

# **2025 International Conference on Computer Vision, Image Processing and Computational Photography (CVIP 2025)**

**Hangzhou, China  
17-19 October 2025**



**IEEE Catalog Number: CFP250AK-POD  
ISBN: 979-8-3315-8630-0**

**Copyright © 2025 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:	CFP250AK-POD
ISBN (Print-On-Demand):	979-8-3315-8630-0
ISBN (Online):	979-8-3315-8629-4

**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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

## Table of Contents

### ● Chapter 1- Object Detection and Tracking

<b>DTDNet: Defect Transfer Detection for Cross-Background Detection of Morphologically Consistent Defects</b>	1
Jiangtao Cheng, Guojun Wen, Zhaolei Diao, Shuang Mei	
<b>Railway Construction Pollutant Detection Algorithm Based on Improved YOLOv8n</b>	5
Zhen Yan, Xu Zhao, Yalong Xie, Yarui Wei	
<b>TSMN: Spatiotemporal Multi-branch Module for Deepfake Detection</b>	9
Junsen Lou, Ben Zhao, Tianhao Li, Haotian Bai, Qing Liu, Zhuoya Zhao, Yawen Zhang, Yaqi Wang, Suiyu Zhang, Man Zhao	
<b>Fusion Mechanism for Text Detection and Tracking in Video Scenes</b>	16
Mingxuan Pan, Kun Xu, Yaru Chen	
<b>Enhanced Bee Detection Using YOLOv8n with FasterNet and ODConv Integration</b>	22
Runxia Ding, Hui Zhang, Yan Yan, Wanjuan Wang, Conghan Zhong	
<b>An X-ray Imaging-Based Tire Defect Detection Framework with Enhanced YOLOv11</b>	26
Maowei Liu, Hao Zheng, Wei Xiong, Xinyue Yu	
<b>Research on Visual Inspection Methods for Surface Weld Seams of Stainless Steel Thermos Cups</b>	31
Song Wang, Jinqian Li	
<b>Infrared Small-Scale Human Target DETR Recognition Algorithm for Complex Environments</b>	37
Jiayin Liu	
<b>Non-destructive defect detection of OLED surface mura based on background reconstruction</b>	42
Yuan Cao	
<b>YOLOv8-MAH: A Vehicle Detection Method Based on Mosaic Augmentation, Attention Mechanism, and Heatmap-based Interpretability</b>	47
Yirong Zhou, Siyu Zhou, Yunxiao Chang	
<b>Research on Small Object Detection in UAV Aerial Imagery Based on YOLOv11n</b>	51
Junjie Xiao, Dan Yu	
<b>Forest Fire and Smoke Detection Algorithm Based on YOLOv13-EDE</b>	59
Wenjun Wang, Yuduo Liu, Lanfang Lei, Ling Zhang, Xinli Zheng, Changyu Liao, Jianwei Wu	
<b>Double-sided braille dot recognition based on improved YOLOv8</b>	64
Jiayan Wang, Yuhua Chen	
<b>Industrial Defect Detection Based on Attention-Enhanced Feature Pyramid and Deep Feature Reconstruction</b>	70
Haitao Wang, Haichun Hong, Jifu Xu, Wenliang Yang, Kaili Deng	
<b>High-Spatial-Resolution Modal Detection of Structural Vibrations from Undersampled Video Images</b>	78
Xuefen Xiong, Mingguang Shan	

