

PROCEEDINGS OF SPIE

Sixth International Conference on Computer Vision and Pattern Analysis (ICCPA 2026)

**Xinnian Wang
Weiwei Zhang**
Editors

**8–10 May 2026
Dalian, China**

Organized by
Dalian Artificial Intelligence Industry Association (China)

Sponsored by
AEIC—Academic Exchange Information Centre (China)

Published by
SPIE

Volume 14296

Proceedings of SPIE 0277-786X, V. 14296

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 *Sixth International Conference on Computer Vision and Pattern Analysis (ICCPA 2026)*, edited by Xinnian Wang, Weiwei Zhang, Proc. of SPIE 14296, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9798902326922

ISBN: 9798902326939 (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 3D GEOMETRIC PERCEPTION

- 14296 02 **Real-time digital image correlation analysis based on 3D-printed anisotropic stripe speckles** [14296-43]
- 14296 03 **Low-light image enhancement network based on multiscale dynamic frequency domain optimization** [14296-9]
- 14296 04 **A dual-camera cooperative localization method for PTZ with unknown spatial position** [14296-5]
- 14296 05 **A unified geometry-aware framework for calibration-free cross-modal image-to-point cloud registration** [14296-17]
- 14296 06 **Study on monocular ranging and scale recovery method based on UniK3D** [14296-45]
- 14296 07 **A physically based transient simulation framework for non-line-of-sight imaging** [14296-23]
- 14296 08 **Physically-informed neural fields for non-line-of-sight transient reconstruction** [14296-37]
- 14296 09 **A robust point cloud registration method based on maximal clique screening and RANSAC-ICP** [14296-3]
- 14296 0A **Geometry-guided neural flow for perspective-robust crowd clustering** [14296-20]
- 14296 0B **High-quality underwater 3D Gaussian splatting reconstruction with multiview consistency constraints** [14296-28]
- 14296 0C **Real-time image super-resolution for anime: a deployment-oriented network with a depth-oriented design** [14296-8]
- 14296 0D **Super-resolution reconstruction of ocean dissolved oxygen via conditional diffusion modeling with auxiliary physical fields** [14296-33]
- 14296 0E **Geometry-constrained BEV feature alignment for vehicle-infrastructure cooperative LiDAR perception** [14296-39]
- 14296 0F **PDSC-Net: a physics-guided dual-domain spectral collaborative network for hyperspectral image super-resolution** [14296-30]

14296 OG **Sparse-cohesion: mitigating quality degradation in multiview synthesis via periodic inter-frame synergy** [14296-51]

VISUAL RECOGNITION AND APPLICATIONS

14296 OH **Topology-aware deep learning for brain tumor segmentation in MRI** [14296-25]

14296 OI **Identifying crucial facial areas with gated fusion for early ASD detection** [14296-47]

14296 OJ **Self-supervised learning-driven medical image recognition: a mask image modeling method for scenes with sparse annotations** [14296-15]

14296 OK **Bulk pile shovel region segmentation under physics-driven loading simulation augmentation: quality evaluation and sim-to-real domain gap analysis** [14296-24]

14296 OL **Research on highway area recognition in remote sensing images based on deep learning** [14296-32]

14296 OM **Leveraging side-view feature guidance for robust vascular volume analysis in OCTA** [14296-27]

14296 ON **SemanticAdapter: parameter-efficient adaptation for label-efficient 3D point cloud semantic segmentation** [14296-31]

14296 OO **A dual-enhanced YOLO11 for weed detection in cotton fields** [14296-14]

14296 OP **Lightweight semantic parsing for liquid level measurement** [14296-13]

14296 OQ **Research on pavement distress detection based on the YOLO26 model** [14296-48]

14296 OR **Research on anti-forgetting training strategies for open-vocabulary object detection systems** [14296-2]

14296 OS **Intelligent identification of cut tobacco types based on improved YOLOv4-tiny** [14296-40]

14296 OT **Improved YOLO11n for traffic light small object detection** [14296-16]

14296 OU **Scale-normalized YOLO for small-object detection in large-oblique-angle UAV imagery** [14296-42]

14296 OV **Research on remote sensing landslide image recognition method based on dual-model ensemble swin transformer** [14296-49]

14296 OW **Lightweight dynamic gesture recognition method for complex background based on improved transformer** [14296-34]

14296 OX **Cattle body size measurement based on MSAF-PointNet++** [14296-21]

14296 OY **Transformer-based heterogeneous attention and multiscale feature extraction network for video saliency prediction** [14296-7]

- 14296 07 **CFRM-MERNet: a multimodal fusion network with cross-modal feature matching and reconstruction for micro-expression recognition** [14296-22]
- 14296 10 **Based on the improved YOLO11n method for automatic reading of pointer-type instruments** [14296-50]
- 14296 11 **Improved lightweight YOLOv11 for foreign object detection in transmission line inspection** [14296-12]
- 14296 12 **Design of image recognition and monitoring system for OpenHarmony agricultural greenhouses incorporating deep learning** [14296-19]
- 14296 13 **LSA-YOLOv13n-based intelligent detection method for asparagus** [14296-1]
- 14296 14 **Collision-free dual-robot cooperative navigation via value decomposition networks and a sparsity-regularized lightweight CNN** [14296-35]
- 14296 15 **A multimodal BEV 3D object detection method with depth uncertainty and geometric saliency** [14296-10]
- 14296 16 **Research on the HSV space PAUT image grab cut segmentation algorithm** [14296-11]
- 14296 17 **Computer vision-based microplastic identification in water and analysis of environmental microbial mechanisms** [14296-46]