IEEE NPSS NUCLEAR & PLASMA SCIENCES SOCIETY

IEEE Transactions on Radiation & Plasma Medical Sciences

Extention of submission deadline: 15 July 2025

Special issue on "Deployment of Large Models in Medical Imaging" Call for papers

With the rapid advancement of artificial intelligence, large models have emerged as powerful tools in medical imaging and provided unprecedented solutions to many critical problems such as medical imaging analysis, disease diagnosis, and treatment planning. However, due to the lack of infrastructure, the pursuit of scalability, and the demand for real-time processing, the widespread adoption of large models in medical imaging faces considerable challenges:

1. Computational Requirements

- **High-Performance Computing**: Large models often require powerful GPUs or TPUs for training and inference. Many healthcare facilities may not have access to such hardware, which can limit the model's use.
- **Memory Constraints**: Large models can have millions or even billions of parameters, necessitating substantial RAM. Limited memory can lead to performance bottlenecks or prevent the model from running altogether.

2. Scalability

- **User Load**: In a clinical setting, multiple users may need to access the model simultaneously (*e.g.*, radiologists analyzing images). Scaling the deployment to handle multiple requests without latency or downtime can be challenging.
- **Data Volume**: Medical imaging generates vast amounts of data. Processing data in real time requires robust computational infrastructure that can scale as data volumes increase.

3. Latency and Response Time

- **Real-Time Processing**: Many medical imaging applications require real-time or near-real-time responses, such as diagnostic imaging related to critical care monitoring. High latency due to insufficient computational resources can hinder clinical workflows.
- **Batch Processing Limitations**: If models are run in batch mode due to computational constraints, this can delay important diagnostic outputs.

4. Infrastructure Complexity

- **System Integration**: Integrating large models into existing healthcare IT systems often requires sophisticated infrastructure and middleware, complicating deployment and potentially leading to increased downtime or system failures.
- **Maintenance and Updates**: Keeping computational resources up to date with the latest technology and models requires ongoing investment and expertise.

This special issue aims to address the pressing need for the effective deployment of large models in the healthcare domain by promoting innovative solutions for the efficient deployment of large AI models in medical imaging. By encouraging interdisciplinary contributions from computer science, medicine, and clinical translation, we aim to bridge

the gap between cutting-edge AI research and practical healthcare implementation. Topics of this special issue include, but are not limited to,

- **Compression Techniques for Large Models**: Exploration of novel methods for compressing large medical models including multi-modal large image-language models and segmentation models without significant loss of accuracy, including but not limited to pruning, quantization, decomposition, and knowledge distillation.
- **Accelarating Model Inference**: Strategies to accelarate large models that integrate various types of data (e.g., text, images, and clinical data) to improve efficiency and facilitate real-time applications in medical imaging diagnosis.
- **Co-design of Hardware and Algorithms:** Methods that solve the hardware-related problems in the algorithmic deployment of large models in medical imaging, including cloud and edge devices.
- **Advanced Image Compression**: Development of advanced image compression techniques tailored for different medical imaging modalities and associated medical data that optimize storage and transmission while preserving the diagnostic quality.
- **Evaluation Metrics and Benchmarks**: Establishing standardized metrics for evaluating the performance of compressed models and images in clinical scenarios, including accuracy, speed, and usability in practice.
- **Security and Privacy in Compression**: Addressing risks and proposing secure solutions for using compressed models in sensitive clinical scenarios.

Authors must submit papers digitally to https://mc.manuscriptcentral.com/trpms, using standard IEEE Transactions format, indicating in their cover letter that the submission is aimed for this special issue. Authors are encouraged to contact the guest editors to determine suitability of their submission for this special issue.

