

PROCEEDINGS OF SPIE

International Conference on Computer Vision and Image Computing (CVIC 2025)

Luis Gomez
Zahid Akhtar
Editors

21–23 November 2025
Hong Kong, China

Organized by
Xi'an Jiaotong University (China)
Macao Polytechnic University (China)

Sponsored by
AEIC—Academic Exchange Information Centre (China)

Published by
SPIE

Volume 14070

Proceedings of SPIE 0277-786X, V. 14070

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 Image Computing (CVIC 2025)*, edited by Luis Gomez, Zahid Akhtar, Proc. of SPIE 14070, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9798902320999

ISBN: 9798902321002 (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*

VISUAL DETECTION AND RECOGNITION

- 14070 02 **Digital tool for counting coffee plants and economic study of alternative crops** [14070-52]
- 14070 03 **Zero-shot assembly anomaly detection based on virtual-real difference comparison**
[14070-5]
- 14070 04 **Drone-based road disease detection** [14070-30]
- 14070 05 **Robust vehicle re-identification under strong illumination and haze conditions via multi-scale feature enhancement** [14070-59]
- 14070 06 **GA-YOLO: an improved YOLOv8-based model for small target detection in UAVs** [14070-21]
- 14070 07 **Tiny object detection algorithm based on edge information reconstruction** [14070-8]
- 14070 08 **Design and implementation of a weed image recognition tool** [14070-35]
- 14070 09 **An improved YOLOv11 for defect detection on steel pipe inner walls** [14070-57]
- 14070 0A **Research on personnel detection in cornfields based on improved YOLOv8n** [14070-25]
- 14070 0B **Research on field weed detection method based on Mc-Yolov8** [14070-14]
- 14070 0C **Research on forest fire smoke object detection based on deep learning** [14070-51]
- 14070 0D **Research on UAV identification and tracking technology for water surface moving targets**
[14070-20]
- 14070 0E **Steel surface defect detection algorithm based on Tcsc-Yolov8** [14070-4]
- 14070 0F **An improved YOLOv8-based framework for small object detection in UAV images**
[14070-55]
- 14070 0G **Research on intelligent detection algorithm for elevator door locks based on improved YOLOv11** [14070-19]

IMAGE PROCESSING AND OPTIMIZATION

- 14070 OH **Self-supervised masked autoencoders pretraining improves pediatric brain tumor segmentation** [14070-15]
- 14070 OI **Multi-layer feature fusion-based data augmentation of defect images** [14070-49]
- 14070 OJ **Mean-teacher-based relationship learning with inter- and intra-category optimization for source-free domain adaptation** [14070-41]
- 14070 OK **Structure feature recognition technology based on extended attribute adjacency graph** [14070-24]
- 14070 OL **A comparative study of encoding schemes for long-sequence skeleton-based action segmentation** [14070-6]
- 14070 OM **Efficient event-based visual recognition with resonate-and-fire spiking neurons** [14070-29]
- 14070 ON **Enhancing LLM inference via training-free anchor compression: a layer-wise representation projection framework** [14070-53]
- 14070 OO **Research on image recognition and edge detection methods based on attention enhancement algorithm** [14070-40]
- 14070 OP **Research on defect imaging and detection technology for brazed copper-aluminum transition clamps based on multiple channels ultrasonic phased arrays** [14070-34]
- 14070 OQ **Extraction of Xiuyan embroidery patterns based on vision transformer** [14070-32]
- 14070 OR **Design of variable channel number for YOLOv3-finy based on FPGA** [14070-54]
- 14070 OS **Research on high-precision visual recognition technology for intelligent waste sorting robots in complex scenarios** [14070-47]
- 14070 OT **CNN-LSTM optimization model for power equipment monitoring and its visualization implementation** [14070-28]
- 14070 OU **Few-shot coral recognition via prototype refinement and segmentation-guided feature enhancement** [14070-42]
- 14070 OV **Self-supervised pretraining-based zero-shot target detection for synthetic aperture sonar** [14070-37]

3D VISION AND SYSTEM APPLICATIONS

- 14070 OW **Creating a multimodal dataset based on non-cooperative spacecraft in blender** [14070-11]

- 14070 0X **Research on the response of different prism materials to ultrasound in an optical reflection system** [14070-26]
- 14070 0Y **Root cause tracing of helicopter quality issues based on Bayesian networks** [14070-23]
- 14070 0Z **Application of 3D point cloud completion for underwater structures** [14070-46]
- 14070 10 **An OPM-MEG source localization evaluation method based on dry phantom** [14070-50]
- 14070 11 **A dynamic dense visual SLAM algorithm based on YOLOv8 and SAM** [14070-10]
- 14070 12 **TurbPredRNN: a study on turbulence-guided radar extrapolation** [14070-13]
- 14070 13 **Design and implementation of visual precision positioning system for substation camera** [14070-27]
- 14070 14 **Quantity calculation of rural construction projects based on backpack 3D laser scanning technology** [14070-31]
- 14070 15 **KungFuCap: monocular martial 3D reconstruction via Gaussian splatting and diffusion-guided view synthesis** [14070-22]
- 14070 16 **Research on interactive motion generation of character animation based on DDPG algorithm** [14070-16]
- 14070 17 **Rasterizing 3D point clouds into depth images via pixel clustering and outlier suppression** [14070-2]
- 14070 18 **Design and implementation of remote control and management system for oilfield workover operations** [14070-44]
- 14070 19 **Remote sensing image captioning: a survey from encoder-decoder model to large vision-language model** [14070-7]