

PROCEEDINGS OF SPIE

Eleventh International Conference on Multimedia and Image Processing (ICMIP 2026)

Yen-Wei Chen
Editor

25–27 April 2026
Sapporo, Japan

Co-sponsored by
Ritsumeikan University (Japan)

Supported by
Niigata University (Japan)
University of the Ryukyus (Japan)
Universiti Malaysia Sarawak (Malaysia)
Wawasan Open University (Malaysia)
Gebze Technical University (Turkey)

Published by
SPIE

Volume 14298

Proceedings of SPIE 0277-786X, V. 14298

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 *Eleventh International Conference on Multimedia and Image Processing (ICMIP 2026)*, edited by Yen-Wei Chen, Proc. of SPIE 14298, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

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

ISBN: 9798902326960
ISBN: 9798902326977 (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*

OPTIMIZATION AND INTELLIGENT APPLICATION OF IMAGE DETECTION MODEL BASED ON DEEP LEARNING

- 14298 02 **Utilization of geometric and photometric image augmentation techniques to enhance YOLOv5 performance in detecting pig lesions** [14298-9]
- 14298 03 **Evaluating YOLO versions and hyperparameter optimization for object detection using an OFAT-guided grid search method** [14298-8]
- 14298 04 **GPE-YOLO: a gradient-prior enhanced detector with dynamic sampling for robust object detection in adverse weather** [14298-25]
- 14298 05 **HASO-DETR: hybrid attention small object detection based on RT-DETR** [14298-4]
- 14298 06 **An object detection-based multimedia system for intelligent classification and grading of unmilled local rice** [14298-5]
- 14298 07 **YOLOv12-S with self-supervised pretraining for Robusta green coffee bean defect detection** [14298-29]
- 14298 08 **Comparative evaluation of deep learning classifiers for fine-grained Robusta coffee bean defect classification** [14298-30]
- 14298 09 **Detection of neutral and rounded back postures in golf swing practice using computer vision** [14298-1]

MULTIMODAL FUSION DIGITAL IMAGE ANALYSIS AND COMPUTER VISION PERCEPTION

- 14298 0A **REDO AI: a multimodal generative pipeline for minimalist logo design revision via text-to-style transformation and dual-GAN refinement** [14298-32]
- 14298 0B **A CNN-based biological image processing and multimedia framework for rice grain purity analysis using enhanced mobile imaging** [14298-7]
- 14298 0C **Automatic tongue motion analysis using segmented images** [14298-2]
- 14298 0D **Blind SVD-based watermarking for color images using block-wise quantization** [14298-10]
- 14298 0E **Toward efficient directional rendering of bokeh effect** [14298-6]
- 14298 0F **3D-to-2D embedding-driven object classification with multiview robustness** [14298-22]

- 14298 OG **Expert-validated automated grading of Robusta green coffee beans using a hybrid deep learning pipeline** [14298-31]
- 14298 OH **Task-aware dual prompt tuning for all-in-one rain restoration with patch-based diffusion** [14298-24]
- 14298 OI **Multimodal IoT and edge computing-based detection of Panama disease in banana plants using EfficientNet-B3 and soil sensor analysis** [14298-33]

MULTIMODAL MEDICAL IMAGE FUSION AND INTELLIGENT ASSISTED DIAGNOSIS SYSTEMS

- 14298 OJ **Segmentation of gray matters and white matters from brain MRI data** [14298-15]
- 14298 OK **Systematic comparison of hyperparameter optimization strategies for Winter's classification of impacted third molars using EfficientNet-V2-S** [14298-17]
- 14298 OL **Performance evaluation of efficientNet-V2-S across single-task and multitask configurations for mandibular impacted third molar classification** [14298-18]
- 14298 OM **Deep learning-based detection of viral skin diseases using ResNet-152 with web deployment** [14298-12]
- 14298 ON **Research on kernel reconstruction methods for all-digital PET image reconstruction** [14298-14]

DATA MODELING AND OPTIMIZATION ALGORITHMS IN DIGITAL IMAGE AND MULTIMEDIA SCENARIOS

- 14298 OO **BANG-SAK: learning the importance of a Filipino street game through a 3D digital game** [14298-11]
- 14298 OP **Research on restoring Guqin music through deep learning** [14298-3]
- 14298 OQ **SALU-SALO: a virtual reality cooking game for simulating Filipino cuisine preparation** [14298-16]
- 14298 OR **Utilization of Unreal Engine 5 as a primary development medium for 2.5D games** [14298-21]
- 14298 OS **Optimizing Linux CPU governors for batch audio preprocessing** [14298-28]
- 14298 OT **KINO: design and development of a kinesthetic immersive navigation object for multi-user XR haptic theater** [14298-13]
- 14298 OU **Multimodal hate speech detection** [14298-26]