

PROCEEDINGS OF SPIE

Seventeenth International Conference on Graphics and Image Processing (ICGIP 2025)

Liang Xiao
Editor

7–9 November 2025
Nanjing, China

Sponsored by
Nanjing University of Science and Technology (China)

Organized by
Nanjing University of Science and Technology (China)
Jiangsu Computer Society (China)
Jiangsu Association of Artificial Intelligence (China)

Co-supported by
Nanjing University of Information Science and Technology (China)
Northwest A & F University (China)
Ocean University of China (China)
University of Portsmouth (United Kingdom)
China Simulation Federation (China)

Published by
SPIE

Volume 14124
Part One of Two Parts

Proceedings of SPIE 0277-786X, V. 14124

SPIE is an international society advancing an interdisciplinary approach to the science and application of light.

The papers in this volume were part of the technical conference cited on the cover and title page. Papers were selected and subject to review by the editors and conference program committee. Some conference presentations may not be available for publication. Additional papers and presentation recordings may be available online in the SPIE Digital Library at SPIDigitalLibrary.org.

The papers reflect the work and thoughts of the authors and are published herein as submitted. The publisher is not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

Please use the following format to cite material from these proceedings:

Author(s), "Title of Paper," in *Seventeenth International Conference on Graphics and Image Processing (ICGIP 2025)*, edited by Liang Xiao, Proc. of SPIE 14124, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9798902322078

ISBN: 9798902322085 (electronic)

Published by

SPIE

P.O. Box 10, Bellingham, Washington 98227-0010 USA

Telephone +1 360 676 3290 (Pacific Time)

SPIE.org

Copyright © 2026 Society of Photo-Optical Instrumentation Engineers (SPIE).

Copying of material in this book for internal or personal use, or for the internal or personal use of specific clients, beyond the fair use provisions granted by the U.S. Copyright Law is authorized by SPIE subject to payment of fees. To obtain permission to use and share articles in this volume, visit Copyright Clearance Center at copyright.com. Other copying for republication, resale, advertising or promotion, or any form of systematic or multiple reproduction of any material in this book is prohibited except with permission in writing from the publisher.

Printed in the United States of America by Curran Associates, Inc., under license from SPIE.

Publication of record for individual papers is online in the SPIE Digital Library.

**SPIE. DIGITAL
LIBRARY**

SPIDigitalLibrary.org

Paper Numbering: A unique citation identifier (CID) number is assigned to each article in the Proceedings of SPIE at the time of publication. Utilization of CIDs allows articles to be fully citable as soon as they are published online, and connects the same identifier to all online and print versions of the publication. SPIE uses a seven-digit CID article numbering system structured as follows:

- The first five digits correspond to the SPIE volume number.
- The last two digits indicate publication order within the volume using a Base 36 numbering system employing both numerals and letters. These two-number sets start with 00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 0A, 0B ... 0Z, followed by 10-1Z, 20-2Z, etc. The CID Number appears on each page of the manuscript.

Contents

xi *Conference Committee*
xiii *Introduction*

Part One

PATTERN RECOGNITION

- 14124 02 **Single-stage automatic license plate recognition with contrastive denoising learning** [14124-126]
- 14124 03 **Research on commodity recognition method based on YOLOv8_SMB** [14124-25]
- 14124 04 **Source camera model identification in smartphone videos using fused features from separated noise residuals** [14124-30]
- 14124 05 **A horse face available weighted-constructed facial image recognition approach in low-dimensional space** [14124-118]
- 14124 06 **A multisource human pose stability recognition method based on a seven-point structural mode** [14124-107]
- 14124 07 **Temporally-aware differentiable frame compression network for action recognition** [14124-120]
- 14124 08 **CLITrans: cross-layer feature interaction vehicle reidentification research with vision transformer** [14124-85]
- 14124 09 **Spatial-temporal channel topology-optimized graph transformer for skeleton-based wild terrestrial mammalian behavior recognition** [14124-79]

IMAGE SEGMENTATION

- 14124 0A **FCMamba: frequency-aware CNN-Mamba hybrid model for skin lesion segmentation** [14124-4]
- 14124 0B **A brain tumor MR image segmentation algorithm based on logical stick-breaking and Markov random field** [14124-15]
- 14124 0C **DiffSeg: text-guided diffusion-based image editing with semantic segmentation** [14124-92]
- 14124 0D **A region-constrained asymmetric geodesic voting method for image segmentation** [14124-113]

- 14124 0E **CNN-transformer collaborative learning for remote sensing semantic segmentation**
[14124-142]
- 14124 0F **DMF2Net: dual-branch multistage feature fusion network for road crack segmentation**
[14124-22]
- 14124 0G **Lightweight semantic segmentation via hierarchical scale aggregation and weighted squeezed axial attention** [14124-115]
- 14124 0H **Local-global consistency guided test-time domain adaptation fundus image segmentation**
[14124-138]
- 14124 0I **DSGE-UNet: a lightweight colony segmentation algorithm based on improved UNet**
[14124-147]

