuqing Liu, Huaping Zhu

Group Intention Recognition via Memory-Extended Activity Forecasting 372
Longteng Kong, Yuhang Rao, Xin Fan, Chuanwen Luo

Hyperspectral Anomaly Detection: A Benchmark 376
Xianfei Qiu, Qiulin Yu, Shaocong Liu, Zhen Li, Sida Tang, Puhong Duan, Xudong Kang

Direction-Context and Quality-Aware YOLOv8 for Steel Surface Defect Detection 380
Huihui Chen

Research on Visual-Based Few-Shot Learning Methods for Industrial Product Appearance Defect Detection 385
Lin Su, Lelun Xu, Jianzhong Wang, Ruosong Liu, Jianjun Yi, Hu Hui

AGC-Net: Adaptive Guided-Context Network for Infrared Maritime Small Target Detection 390
Jiaqi Lin, Xinghua Xiao, Xiaomao Li

RAM: Retinex-Guided Affine Modulator for Low-Light Object Detection 397
Yaoyao Song, Yajun Liu, Zenghui Zhang

Pattern-Enhanced RT-DETR for Multi-Class Battery Detection 403
Xu Zhong, Enyuan Hu

Hierarchical Bidirectional Feature Pyramid Fusion for Small Target Detection in Unmanned Aerial Vehicles 408
Yanshan Zhang, Chenyang Gou, Yuanzhang Fan

A Polar Coordinate Lane Detection Algorithm Fusing Geometric and Structural Features 413
Shasha Tian, Jiaxing Hou, Zhilin Wang, Chongwei Ruan, Mian Wu

DFAST-DETR: An End-to-End Small Object Detection Model for Drone Images Based on Dynamic Filtering and Adaptive Sparse Transformer 421
Feifan Zhao, Guangxi Chen

Road Defect Detection with Transformer Using Contrastive Learning and Adaptive Computing 425
Shichun Wang

A Context-Based Keyframe Extraction Method for Action Videos 429
Ruimin Wang, Qin Wei

Facial Action Unit Detection via Orthogonal Global-Local Disentanglement 434
Haoyu Li, Xiongtao Zhang

DMF-DETR: A Density-Guided Multimodal Fusion Framework for RGB-IR Object Detection 439
Baolin Ma, Jun Zhu, Xinjun Liu

Long-Distance Water Surface Garbage Detection Method Based on Improved RT-DETR 444
Chenglong Lu, Xiangguo Sun

Spatially Adaptive Expert Routing and Multi-Scale Semantic Alignment for Cross-View Geo-Localization 450
Weishan Zhao, Jiajia Huang

Comprehensive Performance Evaluation of Marine Organism Object Detection Models Based on Combined Subjective-Objective Weighting 455
Xicheng Huang

Robust Unified Anomaly Detection Via Spatially Gated Semantic Priors 459
Huiyuan Shi

AERO-DETR: Alignment and Enhancement via Recalibration for Object Detection in Low-SNR Aerial Images 466
Xinyan Mei, Qiyuan Liu, Zhengeng Yang

HCD-YOLO: A Conservation-Aware Detection Framework for Multi-Material Crack Diagnosis in Historic Building Restoration 470
Yujia Li, Zhenyuan Wang, Jiachun Liu, Leihua Xia

DynamiFormer: A Dynamics-Aware Transformer for End-to-End UAV Tracking 477
Wanyi Li

An Improved RT-DETR-Based Conveyor Belt Crack Detection Method with Transfer Learning 482
Wenyao Yan, Jing Luo, Abudhahir Buhari

Lightweight Conveyor-Belt Foreign Object Detection Algorithm Based on YOLOV11 with Spatial-Channel Dual-Branch Attention 490
Wenyao Yan, Jing Luo, Abudhahir Buhari

SKC2PSA-SimLite: A Lightweight Tail Semantic Enhancement for YOLOv11n in Shore-Based Water-Surface Floating Object Detection 496
Fangyu Chang, Songling Fu, Xuan Yuan, Ying Yang

When Physics-Informed Priors Help and Hurt in Highway MOT Under Occlusion: a Controlled Diagnostic Study 500
Mingkang Liu, Xin Zhang, Yunyong Li, Dong Yang

SELayerV2: Grouped Multi-Branch Channel Attention for Fine-Grained Fire Detection 506
Chenglong Wang

LSCD-LQE-YOLO: Channel-Attended Features and Quality-Weighted Localization for Fine-Grained Oriented Aircraft Detection in Optical Remote Sensing Imagery 514
Gezhi Yuan, Mingyong Jiang

