PROCEEDINGS OF SPIE

Eighteenth International Conference on Digital Image Processing (ICDIP 2026)

Ting-Chung Poon
Xudong Jiang
Jindong Tian
Shihui Guo
Editors

24 - 26 April 2026
Xiamen, China

Sponsored by
Xiamen University (China)

Hosted by
School of Informatics, Xiamen University (China)

Published by
SPIE

Volume 14351

Part One of Two Parts

Proceedings of SPIE 0277-786X, V. 14351

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 SPIEDigitalLibrary.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 *Eighteenth International Conference on Digital Image Processing (ICDIP 2026)*, edited by Ting-Chung Poon, Xudong Jiang, Jindong Tian, Shihui Guo, Proc. of SPIE 14351, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X
ISSN: 1996-756X (electronic)

ISBN: 9798902328094
ISBN: 9798902328100 (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
SPIEDigitalLibrary.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

IMAGE ENHANCEMENT AND IMAGE RESTORATION

14351 02 **DKS-Net:a depthwise kernel selective network for single image dehazing** [14351-49]

14351 03 **Low illumination image enhancement method based on U-shaped multi-scale network** [14351-12]

14351 04 **EHENet: edge-guided hybrid enhancement network for low-light light field image enhancement** [14351-3]

14351 05 **Progressive prompt-guided network for unsupervised low-light image enhancement** [14351-93]

14351 06 **Boosting defocus deblurring via learning from all-in-focus images** [14351-81]

14351 07 **FreqMamba: exposure-aware frequency decomposition with interleaved state scanning for HDR imaging** [14351-94]

14351 08 **Enhanced adaptive histogram specification with multi-metric balancing for low-light image and video enhancement** [14351-67]

14351 09 **Adaptive weight bilateral filtering for image denoising model with preserving edge detail** [14351-20]

14351 0A **A multi-scale dual-branch spatial attention network with dynamic training for zero-shot image denoising** [14351-79]

MULTIMODAL IMAGE FUSION AND PROCESSING METHODS

14351 0B **A personalized federated learning-based CSI localization algorithm using multi-class knowledge distillation for heterogeneous data** [14351-29]

14351 0C **Lightweight infrared and visible image fusion via global selective state-space modeling** [14351-96]

14351 0D **A noise assessment-based quality evaluation method of photometric data for space objects** [14351-35]

14351 0E **Robust and efficient UAV tracking via adaptive depth gating and prompt-guided autoregressive decoding** [14351-52]

14351 0F **PNTAFormer: polarity-aware neighborhood token attention for spiking transformers** [14351-50]

14351 0G **A digital image processing method for quantitative study on conformational transitions of DNA with cisplatin** [14351-75]

IMAGE COMPRESSION, ENCODING, AND IMAGE SECURITY

14351 0H **A deep learning based indoor localization algorithm using CSI measurement and time series decomposition** [14351-121]

14351 0I **Invisible medical image watermarking: a biometric–deep learning framework for security and verification** [14351-123]

14351 0J **Image tampering detection based on two-branch gating-augmented transformer encoder with attention-based feature fusion** [14351-16]

14351 0K **HiDD-Net: a hierarchical dual-domain distillation network for efficient deepfake detection** [14351-101]

14351 0L **Robust 3D Gaussian splatting watermarking via hybrid optimization** [14351-60]

14351 0M **Learned screen content image compression with frequency-domain modulation** [14351-112]

14351 0N **Scalable and perception-aware deep joint source–channel coding for underwater acoustic image transmission** [14351-42]

14351 0O **Spatial-channel dual autoregressive model for remote sensing image compression** [14351-37]

MACHINE LEARNING-BASED IMAGE CLASSIFICATION AND INTELLIGENT ALGORITHMS

14351 0P **Evaluation of DSP-driven and ML methods for radar-based classification of human, vehicles, and UAVs** [14351-118]

14351 0Q **AART: image classification with adaptive attention region trans-former** [14351-59]

14351 0R **Enhancing discriminative features: a pneumonia classification network for complex backgrounds and insignificant lesions** [14351-98]

14351 0S **Automatic classification of alveolar bone defects via 3D CBCT for dental implant preoperative evaluation** [14351-64]

14351 0T **Do saliency maps tell the truth? Quantitative reliability evaluation of XAI in multi modal MRI tumor classification** [14351-115]

14351 0U **MedXAIgnosis: a metadata-enhanced graph-fusion pilot for thorax disease classification** [14351-57]

14351 0V **ConvECTNet: efficient fatty liver ultrasound classification using ConvNeXt with Top-T enhanced channel attention** [14351-80]

OBJECT DETECTION MODELS AND INTELLIGENT DETECTION TECHNOLOGIES FOR COMPLEX SCENES

14351 0W **HRT-DETR: efficient traffic sign detection** [14351-66]

