

Wenxuan (Eden) Tan

github.com/Edenzzzz | wtan45@wisc.edu | (+1)608-4405-251

Education

Shenzhen Nanshan Concord College of Sino-Canada

Shenzhen, China

University of Wisconsin-Madison

Sep 2019 – May 2022

B.S. in Computer Science & Mathematics

Madison, WI, USA

- GPA: 3.95
- Advisor: Prof. Fred Sala, Postdoc Stanford University

Sep 2022 – May 2025 (Expected)

Experiences

Foundation models research assistant

Sprocket Lab, UW-Madison

Advisor: Prof. Fred Sala

Sep 2023 - Now

- Working on Parameter efficient finetuning of LLM via expressive, sparse, low-rank and hardware-efficient Monarch matrices (learnable FFT). Conduct large-scale architecture & hyperparam search using Async Successive Halving Algorithm (ASHA). **Surpassing LoRA in performance and efficiency. Our team is actively pursuing ICML/ACL.**

Computer Vision research assistant

Computer Vision Group, UW-Madison

Advisor: Prof. Vikas Singh

Nov 2022 - Sep 2023

Project 1 (Open sourced with 167 ★)[[Link](#)]:

- TensorFlow re-implementation of ECCV 2022 best paper *On the Versatile Uses of Partial Distance Correlations*

Project 2: *Improving compositional generation of diffusion models* [[Link](#)]

- Aim to improve complex prompts with composed attributes/objects. Wrote and maintained code for all experiments. Implemented a GNN-augmented Latent Diffusion Model using syntax parser. Set up evaluation pipeline to detect objects using DETA
- Trained MLPs discriminators constrained by Category Theory axioms to detected poorly composed image embeddings, by how structure/distances across modalities are preserved. (**distribution score estimation & score matching**)
- Implemented distance metrics (Distance Correlation, Wasserstein distance etc.) to optimize intra-phrase attention overlap and discourage that between phrases. Improved samples over Attend and Excite.

Project 3: *Operator-theoretic Implicit Neural Representations*

- Viewed 3D Neural modeling as learning **function-space to function-space instead of as position-encoding to signal/data mapping**. Trained purely convolutional INR on multiple multi-resolution volumes to represent data.
- Adapted the NeSVoR slice-to-volume reconstruction pipeline (MLP + Convolutional INR as density net) with monte-carlo sampled losses via trilinear devoxelization & voxelization. **Achieved 5e-4 MSE on the FeTA MRI dataset**
- Proposed a method combining continuous convolution with Octree CNN to train on high resolutions with reduced compute.

Robotics Club Team Leader

Sep 2020–July 2021

- Organized training; held fund raisers; programmed PID algorithm and detection and shooting algorithm for autonomous period.

Technical Skills

Languages Python, Java, C, C++

Frameworks and Tools PyTorch, Git, TensorFlow, CUDA, OpenMP, Docker, OpenCV, NCCL, Pytorch3D, Docker, SQL, HDFS, Spark, Cassandra, Kafka, MPI

Relevant courses

Undergrad: Neural Nets, Honors Calculus & Diff eq & Linear Alg, DL for CV, Big data systems, Probability, Computer Organization

PhD level: Parallel Programming, Foundation Models

Self-studied: Stanford CS231n;

In-progress: CMU 15-445 Database, CMU 10-414 DL Systems

Research statement

I study foundation models and MLSys. I'm also peripherally interested in Vision problems such as compositional 2D/3D generation and perception.

- On the ML side, I'm interested in building and interpreting multimodal foundation models, especially machines that can interact with the world in a cognitively human-like and mathematically founded way.
 - On the systems side, I'm interested in efficient training and serving of foundational models at scale, via methods like ZeRO and Alpa. I'm always excited about new methods for parallelism, reducing/hiding comm cost and memory management.
-

Side Projects

Parallel matrix multiplication for memory efficiency and performance in ML [\[Link\]](#)

Dec 2023, CS759 HPC for Eng Applications Project

- Implemented non-square tiled matrix multiplication and benchmarked it on a NVLinked cluster of 8 * A100 for FLOPS etc. Implemented a 3-layer MLP trained on MNIST in **Tensor Parallel** in CUDA from scratch with **online softmax**.

Text to Image Book Cover Generation

Nov 2022, CS539 Intro-Neural Nets Project

- Tested the transfer learning capability Stable Diffusion. Employed numerous memory efficient tricks(8-bit Adam, grad accumulation, grad checkpointing) to allow its training on a single 12GB GPU on 60000 text-image pairs. Improved on generated book cover quality/specitivity.

Edu-X: An educational AI app

August 2021

- Led a team of undergrad students to designed a system consisting of multiple models(YoloV3, eye status classifier, mouth status classifier, emotion classifier) for analyzing student response to evaluate the teacher's teaching efficacy.
- Trained tens of emotion classification models to achieve SOTA on FER-2013 dataset and near human accuracy; enhanced performance by collecting and labeling extra (~2000) real-classroom images

Fantaface

August 2022

- Designed and implemented various functionalities including the detection and facial feature analyzing pipeline, emoji generation, setting emoji as wallpaper, and copying to the clipboard.
-

Awards/Test Scores

- SAT 1560
- AP 9*5
- Undergraduate Scholarship for Summer Study, UW-Madison
- FRC Technology Innovation Award
- DJI RoboMaster Youth regional second place
- Kaggle Feedback Prize top 13%