A High-Precision YOLOv8 Detection Model for Detecting Microscopic Surface Defects on PCBs 520
YinPing Liu

Step Wear Analysis Based on Improved Computer Vision and Mathematical Modeling for Infrastructure Condition Assessment 524
Minghong Shen, Yufei Cui, Jingjiang Qiu

Underwater Object Detection via Fusing Spatial Depthwise Convolution and Multi-Kernel Attention 528
Junkai Ma, Jianke Du, Chengjie Feng, Jun Ma

APS-YOLO for UAV-Based PPE Small Object Detection in Complex Industrial Scenes 532
Lin Miao, Haiyan Zeng, Min Yang, Zhuofan Yu, Chaofeng Qu, Tao Wang

A Comparative Study of YOLO and DEIM Bird Detection Algorithms for Airport Bird Deterrence Scenarios 538
Hongyu Xiao, Bin Tao, Xue Gao, Shixia Sun, Han Zhang

Multi-Task Deep Learning Framework for Simultaneous Recognition of Tomato Maturity and Biotic Adversities in Greenhouse Robot Vision Systems 542
Zihan Meng, Huacai Chen

Improved YOLOv11n-based Instance Segmentation for Tiny Objects in UAV Imagery 548
Zefeng Zhao, Tao Wu, Jinfei Zhang, Yunyun Dong

An Auxiliary Diagnosis System for Ocular Diseases Based on Dual-Fundus Feature Fusion Architecture and Genetic Algorithm Optimization 557
Shengjie Liu, Yu Zeng, Zekai Gan, Caihong Xiang, Sihan Jiang, Wanjia Chen

Scene Text Image Super-Resolution Based on Semantic Enhancement and Cross-Modal Information 562
Xiaomei Du, Hai Nan

Cognitive State Recognition Model and Adaptive Control Mechanism of Elderly-Friendly Immersive Interactive System 566
Mengge Chang, Yingxin Liu, Tengfei Mo

A Photo-to-Cartoon Conversion Method Based on Edge-Aware Lightweight Networks 573
Pengfei Jia, Mingxia Li, Na Liu

Cross-Task Consistent Multi-Task Learning for Fine-Grained Crop Leaf Disease Diagnosis and Severity Grading 577
Jin Li, Yuqiu Zhang, Shucen Guo

A Convolutional Neural Network-Based Garbage Classification Method with Overfitting Prevention 583
Ruixiang Guan

Multi-Scale Feature Fusion Diffusion Network for High-Quality Image Super-Resolution Reconstruction 587
Wei Fu, Ze Zhang

A Semi-Supervised Medical Image Segmentation Framework Based on Masked Image Modeling and Sequential SAM Feature Fusion 591
Yinjun Liu, Yueyan Bian, Mingrui Zhuang, Jinying Tong, Xiuqin Jia, Hongkai Wang

Traditional Residential Architecture Style Recognition and Digital Extraction of Contextual Features Based on Deep Learning 596
Ronger Huang

A Cost-Oriented Study of Single-Model and Ensemble Learning for Five-Class Surface Defect Classification in Automotive Production Lines 601
Jingyang Chen

Research on High-Resolution Reconstruction of Diffusion Magnetic Resonance Images Based on Graph-Structured Autoencoders 608
Jingyi Zhang, Jiquan Ma

Fabric Defect Classification Based on Lightweight CNN Under Class Imbalance 612
Motian Zhang

An Optimized ResNet18-Based Approach for Cassava Leaf Disease Recognition 617
Zijun Weng

Bridge Crack Segmentation Network Based on Cascaded Feature Fusion and Attention Mechanism 622
Yang Wang, Wei-Yu Wang, Qi-Yang Liu, Hui Du, Meng Zhu

Adaptive Dynamic Loss Threshold-Based Approach for Metallographic Semi-Supervised Semantic Segmentation 627
Tao Li, Huakang Bian

Active Evo-Aug: Feedback-Driven and Boundary-Aware Data Augmentation for Multi-Label Emotion Classification 633
QiaoQiao Yang, JiangHua Gao, Zhenyu Hao

Multi-Scale Prototype Neighbor Network for Few-Shot Food Image Classification 637
Xue Wang, Zeyu Yang, Jiyong Shi

Bamboo Slips Scribe Verification Based on Multi-Scale Dynamic Routing and Dual-Sparse Channel Attention 641
Qihang Cui, Mingjun Li, Yangchen Guo, Yanbo Zhang

