Muxin Wei

School of Medicine and Health,
Harbin Institute of Technology,
No.92, West Da-Zhi Street,
Harbin, Heilongjiang, 150001 China.

muxin-wei.github.io
github.com/muxin-wei
muxin.wei@foxmail.com

Research Interests

My research interests lie in multimodal learning and computer vision, with a focus on the interpretation and downstream application of deep models in healthcare. The goal is to develop trustworthy and interpretable AI models that improve generalization for real-world healthcare tasks.

Education

Harbin Institute of Technology Sep 2023 – Jan 2026
M.Sc., Biomedical Engineering Harbin, China
– *Thesis: A General-purpose Model for Medical Image Segmentation*

Dongguan University of Technology Sep 2018 – Jun 2022
B.Sc., Computer Science and Technology Dongguan, China
– *Thesis: A Real-Time Vehicle Detection and Localization System Based on YOLOV5*

Publications

– **Wei, Muxin**, et al. 'Rep-MedSAM: Towards Real-Time and Universal Medical Image Segmentation'. Lecture Notes in Computer Science, 2025, pp. 57–69.

Research Experience

World Modeling Patient Trajectories with Longitudinal Chest CT Aug 2025 – Present
Collaborator: ***Dr. Zhicheng Zhang***

– Objective: Develop a **digital twin** to simulate and predict the future chest CT trajectories of patients through the autoregressive generation of longitudinal CT and clinical reports.
– Evaluate existing generative methods including Text-to-CT Generation, GenerateCT, MAISI-v1/v2.
– Validate the feasibility of 2D and 3D VQ-GAN on CT dataset scaling from 1 to 30K scans.
– Bridge VQ codebook with text prompts via pretrained **CT-CLIP**, **BiomedCLIP** and **BiomedGPT**.
– Explore RL techniques and adaptation from LLM to VLM, enabling System 2 thinking (**Self-Play**, **Self-reward**).

CCN '25 Challenge: Brain Response Prediction with Multimodal Movies Mar 2025 – Jul 2025

– Align stimulus of different modalities to fMRI responses.
– Evaluate the performance of multiple visual encoding models to predict fMRI brain activity.
– Model building and training for cross-modality with V-JEPA 2.

Medical Images Segmentation with Open-vocabulary Sep 2024 – Mar 2025

– Explore interpretability of CLIP using Text-Based Decomposition via *CLIP-Text-Span* and attempt to apply to volumetric medical images.
– Improve medical image segmentation with train-free strategies
– Bridge Med-SAM with CLIP into text-guided segmentation model.

CVPR '24 Challenge: Segment Anything In Medical Images On Laptop Apr 2024 – Jun 2024

– Achieved the **Winner Finalist Award (#3/102)**.
– Proposed an efficient knowledge distillation framework that achieved better results in **under one week** with a single GPU, compared to the baseline (20 A100 GPUs for **two weeks**).
– Increased model inference speed by **2**× for 2D and **2.7**×for 3D multi-object instance segmentation under computationally restricted Docker environments
– Achieved an average DSC of **85.90**% and an average NSD of **87.07**% on the validation set, reflecting a significant improvement of +2.68 and +4.36, respectively.
– Curated more than 10 diverse datasets with uniform preprocessing pipelines.

Relvant Project

- Implementation of two stage post-training (SFT and RLHF) on LLaMA2 7B with 100k medical conversations.
- Implementation of `MLPs`, `VGGNets`, `GoogLeNet`, `U-Nets`, `RNNs`, `LSTMs`, `Transformers` (Deep Learning Assignments)
- Implementation of `Autoencoders`, `VAEs`, `GANs` on MNIST
- Implementation and analysis of `Forward Sampling`, `MCMC Sampling` and Likelihood Weighting for Approximate inference (Machine Learning Course)

Conference Activity

CVPR 2024 Workshop on Foundation Models For Medical Vision

- Presentation for the summary of proposed method and results analysis for the challenge.

Skills

Coding: Python (PyTorch, NumPy, Scikit-Learn, OpenCV, ITK), C/C++, LaTeX, Bash Scripting
Tools: Git, Docker, Linux, MySQL, SQLite
Language: Chinese, English

Miscellaneous

Jul 2022 – Mar 2023: Perplexed, rethought of "what's next", and preparation for postgraduate entrance exam.