Daniel Choi

647-336-3248 | danielchoi360@gmail.com | Linkedin | Website

Experience

Undergraduate Thesis Researcher Sept 2023 – May 2024
University of Toronto, Autonomous Systems and Biomechatronics Lab *Toronto, ON*

- Built a social robot navigation training environment in **Isaac Gym**, achieving 75% increase in training efficiency
- Improved **reinforcement learning** policy performance by 50% using the Eureka framework by Nvidia
- Reduced training times by 60% using **CUDA** parallel programming for robot policy optimizations
- Optimized NN models for CPU/GPU **inference** with **Pytorch distributed training**, cutting latency by 25%

ML/Software Engineer Jan 2023 – Dec 2023
ONE800 *Toronto, ON*

- Spearheaded **user engagement** optimization project, increasing user interaction by 20% with predictive modeling
- Enhanced an **LLM chat-bot** with **long-term memory** and increased user base by 12%
- Integrated **OCR** feature with **GCP** & **OpenAI API**, increasing daily active users by 10% and feature use by 15%
- Deployed **autonomous agents** for customer support, reducing response time by 25%

Mechatronics Engineer July 2021 – Aug 2022
Thornhill Medical *Toronto, ON*

- Led the ventilator algorithms team using **predictive learning**, contributing to UHN research
- Enhanced ventilator airflow by 20% with an **RL-based PID Controller** optimization in **C++/Linux**
- Performed **quality validation** with physicians, researchers, and military professionals for **Ukraine deployment**
- Utilized **Python** to engineer and visualize ventilator flow data, aiding hypoxia emergency responses of 10 patients

Artificial Intelligence Researcher May 2019 – Sept 2019
University of Toronto Robotics and AI Lab *Toronto, ON*

- Engineered **IR detection** algorithm using **ROS**, **C++**, **Python**, amassing 50GB data points
- Transformed **U-Net** with **attention mechanisms** for segmentation, achieving 90% accuracy
- Utilized **MLPerf** benchmarks to evaluate and improve model performance, achieving a 20% increase in efficiency
- Implemented Google's **Inception-V3 CNN** and achieved 94.4% accuracy in classifying farm animal species
- Optimized models for **low precision inference** using **Horovod**, improving inference speed by 30%

Projects

PolySumm | *XGBoost, C++, DVC, MLFlow, BERT, TensorFlow, CUDA, MXNet* April 2024 – Present

- Developed Real-Time Multilingual Document Summarization System with active **RAG**
- Integrated **BERT** with **Llama 3**, improving summarization coherence by 34%
- Incorporated **DVC** and **MLflow** to cut development time by 15% and enhance model iteration efficiency by 20%

Face Tracker | *Python, TensorFlow, OpenCV, CUDA, Arduino* June 2020 – Sept 2020

- Designed a **face-tracking camera** with **3D-printed** platform from scratch with a custom loss function
- Increased **frame processing** speed 5x with a processing pipeline using **CUDA**
- Trained a custom model with in-house data augmented using **Albumentations**, **Labelme** for 75% accuracy

Education

University of Toronto Toronto, ON
Bachelor of Applied Science & Engineering, Minor in Robotics, Certificate in AI *May 2024*

University of Toronto Toronto, ON
Master of Applied Science & Engineering, AI/Robotics *June 2026*

Technical Skills

Languages: Python, C/C++, SQL, JavaScript, Java, R, MATLAB
Frameworks: TensorFlow, Pytorch, Keras, JAX, XGBoost, DVC, MLflow, CUDA, Hugging Face
Tools: AWS, ROS/ROS2, Isaac Gym, GCP, Langchain, Pinecone, Shell, Git, Docker, MongoDB
Libraries: Scikit-learn, SciPy, MXNet, Open3D, OpenCV, ONNX, CGAL