Zero-Shot, Linear Probe, or Full Fine-Tune? Benchmarking Vision Foundation Models on Sentinel-2 Land-Use Classification 645
Haijia Zhong, Silalai Phadboke, Fucai Xie, Ling Zhu, Xianpeng Wang, Pengfei Zhang

A Garbage Classification Method Based on Feature Extraction and SVM Clustering 650
Xiangshi Huang

Attention-Guided Hybrid Gaussian Adversarial Testing with AHG-Adv 654
Dekang Ma, Ronghao Zhang, Zhigang Zhou

FJID: Fast Joint Intra-Channel and Inter-Channel Variation Demosaicing for Bayer Pattern 660
Zheng Yang, Wen Liu, Huichao Liu

Intrinsic Geometric Prior Bias in Edge Detection 665
Haokai Gao, Junwei Tian, Xingang Wang, Yupeng Feng, Shifan Yu

Few-Shot Cross-View Gait Recognition Based on Meta-Learning and Generative Adversarial Networks 669
Lingui Xiao, Tianfang Xu, Hongru Wang, Qiang Guo

SER: Semantic-Enhanced Re-Ranking for Cross-View Geo-Localization 673
Zhuangzhuang Han, Xudong Kang

Advancing Acute Lymphoblastic Leukemia Classification Through Lightweight Vision Models and Mamba3 Architecture for Efficient Computational Pathology 679
Runyang Zhang

AIGC Style Transfer and High-Fidelity Image Generation Method for Zhuang Brocade Patterns Based on Controlnet and Style-Decoupled Loss Function 687
Xiaozhen Wang, Yingying Guan, Dangqing Xu, Jiahui Guo

TCPA-MedMTL: A Task-Conditioned Prototype Adapter for Multi-Task Medical Image Classification 691
Jiarou Lu, Haibiao Liu, Wanyi Li, Qian Zhang

A Hybrid Transformer Network with Pyramid Attention and Content-Aware Upsampling for Breast Lesion Segmentation 698
Xinquan Chen

Learnable Evidence Fusion With Teacher-Guided Distillation for Urban-Scene Semantic Segmentation 703
Zihao Chen, Lei Tian

A Multi-Source Self-Supervised Feature Fusion Method for Enhancing Unsupervised Image Clustering 708
Jinzhong Ma, Yanshan Tian

Prompt-Aware Progressive Decoder with Dual Semantic Branches for Dunhuang Mural Damage Segmentation 713
Chenkai Ding

An Adaptive Enhancement Method for White Blood Cell Micrographs Based on the Lab Colour Space 717
Zhitong Lv, Jie Li, Xiaoli Wang

Blind Deconvolution via Game-Theoretic Framework: Adaptive TV-NL and Fractional-Order Derivatives 721
Hong Liu, Lei Chen, JinLi Li, ZongXuan Wei

Research on Forward-Looking Sonar Image Feature Enhancement Method Based on SSLSonarNet 727
Hao Wang

A Comparative Analysis of Convolutional Neural Networks and Spiking Neural Networks for MNIST Handwritten Digit Classification 732
Jinchao Xiang

A Color Image Steganalysis Network Based on Grouped SRM and Channel Attention (DCA-Net) 737
WenShuai Liu, JunPeng Qi, XiaoJie Xin, Min Zhan

Design of a Raspberry Pi-Based Carbonated Beverage Bottle Recognition System 743
Yu Tang, Yan Yang, Zhen Zhou

Clustering-Based One-Class Classification of Remote Sensing Imagery 750
Shukui Bo, Yiran Lu, Zheng Wei, Xiaoyu Hao, Yakuan Wang, Yongju Jing

Hybrid Hypergraph and HOG-Gated Network for Hyperspectral Image Classification 758
Jiaqi Lv, Jiayi Shi, Boyuan Zhang, Xuewei Liu, Fan Zhang, Xiaofeng Zhao

Residual Spatial Gating with AU-Assisted Supervision for Facial Expression Recognition 762
Zhongyu Wang, Lucai Wang

Dual-Channel Cloud Particle Image Classification Based on Transfer Learning and U-NET 767
Guodong Jing, Xiudian Zhang, Guiyu Sun, Bao Lei, Baopeng Zhou, Rong Zhang

Low-Light Dynamic Decoupling Enhancement Network Based on Feedback-Guided Dynamic Anisotropic Convolution 774
Shihao Liang, Kaiping Sun, Jianping Xing, Chuangxin Han, Zijie Tang, Tong Liu

