

PROCEEDINGS OF SPIE

# ***International Conference on Computer Vision and Image Processing (CVIP 2024)***

**Xin Xu**  
**Zhenghao Shi**  
*Editors*

**15–17 November 2024**  
**Hangzhou, China**

*Organized by*  
Xi'an University of Technology (China)  
Beihang University (China)  
SouthWest Petroleum University (China)

*Published by*  
SPIE

**Volume 13521**

Proceedings of SPIE 0277-786X, V. 13521

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](http://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 Processing (CVIP 2024)*, edited by Xin Xu, Zhenghao Shi, Proc. of SPIE 13521, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9781510688384

ISBN: 9781510688391 (electronic)

Published by

**SPIE**

P.O. Box 10, Bellingham, Washington 98227-0010 USA

Telephone +1 360 676 3290 (Pacific Time)

[SPIE.org](http://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](http://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](http://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*
- ix *Preface*

---

## COMPUTER VISION AND IMAGE PROCESSING TECHNOLOGIES

---

- 13521 02 **Video object segmentation quality assessment without groundtruth** [13521-8]
- 13521 03 **YOLOv8's layout analysis algorithm integrates multiscale features** [13521-4]
- 13521 04 **Daytime overexposed image processing for autonomous driving based on improved multiscale Retinex with color restoration** [13521-28]
- 13521 05 **Remote sensing image super-resolution reconstruction using hybrid dilated convolution and adaptive pruning** [13521-5]
- 13521 06 **ADLT: a novel method for effective detection of adversarial attacks in image retrieval** [13521-23]
- 13521 07 **CAU-Net: U-shaped medical image segmentation network based on compound attention** [13521-7]
- 13521 08 **B-H-deformable-DETR: H-deformable-DETR model based on Bayesian neural network optimization in few-shot object detection** [13521-17]
- 13521 09 **Efficient LED image reflection removal via mixture of experts** [13521-18]
- 13521 0A **A zero-reference nighttime road image enhancement method** [13521-12]
- 13521 0B **SFPCDDFuse: an enhanced CDDFUSE infrared and visible image fusion model** [13521-29]
- 13521 0C **Recognition of side-scan sonar images under long-tail distribution** [13521-33]
- 13521 0D **Comparison of different masking strategies of MAE** [13521-43]
- 13521 0E **Real-time traffic sign recognition for driving: a hybrid approach integrating efficient Mamba models with dilated convolution** [13521-13]

## WIRELESS COMMUNICATION AND NETWORK TECHNOLOGIES

---

- 13521 OF **Mask R-CNN-based method for recognizing external breakage of transmission lines** [13521-11]
- 13521 OG **Light-aware luminance adaptive enhancement network for RGBT video object detection** [13521-40]
- 13521 OH **Wireless multipath VR audio and video file synchronous transmission technology based on two-way transmission** [13521-38]
- 13521 OI **Research on fine recognition method of electric power equipment safety hazards based on visual processing** [13521-39]
- 13521 OJ **Workpiece inspection system based on zoom lens** [13521-3]
- 13521 OK **TSEP: text spotting based on transformer with bidirectional explicit points sampling** [13521-2]
- 13521 OL **Research on anti-breakout system of power optical fiber communication network based on remote sensing image target detection** [13521-41]
- 13521 OM **DSPR-DoubleU-Net: a network for tongue crack segmentation** [13521-32]
- 13521 ON **Lightweight insulator defect detection algorithm based on improved Deeplabv3+ and YOLOv5s** [13521-9]
- 13521 OO **Deep difference convolutional network for RGBD instance segmentation** [13521-25]
- 13521 OP **Research on instantaneous feature recognition models of finger movement based on surface electromyography signals** [13521-10]
- 13521 OQ **IHD-Former: infrared hybrid denoising transformer for removing infrared mixed noise** [13521-24]
- 13521 OR **Swin-BFB-UNet: a Swin transformer-based UNet combined with bottleneck fusion block for vertebra segmentation** [13521-6]

## INTELLIGENT SYSTEMS AND DETECTION TECHNOLOGIES

---

- 13521 OS **SparseNetYOLOv8: integrating vision transformers and dynamic probing for enhanced sparse object detection** [13521-20]
- 13521 OT **3D lane detection method based on M-BEV-LaneDet** [13521-31]
- 13521 OU **A robust lane detection and tracking method based on vanishing point for multi-environment situation** [13521-27]

- 13521 0V **Vehicle detection in different traffic scenarios based on YOLOv5** [13521-15]
- 13521 0W **A method for stochastic human action prediction based on denoising diffusion probability model** [13521-21]
- 13521 0X **Facial expression recognition based on YOLOv9 improvement** [13521-19]
- 13521 0Y **Improved YOLOv8-based PCB defect detection** [13521-26]
- 13521 0Z **Early detection and diagnosis of skin diseases based on YOLOv9 improvement** [13521-16]
- 13521 10 **Multiphase difference enhanced fusion for tumor and pancreas segmentation** [13521-30]
- 13521 11 **Attention synergy adaptive fusion convolutional neural network for multilabel skin lesion classification** [13521-14]
- 13521 12 **Laser spot array detection for structured light calibration** [13521-42]
- 13521 13 **Research on visual monitoring of natural gas leak based on spectral video technology** [13521-37]
- 13521 14 **Binocular vision measurement of robot end position based on orthogonal iteration to suppress covariance error** [13521-34]