Guest Editors:

- Fenglei Fan (Executive Guest Editor), Assistant Professor (Incoming)
  Department of Data Science, City University of Hong Kong, Hong Kong
  hitfanfenglei@gmail.com
- Joyita Dutta, Full Professor
  Department of Biomedical Engineering, University of Massachusetts Amherst, US
  jdutta@umass.edu
- Yi Zhang, Full Professor
  School of Cyber Science and Engineering, Sichuan University, China
  yzhang@scu.edu.cn
- Dong Zeng, Associate Professor
  School of Biomedical Engineering, Southern Medical University, China
  zd1989@smu.edu.cn
- Dufan Wu, Assistant Professor

Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Harvard University, US
dwu6@mgh.harvard.edu

- Xi Peng, Assistant Professor
  Department of Computer and Information Sciences, University of Delaware, US
  xipeng@udel.edu

Provisional schedule:

Submission deadline: July 15th 2025
Acceptance/rejection notification: October 2025
Revised manuscripts due: December 2025
Publication: May 2026

Supplementary Information (Biography+5 Representative Papers Related to This Proposal):

Dr. Fenglei Fan is currently a Research Assistant Professor with Department of Mathematics, The Chinese University of Hong Kong. He will join the City University of Hong Kong as a tenure-track assistant professor in Spring 2025. His primary research interests lie in NeuroAI and its applications in model compression and medical imaging. He has authored 26 papers in flagship AI and medical imaging venues such as JMLR, TNNLS, TMI, CVPR, TCI, and TRPMS. He was the recipient of the IBM AI Horizon Scholarship. He was also selected as the recipient of the 2021 International Neural Network Society Doctoral Dissertation Award. His primarily-authored papers were selected as one of few 2024 CVPR Best Paper Award Candidates (26 out of 1W+ submissions) and won the IEEE Nuclear and Plasma Society Best Paper Award, respectively. The following work by Dr. Fenglei Fan are directly relevant to this special issue proposal:

F. L. Fan, J. Fan, D. Wang, J. Zhang, Z. Dong, S. Zhang, S., G. Wang and T. Zeng, Hyper-Compression: Model Compression via Hyperfunction. *arXiv preprint arXiv:2409.00592*, 2024.

F. L. Fan, M. Li, F. Wang, R. Lai, G. Wang. Expressivity and trainability of quadratic networks. IEEE transactions on Neural Networks and Learning Systems, 2023, in press.

F. L. Fan, R. Lai, & G. Wang, Quasi-equivalence between width and depth of neural networks. Journal of Machine Learning Research, *24*(183), 1-22, 2023.

F. L. Fan, J. Xiong, M. Li, G. Wang. On interpretability of artificial neural networks: A survey. IEEE Transactions on Radiation and Plasma Medical Sciences, 5(6), 741-760, 2021.

F. L. Fan, H. Shan, M. K. Kalra, R. Singh, G. Qian, M. Getzin, Y. Teng, J. Hahn, G. Wang. Quadratic autoencoder (Q-AE) for low-dose CT denoising. IEEE transactions on medical imaging, 39(6), 2035-2050, 2019.

Dr. Joyita Dutta is a tenured Professor in the Department of Biomedical Engineering at the University of Massachusetts Amherst. She received her B.Tech. (Honors) from the Indian Institute of Technology (IIT) Kharagpur and M.S. and Ph.D. from the University of Southern California. She directs the Biomedical Imaging and Data Science Laboratory (BIDSLab) at UMass Amherst, which develops signal processing and artificial intelligence (AI) techniques for image, graph, and time-series datasets. Her scientific contributions include the development of a broad range of tools for medical image enhancement and reconstruction with a focus on multimodality information integration. Her current research interests include developing AI approaches for the diagnosis and prognosis of Alzheimer's disease. Dr. Dutta was the recipient of the 2016 Tracy Lynn Faber Memorial Award from the Society of Nuclear Medicine and Molecular Imaging (SNMMI) and the 2016 Bruce Hasegawa Young Investigator Medical Imaging Science Award from the IEEE. As a postdoc/junior faculty, she received an SNMMI Young Investigator Award, the 2013 SNMMI Mitzi & William Blahd MD Pilot Research Grant, the 2013 American Lung Association Senior Research Training Fellowship, and an NIH K01 Career Development Award. Her research is supported by multiple grants, including active NIH R01, R21, and