Low-Light Image Decoupling Feedback Enhancement Network Based on Channel-Independent Windmill Convolution 778
Shihao Liang, Kaiping Sun, Jianping Xing, Tianyi Xi, Zixuan Liang, Chenxi Li

Underwater Image Enhancement Model Based on Multi-Module Cascaded Optimization 782
Zizheng Zhang, Haixia Wang, MengAn Shi, Zhenyu Lu

ASRM-Net: Multi-Scale Feature Enhancement with Adaptive Spatial Recalibration for Medical Image Segmentation 786
Xinyang Chen

Texture-Prior-Attentive ROI Extraction for Wrist Vein Recognition 793
Hongkun Han, Zhiyuan Wang

DRS-Net: Spatially Calibrated Adaptive Normalization and Hierarchical Spatial Dropout for Retinal Vessel Segmentation 798
Peiying Chen, Huachao Chen, Minghui Zhuo

SGA-Mamba: A Spectral-Guided Asymmetric State Space Model for Hyperspectral Image Classification 803
Chenxiang Xiong, Bosheng Wei

Ground-to-Satellite Geo-Localization with Intermediate Drone-View Feature Fusion 808
Xianghui Dai, Shutao Li, Puhong Duan, Xudong Kang

A Spatial-Frequency Enhanced U-Net With Residual Gated Fusion for Hyperspectral Image Denoising 814
Shiyu Liu

Structure Tensor and Retinex Model Based Enhancement for Microscopic Images of Micro-Holes 818
Ziwei Tong, Zeyu Yang, Weishi Li

Degradation State-Guided Mixture-of-Experts Network for Marine Snow Removal 829
Haoru Zhao

BiomedQSD: Enhancing Few-Shot Adaptation in Biomedical Images Classification 833
Weihong Zhu, Rongsheng Xie

WAFD: Joint Structural and High-Frequency Texture Modeling for Quality Assessment of Generated Lace Images 838
Shuyun He, Shangping Zhong, Yunbing Wu, Kaizhi Chen

E2SMoE for Breast Ultrasound Image Classification 842
Jiayuan Li

STAR-UNet: Multi-Branch Parallel Fusion and Global Sequential Spatial Adaptive Learning for Retinal Vessel Segmentation 846
Haibo Xu

SegLite Decoder (SLDC): A Lightweight Decoder With Compressed Channel Attention for Real-Time Egocentric Hand Segmentation 851
Xiaoxiao Zhang, Xin He, Kaile Wu, Yu Guo

Transfer Learning for Fine-Grained Image Classification Under Controlled Fine-Tuning Settings 856
Siying Li

C2R: A Coarse-to-Refine Centerline-Aware Multi-Task Learning Framework for Coronary Artery Segmentation in CCTA 863
Lele Yue, Shuang Song, Lixing Tan, Ping Wang, Zheng Zhang, Xin Zhang, Ruoxiu Xiao

Image Foreground Segmentation Algorithm Based on Hierarchical Coding U-Net 867
Shenghui Wang, Jian Chen

CAE²-CRC: Risk-Aware Early Exit with Calibrated Reliability Control 872
Yixuan Li, Arlo Octavia

HAM-UNet: A Lightweight Hybrid Attention Network with Boundary-Aware Priors for Fetal Head Ultrasound Segmentation 876
Jialu Lu

Feature Decorrelation and Source-Aware Detail Compensation for Lightweight SwinIR Image Super-Resolution 883
Liyi Wu

Transformer-Based Content-Style Fusion for Image Style Transfer 887
Zhiheng Cui, Haoran Deng, Xiaoming Guo

Research on Lightweight Semantic Segmentation Networks for Edge Deployment 891
Xixi Jia

MF-RWKV: A Multi-Scale Feature Fusion Network Based on RWKV for Brain Tumor Image Segmentation 895
Weiman Xiao, Gang Wang, Tao Zhang, Yugui Zhang

DM-FCMNet: Direction-Aware Morphology-Guided Fuzzy Clustering Network for Robust Magnetic Tile Surface Defect Segmentation 899
Haoran Tang

Visually Enhanced Fundus Image Classification Network Fusing Contour Cues and BiLSTM 906
Heng Zhang, Yijie Xue, Yanli Liu

FDP-PVT: Frequency-Decoupled Pyramid Vision Transformer for Polyp Segmentation 910
Siqi Huang

MBDNet: Multi-Branch Dynamic Routing Denoising Network for Efficient Image Restoration 916
Xueqiang Li, Siyu Li, Yuhang Lu, Yongjiang Xue