<b>Small Target Detection in UAV Remote Sensing Images Using Deep Learning</b>	82
Haiyang Bo, Jixiang Li, Wei Wang, Xu Liu	
<b>Transformer image recognition algorithm integrating PDE</b>	88
Yujing Wang, Lifen Wang, Qiuyi Lu, Yunxiang Zhang, Yue Zhao, Yu Chen	
<b>ResFPNNet: A Lightweight Multi-Object Detection Network for Embedded Systems</b>	92
Xinran Yang, Zhong Wang, Yanan Xu	
<b>DCA-YOLOv8n: A Lightweight Multiscale Attention Model for Efficient Chinese Food Recognition</b>	99
Ziyu Xu, Qiang Cai, Yanzhao Ren, Yating Huang, Haisheng Li, Yuxiao Wu	
<b>A Synergistic Paradigm of Heterogeneous Attention and Curriculum Learning for Small Object Detection</b>	105
Zhaoyang Zhang, Qikun Shi, Xiao Wang, Jiayi Lai	
<b>SRM-based parallel compressive sensing for image tampering detection and localization in encrypted domain</b>	110
Peng Wang, Hui Wang	
<b>Chapter 2- Image Enhancement and Generation</b>	
<b>Research on the Innovative Design of traditional paper-cut patterns in Yangzhou using the CycleGAN Algorithm</b>	114
Wenming Liu, Wenya Chu, Xiudan Wang	
<b>Study on evaluation indexes for deep learning-based image enhancement of aero-engine borescope probing</b>	119
Shimeng Cui, Shuo Wang, Zhenghong Han	
<b>The Key is the Prior: A Perspective on 2D Synthesis and the Future of Generative Vision</b>	126
Liuqing Zhang, Fan Zhang, Naye Ji	
<b>DSTCM-VHNet: A Robust Video Hiding Network with Dynamic Spatio-Temporal Compensation Module</b>	134
Yihang Lu, Ke Niu, Lichen Zhou, Shunzhe Zhu	
<b>Dual Structure-Texture Aware Mapping for Retinex-Based Low-Light Image Enhancement</b>	138
Haocheng Lin, Zhihui Wei	
<b>End-to-End Parametric Avatar Generation from Monocular Image</b>	145
Taehyeon Kim	
<b>Phase unwrapping algorithm based on mass-guided four-way least squares slope image</b>	149
Ziming Huang, Yan Zhang, Guoliang Peng, Xiaohui Tao, Rui Cao	
<b>Research on a Style Transfer Algorithm from Lung UTE Images to CT Images Based on Improved CycleGAN</b>	153
Yanbin Cui, Chaoying Tang	
<b>A Lightweight Preprocessing Pipeline for Orthophoto Generation of Grotto Murals</b>	158
Kaiyi Fang, Zhongyu Min, Changyu Diao	

<b>Secure3DGS: A Robust and Imperceptible Digital Watermarking Framework for 3D Gaussian Splatting</b>	163
Yanfeng Jiang, Naye Ji	
<b>DWS-Net: A Lightweight Image Super-Resolution Network via Synergized Dynamic Hybrid Attention and Wavelet Convolution</b>	169
Luojun Liu, Yuhua Chen	
<b>Efficient Carrier Frequency Estimation for High-Quality Off-Axis Digital Holographic Reconstruction</b>	175
Yanjun Xu, Mingguang Shan	
<b>Navigation map real-time inlaid photoelectric video technology</b>	179
Zhoufeng Yu, Fei Guo, Fei Ning, Libing Hou	
<b>DiM-Gestor: Co-Speech Gesture Generation with Adaptive Layer Normalization Mamba-2</b>	183
Fan Zhang, Yi Wei, Naye Ji, Jingmei Wu, Zhaohan Wang, Fuxing Gao, Liuqing Zhang, Zhenqing Ye, Leyao Yan, Lanxin Dai, Xin Lyu, Dingguo Yu, Yanbing Jiang, Hui Du, Bozuo Zhao	
<b>BookScanNet: Geometrical Rectification for Scanned book Images Using CoupledTPS</b>	196
Yao Lu, Zhiruo Zhang, Xunran Zhang, Aihong Qin, Zhengrong Li, Yun Zhang	
<b>A Canny-ORB Registration and GAN-based Fusion Method For Infrared and Visible Images</b>	204
Mengdan Zhao	
<b>Solving Imaging Inverse Problems via Diffusion Models with Noise-Refined Likelihood Guidance</b>	211
Zhen Wang, Hongyi Liu, Zhihui Wei	
<b>Fast CU Partitioning Decision for VVC Intra Coding</b>	216
Jiacheng Li, Qingkun Chen	
<b>Attention-Guided Multi-Scale Fusion for Accurate Image Stitching</b>	223
Zongkai Huang, Yiliang Wu	
<b>A Neural-Snapshot Hybrid Framework for Bio-Inspired UAV Navigation</b>	227
Runheng Huang, Yiliang Wu, Yendo Hu	
 <b>Chapter 3- Image Segmentation and Classification</b>	
<b>Coarse-to-Vol: A Two-Stage Framework for Coronary Artery Segmentation in 3D Computed Tomography Angiography</b>	231
Fei Jing, Zhuang Fu, Zi Fang, Jianfeng Yao, Feiyu Meng, Bang Liu	
<b>RCN-Net: A ROI-Guided Co-Attention Normalization-Free Network for Cataract Surgery Phase Recognition</b>	238
Haotian Bai, Tianhao Li, Ben Zhao, Yawen Zhang, Junsen Lou, Qing Liu, Xuze Shi, Yitong Wang, Yaqi Wang, Suiyu Zhang	
<b>High-frequency-aware Multi-task Learning and Transformation Consistency for semi-supervised Electromagnetic Shielding Optical Window Image Segmentation</b>	244
Qiong Zhou, Tao Lei, Bo Liu, Tongfei Liu, Yan Lu, Hongying Meng	