14351 0X **Vehicle and driver detection in real-world traffic scenarios with an improved YOLO model** [14351-88]

14351 0Y **SRDet: self-regularized object detection framework for UAV perspective** [14351-51]

14351 0Z **Open-vocabulary object detection based on fine-grained alignment** [14351-63]

14351 10 **Privacy-aware multi-object tracking for retail environments** [14351-106]

14351 11 **DMONet: deep multi-region joint fusion optimization network for camouflaged object detection** [14351-43]

14351 12 **Anchor-free Gaussian heatmap ball detection network based on dynamic soft labels** [14351-105]

14351 13 **Investigating camouflaged object detection difficulty through human visual perception** [14351-111]

14351 14 **Structured lane instance detection with instance activation query and adaptive cross-attention** [14351-73]

14351 15 **Cross-view consistent teacher for source-free object detection** [14351-32]

14351 16 **SAR ship detection in range-compressed domain based on adaptive Gabor filter and wavelet convolution** [14351-19]

14351 17 **A multimodal remote sensing object detection model based on the Swin-MoE framework** [14351-31]

14351 18 **Improved RT-DETR for remote sensing image object detection** [14351-117]

DEEP LEARNING-BASED MEDICAL IMAGE SEGMENTATION AND OPTIMIZATION ALGORITHMS

14351 19 **Transferring SAM-pretrained 2D ViTs for Semi-supervised 3D medical image segmentation** [14351-9]

14351 1A **QVP: quaternion visual prompt-based frequency-enhancing and hierarchical learning for blurred endoscopic image segmentation** [14351-92]

14351 1B **Self-correcting contrastive learning with pretrain-free copy-paste for semi-supervised medical image segmentation** [14351-72]

14351 1C **GPG-Net: graph-prior-guided breast ultrasound image segmentation network** [14351-99]

14351 1D **Query-adaptive prototype refinement for few-shot aortic CTA segmentation** [14351-76]

14351 1E **PFT-Net: prototypes filtering and transformation for few-shot medical image segmentation** [14351-65]

14351 1F **Conditional flow matching with coordinate-aware patch training for efficient 3D medical image segmentation** [14351-107]

14351 1G **Research on breast cancer image segmentation method based on improved ResNet34-UNet hybrid loss function** [14351-13]

14351 1H **BE-Unet: a boundary-enhanced UNet for skin lesion segmentation** [14351-113]

14351 1I **Optimizing multi-dimensional retinal segmentation: a performance-practicality analysis** [14351-27]

14351 1J **HiT: holistic-integration transformer network for efficient segmentation in endoscopic medical imaging** [14351-78]

Part Two

IMAGE GENERATION MODELS AND 3D VISION RECONSTRUCTION

14351 1K **A method for 3D reconstruction of train carriages based on LiDAR point clouds** [14351-7]

14351 1L **Enhancing 3D semantic scene completion via efficient attention and feature augmentation** [14351-38]

14351 1M **Data-driven 3D face point cloud segmentation** [14351-28]

14351 1N **Efficiency optimization for Gaussian surface reconstruction via subject-aware modeling and adaptive resource allocation** [14351-68]

14351 1O **Lightweight and high-fidelity 3DGS fruit reconstruction via geometric-semantic joint constraints** [14351-44]

14351 1P **Research and application of virtual staining in renal cancer based on controllable Brownian bridge diffusion model** [14351-84]

14351 1Q **Prior-guided noise alignment diffusion model for fundus image generation with peripapillary atrophy mask** [14351-86]

14351 1R **Research on super-resolution image reconstruction method based on multi-axis attention** [14351-55]

14351 1S **Vortex spatio-temporal reconstruction model based on flow phase topology** [14351-36]

14351 1T **Layered refinement algorithm for coastal wind speed reconstruction from sparse buoy observations** [14351-91]

GENERATIVE AI-BASED IMAGE SYNTHESIS AND VIDEO EDITING

14351 1U **EEG-based validation of a neuroaesthetic color model for educational animation** [14351-95]

14351 1V **Survey on video restoration and enhancement technologies** [14351-24]

14351 1W **Mamba-guided feature enhancement and fusion network for fake video detection** [14351-90]

14351 1X **Design and simulation research of an intelligent stage interaction system based on real time motion capture** [14351-39]

14351 1Y **HierDiff: hierarchical saliency-guided diffusion for template synthesis** [14351-41]

14351 1Z **Gags: graph-guided adaptive Gaussian splatting for scene stylization** [14351-82]

14351 20 **Improving the quality of synthetic echocardiogram images using hybrid GAN and attention mechanisms** [14351-74]

