

## **Abed Matinpour**

Toronto, M2J 1B6, Canada

(647) 979-1124 | [abedmatinpour@gmail.com](mailto:abedmatinpour@gmail.com)

Linkedin: <https://ca.linkedin.com/in/abed-matinpour-8326921ba>

Google Scholar: <https://scholar.google.com/citations?hl=en&user=GFc61u8AAAAJ>

Research Gate: <https://www.researchgate.net/profile/Abed-Matinpour>

---

## **EDUCATION**

**Master of Science in Computer Science (Incoming), York University, Toronto, Ontario, Canada**

*Starting January 2026 (01/26)*

**Bachelor of Science, Honours in Computer Science, York University, Toronto, Ontario, Canada**

*September 2022 (09/22) - December 2025 (12/25) (Anticipated)*

*First Class with Distinction*

**Bachelor of Science, Honours in Computer Science, University of Saskatchewan,**

**Saskatoon, Saskatchewan, Canada**

*September 2020 (09/20) - April 2022 (04/22)*

---

## **RESEARCH EXPERIENCE**

**Research Assistant, Centre of Excellence for Research in Adaptive Systems (PI: Marin**

**Litoiu), York University**

*February 2025 - Present*

- Fine-tuned language models from 8B to 700B MoE architectures using LoRA-based parameter-efficient tuning and supervised feedback loops on domain-specific and math reasoning datasets, increasing accuracy while minimizing compute and storage overhead.
- Experimented with adaptive context windowing strategies using similarity-based ranking, boosting math reasoning performance on smaller models with reduced context length and latency.
- Built scalable retrieval-augmented generation (RAG) pipelines using LangChain and Llamaindex, enabling low-latency, context-aware QA and summarization across distributed and hybrid deployments.
- Leveraged TensorRT-LLM to optimize inference performance for large models and achieving significant latency and throughput improvements for real-time applications.
- Developed an automated, adaptive stress and load testing framework powered by LLM-driven agents, cutting test iteration counts by over 40% compared to baseline load-testing approaches.
- Designed a benchmarking framework for measuring accuracy, latency, GPU utilization, and energy efficiency across edge, cloud, and hybrid environments; two publications (on benchmarking and adaptive LLM-agent testing) are in progress for submission to FSE as First Author.

**Research Assistant, Laboratory of Advanced Biotechnologies for Health Assessment (PI: Razieh Salahandish), York University**  
*September 2024 - Present*

- Applied U-Net++, DeepLabV3+, and FPN for tumor detection on mouse brain MRI datasets, improving segmentation accuracy through channel optimization and output fusion.
- Integrated Particle Swarm Optimization (PSO), Dynamic Whale Optimization Algorithm (DWOA), and fuzzy logic-driven methods for channel selection and weight tuning, enhancing clinical imaging pipelines.
- Introduced game theory-based output concatenation techniques, increasing prediction accuracy across multi-model tumor segmentation workflows.
- Co-authored and contributed to research papers, including:
  - Advanced PSO-Based Channel Selection for DeepLabv3+ in Clinical Data Analysis
  - Breast Cancer Segmentation Using a Modified U-Net with DWOA Channel Selection
  - Fuzzy Logic-Driven Optimization for MRI-Based Neurodegenerative Disease Diagnosis (Not Accepted)
- Fine-tuned DistilHuBERT for stuttering detection and utilized Whisper for word-by-word speech transcription, achieving improved detection precision and fluency assessment.
- Developed a bi-directional GAN for crystal bandgap prediction, advancing machine learning applications in materials science.
- Designed a wearable device for heart attack prediction by converting PCG signals to ECG-like signals using dual ML models; contributed to both software and embedded hardware development.

**Machine Learning Research Engineer, Sarit International (PI: Andrew Maxwell), York University**  
*January 2025 - Present*

- Developed and deployed **real-time pedestrian detection, trajectory prediction, and collision avoidance systems** on SARIT micro-mobility vehicles using **YOLOv9** and **MobileNetV2**, achieving **95%+ detection accuracy** in dynamic campus environments.
- Integrated **LiDAR with OAK-D Pro camera pipelines** to enhance **3D perception, depth estimation, and object localization**, enabling more accurate detection in complex traffic scenarios.
- Optimized deep learning inference by using **TensorRT with FP16 and INT8 precision calibration**, deploying on **NVIDIA Jetson** to reduce latency by over **60%** and support real-time decision-making.
- Designed a **trajectory prediction module** that forecasts pedestrian and vehicle paths, improving proactive collision avoidance and route planning.
- Collaborated with software and mechanical teams to **redesign the computer vision and sensor fusion pipeline**, boosting frame throughput, synchronization, and processing speeds.
- Implemented a **low-latency communication protocol** for transmitting detection and trajectory alerts to user interface modules, reducing detection-to-alert time to **under 50ms**.
- Collected and annotated **custom datasets** across varied conditions (low light, occlusions, and heavy traffic) to improve model robustness and accuracy for real-world deployment.

