PROCEEDINGS OF SPIE

Fourth International Conference on Computer Vision, Application, and Algorithm (CVAA 2024)

Lu Leng
Hui Yuan
Editors

11–13 October 2024
Chengdu, China

Organized by
Wuhan University of Science and Technology (China)

Sponsored by
AEIC—Academic Exchange Information Centre (China)

Published by
SPIE

Volume 13486

Proceedings of SPIE 0277-786X, V. 13486

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 *Fourth International Conference on Computer Vision, Application, and Algorithm (CVAA 2024)*, edited by Lu Leng, Hui Yuan, Proc. of SPIE 13486, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

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

ISBN: 9781510687615
ISBN: 9781510687622 (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

xi *Conference Committee*

COMPUTER VISION AND IMAGE PROCESSING

13486 02 **Soft shadow rendering algorithm based on light characteristics of rectangle regions** [13486-25]

13486 03 **Infrared and visible image fusion method based on fractional-order differentiation** [13486-101]

13486 04 **Spatiotemporal attention fusion network for multiple object tracking of unmanned aerial vehicle** [13486-74]

13486 05 **Based on improved YOLOv7 tomato fruit recognition method in natural environment** [13486-99]

13486 06 **Multi-label image classification with semantic information and upsampling** [13486-63]

13486 07 **3D reconstruction model of atlantoaxial joint gap based on Cxy-Net** [13486-33]

13486 08 **Classification of hand rehabilitation surface electromyographic signals based on FLO-SVM classifier** [13486-155]

13486 09 **Fast localization method for UAV images in GNSS-denied environments utilizing image sequence relationships** [13486-39]

13486 0A **DFrFT-ES model for emotion recognition based on fractional Fourier transform of EEG signals** [13486-34]

13486 0B **Research on traffic sign recognition under complex meteorological conditions** [13486-12]

13486 0C **Diffusion-augmented learning for long-tail recognition** [13486-67]

13486 0D **Apple leaf scab recognition using CNN and transfer learning** [13486-130]

13486 0E **Container image management in cloud-edge environments: an image deletion method based on layer affinity** [13486-138]

13486 0F **Computer graphics and image processing techniques based on visual communication design** [13486-14]

13486 0G **Dynamic fusion and non-negative matrix factorization-based multiview clustering method** [13486-61]

13486 0H **Convolutional recurrent neural network-based EEG signal classification in motor imagery** [13486-152]

13486 0I **Sentiment classification of MOOC courses by merging local context focus and bi-directional gated recurrent unit** [13486-137]

13486 0J **Optimization of visual SLAM in dynamic environments using object detection** [13486-7]

13486 0K **MLCDT: fine-grained multi-task learning for enhanced cognitive assessment in the clock drawing test** [13486-50]

13486 0L **Improving the quality of facial attribute transfer images based on generative adversarial networks** [13486-136]

13486 0M **Binary neural network-based lightweight ECG recognition for edge devices** [13486-100]

13486 0N **Texture image classification based on persistent homology** [13486-139]

13486 0O **Robust RGB-D visual odometry with adaptive selection of point, line, and plane features** [13486-18]

13486 0P **Depth map estimation method based on deformable convolutional neural network** [13486-165]

13486 0Q **Grading of wheat powdery mildew based on pixel point** [13486-23]

13486 0R **Behavior recognition method based on short-term spatiotemporal information fusion attention** [13486-73]

13486 0S **Enhancing pedestrian re-identification in low-light environments with multiscale feature attention extraction network** [13486-75]

13486 0T **Vision-assisted applications based on deep learning** [13486-11]

13486 0U **The UAV employs U-Net model for identification of forest fires** [13486-93]

13486 0V **Semantic segmentation for soccer robots based on the U-Net model** [13486-120]

13486 0W **Bone tumor boundary demarcation based on multimodal image fusion** [13486-134]

13486 0X **Identification method of bolt looseness of power tower based on laser vibration measurement technology and time inversion** [13486-114]

13486 0Y **Point cloud denoising method based on iterativePFN** [13486-8]

13486 0Z **Image captioning based on low-level grid features, segmentation features, and high-level fusion features** [13486-10]

13486 10 **Bio-inspired visual neural network for stop-and-go wave perception in high-density crowd** [13486-96]

13486 11 **Application of LSTM emotion recognition based on deep learning and attention mechanism in patient relationship mediation** [13486-71]

13486 12 **A comprehensive survey of semantic segmentation based on computer vision** [13486-140]

