

PROCEEDINGS OF SPIE

International Conference on Computer Vision and Machine Learning (CVML 2025)

**Massimo Tistarelli
Rui Fan**
Editors

**24–26 October 2025
Chengdu, China**

Organized by
Chengdu University of Information Technology (China)

Sponsored by
Wuhan University (China)

Published by
SPIE

Volume 14065

Proceedings of SPIE 0277-786X, V. 14065

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 and Machine Learning (CVML 2025)*, edited by Massimo Tistarelli, Rui Fan, Proc. of SPIE 14065, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9798902320890

ISBN: 9798902320906 (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

v *Conference Committee*

INTERNATIONAL CONFERENCE ON COMPUTER VISION AND MACHINE LEARNING (CVML 2025)

- 14065 02 **Future-ready smart grids harnessing predictive analytics and optimization for sustainable energy** [14065-4]
- 14065 03 **Transformer-based energy consumption prediction and optimization framework for BIM-enabled green buildings** [14065-26]
- 14065 04 **Industrial equipment energy consumption anomaly detection and optimization based on vision transformer** [14065-23]
- 14065 05 **Collaborative optimization of ad-hoc network topology and combat formation** [14065-20]
- 14065 06 **KG-VLM: visual analytics for network fault diagnosis via knowledge graph-enhanced vision-language model** [14065-25]
- 14065 07 **Computer vision-enabled knowledge and skills mapping for network security operations: a practical framework under industry–education integration** [14065-30]
- 14065 08 **Vision-based skill assessment and personalized guidance for modern apprenticeship** [14065-12]
- 14065 09 **Mask-guided multistage food recognition via adaptive segmentation and efficient classification** [14065-31]
- 14065 0A **Automatic piano performance correction using a cross-modal visual network integrated with optical music recognition** [14065-16]
- 14065 0B **Nowcasting of cloud images based on dynamic multiscale voxel flow network and FY-4B satellite data** [14065-10]
- 14065 0C **Dual-input forecasting: leveraging image and numerical data with a transformer-RNN approach for 10-minutes solar irradiance prediction** [14065-3]
- 14065 0D **Transformer-based single-image 3D panoramic reconstruction for table tennis match scenes** [14065-28]
- 14065 0E **Interactive 3D scene reconstruction for virtual reality based on vision transformer** [14065-24]
- 14065 0F **Vision transformer-based adaptive optimization for CNC machining with real-time quality monitoring** [14065-22]

- 14065 0G **Employee tardiness prediction using machine learning and large language models: a case study in logistics enterprises** [14065-14]
- 14065 0H **Vision transformer-based student engagement recognition in online economic law education** [14065-29]
- 14065 0I **Transformer-based adaptive learning system for English cross-cultural communication: a web-enabled approach** [14065-27]