R03 grants held as PI. Dr. Dutta has served as a member of the SNMMI AI Task Force and the Program Chair for the 2022 IEEE Medical Imaging Conference in Milan, Italy. She is currently the President of the SNMMI Physics, Instrumentation and Data Sciences Council (PIDSC). Her trainees at BIDSLab have been awarded prestigious extramural training grants from the SNMMI, IEEE, APS, and AAUW. The following work by Dr. Joyita Dutta are directly relevant to this special issue proposal:

Song T-A, Chowdhury SR, Malekzadeh M, Harrison S, Blackwell Hoge T, Redline S, Stone KL, Saxena R, Purcell SM, **Dutta J**. AI-driven sleep staging from actigraphy and heart rate. PLOS One, 2023; 18(5): e0285703.

Song T-A, Yang F, **Dutta J**. Noise2Void: Unsupervised denoising of PET images. Phys Med Biol, 2021; 66:214002.

Song T-A, Chowdhury SR, Yang F, Sepulcre J, Jacobs HIL, Wedeen V, Johnson KA, **Dutta J**. A physics-informed geometric learning model for pathological tau spread in Alzheimer's disease. Medical Image Computing and Computer Assisted Intervention (MICCAI), Lima, Peru, Oct 4-8, 2020.

Song T-A, Chowdhury SR, Yang F, Dutta J. PET image super-resolution using generative adversarial networks. Neural Netw, 2020;125: 83-91.

Song T-A, Chowdhury SR, Yang F, **Dutta J**. Super-resolution PET imaging using convolutional neural networks. IEEE Trans Comput Imaging, 2020;6: 518-28.

Dr. Yi Zhang received the B.S., M.S., and Ph.D. degrees in computer science and technology from the College of Computer Science, Sichuan University, Chengdu, China, in 2005, 2008, and 2012, respectively. From 2014 to 2015, he was with the Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA, as a Post-Doctoral Researcher. He is currently a Full Professor with the School of Cyber Science and Engineering, Sichuan University, and is the Director of the Deep Imaging Group (DIG). He has authored more than 80 articles in the field of image processing. These papers were published in several leading journals, including IEEE Transactions on Medical Imaging, IEEE Transactions on Radiation, Plasma and Medical Sciences, IEEE

Transactions on Computational Imaging, Medical Image Analysis, European Radiology, and Optics Express, and reported by the Institute of Physics (IOP) and during the Lindau Nobel Laureate Meeting. He received major funding from China's National Key Research and Development Program, the National Natural Science Foundation of China, and the Science and Technology Support Project of Sichuan Province, China. His research interests include medical imaging, compressive sensing, and deep learning. He is the Guest Editor of the International Journal of Biomedical Imaging and Sensing and Imaging and an Associate Editor of IEEE Transactions on Medical Imaging and IEEE Transactions on Radiation, Plasma and Medical Sciences. The following work by Dr. Yi Zhang are directly relevant to this special issue proposal:

Chen, H., **Zhang, Y**., Kalra, M. K., Lin, F., Chen, Y., Liao, P., Zhou, J., Wang, G. (2017). Low-dose CT with a residual encoder-decoder convolutional neural network. IEEE transactions on medical imaging, 36(12), 2524-2535.

He, L., Du, W., Liao, P., Fan, F., Chen, H., Yang, H., **Zhang, Y.** (2024). Solving Zero-Shot Sparse-View CT Reconstruction With Variational Score Solver. IEEE Transactions on Medical Imaging, DOI: 10.1109/TMI.2024.3475516.

