

PROCEEDINGS OF SPIE

***International Conference
on Computer Vision, Pattern
Recognition, and Detection
(ICVPRD 2025)***

Zhuoran Wang
Editor

25–26 October 2025
Chongqing, China

Organized by
Yangtze River Delta Research Institute, University of Electronic Science and
Technology of China (China)

Sponsored by
Nanyang Technological University (China)

Published by
SPIE

Volume 14121

Proceedings of SPIE 0277-786X, V. 14121

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, Pattern Recognition, and Detection (ICVPRD 2025)*, edited by Zhuoran Wang, Proc. of SPIE 14121, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

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

ISBN: 9798902322016
ISBN: 9798902322023 (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

ix *Conference Committee*

INTERNATIONAL CONFERENCE ON COMPUTER VISION, PATTERN RECOGNITION, AND DETECTION (ICVPRD 2025)

- 14121 02 **Deep learning-driven industrial robot visual grasping system modelling and optimization** [14121-1]
- 14121 03 **Research on detection equipment system for key components of new energy vehicles based on Digital Twin and 3D vision** [14121-2]
- 14121 04 **Art creation image generation method based on improved GAN** [14121-3]
- 14121 05 **Intelligent identification of illegal behaviors of offshore construction vessels by integrating AIS data and YOLOv7** [14121-4]
- 14121 06 **Image matching algorithms for VR panoramic solutions for interior design** [14121-5]
- 14121 07 **Visual marketing graphic image processing method based on FR-ATT algorithm and LSTM network** [14121-6]
- 14121 08 **Research on image processing algorithm for unmanned aerial vehicle inspection of damaged power lines in debris flow disasters** [14121-7]
- 14121 09 **Visual positioning algorithm for abnormal positions of book arrangement in university libraries** [14121-8]
- 14121 0A **Research on multi-source data fusion optimization design for oil and gas gathering and transportation pipelines** [14121-9]
- 14121 0B **Remote sensing image building extraction model based on hybrid attention mechanism** [14121-10]
- 14121 0C **Surface defect detection network of power equipment based on DCGAN** [14121-11]
- 14121 0D **Intelligent navigation for power substation inspection based on improved A* and image processing** [14121-12]
- 14121 0E **3D cadastral model construction method based on UAV tilt photography and GIS positioning** [14121-13]
- 14121 0F **Distribution grid planning and surveying based on UAV image acquisition and optimization algorithm** [14121-14]

- 14121 OG **Enterprise financial fraud identification based on Transformer and SMOTE algorithm** [14121-19]
- 14121 OH **Urban land remote sensing surveying method integrating visual attention and SegFormer algorithm** [14121-20]
- 14121 OI **Well logging image enhancement method based on Retinex algorithm and soft cutout** [14121-21]
- 14121 OJ **FreeGazeFormer: a CNN-transformer cross-modal attention framework for head pose-free gaze estimation** [14121-22]
- 14121 OK **Crack avoidance path planning design based on an improved A* algorithm** [14121-23]
- 14121 OL **Research on an intelligent urban fire monitoring system based on multi-source data fusion** [14121-24]
- 14121 OM **Semantic segmentation of remote sensing image point cloud data based on dual channel attention improved encoding and decoding structure** [14121-25]
- 14121 ON **Pattern recognition enhanced electric four-drum synchronous traction equipment** [14121-26]
- 14121 OO **Research on semantic segmentation of high-resolution remote sensing images based on CW SwinUNet** [14121-27]
- 14121 OP **Computer vision-assisted dynamic identification of sealing microfractures in geologic bodies around gas storage wells** [14121-28]
- 14121 OQ **Mine ecological environment evaluation and rehabilitation based on migration learning image recognition algorithm and hierarchical analysis** [14121-30]
- 14121 OR **Video object detection-based intelligent early warning method using BIM technology and YOLOv5** [14121-31]
- 14121 OS **CNN-driven multi-scale image segmentation algorithm for digital architecture** [14121-32]
- 14121 OT **Image defogging algorithm based on multi-scale residuals and coordinate space attention** [14121-33]
- 14121 OU **A mental health assessment model integrating the CNN-BiLSTM-attention mechanism** [14121-34]
- 14121 OV **Optimization of small-sample microscopic residual oil distribution recognition based on Mask R-CNN and migration learning** [14121-35]
- 14121 OW **Method for processing fuzzy images of IoT based on anti-blur image recognition algorithm** [14121-36]
- 14121 OX **A study of ResNet18 emotion recognition algorithm for digital childcare** [14121-39]

