

PROCEEDINGS OF SPIE

International Conference on Computer Vision and Augmented Reality (CVAR 2026)

**Kamarul Hawari
Kannimuthu Subramanian**
Editors

**27–29 March 2026
Changsha, China**

Organized by
Zhengzhou University (China)

Sponsored by
Hunan Talent Group (China)
AEIC Academic Exchange Information Centre (China)

Published by
SPIE

Volume 14265

Proceedings of SPIE 0277-786X, V. 14265

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 Computer Vision and Augmented Reality (CVAR 2026)*, edited by Kamarul Hawari, Kannimuthu Subramanian, Proc. of SPIE 14265, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9798902326199

ISBN: 9798902326205 (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

vii *Conference Committee*

IMAGE PROCESSING AND ENHANCEMENT

- 14265 02 **Perceptual color difference measurement for photographic images** [14265-17]
- 14265 03 **Text-driven stylized image generation via scene decomposition and consistency constraints** [14265-28]
- 14265 04 **Research on real-to-anime facial style transfer based on improved CycleGAN** [14265-10]
- 14265 05 **Nighttime road video enhancement based on Poisson regularization and adaptive CLAHE with CPU parallel acceleration optimization** [14265-7]
- 14265 06 **Semi-supervised image dehazing network with dual-branch learning** [14265-39]
- 14265 07 **Multi-focus image fusion network based on adaptive graph representation and graph-guided spatial refinement** [14265-53]
- 14265 08 **No-reference video quality assessment model based on saliency semantic enhancement** [14265-33]
- 14265 09 **Coreset selection via training dynamics and graph message passing** [14265-12]
- 14265 0A **Feed-forward 3D Gaussian splatting for long-sequence reconstruction** [14265-1]
- 14265 0B **Fast-FAAGC: accelerating geodesic feature augmentation via distribution-aligned initialization** [14265-38]
- 14265 0C **Multi-branch parallel image dehazing network based on multi-scale residuals and fusion attention** [14265-4]
- 14265 0D **CAPS: candidate-aware prompt selection for egocentric video class incremental learning** [14265-50]
- 14265 0E **Fusion of dynamic topology and multi-scale temporal modeling for human action recognition** [14265-25]

- 14265 OF **A hybrid approach to off-screen object visualization and user behavior mapping in mobile augmented reality: a comprehensive study on adaptive cues and in-situ analytics** [14265-51]
- 14265 OG **RGRPD: a local 3D descriptor using reliability guided weighting with robust projection pooling techniques** [14265-20]

OBJECT DETECTION AND SEGMENTATION

- 14265 OH **Optimization of image semantic segmentation based on DeepLabv3+ algorithm** [14265-49]
- 14265 OI **An improved lightweight semantic segmentation model for forest fire source based on U-Net** [14265-24]
- 14265 OJ **Accurate retinal vessel image segmentation with coarse-to-refine U-shaped network** [14265-18]
- 14265 OK **CSDPF-Net: a lightweight local dual-path fusion network for surface defect detection** [14265-44]
- 14265 OL **DSRFNet: a dual-stream resonant fusion network for real-time semantic segmentation** [14265-36]
- 14265 OM **LDO-DETR: a discharge outlet detection algorithm for land-based scenarios** [14265-19]
- 14265 ON **Water pollution level detection based on computer vision technology** [14265-23]
- 14265 OO **The impact of scanning methods on object detection performance with Mamba and RWKV** [14265-2]
- 14265 OP **Monocular 3D object detection based on fusion of dual-target bounding boxes and relative positions** [14265-14]
- 14265 OQ **CBAM-UNet: a dual-attention guided network for fine-grained segmentation of hepatic vessels and tumors** [14265-16]
- 14265 OR **Refined identification of concrete bridge cracks and intelligent inversion of full-field stress based on deep cascade networks** [14265-43]
- 14265 OS **Motion object detection based on adaptive frame difference method and its FPGA implementation** [14265-35]
- 14265 OT **CGES_YOLO: a discriminative detail enhancement model for small target detection in optical remote sensing images** [14265-21]

- 14265 0U **LCE-Det: a language-driven contrast-enhanced detection network for oil stains in power inspection scenarios** [14265-26]
- 14265 0V **Event-based semantic segmentation of space targets with SNN-HRNet** [14265-30]
- 14265 0W **Deep learning-based traffic sign detection and recognition: algorithms, system design, and emerging trends** [14265-8]
- 14265 0X **A vision transformer U-Net for soot instance segmentation in transmission electron microscopy images** [14265-47]
- 14265 0Y **Prohibited item detection in x-ray security images based on an improved YOLOv11** [14265-37]
- 14265 0Z **Optimization of small target detection algorithm in complex UAV perspectives based on improved YOLOv11** [14265-54]

REMOTE SENSING AND INDUSTRIAL APPLICATIONS

- 14265 10 **Identification of potential geological hazard risks in mountain gas fields based on SBAS-InSAR technology** [14265-46]
- 14265 11 **A vision-based method for analyzing the posture positioning of welding clamp electrode bar** [14265-45]
- 14265 12 **Application of 3D geographic information system (3D-GIS) in deformation analysis of the Yaken anticline basin based on unmanned aerial vehicle photogrammetry data** [14265-55]
- 14265 13 **Vehicle detection method from UAV perspective based on improved YOLO** [14265-5]
- 14265 14 **High-precision building damage detection method based on optical-SAR heterogeneous data fusion and a UNet++-based framework** [14265-40]
- 14265 15 **A maturity detection model for greenhouse tomato** [14265-34]
- 14265 16 **Research on visual positioning technology for pin assembly in aluminium alloy enclosures** [14265-27]
- 14265 17 **Research on YOLOv13 based on U-Net_V2 backbone network for anomaly detection in tube railway** [14265-57]
- 14265 18 **Lightweight YOLOv5 for LKJ data management: loss optimization, knowledge distillation, INT8 quantization, and ZYNQ7100 deployment** [14265-32]
- 14265 19 **YOLOv8n-based surface residue detection for metal molds of prefabricated concrete components** [14265-11]

- 14265 1A **A lightweight road distress detection method for UAV aerial imagery** [14265-29]
- 14265 1B **PortVision: spatiotemporal activity modeling for port scene understanding** [14265-52]