Physics-Aware Generative Adversarial Network for SAR Image Generation from Optical Imagery 922
Xiao Liu, HuaMin Tao, Shanzhu Xiao, Qiuqun Deng

Replacing Layer Normalization with Dynamic Tanh in Segment Anything Model to Reduce Computational Resource 928
Zhenhong Guo

Surgical-AGPSAM2 for Auto-Generated Prompts in Surgical Instrument Segmentation 933
Chengbo Duan, Shuang Song, Xin Zhang, Tongyu Jia, Huaikang Li, Ruoxiu Xiao

Emotional Mixture of Experts for Dynamic Facial Expression Recognition 937
Yuan Jin, Guoyong Cai

BronzeNet: A Multi-Task Network for Chinese Bronze Artifact Classification and Dating 942
Bujin Zhang, Siyu Ye, Zhongling Luo, Lei Huang

Wavelet Domain Sparse Transformer for Single Image Deraining 947
Xiaoran Weng

DPFE-Net: Dual-Pathway Feature Extraction with Class-Balanced Focal Loss for Alzheimer's Disease Staging on Brain MRI 953
Yusheng Shen, Zhujiang Li, Pingping Lu, Mingwang Kong

Cylinder Head Defect Segmentation Algorithm Based on Deeplabv3+ 957
BingZhuo Li, Qingbo Ji, Heyang Sun, Zhihao Li

AGTrack: A 3D Multi-Object Tracking Framework With Adaptive Kalman Filtering and Geometric Similarity Perception 962
Biao Zhang, Weihan Wang, Kai Yu, Chengding Bi, Yupeng Sun, Ye Lin, Jiwei Jia

Adapting Monocular 3D Object Detectors to Fisheye Images 971
Yue Cao, Ya-Li Hou

FXD-3D: Geometry- and Texture-Preserving Image- to-3D for Fengxiang Dyeing 976
Size Wang, Zhiqi Shen, Yudian Wang, Yurong He, Haotian Liu, Zhu Zhu, Ziliang Wang, Honggang Qi

Polarization-Guided Multimodal Semantic Segmentation with a Shared Encoder 984
Zixuan Ge, Maichen Xiong, Zhixin Dong, Chengdong Wu, Luxiu Li

MGP-LoRA: Progressive Multi-Granularity Prototype Cascade for Few-Shot Vision-Language Model Adaptation 989
Zelong Huo, Guodong Zhao

A Lightweight Network for Event-Camera-Based Hand Pose Estimation 994
Yu Guo, Xin He, Xiaoxiao Zhang, Sishun Song

Asynchronous Speculative Decoding for Low-Latency Vision-Language Navigation 999
Lin Feng, Wei Luo

Multi-Task Cascaded 3D-UNet and Multi-View Graph Neural Networks for Pulmonary Nodule Analysis on LIDC-IDRI CT Scans 1005
Yusen Fu

A Multi-Scale Attention Feature Fusion-Based Method for OCT-to-OCTA Image Generation 1012
Shenao Zhang, Jinke Wang, Tianfu Li

A Comparative Study on Sentiment Analysis of E-Commerce Reviews Based on Attention-Enhanced Single Model and Hybrid Model 1016
Chao Deng, Jiao Peng, SiYi Xie

A Dual-Branch Multimodal Feature Fusion Network Based on Cross-Attention for Bearing Fault Diagnosis 1020
Ruyu Liu, Xuehong Sun, Liping Liu, Siqi Luo

Multi-Source Image Processing Based Intelligent Perception Technology for Seagrass Restoration 1024
Yu Lu, Jie Ai, Dalei Song, Yijia Sun, Wenjing Yang, Huanzheng Hu, Xianhao Pan, Jiangzihui Yuan

A Study on Multi-Step Influenza Forecasting Using a TCN-XGBoost-Prophet Ensemble Model 1028
Qiwen Yuan, Jia Mu, Zhen Zhao

Research on Bp Neural Network Image Compression Algorithm Based on Stochastic Gradient Descent 1032
Li Wang, Meng Yan, Yilei Chen, Dan Feng, Yuming Niu

Research on Speech Signal Separation and Restoration Based on Fast Independent Component Analysis 1039
Wei Wang, Na Li, Meng Yan, Li Wang, En Kang

A Review of Data-Driven Co-Speech Gesture Generation: Methods, Datasets, and Challenges 1047
Xinzhi Cai, Yunqi Guo, Xinyue Xu, Di Chen

