Alan (Jia Lin) Yuan

alan.yuan.jly@gmail.com | linkedin.com/in/jalnyn | github.com/jalnyn | scholar.google.com/

Education

University of Toronto 4.0 cGPA
MSc in Applied Computing (A.I.) *Sept 2023 – Dec 2024*

University of Toronto 3.84 cGPA
HBSc Computer Science Specialist, Major in Mathematics *Sep. 2018 – May 2023*

- **Relevant Coursework**: Parallel Programming (A+), Neural Networks and Deep Learning (A+), Algorithm Design, Analysis & Complexity (A), Operating Systems (A+), Intro to AI (A+)

Publications

M. Skreta*, Z. Zhou*, **J. L. Yuan***, K. Darvish, A. Aspuru-Guzik, A. Garg. RePLan: Robotic Replanning with Perception and Language Models. *arXiv:2401.04157*, 2024.

M. Mittal, C. Yu, Q. Yu, J. Liu, N. Rudin, D. Hoeller, **J. L. Yuan**, R. Singh, Y. Guo, H. Mazhar, A. U. Mandlekar, B. Babich, G. State, M. Hutter, A. Garg. ORBIT: A Unified Simulation Framework for Interactive Robot Learning Environments. *IEEE Robotics and Automation Letters (RA-L)*, 2023.

Work Experience

Instacart Feb 2025 – Present
Software Developer *Toronto, Ontario*

Unilever April 2024 – Dec 2024
Machine Learning Engineer - Intern *Toronto, Ontario*

- Trained **time-series transformer** models to **forecast sales** with **PyTorch**, increasing accuracy by 5%.
- **Optimized promotion times and prices** with augmented genetic algorithms **GNGSA**.
- Utilized novel algorithm **GNGSA** increasing convergence iterations 70% and hypervolume AUC by 50%.
- Implemented multi-thread processing to increase data pipeline speeds by **50%**.

Amazon Jun 2023 – Aug 2023
Software Developer - Intern *Toronto, Ontario*

- Designed a **precompute** layer using **Apache Spark** to increase recommendation speed by **99%** from 150ms.
- Built data pipelines to ingest updated data and retrain models on a scheduled basis.
- Created automated data analysis tool to ensure predictions are above **75%** accuracy.

Amazon May 2022 – Aug 2022
Software Developer - Intern *Vancouver, British Columbia*

- Engineered a modular microservice in **Java** to send cashback notifications to customers on select products.
- Utilized **AWS** web services such as **Lambda**, **SQS** and **SNS** to ensure scalability.
- Integrated service into data pipeline utilizing Amazon internal language **Datapath**.

Intel May 2021 – May 2022
Software Engineer - Intern *Toronto, Ontario*

- Developed support software to generate 4000+ completely random test-cases for edge-case testing.
- Optimized support tool's RAM templates to reduce false positives and failing cases by around **70%**.

Centivizer Apr 2020 – Sep 2020
Software Developer - Part-time *Toronto, Ontario*

- Designed backend application using **Node.JS** and **SimplePeer** to facilitate **real-time video streams**.
- Integrated video system with main **React** site and **NoSQL** database.

Projects

WIDE: Websocket-IDE | GitHub — Sep 2024 – present

- Architecting and developing a websocket-based IDE using **Rust** and **React** frontend (58 GitHub stars).

Decomposed Face Generation | GitHub — Apr 2024 – Aug 2024

- Decomposed faces into **id, pose, emotion embeddings** using **contrastive learning** approaches.
- Developed architecture combining **StyleGAN** with multiple embedding spaces for controllable generation.

RePlan: Robot Planning with LLMs | arXiv:2401.04157 — Sep 2023 – Aug 2024

- Developed novel system combining **MPC** and **MultiModal LLM** for completing high-level robot tasks.
- Created **MuJoCo** environments for training and evaluating manipulation policies.
- Implemented meta-controllers for **multi-policy** agents using **sparse rewards**.

MultiModal AI Story Teller | private repo — Jul 2023 – Jul 2024

- Built interactive **socket-based** streaming service serving up to 200+ users a day.
- Managed **auto-scaling GPU** resources with **Kubernetes** saving up to 90% on **AWS** costs.
- Implemented **Multimodality** using **LLMs** and **Latent Diffusion Models** to build an interactive story teller.
- Utilized NLP techniques to summarize context to reduce context size, reducing inference time by up to 10%.

Orbit: Robot Learning Framework | project-site — Jan 2022 – May 2023

- Core contributor to robotics framework built on **NVIDIA IsaacSim**, published in *IEEE RA-L*.
- Designed **GPU-accelerated** state systems achieving a **4x** speedup over CPU solutions.
- Implemented **PPO** training pipelines with **domain randomization** for manipulation tasks.
- Wrote a low-overhead semantics system to represent semantics and physics simulation.

CaNetDa: Deep Learning for GeoGuesser in Canada | GitHub — Jan 2021 – Apr 2021

- Mined dataset and trained an ensemble of **ResNet**, **EfficientNet** and **Vision Transformer** models.
- Achieved **47%** improvement over random agent in predicting province of image in Canada.
- Scraped **Google Street View** images and used **Google Maps API** to preprocess images for training.

Deep Q-Learning Snake Agent | GitHub — May 2021 – Dec 2021

- Utilized **PyTorch** to write a Deep Q-Learning snake agent reaching a high score of **40** after **5** minutes of training.

Tron UDP Multiplayer | GitHub — Sep 2019 – Dec 2019

- Created a four-player game for local networks using the **UDP network protocol** and C++.
- Utilized **epoll** for both client and server to monitor the socket, timer (server) and stdin (client).

Technical Skills

Tech: Deep Neural Networks, Large Language Models, Latent Diffusion Models, Reinforcement Learning, Python, C++, JavaScript, Java, Rust
Tools: Huggingface, PyTorch, Git, React, Node.js, MongoDB, SQL, Numpy, GraphQL, NVIDIA IsaacSim, MuJoCo, Kubernetes, AWS, Apache Spark