FEATURE EXTRACTION AND FEATURE FUSION

- 14124 0J **HISFA: hierarchical scale-aware feature aggregation approach for the lightweight fine-grained visual classification** [14124-26]
- 14124 0K **MEPNet: a multielement prediction model for the Arctic sea area catering to spatiotemporal heterogeneity** [14124-108]
- 14124 0L **Context-guided feature enhancement for robust monocular 3D lane detection** [14124-95]
- 14124 0M **Dual-modality adaptive feature fusion for object detection guided by illumination awareness** [14124-21]
- 14124 0N **Texture and shape enhanced feature reconstruction network for few-shot fine-grained aircraft classification** [14124-48]
- 14124 0O **Few-shot learning via dual hypergraph graph convolutional representation** [14124-20]

TARGET DETECTION MODELS AND ALGORITHMS

- 14124 0P **Evolutionary trajectory and multifaceted applications of YOLO object detection algorithms**
[14124-10]
- 14124 0Q **A bionic compound eye-based 3D object detection method for intelligent driving** [14124-2]
- 14124 0R **EMCAM-YOLOv8: enhancing screen defect detection with multi-channel attention** [14124-34]
- 14124 0S **Retinex-based feature-consistent domain adaptation for nighttime object detection**
[14124-129]

- 14124 OT **Lightweight multiscale object detection in challenging construction site environments** [14124-137]
- 14124 OU **Robust elderly fall detection using CARE-light-enhanced YOLOv8 with triplet attention and BMCNet** [14124-125]

TARGET TRACKING AND ALGORITHM DESIGN

- 14124 OV **Multiobject tracking via multicue spatiotemporal fusion in complex visual scenes with cluttered illumination** [14124-42]
- 14124 OW **Adaptive target-aware semantic network for vehicle tracking in satellite videos** [14124-111]
- 14124 OX **An RGB-T multimodal information fusion visual object tracking algorithm based on correlation filtering** [14124-88]

3D IMAGE RECONSTRUCTION AND VISUALIZATION

- 14124 OY **Improving radar echo reconstruction via deep learning of Himawari-8 and Goes-R** [14124-11]
- 14124 OZ **Weakly supervised multiview 3D face reconstruction with more details** [14124-3]
- 14124 10 **Underwater 3D reconstruction based on 3D filter Gauss splatting** [14124-24]
- 14124 11 **4DM: 4D photorealistic mapping from monocular video with 3D Gaussian splatting** [14124-8]
- 14124 12 **Large-scale 3D scene reconstruction via gradient field scaling and pixel-aware adaptive density control** [14124-99]
- 14124 13 **DiffScene3D: sparse diffusion for high-fidelity indoor scene geometry generation** [14124-28]
- 14124 14 **Exposure compensation and computational resource optimization in 3DGS-SLAM** [14124-39]
- 14124 15 **D-NeSF: dynamic neural shadow fields for two-bounce non-line-of-sight stereo reconstruction** [14124-65]
- 14124 16 **Antialiased 3D Gaussian splatting for human reconstruction: a dual-filter approach with K-NN pruning** [14124-97]
- 14124 17 **Edge-embedded point cloud sampling with DINOv3 constraints** [14124-70]

MEDICAL IMAGE MODELING AND IMAGE ANALYSIS

- 14124 18 **Structure-aware neural modeling of Circle of Willis vessels from noncontrast MRA data** [14124-96]
- 14124 19 **Crowd simulation with fear and infection awareness for Ebola transmission** [14124-47]
- 14124 1A **Brain age estimation based on age-specific graph structure learning** [14124-64]
- 14124 1B **A lesion-prior guided hyperedge optimization algorithm for early Alzheimer's disease prediction** [14124-62]
- 14124 1C **mBMI: a multimodal brain multiplex interaction network for multimodal brain network analysis** [14124-117]
- 14124 1D **GPU-accelerated procedural texture generation of liver and its cut surfaces based on Worley noise** [14124-130]

INTELLIGENT IMAGE DETECTION AND OPTIMIZATION ALGORITHMS

- 14124 1E **Dual-correlation graph attention network for video anomaly detection** [14124-12]
- 14124 1F **Dynamic generalized IoU threshold: optimizing AP and mAP for object detection using a data-driven approach** [14124-93]
- 14124 1G **EdgeFireNet: a lightweight anchor-free fire detection network for edge devices** [14124-143]
- 14124 1H **Optimization of object detection models for low-altitude scenarios** [14124-105]

Part Two

- 14124 1I **RobustMMD: towards robust multimodal 3D detection via efficient multilevel adapter and spatial denoiser** [14124-46]
- 14124 1J **A two-stage contrastive learning paradigm for fine-grained SAR object detection with limited labeled samples** [14124-91]
- 14124 1K **Joint gaze detection based on head pose and depth estimation using a 360-degree camera** [14124-54]
- 14124 1L **Small object detection algorithm for remote sensing images based on multiscale feature fusion** [14124-106]
- 14124 1M **Urban traffic flow analysis via cross-scale cross-attention mechanisms** [14124-38]