---

## WORK EXPERIENCE

### Machine Learning Intern, Deep Dream Studios, Toronto & Dubai

*July 2023 - Sep 2024*

- Designed and trained **diffusion models and GAN architectures** for **high-fidelity, multi-style image generation**, improving visual realism and diversity while optimizing inference efficiency for large-scale deployment.
- Researched and implemented **Vision Transformers (ViT) and Video Vision Transformers (ViViT)** to generate **context-aware images and short video sequences**, enhancing **temporal coherence and object consistency** across frames.
- Applied **Variational Autoencoders (VAEs)** for **latent space manipulation and controlled synthesis**, enabling **parameterized, user-driven image and video generation**.
- Fine-tuned **transformer-based generative models** on proprietary video datasets using **distributed training and performance engineering** techniques, achieving **smoother transitions, improved frame-to-frame consistency, and reduced training time**.
- Built **scalable evaluation pipelines** leveraging **FID, PSNR, LPIPS, and throughput benchmarks** to systematically compare generative models, driving **data-informed architectural optimizations** and accelerating experimentation cycles.

### Technical Director (Volunteer), Google Developers Student Club, Saskatoon, Saskatchewan, Canada

*July 2021 - April 2022*

- Developed a **WebSocket chat application**, improving project completion times by 10%.
- Conducted workshops on modern tech stacks, advancing club members' skills in web and software development.
- Mentored **students** on full-stack development, cloud computing, and AI/ML integration.
- Led a team in **developing open-source projects**, fostering collaboration and hands-on learning.
- Organized **hackathons and coding challenges** to enhance problem-solving skills among members.

---

## TEACHING EXPERIENCE

Teaching Assistant, University of Saskatchewan, Saskatoon

September 2021 - December 2021

- Taught **Introduction to Computer Organization and Architecture** and **Principles of Software Engineering**.
- Led tutorials and facilitated lab sessions, providing support to undergraduate students.
- Conducted review sessions, helping students improve their performance by clarifying complex topics.

---

## Publications

- Norouzi Azad, A., Homam, B., Feygin, A., Fardshad, M. N. G., Rozenblat, S., **Matinpour, A.**, Laly, A., Esmaeildoost, F., & Salahandish, R. (2025). *Optimized DeepLabV3+ for Clinical Data Analysis through Advanced Particle Swarm Optimization-Based Channel Selection*. **Advanced Intelligent Systems**. <https://doi.org/10.1002/aisy.202500282>
- NorouziAzad, A., **Matinpour, A.**, Deljoo, F., Homam, B., Trivedi, B., & Salahandish, R. (2025). *Breast cancer segmentation using a modified U-Net with Dynamic Whale Optimization Algorithm (DWOA) channel selection*. Accepted for presentation and publication in the proceedings of the **11th International Conference on Machine Vision and Machine Learning (MVML 2025)**, Paris, France, August 2025

---

## Projects

- **LangChain and LlamaIndex-Based LLM Benchmarking Framework** (York University, 2025 - Present): Built a scalable orchestration framework for benchmarking large language models across accuracy, latency, GPU/memory utilization, and energy efficiency, with support for dynamic token batching and context window adaptation.
- **Adaptive Stress and Load Testing Framework for AI Web Applications** (York University, 2025 - Present): Developed an automated, LLM-driven testing system that uses adaptive agent-based strategies to generate realistic user traffic and stress scenarios, reducing test iterations by 40% compared to traditional approaches.
- **WOA Optimization for Attention U-Net** (York University, 2025 - Present): Implemented Whale Optimization Algorithm (WOA) on the Attention U-Net model, leading to significant improvements in meanIoU across diverse datasets, including breast cancer imaging. The approach demonstrated enhanced segmentation accuracy by fine-tuning attention mechanisms.

