IEEE Transactions on Medical Imaging

Special Issue on Large Multimodal and World Models for Medical Imaging

Large multimodal models (LMMs), vision–language models, and world models represent a major paradigm shift in artificial intelligence (AI) and machine learning, building upon the rapid advances of large language models (LLMs). By jointly modeling perception, reasoning, and prediction across heterogeneous data modalities, these models provide a powerful and flexible foundation for solving complex real-world problems. World models, in particular, aim to learn structured latent representations of environments that support long-horizon forecasting, planning, and decision-making, while large multimodal models enable rich semantic reasoning over visual, textual, and contextual inputs.

In the domain of medical imaging, these advances are especially impactful. Medical imaging forms the primary source of clinical perception across diagnostic and interventional workflows, including radiology, pathology, endoscopy, ultrasound, and image-guided procedures. The growing availability of large-scale image–text datasets, longitudinal imaging records, and multimodal clinical data has positioned medical imaging as a natural testbed for multimodal foundation models and world-model-based reasoning. Imaging-grounded multimodal models offer new opportunities for unified perception, contextual understanding, predictive modeling, and intelligent assistance in safety-critical clinical environments.

This Special Issue of *IEEE Transactions on Medical Imaging* aims to explore and showcase cutting-edge research on **large multimodal models and world models grounded in medical imaging**. The focus is on methodological advances that treat medical imaging as a central substrate for multimodal reasoning, forecasting, and decision support. While applications in image-guided surgery and interventional medicine are of particular interest due to their complexity and clinical impact, the scope of the Special Issue extends broadly across diagnostic and interventional imaging domains.

The Special Issue calls for **original and innovative methodological contributions** addressing key challenges in the development, training, validation, and deployment of large multimodal and world models for medical imaging. Submissions should emphasize technical novelty and rigorous experimental evaluation on real-world medical imaging datasets, or on high-fidelity simulation platforms closely tied to imaging data. Review papers, surveys, or purely conceptual position papers are considered out of scope. This is an open call for papers. Authors of relevant papers accepted to leading conferences about medical imaging, machine learning, robotics are encouraged to substantially extend their work to meet the scientific depth and validation standards of IEEE TMI.

Topics of interest include, but are not limited to:

- Vision foundation models and large multi-modality models for medical imaging
- Imaging-driven agentic AI with multi-modal LLMs for medical imaging tasks
- Imaging-driven surgical data science
- Action planning and navigation in image-guided surgery
- Digital twin and physical AI grounded in medical imaging data
- Virtual/augmented reality in surgery
- Pre-/intraoperative image registration
- AI-assisted autonomy and embodied intelligence in image-guided surgery
- Medical imaging in robotic surgery
- Imaging-driven human-robot collaboration in medicine

- Intelligent imaging-based assistants in operating rooms
- Safety, robustness, imaging-centric uncertainty modeling, and interpretability of medical multi-modal LLM and world model
- Imaging-driven evaluation protocols, clinical studies, and ethical or regulatory aspects in intelligent medical multi-modal LLM and world model

Authors must submit papers through **IEEE Author Portal** according to the standard IEEE TMI submission instructions. During submission, authors should select **"Special Issue on Large Multimodal and World Models for Medical Imaging"** as the manuscript type. Each paper will typically be reviewed by four expert reviewers following the standard TMI peer-review protocol.

Guest Editors

- Yueming Jin, National University of Singapore, Singapore; Email: ymjin@nus.edu.sg
- Sophia Bano, University College London, UK; Email: sophia.bano@ucl.ac.uk
- Mathias Unberath, Johns Hopkins University, USA; Email: mathias@jhu.edu
- Qi Dou, The Chinese University of Hong Kong, Hong Kong, China; Email: qidou@cuhk.edu.hk

Advisors

- Yutong Ban, Shanghai Jiao Tong University, China; Email: yban@sjtu.edu.cn
- Daguang Xu, NVIDIA, USA; Email: daguangx@nvidia.com

Schedule

- Submission deadline: August 1, 2026
- Initial decision: September 15, 2026
- Revised manuscripts due: November 15, 2026
- Final acceptance: December 20, 2026
- Publication: To be determined