PROCEEDINGS OF SPIE

International Conference on Image Processing, Machine Learning, and Pattern Recognition (IPMLPR 2026)

Lei Chen
Editor

15 - 17 May 2026
Dalian, China

Organized by
International Association for Interdisciplinary Studies (China)

Published by
SPIE

Volume 14304

Proceedings of SPIE 0277-786X, V. 14304

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 Image Processing, Machine Learning, and Pattern Recognition (IPMLPR 2026)*, edited by Lei Chen, Proc. of SPIE 14304, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

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

ISBN: 9798902327080
ISBN: 9798902327097 (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
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*

IMAGE/VIDEO GENERATION AND PROCESSING

14304 02 **Aesthetic-LoRA: a method for generating aesthetically pleasing videos based on custom datasets and adaptive LoRA fine-tuning** [14304-3]

14304 03 **Meaning-aligned poster generation for cultural heritage: a human-AI loop with semantically decoupled prompts and affective closure** [14304-9]

14304 04 **Regional self-attention for layout-to-image diffusion: resolving overlap, scale, and semantic misalignment** [14304-24]

14304 05 **Single-step perspective registration to static template space: an ORB-FLANN based UAV video stabilization approach with zero cumulative drift** [14304-34]

14304 06 **Semantic-aware wireless image transmission based on generative diffusion models in low signal-to-noise ratio environments** [14304-22]

DETECTION AND DEFECT RECOGNITION

14304 07 **Research on lightweight YOLO visual detection method for complex scenes** [14304-1]

14304 08 **Detection of cigarette package appearance quality based on BoTNet network** [14304-2]

14304 09 **MAS-YOLO: a motion-augmented small-target detection network for clay pigeon detection in skeet shooting training** [14304-8]

14304 0A **Early fire detection algorithm based on improved computer vision** [14304-10]

14304 0B **Multiscale context aggregation and residual fusion for robotic grasp detection** [14304-18]

14304 0C **Research on grate defect identification based on semantic segmentation** [14304-23]

14304 0D **DCD-YOLO: an improved YOLO method for surface defect detection of decorative coatings** [14304-33]

14304 0E **Intelligent detection of bridge surface cracks based on an improved convolutional neural network and metric representation learning** [14304-26]

14304 0F **Prior-matched dual-branch network for unsupervised domain-adaptive plant disease classification** [14304-17]

UAV AND ROBOTIC SYSTEMS

14304 0G **UAV transmission line inspection system based on deep learning visual perception and adaptive model predictive control** [14304-11]

14304 0H **Research on the precise mounting algorithm of unmanned aerial vehicle drop-type self-locking pulley based on visual servoing** [14304-19]

14304 0I **A global template-based video stabilization method for UAV aerial footage using ORB and FLANN** [14304-21]

14304 0J **Zero-shot component recognition for transmission lines in industrial park microgrids based on permutation-attention transformer and lightweight VLM** [14304-35]

INTELLIGENT ANALYTICS AND DIGITAL APPLICATIONS

14304 0K **Vision-augmented multimodal analysis for competitive rifle shooting via pose-guided textualization and LLM reasoning** [14304-5]

14304 0L **A hybrid gabriel graph and simulated annealing framework for star map pattern recognition and structural classification** [14304-6]

14304 0M **Digital reasoning and application of Dezhou black pottery cultural elements based on knowledge graph** [14304-29]

14304 0N **Research on intelligent assessment of automobile collision damage degree based on image features** [14304-7]

14304 0O **Consumer purchase behavior prediction based on user-generated content sentiment analysis algorithm** [14304-12]

14304 0P **Detection and analysis of steel heat treatment deformation images based on convolutional neural network** [14304-13]

14304 0Q **Intelligent prediction of academic performance using image recognition and hybrid attention transformer** [14304-15]

14304 0R **Research on the dynamic adaptation of traditional cultural patterns based on generative adversarial networks and their integration with modern aesthetics** [14304-27]

14304 0S **Multiple object tracking and passing network analysis system in football matches** [14304-30]

14304 0T **Research on visual integration design of rural cultural and creative brands based on artificial intelligence algorithms** [14304-31]

14304 0U **Research on emotional interaction design of black pottery collectibles based on multimodal perception** [14304-32]

14304 0V **Research on an attribute-based encryption sharing model for image feature extraction** [14304-36]

MEDICAL IMAGING AND REHABILITATION

14304 0W **LiteBUS-SSM: a boundary-preserving lightweight state-space U-Net for breast ultrasound lesion segmentation** [14304-16]

14304 0X **Design of hand rehabilitation movement evaluation system based on lightweight critical point detection** [14304-20]