PROCEEDINGS OF SPIE

International Conference on Advances in Computer Vision Research and Applications (ACVRA 2025)

Zhonghong Ou
Hui Liu
Editors

28 February – 2 March 2025
Guangzhou, China

Organized by
Southwest Petroleum University (China)

Published by
SPIE

Volume 13690

Proceedings of SPIE 0277-786X, V. 13690

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 SPIEDigitalLibrary.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 2025)*, edited by Zhonghong Ou, Hui Liu, Proc. of SPIE 13690, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X
ISSN: 1996-756X (electronic)

ISBN: 9781510693265
ISBN: 9781510693272 (electronic)

Published by
SPIE
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone +1 360 676 3290 (Pacific Time)
SPIE.org
Copyright © 2025 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
SPIEDigitalLibrary.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*
ix *Introduction*

IMAGE ENHANCEMENT AND UNDERLYING VISION TECHNOLOGY

13690 02 **Research on image enhancement preprocessing techniques** [13690-49]

13690 03 **Detection of agglomerate fog on highway based on panoramic image compressive sensing** [13690-34]

13690 04 **Slim_Unet: a lightweight segmentation model for core images** [13690-36]

13690 05 **Improved low-light image enhancement algorithm based on the Retinex model** [13690-22]

13690 06 **Research on convolutional feature encoding method and retrieval application of ceramic patterns** [13690-23]

13690 07 **Improved distributed video coding scheme based on auxiliary information** [13690-38]

13690 08 **Visual image defect diagnosis of photovoltaic modules under unbalanced samples** [13690-57]

13690 09 **Unsupervised 3D reconstruction from a single face image** [13690-56]

13690 0A **DGTNet: efficient pathology image classification via dynamic axial graph construction and graph-transformer architecture** [13690-29]

13690 0B **Image steganalysis model based on channel transformation and deep residual networks** [13690-52]

13690 0C **GAN adversarial network-based de-fogging of underwater images** [13690-60]

13690 0D **RDA-IC: a data synthesis approach for enhancing network performance in image classification tasks** [13690-59]

13690 0E **Enhanced visual inspection algorithm for mechanical components based on improved YOLOv8 architecture** [13690-53]

13690 0F **Alzheimer's disease enhance coarse-to-fine deformable MR brain image registration** [13690-27]

13690 0G **Semi-supervised medical image classification method based on category pseudo-labels and weighted consistency loss** [13690-19]

13690 0H **Attention-enhanced feature representation and semantic understanding of large-scale outdoor environment** [13690-31]

13690 0I **Automatic dissection of pituitary tumor based on multiple semantic segmentation algorithms** [13690-45]

INTELLIGENT TARGET DETECTION AND RECOGNITION TECHNOLOGY

13690 0J **Research on illumination type recognition based on deep learning using SDGSAT-1 data** [13690-55]

13690 0K **YOLOv10-Turbo: a YOLOv10-based infrared target detection model for unmanned aerial vehicles** [13690-10]

13690 0L **Security situation prediction in UAV swarm networks based on CNN-BiLSTM** [13690-21]

13690 0M **UTD-YOLO: an improved underwater trash detection method based on YOLO11** [13690-50]

13690 0N **Building change detection method based on multiscale feature fusion and uncertainty awareness** [13690-42]

13690 0O **Lightweight algorithm of bearing surface defect detection** [13690-7]

13690 0P **YOLOv8-MD: an enhanced defect detection model for aluminum alloy automotive body panels using multiscale feature fusion** [13690-35]

13690 0Q **Research on automotive adhesive detection method based on PatchDCT and OpenCV fusion** [13690-64]

13690 0R **Identification of few-shot crop pests and diseases based on improved DN4** [13690-47]

13690 0S **YOLOv8n-LSC: a lightweight submarine cable detection algorithm based on improved YOLOv8n** [13690-20]

13690 0T **A study on the sonar image target recognition method based on improved YOLOv5 model** [13690-16]

13690 0U **Design and implementation of novel deep-learning-based system architecture for remote sensing image object detection and recognition** [13690-24]

13690 0V **Improved multiscale feature fusion and small object detection layer optimization for UAV aerial small object detection based on YOLOv8** [13690-28]

13690 0W **Multipath feature fusion for oriented object detection** [13690-11]

13690 0X **A feature-alignment and attention-fusion network for arbitrary-oriented object detection in remote-sensing images** [13690-13]

13690 0Y **Human activity recognition based on sparse reconstruction** [13690-39]

13690 0Z **Dual-granularity attention aggregation framework for VI-ReID** [13690-5]

13690 10 **Few-shot object detection: hierarchical feature processing module** [13690-15]

13690 11 **Masked face detection algorithm based on YOLOv8** [13690-61]

13690 12 **BSTR: a transformer-based framework for robust and efficient basketball object detection** [13690-4]

13690 13 **Mobile volume measurement method based on binocular vision** [13690-46]

13690 14 **A quality inspection method for steel plate welding based on improved YOLOv5** [13690-17]

3D VISION AND CROSS-DOMAIN APPLICATION SYSTEM

13690 15 **Application of computer vision based on deep learning in dynamic graphic design recognition** [13690-44]

13690 16 **AI chess-playing robot** [13690-62]

13690 17 **Image recognition techniques in autonomous driving: a bibliometrics-based study** [13690-32]

13690 18 **A new approach to angle-preserving mapping for quadrilateral meshes** [13690-18]

13690 19 **Design and experiment of micromirror integrating with angular velocity sensor for binocular camera** [13690-26]

13690 1A **Visual analysis of pesticide usage volatility in Guangdong Province empowered by image vision based on GARCH algorithm** [13690-43]

13690 1B **LLM performance in multimodal learning environments: study of integration of text with visual, audio, and sensor data for holistic decision-making** [13690-30]

13690 1C **Automated quality inspection system based on machine vision** [13690-54]

13690 1D **Design and development of a machine vision-based intelligent inspection system for underground spaces** [13690-33]

13690 1E **An improved obstacle avoidance control algorithm based on dynamic window method** [13690-58]

13690 1F **Impact of dynamic colored light interaction in smart cockpits on user emotion regulation** [13690-51]

13690 1G **Correction of OCR results using large language models** [13690-14]

13690 1H **Research on dimension measurement based on 3D point cloud for helicopter rotor blade** [13690-25]

13690 1I **HR-DVGO: a 3D digital reconstruction model for cultural heritage and historical sites** [13690-48]

13690 1J **Visual experience design for food based on eye movement experiments** [13690-2]