- 14121 0Y **Cross-cultural communication assistance: emotional analysis method based on XLNet and Mixup** [14121-41]
- 14121 0Z **Facial recognition access control system based on image enhancement and improved PCA algorithm** [14121-42]
- 14121 10 **Packaging design pattern aesthetic quality evaluation model based on GCN and FCN** [14121-43]
- 14121 11 **Research on vocational education skill supply and demand matching based on recruitment data mining and knowledge graph** [14121-44]
- 14121 12 **Improving the performance of computer vision target detection with deformable convolution** [14121-46]
- 14121 13 **Infrared image recognition method for nuclear power plant electrical equipment based on EFRM** [14121-47]
- 14121 14 **Design of photovoltaic panel intelligent cleaning system based on deep learning and SLAM** [14121-48]
- 14121 15 **Research on image denoising methods based on an improved U-Net and Transformer** [14121-49]
- 14121 16 **A study on Lao text recognition method based on Seq2Seq and BiLSTM** [14121-50]
- 14121 17 **The application of cloud-edge-terminal integrated technology of unmanned aerial vehicles in the survey and design of transmission lines** [14121-51]
- 14121 18 **Research on UAV transmission line survey based on multi-source data integration** [14121-52]
- 14121 19 **Optimized design method for fracture network fracturing in ultra-low permeability fractured reservoirs based on fracture flow simulation** [14121-53]
- 14121 1A **Intelligent identification of water and mud inrush hazards in underground engineering in karst areas based on YOLO v5** [14121-54]
- 14121 1B **Automated interaction in state grid design deliverable review systems based on structured data processing** [14121-55]
- 14121 1C **Intelligent matching and lossless migration technology of state grid design achievement documents for project review** [14121-56]
- 14121 1D **Efficient algorithms for data modeling and recognition of distribution networks based on random forests** [14121-57]
- 14121 1E **Power grid data security detection technology based on unsupervised learning and difference feature extraction** [14121-58]

- 14121 1F **Diagnostic method for wind turbine yaw system misalignment based on a KDE-GPR hybrid model** [14121-59]
- 14121 1G **Research on virtualization-based planning optimization method for distribution networks integrating multi-source power signal feature** [14121-60]
- 14121 1H **Multimodal image analysis of cross-tissue cotton GBLTP gene expression Patterns** [14121-61]
- 14121 1I **Research on intelligent generation of construction progress Gantt chart based on artificial intelligence** [14121-62]
- 14121 1J **Research on water resources remote sensing image segmentation method based on improved UD-Net network** [14121-63]
- 14121 1K **Research on intelligent aircraft recognition in flight command scenarios by fusion of deep CNN and object detection** [14121-64]
- 14121 1L **Research on electrical equipment fault detection by combining object detection and image segmentation algorithms** [14121-65]
- 14121 1M **Air navigation control surveillance information technology based on multi-source radar data fusion** [14121-66]
- 14121 1N **Research on cooperative navigation for UAV swarms integrating enhanced YOLOv5 with inertial ranging information** [14121-67]
- 14121 1O **Optimizing multimodal achievement retrieval algorithms for large-scale technology markets** [14121-68]
- 14121 1P **Intelligent load forecasting of power grid based on Spark-LSTM** [14121-69]
- 14121 1Q **Multi-objective genetic algorithm-driven linear optimization and safety evaluation model for grade-separated interchanges** [14121-70]
- 14121 1R **Multimodal visual health risk assessment enhanced by temporal features and spatial attention** [14121-71]
- 14121 1S **Image processing systems for domestic printers based on image printing algorithms** [14121-72]
- 14121 1T **Research on substation monitoring point selection integrating attention mechanisms with YOLOX** [14121-73]
- 14121 1U **Research on cross-border e-commerce product marketing recommendations based on multimodal signal fusion and deep learning technologies** [14121-74]
- 14121 1V **A study on a student mental health intervention model combining artificial intelligence and smartphone sensors** [14121-75]
- 14121 1W **Virtual reality scene assessment of Lingjiatan culture based on MT-2DCNN-LSTM algorithm in the context of cultural heritage** [14121-76]

- 14121 1X **Automatic correction of grammatical errors in Russian based on hierarchical language model and dependency relations** [14121-77]
- 14121 1Y **Automatic assignment method of educational and teaching resource tags based on active learning** [14121-78]
- 14121 1Z **A study on the collaborative model of computer-assisted grammar checking and human feedback in English writing** [14121-79]
- 14121 20 **Scene evaluation of English teaching video images based on active-learning annotation algorithm** [14121-80]
- 14121 21 **Dynamic graph analysis of intangible cultural heritage based on multimodal alignment of spatiotemporal knowledge graph** [14121-81]
- 14121 22 **Design of an intelligent aesthetic education teaching system based on multimodal signal interaction** [14121-82]
- 14121 23 **Intelligent diagnosis of coastal zone ecological security integrating graph neural networks and multimodal data** [14121-83]
- 14121 24 **Research on the application of YOLOv8-based port cargo image recognition in trade supervision** [14121-84]
- 14121 25 **Detection and layout rationality evaluation of blackboard areas in teaching videos by integrating Faster R-CNN and 3D convolutional networks** [14121-85]
- 14121 26 **Construction of an intelligent recognition model for elderly fitness behavior preferences based on decision tree algorithms** [14121-86]
- 14121 27 **Highway tunnel crack recognition algorithm based on improved DeepLabv3+ and attention mechanism** [14121-87]
- 14121 28 **Four-drum synchronous electric traction detection system based on computer vision** [14121-88]
- 14121 29 **Online monitoring system of resource status based on segmented linear trend compression algorithm** [14121-89]