

PROCEEDINGS OF SPIE

International Conference on Advances in Computer Vision Research and Applications (ACVRA 2026)

**Kamarul Hawari
João Manuel R. S. Tavares
Guanghui (Richard) Wang**
Editors

**6–8 February 2026
Wuhan, China**

Organized by
Wuhan University of Technology (China)

Published by
SPIE

Volume 14247

Proceedings of SPIE 0277-786X, V. 14247

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 *International Conference on Advances in Computer Vision Research and Applications (ACVRA 2026)*, edited by Kamarul Hawari, Joao Manuel R. Tavares, Guanghui (Richard) Wang, Proc. of SPIE 14247, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9798902325604

ISBN: 9798902325611 (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

OBJECT DETECTION AND SEGMENTATION

14247 02	DilatedToothSegNet++: topology-aware graph neural network for robust 3D dental mesh segmentation [14247-102]
14247 03	Medical image segmentation network based on multigranularity feature extraction [14247-33]
14247 04	DARB-Net: a dynamic attentive residual boundary network for boundary-sharpened automatic polyp segmentation [14247-101]
14247 05	Penetrative AI target detection based on knowledge enhancement [14247-1]
14247 06	Ship fire detection in haze weather based on improved YOLOv11 [14247-49]
14247 07	AGDM-PiMAE and AFP-MonoDETR: a monocular 3D object detection method based on multimodal pretraining and adaptive feature projection [14247-75]
14247 08	Detection of floating objects on water surfaces in complex environments based on YOLOv9c [14247-160]
14247 09	MEUNet: mamba-empowered UNet with edge refinement for remote sensing image segmentation [14247-69]
14247 0A	A method for geometric parameter measurement of railway rigid catenary contact wires based on image segmentation [14247-128]
14247 0B	Water body recognition in remote sensing imagery based on YOLOv5 and DeepLabV3 [14247-5]
14247 0C	One-shot multiscale detection for unseen objects via feature-Level hybrid correlation [14247-96]
14247 0D	BGSNet: boundary-aware and global-context network for road scene semantic segmentation [14247-105]
14247 0E	LFE-RTDETR: an aerial object detection method incorporating Gaussian filtering and efficient multiscale attention [14247-109]
14247 0F	Enhancing small object detection in UAV imagery via detail-aware feature extraction and adaptive cross-scale fusion [14247-133]

- 14247 OG **Point cloud precise registration and grinding feature detection method for workpieces with holes based on hole-position constraints** [14247-44]
- 14247 OH **An occlusion-aware 3D lane detection method based on visibility supervision** [14247-40]
- 14247 OI **Enhancing open-world object detection with multiscale feature adaptation and compact feature encoding** [14247-35]
- 14247 OJ **Research and implementation of remote sensing image object detection algorithm based on RT-DETR** [14247-17]
- 14247 OK **Surface defect detection in surfacing welds based on YOLOv11-R** [14247-82]
- 14247 OL **CMYOLO: visible-infrared cross Mamba fusion for round-the-clock object detection** [14247-6]
- 14247 OM **Small target detection algorithm for road damage based on UAV images** [14247-38]
- 14247 ON **An underwater small target detection algorithm based on YOLOv26** [14247-56]
- 14247 OO **A lightweight fog-aware object detector with attention and deformable convolution** [14247-155]
- 14247 OP **LMSTA: a lightweight, task-aligned detector in UAV imagery** [14247-29]
- 14247 OQ **Hypergraph-enhanced relation perception for robust open-vocabulary detection** [14247-2]
- 14247 OR **A universal disease detection model for multiscenario smart agriculture: research on improved TSBA-YOLO based on foundation model mechanisms and adaptive feature fusion** [14247-7]
- 14247 OS **BMTransUNet: boundary-enhanced multimodal transformer for remote sensing semantic segmentation** [14247-123]
- 14247 OT **YOLO-SARLite: efficient lightweight YOLOv8 framework for SAR ship detection** [14247-16]
- 14247 OU **Ship detection in defocused SAR images via feature-level adversarial transfer learning** [14247-89]
- 14247 OV **Prior-driven keypoint detection and bare footprint reconstruction for forensic investigation** [14247-79]
- 14247 OW **YOLOv11n-DWRHS: a novel high-precision algorithm for water surface object detection** [14247-84]
- 14247 OX **LP-GroundingDINO: learnable prompts for fine-grained language-guided detection** [14247-150]

