

PROCEEDINGS OF SPIE

Seventh International Conference on Image, Video Processing, and Artificial Intelligence (IVPAI 2025)

Ruidan Su
Editor

18–20 May 2025
Bangkok, Thailand

Organized by
Sun Yat-sen University (China)
Universiti Teknologi Malaysia (Malaysia)
Shanghai Jiao Tong University (China)

Published by
SPIE

Volume 13731

Proceedings of SPIE 0277-786X, V. 13731

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 *Seventh International Conference on Image, Video Processing, and Artificial Intelligence (IVPAI 2025)*, edited by Ruidan Su, Proc. of SPIE 13731, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9781510694170

ISBN: 9781510694187 (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	<i>Conference Committee</i>
ix	<i>Introduction</i>

COMPUTER VISION AND OBJECT RECOGNITION SYSTEMS

13731 02	GIPPE-YOLO: a global-aware and IoU-optimized method for personal protective equipment detection [13731-9]
13731 03	Bridge deck pavement debonding detection system using impact sound and deep learning approach [13731-29]
13731 04	Improved YOLOv8 for oil manhole covers detection using attention mechanism and depthwise convolution [13731-21]
13731 05	CricXpert: a hybrid spatial fusion model for enhanced player recognition [13731-12]
13731 06	3D semantic segmentation for autonomous driving using late fusion of LiDAR range images and camera images with simple 2D transformation [13731-17]
13731 07	Research on gap measurement method based on 3D laser scanning technology [13731-14]
13731 08	Mamba-SaCIF Net: a lightweight hybrid model for cardiac MRI segmentation using spatial and channel fusion [13731-18]
13731 09	Single-image water surface reflection and shadow removal via gradient-color guided region-growing segmentation and inpainting [13731-23]
13731 0A	Shufflenet-le: an efficient shuffle attention-convolutional model for visual place recognition [13731-11]
13731 0B	Research on object tracking algorithm based on Siamese network [13731-2]
13731 0C	Research on high-resolution high-definition imaging algorithm for forward-looking sonar [13731-47]
13731 0D	Analysis on adversarial attacks in optical satellite image recognition [13731-42]
13731 0E	Intelligent image processing and cultural style transfer for brand visual design with Yanchuan cloth paste art [13731-71]

MULTIMODAL AND CROSS-MODAL AI

- 13731 0F **Phonetic and respiratory prosodic analysis of Cantonese Song Ci chanting** [13731-10]
- 13731 0G **MMF-Net: multi-model fusion network for hyperspectral image classification based on transfer learning** [13731-32]
- 13731 0H **Edge computing model for face detection and recognition based on federated learning for university** [13731-34]
- 13731 0I **Multimodal recognition of visually similar objects using images and impact sounds** [13731-35]
- 13731 0J **A multimodal attention-based system for real-time listening comprehension evaluation and feedback** [13731-55]
- 13731 0K **Multimodal dialogue generation and pronunciation feedback system based on AI for intelligent English tutoring** [13731-60]
- 13731 0L **MM-OpinionNet: a cross-modal deep learning framework for multimodal news mining and intelligent public opinion analysis** [13731-65]
- 13731 0M **Attention-guided generative adversarial networks for semantic extraction and synthesis of cultural heritage imagery** [13731-62]
- 13731 0N **CNN-Transformer architecture for piano performance style recognition and AI-based real-time music accompaniment** [13731-69]

AI FOR PREDICTIVE MODELING AND FORECASTING

- 13731 0O **Multi-model ensemble deep learning framework for robust stock market forecasting: integration of economic indicators and temporal dependencies** [13731-27]
- 13731 0P **Graph-based multitask learning framework for property price prediction and data imputation in the Sri Lankan real estate market** [13731-13]
- 13731 0Q **Machine-learning-based predictive maintenance: data-driven equipment failure forecasting** [13731-22]
- 13731 0R **Predicting stock prices: an empirical study using ARIMA and LSTM with model evaluation** [13731-6]
- 13731 0S **Optimizing pre-pickup ride retention in ride-hailing platforms using behavioral features with reinforcement learning** [13731-8]
- 13731 0T **A hybrid deep learning model for tourism trend forecasting based on big data and spatiotemporal pattern analysis** [13731-63]

