

PROCEEDINGS OF SPIE

Second International Conference on Image Processing and Deep Learning (IPDL 2026)

Jun Wang
Lu Leng
Editors

6–8 March 2026
Chongqing, China

Organized by
Chongqing University (China)
Tianjin Normal University (China)

Sponsored by
AEIC—Academic Exchange Information Centre (China)

Published by
SPIE

Volume 14181

Proceedings of SPIE 0277-786X, V. 14181

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 *Second International Conference on Image Processing and Deep Learning (IPDL 2026)*, edited by Jun Wang, Lu Leng, Proc. of SPIE 14181, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9798902324164

ISBN: 9798902324171 (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

ix *Conference Committee*

IMAGE PROCESSING AND OBJECT DETECTION

- 14181 02 **Automatic recognition and classification system for multilingual signage in Dongguan historical and cultural districts based on YOLO-v8 and attention mechanism [14181-85]**
- 14181 03 **Attention-enhanced YOLOv11N for traffic small object detection on VisDrone dataset [14181-29]**
- 14181 04 **Improved YOLOv11 algorithm for rice disease detection [14181-84]**
- 14181 05 **An improved YOLOv11 small object detection algorithm for UAV imagery [14181-19]**
- 14181 06 **Road traffic sign detection network combining compound scaling and attention mechanism [14181-6]**
- 14181 07 **Research on deep learning-based face detection and recognition [14181-47]**
- 14181 08 **A lightweight QR code detection method for printed circuit boards based on improved YOLOv8n [14181-73]**
- 14181 09 **Automatic water column deflection method of fire control radar based on image peak recognition [14181-86]**
- 14181 0A **Research on a movie product placement detection model integrating attention mechanism and edge features [14181-32]**
- 14181 0B **GLDNet: a global-local collaborative and dual-domain interactive network for camouflaged object detection [14181-46]**
- 14181 0C **GAT-LSD: geometry-aligned transformer with gated attention for robust line segment detection [14181-87]**
- 14181 0D **Real-time binocular tennis line-call via distributed geometry-constrained event streaming [14181-76]**
- 14181 0E **Enhancing field-based tea disease detection via generative diffusion priors with semantic consistency and task-specific YOLOv26 adaptation [14181-54]**
- 14181 0F **Fruit recognition research based on AIoU loss function [14181-38]**

- 14181 OG **An underwater sea cucumber detection method based on an improved YOLOv5 algorithm** [14181-37]
- 14181 OH **Edge-deployable parameter-efficient YOLOv8n for robust QR code detection in PCB inspection** [14181-34]
- 14181 OI **CSA-RENet: replacing bi-directional calibration with cross-spatial attention for efficient RGB-event moving object detection** [14181-61]
- 14181 OJ **VIM-DETR: efficient anti-UAV detection via hybrid vision Mamba and frequency-aware fusion** [14181-35]
- 14181 OK **Shoe upper pattern contour recognition based on the canny operator and pattern nesting positioning algorithm** [14181-21]
- 14181 OL **Deep learning-driven optimization technology for UAV low-altitude aerial survey** [14181-75]
- 14181 OM **A cascaded feature matching network based on spatially constrained bipartite graph optimization** [14181-60]
- 14181 ON **Mixed precision matrix processing array design and implementation for AI SAR image processing** [14181-11]
- 14181 OO **Intelligent translation quality assessment model for tourism scenic area linguistic landscape based on transformer and multimodal fusion** [14181-92]
- 14181 OP **Research on short text semantic similarity method with multi-perspective alignment: a Chinese dual-tower model combined with terminology anchor constraints** [14181-89]
- 14181 OQ **PIT-Net: disentangling local nuance in global context for multivariate time series classification** [14181-56]
- 14181 OR **ENGformer: an enhanced non-linearly gating model for long-term time series forecasting** [14181-65]

DEEP LEARNING FOR VISUAL PERCEPTION AND ENHANCEMENT

- 14181 OS **Research on the analysis of flight personnel's spiral staircase training posture based on visual detection** [14181-81]
- 14181 OT **Automatic fruit recognition using a lightweight CNN model** [14181-7]
- 14181 OU **Research on multiperson sports target detection based on YOLO26s-pose** [14181-79]
- 14181 OV **Research on lightweight crop disease recognition based on improved MobileViT-ECA and its technology adoption potential** [14181-71]
- 14181 OW **A leakage-aware and lightweight framework for root vegetable classification** [14181-70]