An Improved SuperPoint-Based Image Stitching Method for Shipborne Navigation Images 1052
Tao Hu, Feixiang Zhu, Haoyu Wang

A Study on the Performance Complementarity of RGB-D Visual SLAM and LiDAR SLAM in Indoor Navigation Tasks 1057
Weiming Zhang, Shuohuan Wei

An Adaptive System for Two-Lane Overtaking Warning Integrating Visual Perception and Field-of-View Sharing 1064
Yuansheng Shao, Ruihao Wang, Xiang Liu, Han Wang

Addressing Data Scarcity and Domain Shift in Computational Pathology: From Pixel Harmonization to Foundation Models 1070
Yixuan Guo

Frozen Vision-Language Models with Adaptive Voting and Verification for Visual Illusion Reasoning 1079
Zhuofan Zhong, Yehan Xia, Yiqiang Wan, Yudong Fang

A Study on Laser SLAM Technology and Interactive Guided Vehicle Systems in Complex Campus Environments 1085
Zuowei Lin, Fan Gong, Yicheng Fan, Jincheng Guo

GHA-DHNN: Achieving EEG Emotion Recognition by Integrating Gated Attention and Dynamic Hypergraph Neural Networks 1091
Chunjiao Wu, Xiaoyu Fan, Xin Liu, Axu Hu

Research on Visual Perception and Camera Calibration Based Tracking and Localization for XR 1098
Jia Wang, Si Wei Peng

A Transformer-Based Siamese Network with Semantic Enhancement for RGBT Tracking 1102
Ninghao Ma

HARPFormer: Adaptive Position-Aware and Heterogeneous-Response Modeling for Traffic Flow Forecasting 1106
Xin Zhang, Qun Fang, Xiaofang He, Jin Zhi

A Multimodal Image Segmentation Framework for Embodied Industrial Scenarios Based on SAM2-Assisted Annotation and SAM3 Prompting 1111
Yongjian Chen, Jiwei Yu, Ting Wang, Hongxiao Gan, Kecheng Zhou, Rui Wang, Huaizhen Wang Shuzhi Gao, Jun Liu

FlashSandwich: A Hardware-Aware Hybrid SSM-Transformer Architecture with Optimized KV Cache Management for Efficient LLM Inference 1118
Zhizhong Wu

Footprint-Based Height Prediction Using U-Net and Morphological Features 1125
Dongliang Yang, Changjiang Song, Xiaodan Cong

Multidimensional Time-Series Data Imputation and Forecasting Based on Non-Negative Tensor Decomposition and the Holt-Winters Model 1129
Haoyu Wei

Cross-Modal Feature Enhancement Based DIBR 3D Image Watermarking 1133
Jinxian Zhang, Zhang Wang, Qingmo Chen

Robust Localization Method for Unmanned Surface Vehicle Binocular Vision 1138
Mingchun Zhou

A Joint Feature Space Based Visualization Platform for Amplitude Estimation Results of Mixed Signal 1145
Nanxi Zhao, Yupeng Wang, Jinru Fan

Optimization of High-Pressure Water Jet Cutting Process Parameters for HTPB Propellant Based on Kriging–Bayesian Optimization 1150
Zhen Zhang, Dayong Jiang

Reliable and Efficient Decision-Making for Large Language Model Agents Through Uncertainty-Guided Adaptive Reasoning 1157
Tian Liao, Xiaotian Fang, Yihui Feng, Shuyuan Wang

Research on Intelligent Evaluation Model Based on L-M Algorithm Optimized BP Neural Network 1162
Jia Liu, Wenjing Li, Jinfeng Chen, Jingqian Wang

A COLMAP-Constrained 3D Gaussian Splatting Framework for UAV Multi-View Real-Scene Reconstruction 1167
Huanyu Hou

The Impact of Classifier Structure and Data Imbalance on Conditional Diffusion-Based Procedural Game Generation 1172
Chengrui Zhi

Adaptive Multi-Scale Feature Fusion Network for Dense Multi-View Stereo 3D Reconstruction 1177
WeiYe Xing, Jin Huang

Deep Learning Classification of ECG Signals to Diagnose Heart Disease 1181
Hongwei Gao

A Study on Wireless Communication Signal Recognition Technology Based on Improved Convolutional Neural Networks 1186
Ao Li, Te Du, Yinqin Cheng

Research on a Lightweight Dual-Modal Target Detection Method for Complex Scenes 1191
Xuanyu Li, Zhisong Qing, Chongshuangma Tian