13486 13 **A study on lightweight facial expression recognition based on enhanced network models** [13486-151]

COMPUTER APPLICATION AND DETECTION TECHNOLOGY

13486 14 **Generate vector graphics of fine-grained pattern based on the Xception edge detection** [13486-36]

13486 15 **Mixed local channel attention-based YOLOv8 for pedestrian detection in infrared images** [13486-52]

13486 16 **Application of artificial intelligence online monitoring technology in fault diagnosis of cigarette packaging equipment** [13486-166]

13486 17 **Software heterogeneity technology based on application address randomization during software runtime** [13486-46]

13486 18 **Fault diagnosis method of soft-switching inverter based on deep learning** [13486-77]

13486 19 **Optimized 3D object detection algorithm for complex road scenes** [13486-97]

13486 1A **Detection of human fall behavior based on YOLOv5** [13486-37]

13486 1B **Enhancing traffic light recognition in autonomous vehicles: a novel convolutional neural network approach for fog removal using efficient channel attention** [13486-123]

13486 1C **Research on improvement strategy of DETR real-time object detection algorithm for small devices** [13486-91]

13486 1D **Faster-YOLOv8n-based lightweight road sign detection for intelligent unmanned vehicles** [13486-41]

13486 1E **Cigarette detection in diverse environments using YOLO-CD with adaptive multiscale feature extraction** [13486-148]

13486 1F **Research on underwater object detection based on frequency domain attention mechanism** [13486-5]

13486 1G **Revolutionizing power line asset detection: a YOLOv8-driven lightweight algorithm for UAV image recognition** [13486-45]

13486 1H **Machine vision and image processing for intelligent online monitoring and control of pharmaceutical fluidized bed granulation processes** [13486-160]

13486 1I **Improved YOLOv8-based YOLO-DCNet for road surface damage detection** [13486-15]

13486 1J **Lightweight saliency target detection algorithm with semantic-guided feature fusion** [13486-51]

13486 1K **Enhancing grid reliability through Transformer-LSTM model in smart grids** [13486-69]

13486 1L **Hierarchical integration transformer for marine snow removal** [13486-108]

13486 1M **Fire location detection method based on deep learning** [13486-124]

13486 1N **Application of an improved YOLOv9 algorithm for stem node detection in high-resolution global sugarcane images** [13486-27]

13486 1O **Saliency detection method for panoramic images based on GCN-ELM** [13486-19]

13486 1P **Multi-stage motion excitation network for fine-grained action recognition** [13486-81]

13486 1Q **Research on pose measurement system of space container based on multi-line structured light** [13486-31]

13486 1R **An online monitoring method for fiber optic cables based on optical signal reconstruction algorithm and wavelength conversion** [13486-104]

13486 1S **Detection of tennis ball landing points in videos using enhanced ppTSM v2 recognition model** [13486-76]

13486 1T **Research on object detection methods based on deformable convolutional networks** [13486-122]

13486 1U **Research on instrument fault detection based on Internet of Things and machine learning** [13486-89]

13486 1V **AttnEdge: an enhanced edge detection method based on self-attention mechanism** [13486-111]

13486 1W **Lightweight underwater garbage target detection algorithm based on improved YOLOv7-Tiny** [13486-85]

13486 1X **Research on detection of cigarette empty based on GAN-YOLOv4** [13486-162]

13486 1Y **Adaptive pruning techniques in YOLOv5 for industrial training on equipment operation** [13486-156]

13486 1Z **Development and application of virtual display system for interior decoration design based on Unity3D** [13486-26]

13486 20 **Lightweight design and deployment of object detection model for low-light scenes** [13486-95]

13486 21 **Research on the application of selective harmonic elimination pulse width modulation technology in fire emergency power supply** [13486-47]

13486 22 **Rolling bearing fault diagnosis based on VMD-CNN-BiLSTM** [13486-135]

13486 23 **Non-destructive detection of components using x-ray based on unsupervised learning** [13486-58]

13486 24 **Evaluation and optimization of image segmentation models in pavement crack detection** [13486-29]

13486 25 **Research on the development technology of dual-spectrum fusion hardware module for high-definition visible light and high-resolution infrared thermal imaging** [13486-32]

13486 26 **Power component detection in UAV via virtual assisted reality mechanisms** [13486-132]

13486 27 **Detection method of abnormal load state of diesel generator set based on PANet network and IMF** [13486-115]

13486 28 **LDIRNet: ceramic bearing end face defect detection based on lightweight dynamic inverted residual network** [13486-22]

