

PROCEEDINGS OF SPIE

Seventeenth International Conference on Digital Image Processing (ICDIP 2025)

**Ting-Chung Poon
Xudong Jiang
Zhaohui Wang
Jindong Tian**
Editors

**25–27 April 2025
Haikou, China**

Organized by
Hainan University (China)

Hosted by
School of Computer Science and Technology, Hainan University (China)
Center for Information Technology Services, Hainan University (China)

Published by
SPIE

Volume 13709

Part One of Two Parts

Proceedings of SPIE 0277-786X, V. 13709

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 Digital Image Processing (ICDIP 2025)*, edited by Ting-Chung Poon, Xudong Jiang, Zhaohui Wang, Jindong Tian, Proc. of SPIE 13709, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9781510693708

ISBN: 9781510693715 (electronic)

Published by

SPIE

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

Telephone +1 360 676 3290 (Pacific Time)

SPIE.org

Copyright © 2025 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*

Part One

MULTIMODAL IMAGE INFORMATION EXTRACTION AND DATA VISUALIZATION

- 13709 02 **Inception-residual-based network for TDLAS temperature imaging** [13709-23]
- 13709 03 **Blockchain-enhanced spectrum sensing with PoAS and game-theoretic detection**
[13709-83]
- 13709 04 **ESC-SORT: enhanced similarity costs lead to an improved multi-object tracker** [13709-75]
- 13709 05 **A data hiding algorithm for the encrypted satellite image based on quantization
compression** [13709-24]
- 13709 06 **Hierarchical hypergraph parallel networks for 3D human pose estimation** [13709-56]
- 13709 07 **A survey on expressive action quality assessment driven by aesthetics and biomechanics**
[13709-95]
- 13709 08 **Online signature verification based on Siamese networks and LSTM** [13709-73]
- 13709 09 **Frequency-domain and spatial-domain fusion for robust and invisible backdoor attack**
[13709-8]
- 13709 0A **OST-SLAM: a relative pose estimation for non-cooperative satellite combining point and
line features with scene semantic constraints** [13709-5]
- 13709 0B **Real-time visual odometry for resource-constrained systems: a heterogeneous computing
approach** [13709-108]
- 13709 0C **Risk AD classification by using ensemble 3D DenseNets based on multi-modal data**
[13709-93]

POINT-CLOUD-BASED IMAGE MODEL AND IMAGE RECONSTRUCTION

- 13709 0D **Automatic extraction of pole-like objects from different scene point clouds** [13709-125]
- 13709 0E **A double-layer structure for pseudo-CT generation from PET images** [13709-112]

- 13709 OF **FSAM-DiffSR: frequency-assisted structural-modulated diffusion model for image super-resolution** [13709-131]
- 13709 OG **Recreating retro pixel art using digital image processing techniques** [13709-94]
- 13709 OH **End-to-end generative adversarial network model for in-air handwritten character generation** [13709-4]
- 13709 OI **Building reconstruction based on Segment Anything Model guided by structural feature** [13709-101]
- 13709 OJ **Optimizing dairy cow condition scoring with YOLO-HGStar: a lightweight deep learning model** [13709-14]
- 13709 OK **FD-GAN: generative adversarial network for reconstructing CT from biplane x-rays based on frequency decoupling and diffusion** [13709-39]
- 13709 OL **Coarse-to-fine real-time GS-SLAM from endoscopic RGBD data** [13709-33]

IMAGE SEGMENTATION

- 13709 OM **LayerScore-CAM: hierarchical score-weighted class activation maps for defect segmentation** [13709-55]
- 13709 ON **Automatic building segmentation based on DOM and DSM** [13709-6]
- 13709 OO **BUMP: a boundary-aware active learning framework for point cloud semantic segmentation** [13709-63]
- 13709 OP **Cross-domain spacecraft component segmentation based on edge consistency generative neural network** [13709-60]
- 13709 OQ **MIFNet: multilevel interactive fusion network for real-time semantic segmentation** [13709-120]
- 13709 OR **Moving object segmentation with SNGAT: a graph attention network enhanced by adaptive similarity** [13709-117]
- 13709 OS **Accurate airway tree segmentation using feature extraction fusion and the selection state space method** [13709-72]
- 13709 OT **Advancing medical image segmentation with multi-directional feature extraction and contextual attention** [13709-130]
- 13709 OU **SSCA-UNet: improved U-Net for nuclear segmentation by 2D-selective-scan and channel attention** [13709-110]

