Usage Report for Fiscal Year 2024

Project Title:

Reliable Machine Learning Approach for Medical Imaging Processing

Name:
○Li Chao (1), Chang Ziqing (1), Li Binghua (1), and Wang Minghao (1),
Laboratory at RIKEN:
(1) RIKEN Center for Advanced Intelligence Project, Tensor Learning Team.

1. Background and purpose of the project, relationship of the project with other projects

The utilization of the supercomputer is closely linked to the Kiban-B project (No. 24K03005), which aims to enhance machine learning-based data fusion methods within the domain of medical informatics. A key objective of this project is to introduce novel benchmark datasets tailored for the medical field, providing a valuable platform for computer scientists to evaluate machine learning techniques using real-world medical data—an aspect that is often overlooked in cutting-edge ML research. To achieve these goals, large-scale processing of brain imaging data necessitates the use of supercomputers. Specifically, CPU nodes play a critical role in data preprocessing and feature extraction, while GPU nodes are indispensable for training deep neural networks efficiently.

2. Specific usage status of the system and calculation method

At the current stage, our primary usage has been GPU-based, accounting for approximately 40% of our allocated resources. The total computing cost amounts to 21,893.3 core hours, as reflected in the system usage portal.

3. Result

By leveraging the Hokusai supercomputing system, we have successfully implemented novel approaches in two key areas: 1) our progress in this area has been documented in a preprint on arXiv and is currently undergoing revisions for journal submission; 2) Diffusion Model-Based MRI Image Generation: Research in this area is still in progress, with ongoing development efforts to refine the methodology and improve the generated image quality. Looking ahead, we plan to further enrich our training datasets by incorporating multiple modalities, which will enhance the robustness and generalizability of our models.

4. Conclusion

The Hokusai supercomputer has been crucial for our medical AI research, enabling interpretable brain image classification and MRI image generation using GPU nodes. Its computational power accelerates large-scale data processing, advancing robust AI models. Further improvements in data preprocessing and model optimization are needed. Increased CPU utilization will enhance multi-modal data handling, improving model accuracy and generalizability. We will continue refining methodologies, expanding datasets, and optimizing efficiency to push medical AI forward.

5. Schedule and prospect for the future

In the coming year, we plan to expand our computational usage by increasing reliance on GPU nodes to further enhance the performance of our learning models. As the data collection phase concludes, CPU utilization will become increasingly significant, particularly for large-scale data preprocessing and feature extraction.

6. If no job was executed, specify the reason.

No CPU jobs have been executed as our workload relies on GPU acceleration. CPU-based processing is in early testing on local machines and will be used for data preprocessing after data collection.

Fiscal Year 2024 List of Publications Resulting from the Use of the supercomputer

[Paper accepted by a journal]

None.

[Conference Proceedings]

None.

[Oral presentation]

None.

[Poster presentation]

1. Chao Li, Symmetry Breaking: Efficient Structure Search for Quantum Tensor Network Systems, 7th R-CCS International Symposium, Jan. 23-24, Kobe, Japan
2. Chao Li, Yuchen Cong, Tensor Network and its Structure Search Problem, Mechanism of Brain and Mind Winter workshop 2025, March. 10-12, Hokkaido, Japan

[Others (Book, Press release, etc.)]

Li, Binghua, et al. "MAProtoNet: A Multi-scale Attentive Interpretable Prototypical Part Network for 3D Magnetic Resonance Imaging Brain Tumor Classification." arXiv preprint arXiv:2404.08917 (2024).