A Degradation-Onset-Aware Multi-Scale Time-Frequency Network for Bearing Remaining Useful Life Prediction 1198
Siqi Luo, Xuehong Sun, Liping Liu, Ruyu Liu

Visibility Prediction Model Based on Recursive Feature Elimination and iTransformer Framework 1202
Tao Liu, Buling Zheng, Xue Gao, Wenjing Gao, Minglin Yu

A Multi-Modal Fusion Deep Learning Method for Radar Target Recognition 1208
Xu Cai, Cangzhen Meng, Shi Qu, Hongtao Li

A Geometry-Enhanced 3D Gaussian Splatting Reconstruction Method with Depth Priors 1213
Shou Shan, Haidong Wu

Research on Servo Technology for Laser Direct Writing Systems 1218
Yijia Tang, Jingwei Zhang, Wei Liu

Research on Human Blood Pressure Evaluation Based on Multimodal Deep Learning and Cholesterol Metabolomics 1223
Tiancheng Shan, Jiajia Chen, Jufang Lu, Yufang Yuan, Mengge Liu, Xiaoling Li

A Multimodal Spatiotemporal Deep Learning Framework for Yangtze River Water Level Prediction via Remote Sensing and Hydrological Time-Series Fusion 1228
Shengming Zhou

Conversational Emotion Recognition Based on Pre- Extracted Features, Attention-Based Temporal Pooling and Contrastive Learning 1232
Feng Li, Huixin Zhao, Tiantian Wang

From Time Domain to Frequency Domain: A Music Deepfake Detection Model Based on Multi-Branch Convolutional Neural Network 1237
Menglin Li, Yang Qu, Yangguang Ye

Research on Airport De-icing Vehicle Scheduling Method Based on Deep Q-Network 1244
Zhiwei Xing, Guangyu Wang

Path Planning for Mobile Robots Based on the Integration of Improved A* and DWA 1253
Jianghai Li, Chunling Liu

Infrared-Enhanced Multitask Deep Learning for Particle-Size Classification and Concentration Inversion of Mixed Suspended Sediments 1257
Zihan Cao

SGBEV: Semantic Enhancement and Geometric Guidance for 3D Object Detection 1262
Yang Cheng, Xin Shi

Trusted Collaborative Method for Multi-Agent Semantic Strategy Based on Reversible Data Hiding 1266
Yong-jun Kong, Min Li, Wen-chao Liu, Min-Qing Zhang, Shun-Yi Fan, Fu-qiang Di

LLM-GRU-GAN: a Multi-Modal Adversarial Framework for Transactional Fraud Detection 1271
Cameron Syed, Keyu Yuan, Mingxiu Sui, Ebinezer Chambers, John Wilson

Research on UAV Remote Sensing Monitoring of Lightweight Wheat Lodging Segmentation Network Using Multimodal Fusion (RGB+DSM) 1276
Chenchen Yang, Qang Wang, Xiaohong Duan

CGMC: Consistency Graph-Guided Pose Estimation for Cylindrical Parts Via Maximal Clique Search 1280
Xuhao Chen

A Surface-Based Grasp Point Estimation Method for Bulk Material Piles in Complex Industrial Environments 1285
Yonglin Yu, Xiaohua Cao

VoMA: Voxel Mamba Module for 3D Object Detection in Autonomous Driving 1292
Boyi Sun, Yuhang Liu, Yuhang Li

Video-to-Emotional Music Generation Based on Temporal Structure Alignment 1298
Jing Dong, Wenlian Wu, Ling Wang, Xian Wang

MD-ProtoNet: Modality-Dropout Prototypical Networks for Cross-Subject Cognitive Workload Recognition 1302
Yanxi Liu

Physics-Conditioned Diffusion Detection of Mesoscale Eddies 1309
Yunqi Wang

Frequency-Guided Adversarial Modality Augmentation Network for Visible-Infrared Person Re-Identification 1314
Kongzhu Jiang, Shifeng Chen, Xu Zhou

TESA-LLM: A Temporal Knowledge Graph Reasoning Method Integrating Symbolic Alignment and LLM Generation 1324
Jiancun Wang, Yubin Zhou

Spatial Prior Guided Temporal Modelling for End-to-End Continuous Sign Language Translation 1328
Yiming Bo, Jing Dong, Yandong Liu, Zhihang Cao

CoMAP: Cross-Modal Attentive Prompt Generation for Vision-Language Model Adaptation 1333
Hongyu Bai, Haolong Wang, Genke Yang

