Hyperparameter Sensitivity Analysis of ViTs in Computational Fluid Dynamics

Joseph Quinn[1], Pei Zhang[2]

[1]School of Engineering, Vanderbilt University

[2] Computational Sciences and Engineering Division, Oak Ridge National Laboratory

DOE Omni Technology Alliance INTERNSHIP PROGRAM

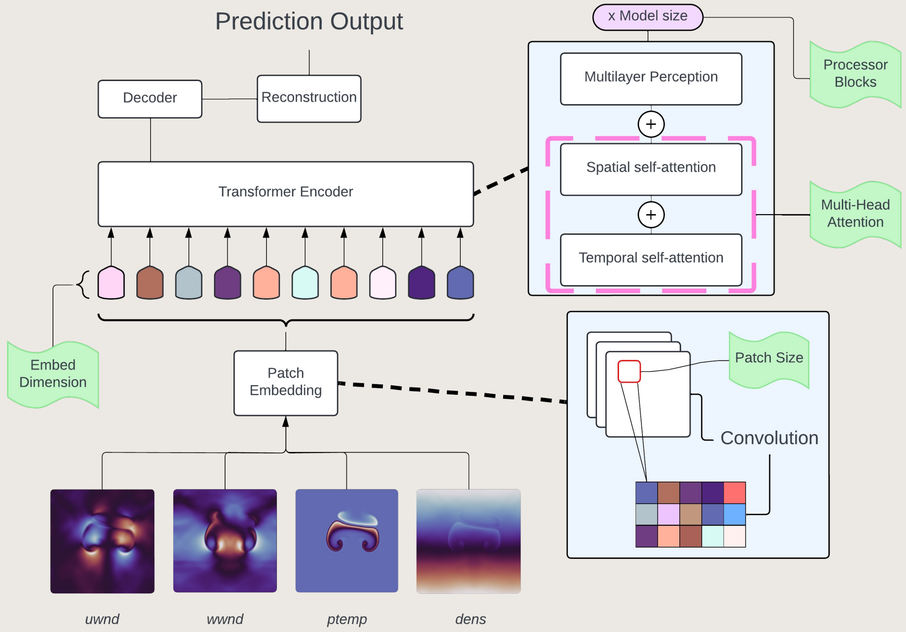
de omnibus dubitandum

Be suspicious of everything

OAK RIDGE National Laboratory

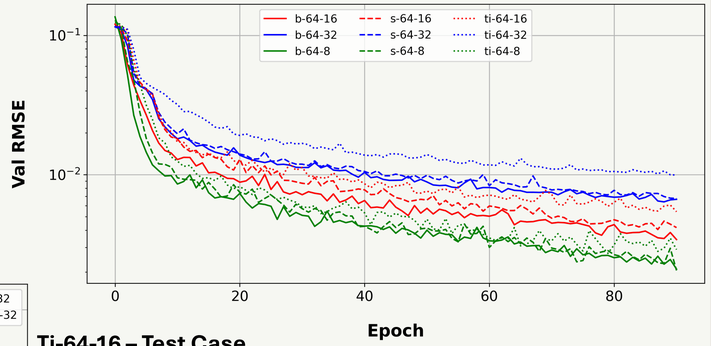
Introduction

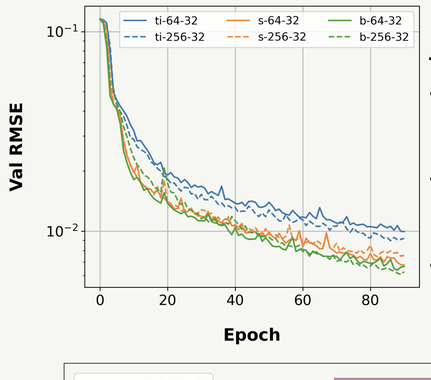
- Physics-based fluid simulations are resource-intensive and time-consuming. Leveraging data-driven machine learning models like vision transformers (ViTs) can offer swift predictive capabilities once deployed.
- However, training these models require significant demands on data, time, and computational resources.
- My work focused on analyzing the hyperparameter sensitivities to enhance performance while minimizing cost during training.



Results

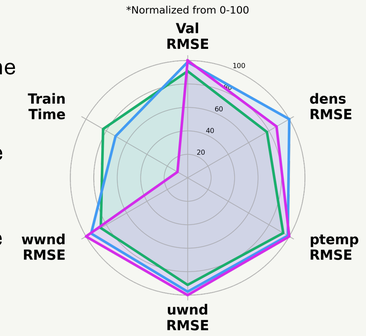
Model shows strong sensitivity to patch sizes while exhibiting low sensitivity to variations in model sizes and batch sizes.