IMAGE DENOISING

- 13709 0V **Blind image deblurring with dynamic dual discriminative prior** [13709-61]
- 13709 0W **Adaptive video denoising based on inter-frame similarity** [13709-132]
- 13709 0X **DenoiseNet: towards advanced denoising with transformer network** [13709-135]
- 13709 0Y **Smoothing-operator-driven single-image dehazing** [13709-38]

IMAGE RECOGNITION

- 13709 0Z **Robust water meter reading recognition in challenging scenarios based on YOLOv11**
[13709-103]
- 13709 10 **Federated-learning-framework-based malware classification algorithm under insufficient
local training data condition** [13709-78]
- 13709 11 **Deep-learning-based method for individual recognition of jadeite** [13709-3]
- 13709 12 **MGLGait: a multimodal gait recognition method based on global-local feature connection**
[13709-67]
- 13709 13 **Cross-domain retrieval based on label dynamic correction and class alignment** [13709-68]
- 13709 14 **Pose-guided occlusion-aware network for occluded person re-identification** [13709-92]
- 13709 15 **Research on rail image acquisition and defects intelligent recognition technology**
[13709-58]
- 13709 16 **Behavior-YOLO: cattle behavior recognition based on improved YOLO11n** [13709-18]
- 13709 17 **SWP-GFR: micro-expression recognition with self-weighted patch and global feature
refinement** [13709-11]

INTELLIGENT IMAGE DETECTION AND APPLICATION

- 13709 18 **Video inter-frame forgery detection based on CNN and ViT** [13709-49]
- 13709 19 **Aluminum surface defect detection method based on improved YOLOv11** [13709-44]
- 13709 1A **Yolov5s-slpd model: a Chinese ship license plate detection algorithm** [13709-20]

- 13709 1B **Single-point supervised infrared small target detection with local affinity feature refinement** [13709-34]
- 13709 1C **Strip attention transformer for crack detection** [13709-16]
- 13709 1D **An improved official document layout element detection transformer based on RT-DETR** [13709-127]
- 13709 1E **Traffic sign detection algorithm based on adaptive enhancement of low-illumination images** [13709-115]
- 13709 1F **SegAuxClsNet: segmentation-task-assisted classification neural network for surface-defect detection** [13709-74]
- 13709 1G **Defect detection of mobile phone back covers under different spectral lighting conditions** [13709-46]
- 13709 1H **The clean sample matters to transferable attacks on object detection models** [13709-30]
- 13709 1I **Online handwritten signature verification method based on attention-BiLSTM for inflection points detecting and stroke segment matching** [13709-121]

IMAGE RESTORATION

- 13709 1J **Hybrid-model-data underwater image restoration and enhancement networks** [13709-10]
- 13709 1K **Cross-frame attention for burst image restoration** [13709-84]
- 13709 1L **A two-stage network for iToF depth image restoration in foggy environments** [13709-88]
- 13709 1M **Deep exploitation of weather information for all-in-one image restoration** [13709-124]

Part Two

MULTIMODAL INFORMATION FUSION AND IMAGE ANALYSIS

- 13709 1N **Multiple MBSs grouping scheme for inter-symbol interference in DSS and static MBSs hybrid networking** [13709-32]
- 13709 1O **Geometry and context encoding cost-volume for deep homography estimation** [13709-62]
- 13709 1P **GVSE++: improved gated visual-semantic embedding for few-shot image and sentence matching** [13709-85]
- 13709 1Q **3D visualization of canned citrus products based on 3D Gaussian splatting** [13709-19]

- 13709 1R **Research on edge shape template matching algorithm based on machine vision** [13709-25]
- 13709 1S **Research on seismic event classification based on multimodal learning** [13709-90]
- 13709 1T **Autonomous localization based on single-channel ground penetrating radar** [13709-27]
- 13709 1U **Color analysis of Bactrian camel pattern textile relic image** [13709-64]

IMAGE DETECTION MODEL AND ALGORITHM

- 13709 1V **Target detection method in sea clutter based on the norm of time-frequency image** [13709-114]
- 13709 1W **Document image tampering detection and localization based on dual-path network with progressive supervision** [13709-50]
- 13709 1X **Dynamic superresolution of DNBs detection based on TRTN** [13709-113]
- 13709 1Y **Real-time object detection and localization for airport aprons** [13709-7]
- 13709 1Z **Apple bruise detection using hyperspectral imaging** [13709-136]
- 13709 20 **A maritime small target detection algorithm based on dual-channel SDP and improved YOLOv11** [13709-118]
- 13709 21 **DR-YOLOv8: a receptive-field-attention-convolution-enhanced object detection model of side scan sonar image** [13709-71]
- 13709 22 **A fine-tuned vision transformer model for efficient plant leaf disease detection** [13709-140]
- 13709 23 **SF-YOLO: a lightweight and precise approach for small object detection in aerial images** [13709-87]