- 13731 0U **Predictive career planning using AI and data-driven models for student development guidance** [13731-58]
- 13731 0V **Deep causal trajectory modeling of public opinion evolution from cross-media information streams** [13731-56]
- 13731 0W **Multidimensional influencing factors of childbearing intention of women in Dongguan City and prediction model based on machine learning** [13731-15]

DEEP LEARNING ARCHITECTURES AND OPTIMIZATION

- 13731 0X **Machine-learning-based prediction of return loss for coaxial probe-fed microstrip patch antennas: a comparative study of SVR, ANN, XGBoost, and Random Forest** [13731-44]
- 13731 0Y **Attention needs intelligence: a human cognitive inspired approach to enhance the reasoning capabilities of large language models** [13731-38]
- 13731 0Z **Improving linear support vector machine classifier using mutual information feature weighting for intrusion detection system** [13731-25]
- 13731 10 **FC-VIB: fact-checking with variational information bottleneck** [13731-26]
- 13731 11 **Artificial intelligence driven transformer through neural machine translation strategy** [13731-30]
- 13731 12 **DRL-MEC-based anti-malicious jamming decision in wireless body area network** [13731-61]
- 13731 13 **The simplified digital predistortion with low-rate ADC for 5G NR signals** [13731-36]
- 13731 14 **Dynamical system autoencoders using piecewise linear transformation** [13731-24]

AI IN EDUCATION AND LANGUAGE TECHNOLOGIES

- 13731 15 **Text analytics of IMDB reviews using latent Dirichlet allocation** [13731-7]
- 13731 16 **Study on the integration of generative artificial intelligence in the modeling design method for preschool children's toy cars** [13731-75]
- 13731 17 **Graph neural network-driven knowledge graph construction and application for Chinese instruction in higher vocational education** [13731-70]
- 13731 18 **Multimodal generation and personalization algorithms for educational content using generative AI** [13731-59]
- 13731 19 **Speech recognition and machine translation technologies in intelligent college English education: a multidisciplinary approach** [13731-68]

- 13731 1A **A deep-learning-based speech recognition and real-time feedback system for English pronunciation training in vocational education** [13731-67]
- 13731 1B **AI-driven personalized learning path optimization and adaptive resource recommendation for foreign language education** [13731-66]
- 13731 1C **Constructing a semantic-driven QA system for TCM cultural archives via deep learning and knowledge graph fusion** [13731-64]

APPLIED AI SYSTEMS AND DOMAIN-SPECIFIC SOLUTIONS

- 13731 1D **Artificial intelligence in solar energy integration for the smart grid: progress, possibilities, and challenges** [13731-5]
- 13731 1E **Enhancing spatial data analysis with generative AI: a novel approach for geospatial intelligence** [13731-31]
- 13731 1F **A methodology for using street view imagery, scene text detection, and GeoDBMS in urban geography studies** [13731-43]
- 13731 1G **Wilderness path planning based on continuous road network environment** [13731-33]
- 13731 1H **A study on the performance of routing protocols for video and voice transmission in MANET-based disaster networks** [13731-40]
- 13731 1I **Research on the fire source location of aircraft cargo compartment based on deep learning** [13731-46]
- 13731 1J **Analysis of hot trends and development vein of AIGC image art design based on big data** [13731-73]
- 13731 1K **A survey of the major applications of transformer-based 100 billion+ parameter large language models** [13731-41]
- 13731 1L **Generative artificial intelligence-assisted Z-generation Sanxingdui tea drink packaging innovation design path construction** [13731-77]