

PROCEEDINGS OF SPIE

***Fourth International Conference on  
Computer Vision and Pattern  
Analysis (ICCPA 2024)***

**Ji Zhao  
Yonghui Yang**  
*Editors*

**17–19 May 2024  
Anshan, China**

*Organized by*  
University of Science and Technology Liaoning (China)

*Sponsored by*  
AEIC—Academic Exchange Information Centre (China)

*Published by*  
SPIE

**Volume 13256**

Proceedings of SPIE 0277-786X, V. 13256

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 *Fourth International Conference on Computer Vision and Pattern Analysis (ICCPA 2024)*, edited by Ji Zhao, Yonghui Yang, Proc. of SPIE 13256, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9781510682528

ISBN: 9781510682535 (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 © 2024 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

ix *Conference Committee*

---

## COMPUTERIZED POSITIONING AND TARGET DETECTION TECHNIQUE

---

- 13256 02 **Object detection algorithm based on residual deformable convolution** [13256-69]
- 13256 03 **Fine-tuned Segment Anything Model based on dual-stage adapter for remote sensing rotate object detection** [13256-24]
- 13256 04 **Arctangent-based nonlinear activation unit: ATLU** [13256-16]
- 13256 05 **Vision language distillation by clustering bitrajectory matching** [13256-17]
- 13256 06 **Infrared-visible cross-domain target recognition based on an improved CycleGAN** [13256-28]
- 13256 07 **Research on cleaning unmanned vessel based on improved dynamic window approach algorithm** [13256-19]
- 13256 08 **Intelligent identification and safety assessment of buildings based on computer vision: a case study of rural villages in Jilin province** [13256-25]
- 13256 09 **Quantification and application of blackness based on Munsell color system** [13256-98]
- 13256 0A **Structured light codec and phase analysis based on Gray code and four-step phase shift** [13256-43]
- 13256 0B **Research on weakly supervised target localization in sonar imagery based on GradCAM** [13256-5]
- 13256 0C **Two-stage self-supervised training vision transformers for small datasets** [13256-50]
- 13256 0D **Remote sensing small object detection algorithm based on contextual information enhancement** [13256-90]
- 13256 0E **A transformer-YOLO-based method for construction machinery monitoring in transmission channels** [13256-87]
- 13256 0F **Research on multifaceted presentation methods for product spectra based on mixed reality** [13256-57]
- 13256 0G **River vessel target detection based on improved Yolov8** [13256-80]

- 13256 OH **Research on operation and maintenance practice of intelligent facilities based on visual recognition technology** [13256-76]
- 13256 OI **Localization through particle filter powered neural network estimated monocular camera poses** [13256-58]
- 13256 OJ **Classification of major depressive disorder based on functional and structural MRI** [13256-15]
- 13256 OK **Pavement disease object detection for UAVs based on improved YOLOv8** [13256-6]
- 13256 OL **Research and analysis of target tracking based on millimeter wave radar** [13256-55]
- 13256 OM **High fidelity point completion for substation equipment point cloud** [13256-9]
- 13256 ON **Study on the model of paddy soil moisture content based on near infrared spectroscopy** [13256-26]
- 13256 OO **A channel estimation algorithm for train wireless communication system based on improved singular value decomposition linear least mean square error** [13256-88]
- 13256 OP **CA-P2PNet: improve P2PNet counting and localization with confidence aggregation** [13256-94]
- 13256 OQ **Small object detection in remote sensing images based on improved YOLOV7** [13256-75]

---

#### NETWORK INFORMATION RECOGNITION AND IMAGE PROCESSING

- 13256 OR **Study on text-to-image generation method based on deep learning** [13256-47]
- 13256 OS **Enhanced feline facial recognition: advancing cat face detection with YOLOv8 and TensorRT** [13256-48]
- 13256 OT **Cow face recognition based on transformer group** [13256-85]
- 13256 OU **A fast bundle adjustment method based on track selection** [13256-82]
- 13256 OV **Underwater SLAM system based on image feature point information enhancement** [13256-83]
- 13256 OW **DS-Swin Transformer HRNet for remote sensing images** [13256-45]
- 13256 OX **Few-shot image recognition based on improved Siamese neural network** [13256-51]
- 13256 OY **Standardized assessment of rock-climbing movements based on posture recognition** [13256-44]