Luo, M., Zhou, N., Wang, T., He, L., Wang, W., Chen, H., ..., **Zhang, Y.** (2024). Bi-Constraints Diffusion: A Conditional Diffusion Model with Degradation Guidance for Metal Artifact Reduction. IEEE Transactions on Medical Imaging, DOI: 10.1109/TMI.2024.3442950.

Lu, Z., Gao, Q., Wang, T., Yang, Z., Wang, Z., Yu, H., … , **Zhang, Y.** (2024). PrideDiff: Physics-regularized generalized diffusion model for CT reconstruction. IEEE Transactions on Radiation and Plasma Medical Sciences, DOI: 10.1109/TRPMS.2024.3471677.

Xia, W., Yang, Z., Lu, Z., Wang, Z., **Zhang, Y.** (2023). RegFormer: A Local–Nonlocal Regularization-Based Model for Sparse-View CT Reconstruction. IEEE Transactions on Radiation and Plasma Medical Sciences, 8(2), 184-194.

Dr. Dong Zeng is currently the Head of the Medical X-ray Advanced Imagimg Algorithm and Instrument (MiX AIAI) lab, and alao an Associate Researcher of medical engineering with Southern Medical University. He is actively involved in developing CT imaging solutions for cutting-edge biomedical research and clinical diagnosis in addition to lecturing undergraduate and postgraduate courses on medical physics and medical imaging. He received the 2022 Young Investigator Award of Guangzhou Association for Science and Technology. He has authored more than 50 peer-reviewed articles in prominent journals. The following work by Dr. Dong Zeng are directly relevant to this special issue proposal:

D. Li, Z. Bian, S. Li, J He, **D. Zeng***; J. Ma, Noise Characteristics Modeled Unsupervised Network for Robust CT Image Reconstruction, IEEE Transactions on Medical Imaging, 2022, 41(12): 3849-3864

D. Li#; **D. Zeng#**; S. Li, Y. Ge, Z. Bian, J. Huang, J. Ma, MDMPCCT: Multiple Dynamic Modulations for High-Performance Spectral PCCT Imaging, IEEE Transactions on Medical Imaging, 2020, 39(11): 3630-3642

D. Zeng, L. Yao, Y. Ge, S. Li, Q. Xie, H. Zhang, Z. Bian, Q. Zhao, Y. Li, Z. Xu, D. Meng, J. Ma, Full-Spectrum-Knowledge-Aware Tensor Model for Energy-Resolved CT Iterative Reconstruction, IEEE Transactions on Medical Imaging, 2020, 39(9): 2831-2843

S. Li#; **D. Zeng#**; J. Peng, Z. Bian; H. Zhang, Q. Xie, Y. Wang, Y. Liao, S. Zhang, J. Huang, D. Meng, Z. Xu, J. Ma, An Efficient Iterative Cerebral Perfusion CT Reconstruction via Low-Rank Tensor Decomposition With Spatial-Temporal Total Variation Regularization, IEEE Transactions on Medical Imaging, 2019, 38(2): 360-370

D. Zeng, Q. Xie, W. Cao, J. Lin, H. Zhang, S. Zhang, J. Huang, Z. Bian, D. Meng, Z. Xu, Z. Liang, W. Chen, J. Ma, Low-Dose Dynamic Cerebral Perfusion Computed Tomography Reconstruction via Kronecker-Basis-Representation Tensor Sparsity Regularization, IEEE Transactions on Medical Imaging, 2017, 36(12): 2546-2556

Dr. Dufan Wu is an Assistant Professor of Radiology in the Center for Advanced Medical Computing and Analysis, Department of Radiology, at Massachusetts General Hospital (MGH) and Harvard Medical School (HMS). He received his PhD in Nuclear Science and Technology from the Department of Engineering Physics at Tsinghua University in 2016 and has been working at MGH since then. His main research interests include AI in medical image processing and reconstruction, photon counting and spectral CT, and translational research of AI in medical image analysis. The following work by Dr. Dufan Wu are directly relevant to this special issue proposal:

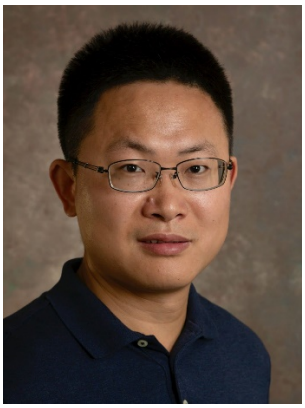
Chen C, Miao J, **Wu D**, Zhong A, Yan Z, Kim S, Hu J, Liu Z, Sun L, Li X, Liu T. Masam: Modality-agnostic sam adaptation for 3d medical image segmentation. Medical Image Analysis. 2024 Aug 22:103310.

Dayan I*, Roth H*, Zhong A*, …, **Wu D** (16th of 99 authors), …, Gilbert F*, Flores M*, Li Q*. Federated Learning used for predicting outcomes in patients with COVID-19. Nat Med. 2021 Sep 15;1–9.

Xu P, Kim K, Koh J, **Wu D**, Lee YR, Park SY, Tak WY, Liu H, Li Q. Efficient knowledge distillation for liver CT segmentation using growing assistant network. Physics in Medicine & Biology. 2021 Nov 26;66(23):235005.

Wang Y, Yoon S, Jin P, Tivnan M, Chen Z, Hu R, Zhang L, Chen Z, Li Q, **Wu D**. Implicit Image-to-Image Schrodinger Bridge for CT Super-Resolution and Denoising. arXiv preprint arXiv:2403.06069. 2024 Mar 10.

Guan Z, Wu Z, Liu Z, **Wu D**, Ren H, Li Q, Li X, Liu N. Cohortgpt: An enhanced gpt for participant recruitment in clinical study. arXiv preprint arXiv:2307.11346. 2023 Jul 21.



Dr. Xi Peng is an assistant professor in the Department of Computer & Information Sciences at the University of Delaware, where he leads the Deep Robust & Explainable AI Lab (DeepREAL). He works on machine learning, computer vision, and safe learning systems. His interest focuses on enhancing the safety and reliability of AI technologies in

high-stakes areas such as science, health, and autonomous vehicles. His research receives support from federal agencies such as NSF, DOD, CDC, and industry partners including the MSK Cancer Center, Google DeepMind, and Snap Research. Dr. Peng has been recognized with several prestigious awards for young investigators, such as the NSF CAREER Award, DOD DEPSCoR Award, Google Faculty Research Award, General University Research Award, and the UD Research Foundation Award. His work won the Best Paper Award at NeurIPSW 2021, the Best Student Paper Finalist at ECCV 2016, and multiple Oral presentation awards at top conferences like CVPR, ICLR, ICCV, and KDD. Dr. Xi Peng earned his Ph.D. degree in computer science from Rutgers University in 2018. The following work by Dr. Xi Peng are directly relevant to this special issue proposal:

Li, T., Ma, M., **Peng, X.** (2024) Beyond Accuracy: Ensuring Correct Predictions With Correct Rationales. In The Thirty-eighth Annual Conference on Neural Information Processing Systems.

Li, T., Ma, M., **Peng, X.** (2024). Deal: Disentangle and localize concept-level explanations for vlms. In European Conference on Computer Vision (pp. 383-401). Springer, Cham.

Qiao, F., & **Peng, X**. (2024) Ensemble Pruning for Out-of-distribution Generalization. In *Forty-first International Conference on Machine Learning*.

Ma, M., Li, T., & **Peng, X**. (2024). Beyond the Federation: Topology-aware Federated Learning for Generalization to Unseen Clients. In *Forty-first International Conference on Machine Learning*.

Qiao, F., & **Peng, X**. (2023) Topology-aware Robust Optimization for Out-of-Distribution Generalization. In *The Eleventh International Conference on Learning Representations*.