- 14247 0Y **An improved YOLOv8-based algorithm for steel surface defect detection in complex industrial scenes** [14247-14]
- 14247 0Z **Small object detection in UAV aerial images based on an improved YOLOv8** [14247-3]
- 14247 10 **Underwater garbage detection algorithm based on improved YOLOv11** [14247-129]
- 14247 11 **RETA-YOLO11: an attention-enhanced YOLO model for UAV-based weed detection** [14247-97]
- 14247 12 **A small object segmentation algorithm based on cross-level feature interaction** [14247-19]
- 14247 13 **SAR incremental object detection via dual-teacher knowledge distillation with dynamic frequency selection** [14247-98]
- 14247 14 **Weld-YOLO: a YOLO11-based model for high accuracy defect detection in weld x-ray images** [14247-68]
- 14247 15 **Research on defect detection technology for pressure vessel welds based on improved U-Net** [14247-77]
- 14247 16 **Frequency-domain decoupling-based channel attention for small infrared target detection** [14247-126]
- 14247 17 **Efficient instance segmentation of UAV components via structure-aware feature modeling** [14247-74]
- 14247 18 **Semantic-artifact collaborative framework for generalizable synthetic image detection** [14247-73]
- 14247 19 **Image target detection for unmanned aerial vehicle autonomous landing based on the improved YOLOv8 algorithm** [14247-130]
- 14247 1A **A multibranch discrepancy consistency learning network for semi-supervised medical image segmentation** [14247-94]
- 14247 1B **Optimization of a 3D U-Net based model for precise segmentation of multimodal MRI brain tumors** [14247-119]
- 14247 1C **YOLOv8 based on attention and multiscale feature fusion** [14247-13]

IMAGE PROCESSING, ENHANCEMENT, AND VISUAL APPLICATIONS

- 14247 1D **ODCSNet: efficient compressed sensing reconstruction via overlapping sampling and diffusion-based refinement** [14247-115]
- 14247 1E **Research on outdoor point cloud completion methods based on deep learning** [14247-148]

- 14247 1F **HSM-SCNet: a hard sample mining-based semantic scene completion network** [14247-147]
- 14247 1G **Strong reflection suppression on shear-sensitive liquid crystal surfaces using full-field uniform illumination** [14247-23]
- 14247 1H **Design of a multimodal medical image intelligent analysis system for radiology** [14247-81]
- 14247 1I **AI image generation empowering visual communication design efficiency** [14247-137]
- 14247 1J **Image generation using computer graphics and enhanced CycleGAN for industrial vision detection of dynamic free objects** [14247-4]
- 14247 1K **High-fidelity plant 3D reconstruction and data augmentation via prompt-driven semantic priors and surfel-based radiance fields** [14247-54]
- 14247 1L **Characterization of laser beam spot size via two-dimensional Gaussian fit** [14247-153]
- 14247 1M **A new cardinal many-knot spline wavelet for image edge detection application** [14247-22]
- 14247 1N **Research on image dehazing algorithm based on semi supervised learning in foggy scenes** [14247-144]
- 14247 1O **Vision-guided PID path tracking for mobile robots in complex environments** [14247-145]
- 14247 1P **Physics-guided learning of inverse reconstruction for structured illumination microscopy** [14247-64]
- 14247 1Q **Research on acupoint localization method for a moxibustion robot based on depth vision** [14247-80]
- 14247 1R **A cascade Transformer-CNN network for efficient edge detection** [14247-11]
- 14247 1S **Synergistic optimization of fidelity and geometric quality for neural geometry field-based mesh simplification** [14247-118]
- 14247 1T **Bidirectional video inpainting via flow-guided feature propagation** [14247-20]
- 14247 1U **Robust hybrid noise removal algorithm based on enhanced dual path adaptive bilateral filtering** [14247-58]
- 14247 1V **Research on precision grasping technology of micro materials by manipulator based on machine vision** [14247-132]
- 14247 1W **An anti-frequency offset synchronization algorithm for OFDM systems based on CAZAC sequences** [14247-121]
- 14247 1X **Application research of low-latency transmission and real-time rendering technology in live broadcast scenario FVS** [14247-120]

- 14247 1Y **Dynamic target perception and motion prediction for mobile manipulators** [14247-122]
- 14247 1Z **High-frequency guided gated feature fusion for underwater image enhancement**
[14247-104]
- 14247 20 **Elderly and disabled care robot system based on VLA large model** [14247-143]
- 14247 21 **Phase error compensation method combining polynomial modeling and LUT correction**
[14247-60]
- 14247 22 **MS-WDNet: a dual-branch image dehazing network with multiscale feature selection and
frequency-domain enhancement** [14247-127]
- 14247 23 **Key frame localization in thyroid ultrasound videos via visual temporal modeling** [14247-15]
- 14247 24 **LMambaFusion: linear Mamba for unsupervised SPECT and MRI image fusion** [14247-70]
- 14247 25 **Research on an underwater target obstacle removal algorithm** [14247-57]
- 14247 26 **Optimization of diffusion model-based edge detection** [14247-52]
- 14247 27 **Image motion deblurring based on prior networks and blur region segmentation** [14247-99]
- 14247 28 **Time-series InSAR deformation monitoring and cause analysis of transportation
infrastructure in Shenzhen** [14247-151]
- 14247 29 **A smart brachytherapy applicator with integrated PVDF pressure sensor array for real-time
tissue contact monitoring** [14247-100]
- 14247 2A **A novel effective multimodal fusion model (PeliCo) for USV water surface multitask
perception** [14247-95]
- 14247 2B **EX-GS: extending Gaussian splatting to 4D via hyper-ellipsoidal projection** [14247-156]
- 14247 2C **Design and application analysis of a plateau lake aquatic ecological restoration and
governance platform based on 3D visualization and digital twin technology** [14247-12]
- 14247 2D **Research on a special algorithm in path planning** [14247-76]
- 14247 2E **Robust rail straightness measurement using line-structured light point clouds** [14247-106]
- 14247 2F **ProtoDiff: a prototype-guided conditional diffusion model for Oracle handprint data
augmentation** [14247-8]
- 14247 2G **AIMS: an adaptive integration of multisensor measurements for quadrupedal robot
localization** [14247-83]