- **Speech Fluency and Stuttering Analysis Framework** (York University, 2025 - Present): Developed a speech analysis pipeline by fine-tuning DistilHuBERT for stuttering detection and customizing Whisper for word-level transcription. Designed real-time processing workflows to assess speech fluency, prosody, and pacing, enabling accurate, low-latency detection for accessibility and clinical applications.
- **PSO Optimization for DeepLabV3+** (York University, 2024 - Present): Enhanced DeepLabV3+ performance by embedding a Particle Swarm Optimization (PSO) algorithm, achieving improved mean Intersection over Union (meanIoU) on epilepsy and tumor MRI datasets.
- **AstraZeneca Tumor Detection Project** (York University, 2024 - Present): Developed tumor detection methods using advanced segmentation models like U-Net++, DeepLabV3+, and FPN on MRI datasets.
- **Bandgap Prediction for Crystals** (York University, 2024 - Present): Implemented a bi-directional GAN model to predict bandgap properties in crystals, advancing machine learning in materials science.
- **Bandgap Prediction Mobile App** (York University, 2024 - Present): Developing a cross-platform mobile application for **iOS and Android** that predicts the bandgap of materials using machine learning algorithms. The app enables users to input material parameters and receive bandgap estimations, assisting researchers and engineers in material selection for semiconductor and photovoltaic applications. Designed with a user-friendly interface and real-time computations, the project enhances accessibility to bandgap predictions without requiring complex simulations.
- **Wearable Heart Attack Detection Device** (York University, 2024 - Present): Designed a device that converts PCG signals to ECG data using machine learning models to predict heart attacks. Involved in hardware and embedded systems design.
- **Real-Time Obstacle Detection and Collision Avoidance System** (York University, 2024 - Present): Integrated LiDAR and camera data using computer vision for real-time obstacle detection.
- **Real-Time Pedestrian Detection with YOLOv7** (June 2024 – Present): Designed and implemented a real-time pedestrian detection system using YOLOv7, achieving high accuracy and low latency. Optimized the model for deployment on edge devices for efficient processing and quick detection in dynamic environments.
- **Telegram Bot for CafeAvaz** (June 2023 – Present): Built an AI-powered Telegram bot that handles user queries and sends corresponding videos from a YouTube playlist. Applied machine learning techniques to analyze users' singing performance.
- **Unix-like OS from Scratch for Real-Time Devices** (March 2022 – Present): Developed a lightweight Unix-like operating system optimized for real-time devices with minimal latency. Implemented core OS components, including process scheduling, memory management, and I/O operations for time-sensitive applications.

- **Personal Budget Tracker** (*April 2020 – July 2020*): Created a Python web application for personal finance management, allowing users to track income and expenses. Included interactive user input and file I/O for data storage.
- **Educational Game for Grade School Students** (*December 2018 – May 2020*): Developed an educational game using DFS, BFS, and Java/Swing, earning accolades for its educational value.

---

## RESEARCH INTERESTS

- Image Processing
- Transformers
- Diffusion models
- LLMs
- Computer Vision
- Machine Learning / Deep Learning / Neural Networks
- Wearable Devices / Biotechnology
- Optimization Algorithms
- Big Data
- AI

---

## SKILLS

- **Programming Languages:** Python, C/C++, Java, JavaScript, PHP, MATLAB, Assembly, R
- **Machine Learning & AI:** Deep Learning, CNNs, Transformers, Diffusion Models, NLP, Retrieval-Augmented Generation (RAG), Image Segmentation, Object Detection, Generative Models, Speech Processing, Causal Inference
- **Frameworks & Tools:** TensorFlow, PyTorch, Keras, OpenCV, LangChain, LlamaIndex, Docker, Kubernetes, AWS, Git, Django, Flask, Spring, Laravel
- **Data Systems & Analytics:** MySQL, PostgreSQL, MongoDB, Data Preprocessing, Performance Engineering, Benchmarking & Optimization
- **Hardware & Embedded Systems:** NVIDIA Jetson, Wearable Devices, LiDAR Systems, Digital & Analog Systems, Signal Processing
- **Software Engineering & Cloud:** System Architecture, Unix/Linux, Algorithm Development, DevOps, Containerized Deployments, Cloud Computing, Mobile Development (Android & iOS)
- **Collaboration & Leadership:** Cross-Functional Teamwork, Research Collaboration, Project Leadership



---

## HONORS & AWARDS

- Council of International Schools Scholarship, University of Saskatchewan, valued at \$20,000 (2020)
- Entrance Scholarship, University of Saskatchewan, valued at \$18,000 (2020)
- First place in University of Tabriz National AI competition among high school students (2018)
- Graduated high school with summa cum laude (2020)
- Head Principle's Scholarship for Top Students, European School Georgia, valued at \$10,000 (2019)
- CEWIL iHub Research Award, York University, valued at \$1800 (2025)
- Google Cloud for Startups Award, valued at \$25,000 (2024)
- Lasonde Undergraduate Research Award, valued at \$10,000 (2025) (Not Accepted)