Nanye (Willis) Ma

Email Website Linkedin Github

Education

New York University **New York, NY**

Honors B.A. in Mathematics; B.A. in Computer Science; 3.9/4.0 GPA *September 2020 - Expected May 2024*

- **Selected Coursework:**
  - **Graduate Level**: Machine Learning, Computer Vision, Probability Theory I, Portfolio and Risk Management, Programming Language.
  - **Undergrad Level**: Honors Analysis, Honors Algebra, Honors Theory of Probability, Honors Numerical Analysis. Operating System, Basic Algorithm, Applied Internet Technology.

Publications / Preprints

- **N. Ma**, M. Goldstein, M. S. Albergo, N. M. Boffi, E. Vanden-Eijnden, S. Xie. SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers. *In submission to CVPR 2024.*

Research Experiences

SiT: Exploring Flow and Diffusion-based Generative Models with Scalable Interpolant Transformers

May 2023 - Present

- Proposed Scalable Interpolant Transformers (SiT), a novel family of generative models built on dynamical transport, effectively merging Flow and Diffusion-based models.
- Conducted a comprehensive study of the SiT design space, focusing on the three critical components: *model prediction and objectives (velocity vs. score)*, *interpolant parameters (choices of α_t and σ_t)*, and *sampling methods (ODE vs. SDE).*
- Employed a methodical approach by independently analyzing each above component within the design space while holding the others invariant, and undertook a detailed analysis of the performance characteristics among different component combinations.
- Innovated a novel backward SDE construction that incorporates a customized diffusion coefficient, enhancing both flexibility and efficacy in the diffusion sampling process.
- Demonstrated that SiT equipped with certain combination of components (*velocity + linear interpolant + SDE with linear diffusion coefficient*) significantly surpasses the performance of baseline DiT model, achieving twice the converging speed across various scales.
- With ODE sampler, the largest SiT-XL model attained a comparable **2.15** FID score with **50** number of function evaluations (NFE), surpassing DiT in both sampling speed and quality.
- With SDE sampler, SiT-XL attained a competitive **2.06** FID score under same NFE as DiT, pushing the performance even further.
- Submit a paper (as 1st author) to CVPR2024.

Text-to-Image Scalable Diffusion Models with Transformers

Feb 2023 - May 2023

- Integrated a pre-trained DistilBERT encoder with Diffusion Transformers (DiT) to enable rapid and high-quality text embedding.
- Achieved a 78% reduction in memory consumption by optimizing training through gradient checkpointing and pre-extracted VAE image features, enhancing training speed to 0.32 steps/sec on a single A100 GPU.
- Fine-tuned text-to-image DiT from pre-trained class-conditional DiT checkpoint by unfreezing Embedders and inject randomly initialized weights to adaLN modules.
- Benchmarking on MS-COCO, the fine-tuned model attained a competitive FID score of 15.49 with the LDM-KL-8-G model with only 200K training steps.

Deep Marching Tetrahedra on Non-Watertight Models

Oct 2022 - Dec 2022

- Re-engineered Deep Marching Tetrahedra (DMTet) network for high-resolution 3D shape synthesis and optimized its training process for non-watertight 3D meshes.
- Replaced the PVCNN Encoder with a fine-tuned Inverse Distance Weighted KD Tree, enabling more detailed and precise spatial encoding of non-watertight models' point cloud.
- Enhanced the Surface Subdivision module with a learnable Signed Distance Field threshold to self-optimize with respect to different surface topologies, extracting more refined surfaces.
- The optimized model significantly outperformed the baseline DMTet on Non-Watertight Models in both L1 and L2 Chamfer Distance on non-watertight meshes, attaining best scores of 5.46 and 3.89, respectively.

Modeling for Green GDP **Feb 2023 - Mar 2023**

- Built the Predicted GDP model using the Cobb-Douglas function, incorporating policy rewards based on carbon neutrality and sustainability.
- Devised a Green GDP model by integrating the Predicted GDP model output with a fine-tuned decaying factor.
- Optimized Green GDP using Euler-Lagrangian conditions, and determined the optimized policy rewards by constructing a convex optimization problem with affine constraints in carbon emission and sustainability factors.
- The optimized model outperformed traditional GDP model, improving the growth rate by 7.1% and reducing global CO2 emission by 14.9% over a ten-year period.

Optimization for Road Cycling Power Distribution **Mar 2022 - Apr 2022**

- Constructed a non-linear model based on existing data from top cyclists and achieved a deviation of less than 3.6% compared to the average podium finish times in each time trial.
- Implemented Genetic Algorithm with Scikit-learn to find the optimal power distribution curve over all time trial courses, enabled the predicted finish time to decrease by an average of 9.3%.

Projects

***Dream Diffusion*, Full Stack Development Project** **Jan 2023 - May 2023**

- Orchestrated the development of a full stack application using Next.js, designed specifically for recording dreams and writing dream journals with DallE backbone.
- Engineered session management from scratch to handle authentication and authorization, utilizing Redis as session store for efficient information retrieval.
- Designed robust authentication middleware using Passport.js and JWT-based strategies to bolster application security and safeguard user privacy.
- Optimized Serverless API architecture to support large-size (1024×1024) images transfers between frontend and backend, maintaining a low latency of up to 120ms for seamless user experience.

***HyperEX*, Full Stack Development Project** **May 2022 - Jan 2023**

- Played a pivotal role in the development team of HyperEx, an innovative online immersive social platform supporting hundreds of users in a shared virtual space.
- Designed a Finite State Machine library and Inverse Kinematics algorithms in JavaScript, enhancing the player's motion system, rag-doll system, and inventory system.
- Developed RESTful API endpoints using Express and Socket.io to facilitate communications between game engine and the application, resulting in seamless interactions between user interfaces and game renderings with a low latency of up to 50ms.
- Engineered a performance optimization module with Draco to ensure the platform's performance on outdated devices, and further increases the game's frame rate by over 300

***Weensy OS*, Operating System Development Project** **Sep 2021 - Nov 2021**

- Implemented a small OS with 2 MB physical and 3MB virtual memory, including one basic shell command ls, multithreaded kernel and userspace, paging and address allocation, and a FUSE file system.

Technical Skills

Coding Languages: Python, JavaScript, C++, Java, C, Standard ML, Scheme, Prolog, Assembly, Julia
Tools/Frameworks: Github, Linux / JAX, Flax, Optax, PyTorch, Numpy, Scipy, Pandas, Node.js, React, Express, MongoDB, Next.js

Honors

Summer Undergraduate Research Experience (SURE) Fund
Dean's Undergraduate Research Fund
Dean's List for Academic Year 2020-2021, 2021-2022, 2022-2023
MCM Success Award 2022, 2023