

PROCEEDINGS OF SPIE

Eighth International Conference on Image, Video, and Signal Processing (IVSP 2026)

Kiyoshi Hoshino
Editor

17–19 March 2026
Tokyo, Japan

Sponsored by
Meiji University

Published by
SPIE

Volume 14268

Proceedings of SPIE 0277-786X, V. 14268

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 *Eighth International Conference on Image, Video, and Signal Processing (IVSP 2026)*, edited by Kiyoshi Hoshino, Proc. of SPIE 14268, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9798902326267

ISBN: 9798902326274 (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

vii *Conference Committee*

CROSS-MODAL INFORMATION FUSION FOR INTELLIGENT TEXT AND IMAGE GENERATION SYSTEMS

- 14268 02 **A Dify-integrated InternVL-based multimodal system for breast cancer report classification and medical diagnostic report question answering** [14268-35]
- 14268 03 **VQA: improving ambiguous question answering through a multi-round clarification mechanism** [14268-23]
- 14268 04 **Artistic dual-message QR code generation by text-guided diffusion models** [14268-2]
- 14268 05 **Candlestick chart representations of musical scores: a cross-domain deep learning study on genre classification and provenance tracing** [14268-15]
- 14268 06 **Automatic seed optimization for personalized text-to-image generation with diffusion models** [14268-22]
- 14268 07 **Real-time speech emotion recognition using Wav2Vec2 and bidirectional GRU for multidataset emotion classification** [14268-28]

3D IMAGE MODELING AND VIRTUAL TECHNOLOGY APPLICATIONS

- 14268 08 **Windowed point cloud manifold harmonic transform** [14268-20]
- 14268 09 **SPAFormer: super-point attention-based point cloud transformer for 3D object segmentation** [14268-1]
- 14268 0A **A study on estimating vertebral positions using three-dimensional analysis of back surface topography** [14268-33]
- 14268 0B **ACANE: asymmetric CLS attention nullification for enhanced classification** [14268-31]
- 14268 0C **Methodological workflow for creating photorealistic neoclassical architecture datasets** [14268-7]
- 14268 0D **An improved beautification method for handwritten character using generative adversarial networks** [14268-44]

DIGITAL IMAGE DETECTION AND INTELLIGENT RECOGNITION TECHNOLOGIES

- 14268 OE **Plastic bottle recognition using an improved DeepLabv3+ [14268-10]**
- 14268 OF **Transformer-augmented EfficientNetV2B3 for robust plant disease identification [14268-11]**
- 14268 OG **Real-time fraudulent license plate detection through high-angle and front-area vehicle images [14268-29]**
- 14268 OH **Fringe detection from optical coherence tomographic images: a hybrid approach [14268-25]**
- 14268 OI **A reachability distance detector and a nonlocal means framework for mixed noise reduction in color images [14268-5]**
- 14268 OJ **VMNet: voxel-based multimodal fusion networks for 3D object detection [14268-40]**
- 14268 OK **Thermal-based gait anomaly scoring for concealed firearm detection using unsupervised deep learning [14268-36]**
- 14268 OL **RINKA: efficient artificial neural network model for drone-based object detection [14268-18]**
- 14268 OM **Auto Vision: a YOLO-based external sedan car damage detection and cost estimation system [14268-12]**
- 14268 ON **Robust real-time facial emotion recognition via adversarial feature learning and GAN-based domain adaptation [14268-27]**
- 14268 OO **Research on the assessment of executive function based on recognition of hand movements during work [14268-24]**

IMAGE ENCRYPTION AND DATA SECURITY

- 14268 OP **Reversible data hiding in encrypted images based on shifting encryption and Huffman coding [14268-4]**
- 14268 OQ **A novel reversible data hiding in encrypted technique for medical image using image interpolation and a two-stage difference preservation encryption [14268-34]**
- 14268 OR **Shallow versus deep features for anomalous/relevant event detection in high-dimensional streaming imagery [14268-6]**
- 14268 OS **Image steganography using XOR transformation and Huffman coding [14268-3]**

KEY TECHNOLOGIES IN COMPUTER VISION AND IMAGE PROCESSING

- 14268 OT **Deep learning for cell-level segmentation of canine mast cell tumors in histopathology: dataset and baselines toward automated grading** [14268-41]
- 14268 OU **An android facial expression control system trainable with a small dataset** [14268-42]
- 14268 OV **Integrating image information into Text2Loc: an extension for natural language-based 3D point cloud localization** [14268-14]
- 14268 OW **Image enhancement with recursively separated fuzzy histogram equalization** [14268-26]
- 14268 OX **Robust supervision for NeRD-Rain: stable single-image deraining under a limited training budget** [14268-38]
- 14268 OY **A two-stage transformer-based framework for view synthesis from a single image** [14268-39]
- 14268 OZ **Investigation of driver habituation in virtual environment with functional near-infrared spectroscopy** [14268-21]

AI-BASED SIGNAL MEASUREMENT AND ANALYSIS METHODS

- 14268 10 **CNN-based acoustic spectrogram analysis for quality assessment of laser-welded screws** [14268-43]
- 14268 11 **Style-aware data augmentation for deep learning on symbolic music** [14268-16]
- 14268 12 **Automatic structural tone extraction from monophonic music** [14268-17]
- 14268 13 **Noncontact heart rate estimation using automotive millimeter-wave radar with range-azimuth heatmap and temporal filtering** [14268-8]
- 14268 14 **Synchronized RGB and NIR imaging for heart rate estimation on the thenar eminence** [14268-9]
- 14268 15 **Single-channel-EEG-based sleep stage classification using machine learning algorithms** [14268-37]