- 14181 0X **Research on wheat disease classification based on LMSCAN** [14181-14]
- 14181 0Y **Enhancing spatiotemporal modeling for egocentric 3D human pose estimation** [14181-15]
- 14181 0Z **Skeleton action recognition with adaptive viewpoint calibration** [14181-49]
- 14181 10 **Intent-aware contrastive learning for trajectory prediction under varying observation lengths** [14181-68]
- 14181 11 **Road surface type recognition based on multiblock classification result fusion using GC-DSF** [14181-25]
- 14181 12 **Image dehazing research based on MB-TaylorFormerV2 model** [14181-41]
- 14181 13 **Window-view visual quality assessment using viewshed images from multisource point clouds** [14181-45]
- 14181 14 **CARU-Net: a deep learning framework with CBAM-enhanced residual blocks for high-resolution precipitation nowcasting** [14181-36]
- 14181 15 **A lightweight diffusion-based image generation method via low-rank adaptation** [14181-2]
- 14181 16 **EEM: an entropy-TOPSIS based comprehensive evaluation framework for image edge detection algorithms** [14181-83]
- 14181 17 **Color transfer for omnidirectional images** [14181-4]
- 14181 18 **Research on super-resolution technology of infrared images based on deep learning** [14181-78]
- 14181 19 **Multihop video temporal grounding via hierarchical feature downsampling and explicit timestamping** [14181-13]
- 14181 1A **Enhanced low-light image restoration with fusion attention and multiscale feature fusion** [14181-67]
- 14181 1B **Unsupervised 3D point cloud semantic segmentation via adaptive multidomain feature enhancement** [14181-31]
- 14181 1C **Research and implementation of an NVMe-GPU heterogeneous parallel framework for gigapixel image stitching** [14181-23]
- 14181 1D **Boosting transferability of adversarial examples through spatial-frequency synergistic attack** [14181-8]
- 14181 1E **Vision-conditioned diffusion steganography with enhanced controllability and robustness** [14181-17]

- 14181 1F **MEME modality alignment based on positional encoding and multistrategy cross-attention fusion** [14181-28]
- 14181 1G **Dual local enhancement transformer for street scene semantic segmentation** [14181-52]
- 14181 1H **Design of image de-jittering algorithm based on Sobel operator gradient projection** [14181-53]
- 14181 1I **A comprehensive survey on deep-learning-based image super-resolution computation** [14181-91]
- 14181 1J **Multiview 3D reconstruction of immovable cultural relics based on AMA-PatchmatchNet** [14181-74]

MEDICAL AND INDUSTRIAL IMAGE ANALYSIS

- 14181 1K **Color-to-composition mapping in blood images via PSO-based predictive modeling** [14181-33]
- 14181 1L **An image editing method for substation breather defects based on dual-branch inversion and defect feature prompt** [14181-59]
- 14181 1M **CATIA-DSE-based data cleaning and repair for wing block scanning** [14181-50]
- 14181 1N **High-speed dynamic image acquisition and synchronization system for railway freight inspection** [14181-9]
- 14181 1O **Based on improved strip convolution intelligent reading method for pointer-type instrument** [14181-3]
- 14181 1P **Establishment of aircraft infrared engineering model based on measured data** [14181-72]
- 14181 1Q **Dynamic feature fusion enhancement of YOLOv8 for intelligent pavement distress detection** [14181-22]
- 14181 1R **FSTW-Net: a few-shot SAM-2-based framework for fine-grained multitype tool-wear segmentation** [14181-82]
- 14181 1S **Research on surface inspection algorithm for aluminum capacitors based on feature enhancement and knowledge distillation** [14181-66]
- 14181 1T **Aluminum profile defect detection algorithm based on improved RT-DETR** [14181-16]
- 14181 1U **Research on steel surface defect detection technology based on improved YOLO11** [14181-40]
- 14181 1V **Deep learning algorithm for predicting Gleason grade group based on prostate MRI image set** [14181-1]

- 14181 1W **FDE-Net: an asphalt pavement distress detection algorithm based on frequency domain enhancement** [14181-12]
- 14181 1X **Optimizing lightweight industrial defect segmentation: an empirical study on input resolution and negative sample distribution** [14181-5]
- 14181 1Y **DS-FNet: a dynamic-static collaborative network for brain disease diagnosis** [14181-93]
- 14181 1Z **Research on lightweight rust diagnosis of transmission line dampers based on Rust-MViT** [14181-26]
- 14181 20 **BICBAF-UX-Net: bidirectional cross-guided bilinear attention fusion for 3D MRI esophageal tumor segmentation** [14181-42]
- 14181 21 **A multidomain feature enhancement method for surface defect detection in cold-rolled strip steel under low-light conditions** [14181-44]
- 14181 22 **BENet: a boundary-enhanced network for medical image segmentation** [14181-57]
- 14181 23 **Elimination of rotational errors in Moiré-based lithography alignment technology** [14181-51]
- 14181 24 **Drug-disease association prediction based on multifeature processing and cross-modal attention fusion** [14181-24]
- 14181 25 **Research on semantic segmentation of cataract surgery video** [14181-55]
- 14181 26 **SGNet: structure-aware group-supervised network for spleen segmentation** [14181-20]
- 14181 27 **Improving generalizability in retinal vessel segmentation through dataset fusion and systematic model benchmarking** [14181-90]