13486 29 **Apple maturity detection based on improved YOLOv8** [13486-87]

13486 2A **An improved YOLOv5s model based on DeepSORT for pedestrian detection and tracking** [13486-157]

13486 2B **RST R-CNN: a triplet matching few-shot remote sensing object detection framework** [13486-86]

13486 2C **Simulation of rotating three-bucket loading method** [13486-158]

13486 2D **Lightweight bus driver smoking detection** [13486-72]

COMPUTER ALGORITHMS AND DATA ANALYSIS

13486 2E **Optimization of the online collaboration platform: an event queue management algorithm based on debounce and throttle techniques** [13486-16]

13486 2F **Point cloud classification method based on local dynamic edge graph convolution and global information fusion** [13486-106]

13486 2G **Multi-label text classification method based on ReBERTa-TextCNN** [13486-133]

13486 2H **Mobile-aware placement: optimizing energy consumption and delay of data stream applications in edge computing** [13486-141]

13486 2I **Application research of Bayesian optimization algorithm in DHCPv6 stateful allocation** [13486-146]

13486 2J **Overview of deep learning YOLO algorithm** [13486-3]

13486 2K **Enhanced image-text retrieval based on CLIP with YOLOv10 and Next-ViT** [13486-112]

13486 2L **A lightweight image segmentation algorithm based on MobileNetv3 and global context block attention mechanism** [13486-6]

13486 2M **The comparative study of microbial classification using machine learning and neural network models based on PCA for feature generation** [13486-70]

13486 2N **Diverse human motion prediction based on orthogonal basis vectors** [13486-53]

13486 2O **Human motion prediction based on spatial information and channel attention** [13486-129]

13486 2P **Improvement of the SAM algorithm based on causal structure analysis** [13486-84]

13486 2Q **Intelligent identification of hydropower engineering safety hazards: research on text classification method based on deep learning** [13486-144]

13486 2R **Research on MOOC dropout prediction by combining CNN-BiGRU and GCN** [13486-109]

13486 2S **Application of machine learning in port throughput prediction** [13486-20]

13486 2T **Ensemble learning-based approach for water quality prediction** [13486-88]

13486 2U **Automated execution algorithm for smart contracts in payment systems based on blockchain and big data** [13486-103]

13486 2V **Adaptive weighted variable step size matching tracking algorithm for Dice coefficient matching** [13486-105]

13486 2W **Research on transmission line simulation wind protection technology based on data analysis and artificial intelligence** [13486-42]

13486 2X **Task prioritization offloading scheme for vehicular edge computing networks: a game-theoretic DRL approach** [13486-9]

13486 2Y **Research on identification of domestic commercial block cipher algorithms based on deep learning** [13486-82]

13486 2Z **Research on optimization of ecological restoration design scheme using genetic algorithm** [13486-163]

13486 30 **Remote sensing recognition of roof area of flat-roofed houses based on Gram Schmit algorithm** [13486-65]

13486 31 **Intelligent scheduling task adjustment method based on graph neural network and adaptive weights** [13486-80]

13486 32 **A multisource heterogeneous data fusion method based on time alignment and federated learning for the dual carbon digital intelligence monitoring center** [13486-113]

13486 33 **Fusion of data dimensionality reduction algorithms based on category representation theory** [13486-40]

13486 34 **Automatic lung segmentation via a superpixel-based conditional breadth-first search algorithm** [13486-98]

13486 35 **M2SDF: a novel method for multi-seasonal and multiscale building energy prediction** [13486-126]

13486 36 **Predictive analysis of atmospheric sulfur dioxide concentration based on fusion modeling** [13486-90]

13486 37 **Animation generation algorithm of quadruped gait** [13486-121]

13486 38 **FRPoint: learning spatial correlations in the frequency domain for point cloud classification** [13486-94]

13486 39 **Efficient stochastic human motion prediction via consistency model** [13486-28]

13486 3A **Application of denoising diffusion probabilistic model in the Minecraft environment** [13486-54]

13486 3B **Medium-term wind speed ensemble forecasting using ICEEMDAN and dual attention mechanism** [13486-13]

13486 3C **Research on automatic furniture arranged method based on genetic algorithm** [13486-68]

13486 3D **Enhancement of YOLOv8 model for dense crowd scenes: incorporating an improved feature pyramid with attention mechanisms** [13486-21]

13486 3E **Research on air combat engagement zone based on differential feedback golden section search algorithm** [13486-154]

13486 3F **Mining analysis of traffic accident features based on Fp-growth algorithm and Apriori algorithm** [13486-164]