Semantic Signal-to-Noise Ratio and Power-Exponential Dynamics for Interpreting Deep Neural Networks 1340
Jinming Zhang, Tao Liu

A Framework for Multimodal Intelligent Analysis in Medicine that Integrates High-Order Omics Interactions with Adversarial Decision-Making 1345
Jiamao Zhao

SMFNet: Selective Mamba-Fusion With Illumination-Conditioned Modulation for RGB-Infrared Pedestrian Detection 1351
Yichen Yuan, Kun Li, Mingchuan Wang

FedRPT: a Dual-Layer Federated Learning Framework for Copyright Protection 1360
Haorui Wang, Guorui Feng

Optically Encoded Wearable Memory Systems for Continuous Personal AI 1365
Weitong Zhang, Shiyu Wang

Eeg Emotion Recognition: A 64-Channel Dynamic Graph Attention Spatio-Temporal Transformer Method 1373
Shuyi Wan

Reliability Aware Gaussian Splatting for Dynamic Sensor Platform SLAM 1378
Zhaotian Deng, Ji Li, Weixin Zhang

Spectrum-Adaptive Architecture for Spatio-Temporal Prediction and Ultra-Low Latency Rendering 1384
Guangliang Zhang, Zongxiu Duan, Xingju An, Ming Qiang, Weijian Huang

A Hierarchical Reinforcement Learning Framework with Action Masking and Teacher Guidance for Dynamic Scheduling 1388
Yunyang Zhao, Zijing Guo, Genke Yang

A Systematic Study of Key Factors Influencing Grokking in Tiny Transformers 1395
Yu Liu, Fuguo Xu, Lijie Zhu, Qianqian Zhang, Shujin Qin, Xiaochen Zhang

Multi-dimensional Pathological Feature Enhancement for Cross-Modal Survival Prediction in Lung Adenocarcinoma 1400
Ziyue Ji, Qiwen Chen, Xinxin Shi, Maoqin Tian, Hao Gao, Sheng Wan, Yinghui Yin, Xinyu Zhou Shaoxin Huang

PW-XSimGCL: Popularity-Aware and Warmup-Enhanced Graph Contrastive Learning for Recommendation 1412
Jie Wu, Xiaofeng Wang

Underwater Monocular Pose Estimation with Occlusion-Robust Matching Under a Multi-Layer Refraction Model 1420
Hongxiang Ye, Lei Yu

Spatiotemporal Multimodal Large Language Model for Facial Micro-Expression Spot-then-Recognize 1425
Li Dangen, Pan Hang

Drone-Based Gaussian Splatting and 3D Reconstruction Considering Fine-Grained Instance-Level Architectural Semantics 1429
Yang Zhan, Siyan Gu, Suxiang Hu, Qiqi Li, Meiqin Che

Research on Robust Detection and Pose Estimation Algorithm for Berth Targets 1434
Yunxing Xiao

Tokenizer Compression for VHR-SAR Scattering Centers Constrained by Metric Relations 1440
Siyan Gu, Yang Zhan, Meiqin Che, Yuhan Zhang, Xiaohui Shi

Frequency-Aware Stage-Tail Attention for Aliasing-Aware Stage Transitions in Convolutional Networks 1445
Han Zhang, Nan Wang

Multimodal Electromagnetic Signal Recognition Using Signal Images and Heterogeneous Feature Fusion 1449
Guoyu Li, Yupeng Wang, Xiaoyu Yu, Shuang Liu

Study on the Influence of Defect Shape on the Quantitative Accuracy of Magnetic Flux Leakage in Subsea Pipelines 1454
Jie You, Wei Zhang

A Multimodal Model for Classifying Thyroid Nodules by Integrating Ultrasound Images and Clinical Data 1458
Hannan Huang

APVSA-RCNN: Adaptive Scale-Aware and Multi-Representation Attention for 3D Object Detection 1462
Yunxiang Liu, Shaofei Liu, Hongkuo Niu, Daizheng Li

Short-Code Radar Phase-Coded Waveform Design Based on WGAN-GP and Ambiguity Function Guided Refinement for N=64 1471
Yuhan Xia, Zhengyang Guo, Xiaoxing Feng, Meng Sun

Bridging Vision and Language Prompts with Optimal Transport for Military Open-Vocabulary Detection 1477
Xiaopeng Rao, Wei Li, Song Liu, Jia Gao, Tianhao Shao, Kai Jiang

A Segmented Dynamic Data Pruning Method Based on Sample Value Evaluation 1484
Yuejiao Pei, Yupeng Wang, Guanxiang Jiang, Jinru Fan