14351 22 **Occlusion-free image mosaicing for rodent gait analysis via motion saliency and catadioptric system** [14351-26]

14351 23 **ActionLMM: captioning long-video actions with memory-augmented VLMs** [14351-126]

COMPUTER-AIDED DIAGNOSIS BASED ON MEDICAL IMAGING: TRENDS AND FUTURE RESEARCH

14351 24 **Dynamic fusion skin disease diagnosis network based on space-domain and frequency-domain dual-domain coordination** [14351-114]

14351 25 **Cross-category domain interaction network for tongue image multi-label recognition** [14351-58]

14351 26 **LSAFormer: a low-resolution spatial-aware transformer for computer-aided diagnosis of pneumonia** [14351-23]

14351 27 **Enhancement of the common carotid ultrasound images using non-subsampled shearlet transform and Hessian matrix** [14351-85]

14351 28 **Low-count PET image denoising based on VQVAE: evaluation on total-body datasets** [14351-119]

14351 29 **Tiling to transfer: patch-level contrastive learning framework for cross-dataset generalization on downstream diabetic retinopathy grading tasks** [14351-87]

14351 2A **MR-HNet: A hybrid neural network for parameter inversion of magnetic resonance fingerprinting** [14351-25]

14351 2B **MSDG-YOLO: multi-scale view and difference-guided enhanced YOLO for mandibular fracture detection** [14351-97]

14351 2C **YOLO-MIA: a lightweight multi-scale attentive network for stent detection in IV-OCT grayscale Images** [14351-6]

14351 2D **Lesion-robust longitudinal brain CT registration via MR-guided image translation** [14351-102]

PATTERN RECOGNITION AND ALGORITHMS

14351 2E **A review of intelligent recognition methods for small aerial targets based on deep learning** [14351-45]

14351 2F **Attribute-aware prompt learning for cloth-changing person re-identification** [14351-122]

14351 2G **YOLO-RED: an improved YOLO11-based network for multi-scale classroom student behavior recognition** [14351-124]

14351 2H **An integrated YOLOv11, action recognition and vision-language model framework for intelligent kindergarten safety monitoring** [14351-69]

14351 2I **MASF-Net: efficient linear attention guided few-shot fine-grained image recognition** [14351-125]

14351 2J **Infant and toddler facial expression recognition based on topology-aware pooling graph U-Net** [14351-18]

14351 2K **A multi-task student facial expression recognition model based on soft parameter sharing** [14351-83]

14351 2L **A survey of generative makeup removal techniques for face recognition** [14351-4]

COMPUTER VISION AND MULTIMEDIA APPLICATIONS

14351 2M **Empowering embodied intelligence via a flexible sensing system integrating multi-dimensional tactile information** [14351-116]

14351 2N **SALIA: research on thangka image captioning model fusing salient attention and local interaction aggregator** [14351-56]

14351 2O **DecKAN: enhancing transformer-based semantic segmentation with Kolmogorov-Arnold network decoders** [14351-47]

14351 2P **Collaborative video object segmentation for underwater filter-net inspection** [14351-30]

14351 2Q **Reference-based ordinal metric learning for safety-aware yoga pose assessment** [14351-100]

14351 2R **FCSRS: research on Thangka image captioning model combining feature compression supplement and semantic re-learning** [14351-61]

14351 2S **TopoParse vision-assisted hierarchical parsing of residential community layouts** [14351-33]

SURFACE DEFECT DETECTION AND QUALITY CONTROL IN INDUSTRIAL VISION

14351 2T **Unsupervised anomaly detection method based on discrete feature rectification** [14351-71]

14351 2U **Insulator defect detection method based on angle restoration and similarity matching** [14351-110]

14351 2V **A vision-guided robotic action generation framework for industrial assembly task** [14351-77]

14351 2W **An intelligent detection method for wiring harness components based on a dynamic focusing mechanism** [14351-5]

14351 2X **Surface partitioning and traveling salesman problem-based method for spray path planning of complex workpieces** [14351-46]

14351 2Y **CID-YOLO: enhancing yolo for cotton impurity detection** [14351-8]

14351 2Z **Zero-shot industrial defect detection via domain-adaptive physical priors and semantics-guided Bayesian inference** [14351-70]

14351 30 **AssistAD: assistant-aided distillation for industry image anomaly detection** [14351-103]

14351 31 **Real-time cable defect detection algorithm based on MCN-YOLO** [14351-22]

14351 32 **A lightweight detection network for immature green citrus via frequency domain sparsity and parameter sharing** [14351-34]

14351 33 **AISC-Former: a physics-informed adaptive hybrid sparse attention model for fault diagnosis** [14351-21]

14351 34 **Improved YOLOv5s rail tiny defect detection method integrating CA attention mechanism and transposed convolution** [14351-2]