

PROCEEDINGS OF SPIE

# ***International Conference on AI-Generated Content (AIGC 2024)***

**Feng Zhao  
Duoqian Miao**  
*Editors*

**21–22 December 2024  
Beijing, China**

*Organized by*  
University of Science and Technology of China (China)  
Tongji University (China)

*Published by*  
SPIE

**Volume 13649**

Proceedings of SPIE 0277-786X, V. 13649

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](http://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 AI-Generated Content (AIGC 2024)*, edited by Feng Zhao, Duoqian Miao, Proc. of SPIE 13649, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X

ISSN: 1996-756X (electronic)

ISBN: 9781510692114

ISBN: 9781510692121 (electronic)

Published by

**SPIE**

P.O. Box 10, Bellingham, Washington 98227-0010 USA

Telephone +1 360 676 3290 (Pacific Time)

[SPIE.org](http://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](http://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](http://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 | <i>Conference Committee</i> |
| xi  | <i>Introduction</i>         |

---

## GENERATIVE MODELS AND MULTIMEDIA PROCESSING

---

|          |                                                                                                                                                                |
|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 13649 02 | <b>Image annotation tools and dataset: a comparative analysis in brief</b> [13649-2]                                                                           |
| 13649 03 | <b>Enhancing egg quality control through image processing and standardized assessment</b><br>[13649-4]                                                         |
| 13649 04 | <b>Spatial layout generation via generative adversarial networks</b> [13649-6]                                                                                 |
| 13649 05 | <b>ENF-GRAD: diffusion neural vocoder based on envelope features with data adaptive prior</b><br>[13649-8]                                                     |
| 13649 06 | <b>Improving AdaBoost face detection quality and speed via cascade classifiers</b> [13649-11]                                                                  |
| 13649 07 | <b>Enhancing fitness visualization: application and efficacy of realistic inpainting techniques<br/>using diffusion models</b> [13649-17]                      |
| 13649 08 | <b>Guided latent diffusion for universal medical image segmentation</b> [13649-7]                                                                              |
| 13649 09 | <b>Character recognition based on convolutional neural network and Hough transform</b><br>[13649-24]                                                           |
| 13649 0A | <b>An augmented format of the implicit randomized progressive-iterative approximation for<br/>curve and surface reconstruction</b> [13649-33]                  |
| 13649 0B | <b>Revolutionizing art education: leveraging diffusion models for high-fidelity 360° panoramic<br/>image synthesis in immersive virtual reality</b> [13649-34] |
| 13649 0C | <b>Research on the application of diffusion models in the automation of pavilion design</b><br>[13649-42]                                                      |
| 13649 0D | <b>Research on super-resolution reconstruction and accuracy improvement of earthquake<br/>disaster remote sensing images based on ESRGAN</b> [13649-44]        |

---

## NATURAL LANGUAGE PROCESSING AND LANGUAGE MODELS

---

|          |                                                                                                                            |
|----------|----------------------------------------------------------------------------------------------------------------------------|
| 13649 0E | <b>Artificial intelligence for text generation: copyright analysis and intellectual property<br/>development</b> [13649-3] |
|----------|----------------------------------------------------------------------------------------------------------------------------|

- 13649 OF **Aspect-based sentiment analysis method for Chinese financial texts using named entity recognition** [13649-16]
- 13649 OG **Use of AI-based text analysis tools to determine human emotional and social intelligence** [13649-22]
- 13649 OH **Sign process in natural and artificial intelligences: semiotic impact and algorithm** [13649-38]
- 13649 OI **II MedGPT-MRE: advanced medical large language model enhanced in domain of oncological rehabilitation by retrieval augmentation** [13649-40]
- 13649 OJ **HELM: a hypergraph enhanced language model for relevance modeling in sponsored search** [13649-41]
- 13649 OK **Aero-RAG: enhancing civil aircraft requirements engineering with integrated RAG and conversational AI** [13649-45]

---

#### MULTIMODAL AND CROSS-DOMAIN AI SYSTEMS

---

- 13649 OL **Recreating cultural festivities with artificial intelligence: a journey through the uncanny valley** [13649-9]
- 13649 OM **Towards a GenAI-assisted multimodal workflow for media productions** [13649-10]
- 13649 ON **A benchmark for cross-modal emotion-infused internet meme recommendation** [13649-12]
- 13649 OO **Visual-CoT: a visual guided chain of thought framework for fine-grained control in StarCraft II actions** [13649-18]
- 13649 OP **Evaluating the multimodal capabilities of ChatGPT for kidney disease diagnosis: integrating text and medical images** [13649-21]
- 13649 OQ **Enhancing the aesthetic appeal of AI-generated physical product designs through LoRA fine-tuning with human feedback** [13649-25]
- 13649 OR **Mozart's Touch: a lightweight multimodal music generation framework based on pre-trained large models** [13649-35]

---

#### AI IN INDUSTRY AND PRACTICAL SOLUTIONS

---

- 13649 OS **Adaptability of data security in a changing environment: a strategy based on artificial intelligence** [13649-5]
- 13649 OT **Exploration of the influence mechanism of AIGC on the financial efficiency of publishing and media enterprises** [13649-13]

- 13649 0U **Privacy bubbles for digital twin interaction in the metaverse** [13649-14]
- 13649 0V **Dynamic systems simulation using digital twins: approaches, technologies, and applications** [13649-15]
- 13649 0W **Discussing the product design process assisted by AIGC** [13649-19]
- 13649 0X **Using artificial intelligence to create 3D model of manufacturing equipment** [13649-20]
- 13649 0Y **Gearbox fault diagnosis method based on attention mechanism WDCNN and gated recurrent neural networks** [13649-30]
- 13649 0Z **Shaft orbit generation and fault diagnosis: a new paradigm based on diffusion models and deep learning** [13649-31]
- 13649 10 **Optimization dynamics in heart rate forecasting: attentional LSTM model with a CNN-GRU-based approach** [13649-32]
- 13649 11 **YODASORT: a YOLO-based level zero advanced driver assistance system with deep simple online and real-time tracking** [13649-37]
- 13649 12 **A case-based artificial neural network model for agricultural price prediction with differential evolution** [13649-46]
- 13649 13 **Research on the innovative practice of reshaping college English teaching mode empowered by AIGC** [13649-48]