MACHINE LEARNING-BASED IMAGE DETECTION MODELS AND ENGINEERING APPLICATIONS

- 14124 1N **Investigating image quality impact on remote sensing object detection** [14124-45]
- 14124 1O **Improved UAV aerial photography small target detection algorithm based on YOLOv11** [14124-32]
- 14124 1P **Two-stage progressive screw hole localization method for unstructured screw assembly** [14124-149]
- 14124 1Q **GeoRankNet: geometry-aware feature ranking network for infrared small-target detection** [14124-133]
- 14124 1R **Emulating self-attention with convolution embedded high-resolution shallow layer for detection of tiny bearing defects** [14124-52]
- 14124 1S **Semantic-prior enhanced representation learning for weakly supervised video anomaly detection** [14124-110]
- 14124 1T **Multiscale feature reconstruction contrastive learning for distribution shift anomaly detection** [14124-98]
- 14124 1U **PetroVR Fire: a virtual reality training system for catastrophic petrochemical firefighting with enhanced collision detection** [14124-131]
- 14124 1V **PT-Grasp: coarse-to-fine affordance transfer for robotic grasp point prediction** [14124-145]
- 14124 1W **An advanced real-time traffic sign detection method based on improved YOLOv11n** [14124-35]

COMPUTER VISION AND VISUALIZATION

- 14124 1X **Optimized reverse distillation method for nut and bolts anomaly detection on edge devices** [14124-89]
- 14124 1Y **Movie recommendation system integrating visual poster features and temporal dynamics** [14124-17]
- 14124 1Z **Language as a supervisor: a text-driven visual learning approach for general purpose meter reading** [14124-67]
- 14124 20 **Monocular pose estimation for vision-guided fixed-wing aircraft airport landing** [14124-132]
- 14124 21 **A multisensor fusion visual SLAM mapping method** [14124-60]
- 14124 22 **Structured pruning for vision transformer via multiple linear regression** [14124-77]

- 14124 23 **Practical and ready-to-use water body observation and visualization using Sentinel-1 SAR in the Copernicus Browser** [14124-116]

DIGITAL SIGNAL DETECTION AND DATA ANALYSIS

- 14124 24 **Defocus deconvolution signal restoration based on Bayesian inference** [14124-29]
- 14124 25 **A method with accelerated sampling and spatiotemporal embedding for nowcasting** [14124-7]
- 14124 26 **Satellite-radar-Swin-Unet: a Swin-Unet-based model for synthetic radar generation from satellite data** [14124-66]
- 14124 27 **Adaptive container scheduling strategy for multidimensional perception** [14124-76]
- 14124 28 **Dynamic-attitude-based modeling of UAV ground coverage areas** [14124-127]
- 14124 29 **Towards lightweight noise-robust models via noise-aware pruning** [14124-124]
- 14124 2A **A novel deep learning method for urban flood prediction** [14124-74]
- 14124 2B **TCMD: trajectory-conditioned motion diffusion model** [14124-75]
- 14124 2C **A novel non-convex regularized model for third-order tensor compressive sensing** [14124-44]
- 14124 2D **Comparative evaluation of TTS and vocoder architectures in lip-to-speech synthesis** [14124-53]
- 14124 2E **A fuzzy spatial and temporal adaptive reflectance fusion model with local information** [14124-81]

INTELLIGENT IMAGE ANALYSIS AND MACHINE VISION FOR COMPLEX SCENES

- 14124 2F **Mamba-based multiview spatial and frequency collaborative network for hyperspectral image unmixing** [14124-109]
- 14124 2G **Digital imagery and real-time interaction in the immersive dance performance asking heaven** [14124-27]
- 14124 2H **Gradient low-rank sparse representation: spatial-spectral group sparse total variation for hyperspectral image denoising** [14124-33]
- 14124 2I **Research on image dehazing method based on the generalized Gaussian function** [14124-13]

- 14124 2J **Image encryption algorithm based on compressed sensing and a new four-dimensional chaotic system** [14124-94]
- 14124 2K **Infrared image hierarchical enhancement network via guided decomposition** [14124-102]
- 14124 2L **Underwater image enhancement model based on GAN with enhanced generative capability** [14124-123]
- 14124 2M **RESONANCE: robust synergistic consistent and covariant learning for adversarial few-shot classification** [14124-56]
- 14124 2N **KA-DMSF: kernel attention and dynamic multishape filtering for real-world image dehazing** [14124-148]
- 14124 2O **A multiscale auxiliary difference context-aware network for remote sensing image change captioning** [14124-101]
- 14124 2P **DAEC-UCT: an improved UCTransNet for SAR road extraction** [14124-128]

COMPUTER-AIDED IMAGING SYSTEMS AND EQUIPMENT

- 14124 2Q **Automatic infrared calibration based on physical guiding** [14124-18]
- 14124 2R **A method for building facade measurement using integrated GNSS RTK and photogrammetry** [14124-9]
- 14124 2S **UAV camera pose estimation based on multiscale features and contextual information** [14124-19]
- 14124 2T **Design and implementation of a lightweight dual-lens camera adapted for logistics scenarios** [14124-90]
- 14124 2U **VT-NeRF: view transformation and EM optimization for robust novel view synthesis from low-variation camera trajectories** [14124-37]