- 14247 2H **Influence of surface image characteristics on friction and wear properties of Cu-HAp coatings** [14247-53]
- 14247 2I **Sketch-VSD: controlled variational score distillation for high-fidelity sketch-to-vector synthesis** [14247-88]

VISUAL RECOGNITION AND ANALYSIS

- 14247 2J **HKA-Net: a hand pose estimation model based on multiple aggregation strategy and attention hourglass network** [14247-114]
- 14247 2K **Real-time detection and behavioral analysis of dynamic scene targets based on YOLO and R (2+1) fusion** [14247-63]
- 14247 2L **Attention-enhanced image classification network based on VGG19** [14247-112]
- 14247 2M **Research on video action recognition based on EfficientNet-B2 and temporal attention mechanism** [14247-42]
- 14247 2N **Multiscale hybrid transformer network for no-reference image quality assessment** [14247-41]
- 14247 2O **A quality assessment scheme and application for video transmission and processing systems** [14247-131]
- 14247 2P **Adaptive channel shift video behavior recognition algorithm based on residual network** [14247-138]
- 14247 2Q **A 3D scene recognition algorithm leveraging point cloud saliency** [14247-10]
- 14247 2R **Cross-modal content understanding and interactive interface auto-generation based on vision-language pretrained models** [14247-157]
- 14247 2S **Lightweight multimodal visual question answering model with integrated hierarchical reasoning and knowledge graphs** [14247-142]
- 14247 2T **A GNN-enabled framework for correlated software defect analysis and prediction** [14247-139]
- 14247 2U **PoseTracker: dynamic 3D human pose tracking via multiview point cloud fusion** [14247-159]
- 14247 2V **Graph attention-enhanced multilabel cross-modal hash retrieval** [14247-18]
- 14247 2W **Enhanced template-search attention and fusion network for RGBT tracking** [14247-111]
- 14247 2X **Effects of visual guidance facilities on drivers' eye movements in expressway tunnel transition zones** [14247-78]

- 14247 2Y **YOLO-based UAV detection and classification for anti-drone early warning systems: a real-time approach** [14247-149]
- 14247 2Z **Remote sensing image classification and visual recognition based on multimodal information fusion** [14247-30]
- 14247 30 **MWI-RAG: an agricultural weed identification method based on retrieval augmented generation and multimodal fusion** [14247-46]
- 14247 31 **Enhanced YOLOv8n-PaddleOCR collaboration for real-time vehicle type and license plate recognition in intelligent transportation** [14247-110]
- 14247 32 **Lightweight multiscale feature fusion for flower recognition** [14247-107]
- 14247 33 **A knowledge-driven framework for illegal parking recognition in urban road scenarios** [14247-85]
- 14247 34 **Multistage spatio-temporal fusion network for action recognition** [14247-66]
- 14247 35 **Low-cost injury prevention: real-time fatigue and stability analysis** [14247-9]
- 14247 36 **Dynamic token sparsification and cross-path attention-based dual-path video action recognition method** [14247-51]
- 14247 37 **Robust classification of Miao embroidery patterns based on shape graph structure alignment** [14247-141]
- 14247 38 **EnhancedLocalizer: fine-grained visual-semantic modeling with pseudo-event guidance for video moment retrieval** [14247-158]
- 14247 39 **Distribution of carbon, phosphorus, and nitrogen in anaerobic codigestion systems through machine learning models** [14247-134]
- 14247 3A **IAGM: interaction and aggregation guided module for RGBT tracking** [14247-62]
- 14247 3C **Uniform-split multiprototype covariance metric networks for few-shot learning** [14247-135]
- 14247 3D **Research on an evaluation-assistance model based on natural language processing and graph embedding techniques** [14247-39]
- 14247 3E **Contrastive learning for continuous touch-based authentication** [14247-72]
- 14247 3F **Dual-branch feature distillation network for low-resolution face recognition in wide-angle surveillance** [14247-65]
- 14247 3G **Radar embrace transformer: an end-to-end framework for radar target tracking** [14247-90]