

Xiang Fan

xiangfan@cs.washington.edu | xiangfan.io

EDUCATION

-
- University of Washington** | Ph.D. in Computer Science Sep 2023 – Present
- Advisor: Ranjay Krishna
- Carnegie Mellon University** | B.S. in Computer Science; GPA: 4.0 Aug 2019 – May 2023
- Graduated with University Honors
 - Selected Coursework: Machine Learning with Large Datasets (PhD), Convex Optimization (PhD), Deep Reinforcement Learning & Control (PhD)

RESEARCH INTERESTS

Video Understanding, Video Generation, Multimodal Learning, Long-term Memory, Computer Vision, Vision & Language

PUBLICATIONS

Nano: Nested Human-in-the-Loop Reward Learning for Controllable Text Generation.

Xiang Fan, Yiwei Lyu, Paul Pu Liang, Ruslan Salakhutdinov, and Louis-Philippe Morency.

In *Findings of the Association for Computational Linguistics (ACL)*, 2023.

Quantifying & Modeling Multimodal Interactions: An Information Decomposition Framework.

Paul Pu Liang, Yun Cheng, Xiang Fan, Chun Kai Ling, Suzanne Nie, Richard Chen, Zihao Deng, Nicholas Allen, Randy Auerbach, Faisal Mahmood, Russ Salakhutdinov, Louis-Philippe Morency.

In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

MultiBench: Multiscale Benchmarks for Multimodal Representation Learning.

Paul Pu Liang, Yiwei Lyu, Xiang Fan, Zetian Wu, Yun Cheng, Jason Wu, Leslie Chen, Peter Wu, Michelle Lee, Yuke Zhu, Ruslan Salakhutdinov, and Louis-Philippe Morency.

In *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

HighMMT: Towards Modality and Task Generalization for High-Modality Representation Learning.

Paul Pu Liang, Yiwei Lyu, Xiang Fan, Shentong Mo, Dani Yogatama, LP Morency, and Ruslan Salakhutdinov.

In *Transactions on Machine Learning Research (TMLR)*, 2023.

RESEARCH EXPERIENCE

UW RAIVN Computer Vision Group Sep 2023 – Present

- Extracted optical flow signals from Modelscope Text-to-Video model to show that they contain motion information about the generated video
- Investigated the effect of perturbing hidden states in Stable Video Diffusion model on the generated motion
- Designed quantized regression loss function to improve stability and performance on regression
- Investigated KITTI and Spring optical flow datasets and measured generalization performance on these datasets
- **Advisors:** Prof. Ranjay Krishna and Anand Bhattad

CMU MultiComp Multimodal Lab

Feb 2022 – Mar 2023

- Designed and implemented human-in-the-loop training algorithm, training procedures, and loss functions
- Designed tree search algorithm with latent updates from a critic network
- Created web interface and analysis for large-scale Amazon Mechanical Turk crowd-sourcing experiments
- Created human interface and acquired participants for small-scale experiments
- Wrote the introduction, methods, experiments, and results sections in the paper
- **Advisors:** Prof. Louis-Philippe Morency and Paul Pu Liang

CMU MultiComp Multimodal Lab

Sep 2020 – Jun 2021

- Designed and implemented Early/Late Fusion, Tensor Fusion, Multimodal Transformer, and Gradient Blending methods for robotics and financial time series datasets
- Collected and analyzed financial time series data from companies in different private sectors
- Performed time complexity and robustness analysis for each method-dataset pair and reported the results in the paper
- **Advisors:** Prof. Louis-Philippe Morency and Paul Pu Liang

CMU MultiComp Multimodal Lab

Sep 2021 – Jan 2022

- Applied a Perceiver-based multimodal model to robotics and financial time series tasks
- Conducted experiments with a unified model on multiple datasets for modality/task generalization and reported the results in the paper
- **Advisors:** Prof. Louis-Philippe Morency and Paul Pu Liang

CMU Natural Language Processing Group

Feb 2022 – Jul 2023

- Designed and implemented a content-based soft grouping mechanism that reduces the sequence length by 50% in Transformer models
- Performed large-scale pretraining on the Wikipedia & Books corpus and optimized pretraining performance
- Benchmarked the model against Funnel Transformer and BERT baselines on the GLUE benchmarks
- **Advisors:** Prof. Yonatan Bisk and Prof. Emma Strubell

INDUSTRY EXPERIENCE**NVIDIA NLP Team** | Applied Deep Learning Research Intern

May 2022 – Aug 2022

- Designed and implemented latent adapters to efficiently fine-tune encoder-decoder models
- Performed distributed training experiments on WMT 14 and WMT 20 machine translation datasets using mT5 and Perceiver models
- Investigated several loss functions, latent matching, decoder feature matching, conditional GAN, and WGAN-GP objectives
- Conducted experiment to gain the insight that translated latent space is closer to the source than to the target, contrary to intuition
- Proposed composing tasks in the latent space for efficient task composition, and presented the results
- **Advisors:** Sandeep Subramanian, Micha Livne, and Oleksii Kuchaiev (NVIDIA)

Google | Software Engineering Intern

Jun 2021 – Aug 2021

- Designed model stability monitoring & analysis for the Unified Store Sales team with 1000+ Ads customers and 300M+ monthly transaction value, minimizing model variation impact and improving future launch impact analyses
- Created Borg jobs to run scalable TensorFlow model experiments, read/query data sources from SpannerDB / RecordIO / SSTable / TFX, and aggregate results using parallelized Flume pipelines
- Created a dashboard to display large-scale stability data from 6 internal model types, with custom grouping and slicing

OTHER PROJECTS & CLUB ACTIVITIES**CMU Game Creation Society**

Sep 2019 – Dec 2019

Google Summer of Code

May 2019 – Aug 2019

Efficient Human Pose Estimation on Mobile

Arp 2018 – Nov 2018

User-friendly Arch Linux Installer (Maintainer)

2018 – Present

HONORS & ACHIEVEMENTS**CRA Outstanding Undergraduate Researcher Awards, *Honorable Mention***

2023

Dean's List with High Honors

2019, 2020, 2021, 2022

USA Computing Olympiad, Platinum Division

2018

SKILLS

Programming Languages: Python, C, C++, Rust, R, Dart, Java, JavaScript, HTML, SQL, Standard ML
Frameworks & Tools: PyTorch, CUDA, TensorFlow, NumPy, Docker, Android Dev, Git, Linux, \LaTeX