IMAGE ENCRYPTION AND INFORMATION AUTHENTICATION

- 13709 24 **Reversible data hiding in encrypted images based on two-dimensional pixel mapping rotation** [13709-59]
- 13709 25 **Image encryption scheme based on 2D-DuffLog map and DNA encoding** [13709-48]
- 13709 26 **Discrete random quaternion linear canonical transform and its application to color image encryption** [13709-35]

- 13709 27 **Image encryption algorithm using 3-D non-equilateral Arnold transform combined with chaos matrix** [13709-99]
- 13709 28 **A non-interactive blind signature algorithm based on ElGamal** [13709-111]
- 13709 29 **Writer-independent online handwritten signature verification using stable segments via an autoencoder-based Siamese network** [13709-105]

AI-BASED MEDICAL IMAGE ANALYSIS: IMAGE DETECTION, RECOGNITION, AND SEGMENTATION

- 13709 2A **Wavelet-guided dual-domain collaborative network for sparse-angle CT image reconstruction** [13709-70]
- 13709 2B **OMMIR: a dynamic-optimization-based multi-modal medical image registration framework for CT-MR** [13709-13]
- 13709 2C **Automated detection algorithm for mesiodens in CBCT images based on a two-stage 3D deep attention network** [13709-29]
- 13709 2D **RetinalYOLO: a novel approach to detect retinal diseases** [13709-12]
- 13709 2E **Intelligent grading diagnosis of COPD based on CT lung parenchyma extraction** [13709-128]
- 13709 2F **DHGFN: dual hypergraph fusion network for brain tumor classification** [13709-122]
- 13709 2G **DAHUNet: a dual aggregation hybrid UNet-like network for medical image segmentation** [13709-31]
- 13709 2H **Uncertainty-aware medical image segmentation with semantic interpretation using large language models** [13709-36]
- 13709 2I **AdaptVeinNet: adaptive finger vein segmentation across domains** [13709-47]
- 13709 2J **Research on segmentation, registration, and fusion methods for multispectral tongue images** [13709-80]
- 13709 2K **Polarized-Self-Attention-enhanced TransUNet in medical image segmentation** [13709-21]
- 13709 2L **AI-driven medical image analysis for Alzheimer's disease: classification, segmentation, and beyond** [13709-98]
- 13709 2M **SE TransUNet: combine transformer with UNet to achieve lung nodule segmentation** [13709-123]

DIGITAL IMAGE PROCESSING AND MULTIMEDIA APPLICATION TECHNOLOGY

- 13709 2N **Eye-tracking-based image annotation behavior analysis** [13709-26]
- 13709 2O **Non-rigid point cloud registration with dynamic computing neural deformation pyramid** [13709-126]
- 13709 2P **Precise surgical planning for bronchial stenosis using an SPH-based pressure ratio index** [13709-57]
- 13709 2Q **Multi-channel video stitching system based on China-made FPGA** [13709-109]
- 13709 2R **DCLK-GAN: a dual-branch GAN for thermal infrared image colorization based on DCLK-Unet and attention-Unet** [13709-91]
- 13709 2S **Multiscale bidirectional structural offset features for removing incorrect keypoint correspondences between images** [13709-104]
- 13709 2T **The optical adversarial attack of deep neural network based on shadows** [13709-129]
- 13709 2U **InkSync: dual-YOLO framework for whiteboard digitization** [13709-97]
- 13709 2V **MGIA-ECG: multi-graph-based inference architecture for ECG classification** [13709-54]
- 13709 2W **A dual-branch spectral-spatial Mamba and convolution network for hyperspectral image classification** [13709-116]
- 13709 2X **Autonomous target search in complex and unknown environments using UAV equipped with limited FOV camera** [13709-82]

DEVELOPMENT AND APPLICATION OF IMAGE PROCESSING AUXILIARY TOOLS

- 13709 2Y **Deep-learning-based malware classification algorithm using data augmentation** [13709-76]
- 13709 2Z **Ensemble-learning-framework-based malware classification algorithm using multiple kinds of features** [13709-81]

ENHANCED DIGITAL IMAGE PROCESSING: INNOVATION AND APPLICATION DRIVEN BY ARTIFICIAL INTELLIGENCE

- 13709 30 **Research on underwater image enhancement methods based on perceptual vision** [13709-17]

- 13709 31 **Low-light image enhancement with optical prior and frequency constraint** [13709-69]
- 13709 32 **HyNPhyAttnGAN: underwater image enhancement with Unet-GAN by fusing hybrid normalization, physics simulation, and multi-scale attention mechanisms** [13709-15]