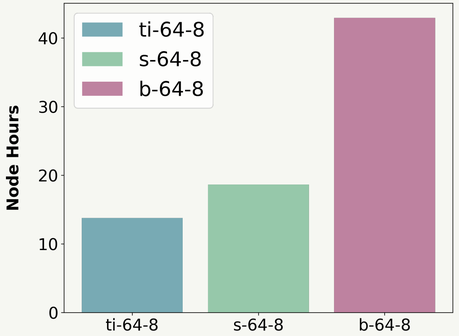




Ti-64-16 – Test Case

- **64 –> 256** Batch Size:
  - 1.91% Increased Acc.
  - 264.04% Increased Time
- **16 –> 8** Patch Size:
  - 45.07 Increased Acc.
  - 94.91% Increased Time
- **Ti –> s** Model Size:
  - 26.64% Increased Acc.
  - 40.12% increased Time





Model Size: Ti to S:

- Acc. Improvement: 21.43%
- Training Time Increase: 35.51%
- Slope: 0.603

Model Size: S to B:

- Acc. Improvement: 4.77%
- Training Time Increase: 130.55%
- Slope: 0.037

Findings

- The model's sensitivity to patch size is consistent with other ViT studies. Smaller patches capture finer spatial details and local interactions, enhancing feature extractions and improving predictive accuracy.
- The smallest model size Ti and its self-attention layers had sufficient capacity to learn and represent the fluid dynamics present in my test case dataset. Increasing model size and the self-attention layers did not capture any new relevant patterns within the fluid data.

Implications

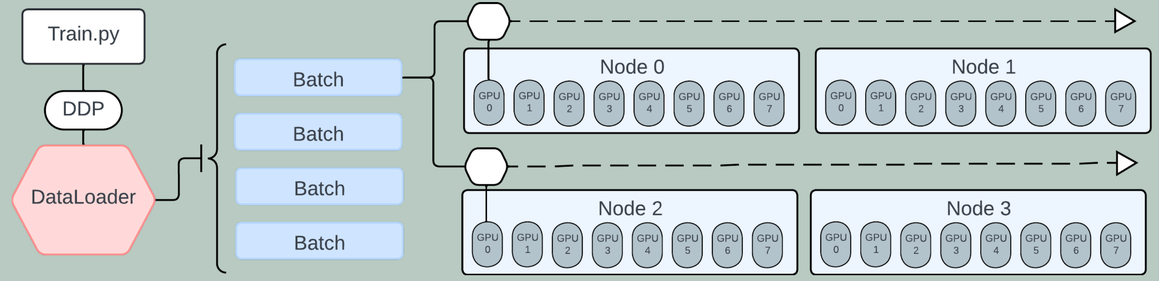
The ability to rapidly and accurately simulate fluid behaviors enables more effective design, control, and optimization of engineering systems. Our work on complex fluid dynamics will advance both academic research and practical solutions in areas such as drug delivery, engine combustion efficiency, climate modeling, and aircraft propulsion.

Methods

Parameter Research → Configuration Scripting → Computing

- Developed a Python script to generate and manage 27 unique experiment cases by varying hyperparameters of the ViT network architecture.
- Submitted 75+ SLURM jobs on the **Frontier** supercomputer.

Patch Size	Batch Size	Model Size
[8,8]	64	Ti
[16,16]	256	b
[32,32]	1024	s



Acknowledgement

This research was supported in part by the U.S. Department of Energy Omni Technology Alliance Program, administered by the Oak Ridge Institute for Science and Education.
This research is sponsored by the Artificial Intelligence Initiative as part of the Laboratory Directed Research and Development (LDRD) Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the US Department of Energy under contract DE-AC05-00OR22725.
This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DEAC05-00OR22725.

References

- Chen, J., Li, Q., & Xu, Y. "Multiple physics pretraining for enhanced generalization." *arXiv:2310.02994*, arXiv, 2023.
- Norman, Matt. *MiniWeather*. GitHub, n.d. Retrieved from https://github.com/mrnorman/miniWeather/tree/main

VANDERBILT UNIVERSITY

OAK RIDGE INSTITUTE FOR SCIENCE AND EDUCATION