- 13256 0Z **Research on the technology of news video layout analysis based on lightweight neural network** [13256-68]
- 13256 10 **Analysis of two handwritten digit recognition methods based on neural network** [13256-78]
- 13256 11 **Research of liveness detection for face recognition on mobile devices** [13256-95]
- 13256 12 **Interventional guidewire segmentation algorithm based on coded deep neural networks** [13256-93]
- 13256 13 **A comprehensive review of interaction methods for mixed reality sandbox applications** [13256-56]
- 13256 14 **A study of automatic classification of news headlines by bidirectional GRU models** [13256-70]
- 13256 15 **Dual-stream information adaptive fusion network for smoke detection in the complex environments of steel mills** [13256-41]
- 13256 16 **A dual-stream spatial-temporal detector for action recognition** [13256-84]
- 13256 17 **Parking lot vehicle recognition system** [13256-14]
- 13256 18 **Octree k-nearest neighbor parallel query** [13256-74]
- 13256 19 **Research on traffic sign recognition algorithms for autonomous vehicles in smart city traffic systems** [13256-32]
- 13256 1A **Multi-dimensional feature fusion network with transformer for image super-resolution** [13256-23]
- 13256 1B **Research on image edge detection algorithm based on matched filtering** [13256-99]
- 13256 1C **A method to improve the computational performance of deep learning models** [13256-46]
- 13256 1D **Research on semantic SLAM algorithm based on deep learning** [13256-61]
- 13256 1E **RoboCup SPL instance segmentation based on improved U-Net** [13256-52]
- 13256 1F **Heterologous image matching based on saliency region using Q-test and kernel density estimation** [13256-39]
- 13256 1G **Enhancing recyclable waste classification based on deep residual learning** [13256-22]
- 13256 1H **I2IP: image to image editing with prompt control using guided diffusion models** [13256-62]

- 13256 1I **Image enhancement algorithm for tunnel construction scenes** [13256-38]
- 13256 1J **The implementation of non-negative matrix decomorization and its optimization algorithm in face recognition** [13256-97]

---

**ADVANCED TECHNOLOGICAL INNOVATION AND INTELLIGENT MONITORING**

---

- 13256 1K **Algorithmic study of multi-sensor fusion-based vehicle navigation: comparative analysis of Kalman filter and machine learning models** [13256-4]
- 13256 1L **Traffic sign detection algorithm based on improved YOLOv5** [13256-67]
- 13256 1M **HISNet: a hybrid instance segmentation network for urban street scenes combining top-down and bottom-up approaches** [13256-81]
- 13256 1N **Research on high-precision detection technology of box trucks based on improved YOLOv5** [13256-7]
- 13256 1O **Building 2D skeletal animation for body control in unity with machine learning** [13256-20]
- 13256 1P **The application of style transfer algorithms in the design of condom packaging** [13256-42]
- 13256 1Q **An improved YOLOv8 safety helmet wearing detection network** [13256-40]
- 13256 1R **Research and application of a mouse device suitable for marine environment** [13256-49]
- 13256 1S **An improved safety helmet detection algorithm based on YOLOv8 in construction site** [13256-65]
- 13256 1T **PC-Yolo: enhanced YOLOv5-based defect detection system with improved partial convolution for ham sausage inspection** [13256-31]
- 13256 1U **Research on the application of a hybrid model based on deep learning and cluster analysis in anomaly detection of cloud platform access** [13256-89]
- 13256 1V **Research on intelligent water cleaning robot based on Raspberry PI and STM32** [13256-13]
- 13256 1W **Design and implementation of small infrastructure digital review system based on BIM** [13256-37]
- 13256 1X **Design of mine air quality evaluation system based on improved random forests** [13256-36]
- 13256 1Y **Fault diagnosis of motor bearing based on Mel-scale frequency cepstral coefficients and 1D-convolutional neural network** [13256-35]

- 13256 17 **Rolling bearing fault feature extraction based on variational mode decomposition and maximum correlated kurtosis deconvolution parameter optimization** [13256-33]
- 13256 20 **A fast and effective U-Net video dehazing network** [13256-73]
- 13256 21 **Research on online monitoring device of capacitive voltage transformer based on deep learning** [13256-71]
- 13256 22 **Research of adaptive three kinds of education robot based on large model framework** [13256-79]
- 13256 23 **A two-step algorithm for detecting helmet wear in electric vehicles based on YOLOv5-FD** [13256-3]
- 13256 24 **Research on lightweight waste-surface detection algorithm based on improved YOLOv5s** [13256-10]
- 13256 25 **Gate gap detection method based on double checkerboard calibration** [13256-77]
- 13256 26 **Aircraft rivet crack defect detection method based on YOLOv5** [13256-54]
- 13256 27 **Steel surface defect detection based on YOLO V8** [13256-53]
- 13256 28 **Drone-based Forest fire detection: a comparison of convolutional neural networks with an emphasis on MobileNetV2 enhancements** [13256-12]
- 13256 29 **Tunnel worker detection method based on improved YOLOv5s** [13256-30]