<b>Automatic Evaluation Framework based on CNN-LSTM for Motion Assessment</b>	252
Lin Huang, Jing Chen	
<b>Multi-scale Feature Fusion Network Integrating KAN and Mamba for Ultrasound Image Segmentation</b>	257
Liming Guo, Shuangquan Li	
<b>Multi-level Convolutional Feature Fusion for Plant Disease Identification</b>	261
Hao Yang, Wenjun Zhang, Kunying Li, Chengzhuan Yang	
<b>A Comparative Study of Land Cover Classification Algorithms Using Hyperspectral Remote Sensing Multimodal Imagery in the Xiong'an New Area</b>	268
Yanli Ma, Yanbing Heng, Jie Zhao, liyan Liu	
<b>A Comparative Study of Convolutional Neural Networks for Jet-Image Classification</b>	272
Tian Dai	
<b>Direction-Aware Spatial Mamba Networks for Hyperspectral Image Classification</b>	277
Xiangqian Ye	
<b>Emerging Trends in Non-Cooperative Satellite Pose Estimation: Deep Learning, Sensor Fusion, and Onboard Readiness</b>	282
Maryam Baghban Kondori, Zhongyi Chu	
<b>Learning Features from Images by Multigrid Enhanced CKAN</b>	287
Junjie Yang, Jiangfeng Yang, Wansheng Wang, YunHao Chen	
<b>A Multimodal Emotion Recognition Framework Based on Multimodal Data</b>	292
Mingyuan Sun, Jiayi Mi, Liang Li, Haozhe Yu, Jiahao Pan, Chang Liu	
<b>A 3D Image Perception Approach based on YOLO and Mamba co-optimisation</b>	298
Huahan Zhang, Jiayi Mi, Jiahao Pan, Haozhe Yu, Chang Liu, Liang Li	
<b>A Light-weight Two-stream Interaction Network for Human Activities Recognition</b>	305
Huimin Qian, Hao Zhang, Luyao Zhang	
<b>CNN and Transformer Fusion for Aerial Human Action Recognition</b>	311
Sumaya Abdulrahman Altuwairqi, Salma Kammoun Jarraya	
<b>Reinforcing Representations and Regularizing Prototypes in Few-Shot Class-Incremental Learning</b>	318
Xiaoyi Tan, Qian Zhu, Qisheng Xu	
<b>Learning for Semantic-Aligned and Detail-Aware Image Feature Matching</b>	326
Kun Huang, Hao Peng, Jianwei Yang, Tielin Zhao, Lei Ji, Yun Zhang, Fanglue Zhang	
Evaluating the effect of heterogeneity on the learning ability of spiking neural network	334
Xu Zheng, Jingfeng Xue, Mingrui Li, Yanjun Zhang	
<b>Author Index</b>	343