

Advancing Sci-ML through large pre-trained video models

IDP/ Master Thesis/ Guided Research/ Semester-Thesis

This project explores how large pre-trained video generation models can be fine-tuned for scientific machine-learning datasets to model complex spatiotemporal phenomena. By adapting foundational video models to complex fluid dynamic problems, the project aims to generate realistic, physically meaningful videos that capture domain-specific dynamics. The work will investigate efficient fine-tuning strategies, including LoRA, adapters, and selective layer freezing, while incorporating scientific constraints such as conservation laws, boundary conditions, and uncertainty estimates. Beyond visual realism, the project will evaluate generated outputs using a wide variety of metrics, including pointwise, spectral, structure-aware, and physics-informed.

Milestones

- Understand Wan ([Wan et al. 2025](#)) / LTX-Video ([HaCohen et al. 2024](#)) architecture and implement the pre-trained checkpoint loading.
- Various experiments on finetuning given CFD Datasets using both full and LoRA finetuning.

Requirements

- Good understanding of pre-training and fine-tuning foundation models
- Strong Python and Pytorch understanding.
- Ability to understand and debug large repositories.
- Ability to work independently.

Contact

Harish Ramachandran: harish.ramachandran@tum.de with the subject "Interested to contribute to finetuning of pre-trained video generation models". Please also attach your CV, current grade report, and link to Github (if worked on any open-source projects).

References

- HaCohen, Y., Chiprut, N., Brazowski, B., Shalem, D., Moshe, D., Richardson, E., Levin, E., Shiran, G., Zabari, N., Gordon, O., Panet, P., Weissbuch, S., Kulikov, V., Bitterman, Y., Melumian, Z. & Bibi, O. (2024), 'Ltx-video: Realtime video latent diffusion', *arXiv preprint arXiv:2501.00103* .
- Wan, T., Wang, A., Ai, B., Wen, B., Mao, C., Xie, C.-W., Chen, D., Yu, F., Zhao, H., Yang, J. et al. (2025), 'Wan: Open and advanced large-scale video generative models', *arXiv preprint arXiv:2503.20314* .