John Zhang

j444zhan@uwaterloo.ca portfolio github.com/jonz9 linkedin.com/in/john-zhang-2665b9236/

TECHNICAL SKILLS

Languages: Python, TypeScript, JavaScript, C++, C, Java, SQL
ML & Data: PyTorch, Tensorflow, NumPy, Jax, Pandas, Keras, spaCy, ONNX, LangChain, Transformers, Sklearn
Frameworks: Next, React, Streamlit, Flask, Django, gRPC, GraphQL, Node, Express, Tailwind, CSS
Tools & Databases: Docker, AWS, GCP, MongoDB, Firebase, SQLite, PostgreSQL, JWT, Git

EXPERIENCE

Machine Learning Research Assistant Jan 2025 – Apr 2025
Alberta Machine Intelligence Institute (Amii), supervised by Prof. Grzegorz Kondrak & Ning Shi (Ph.D.) *Edmonton, Alberta*

- Integrated Transformer architectures, spaCy NER and RAG modules into multilingual, entity-aware machine translation pipeline, resulting in **10% boost in benchmark scores** on SemEval validation sets.
- Achieved **1st place** / **27 research teams** in **SemEval 2025 EA-MT** for Arabic translation and **4th overall**.
- Developed PyTorch-based inference pipeline to extract next-word probabilities from token-level logits in byte-level LLMs.
- *Coauthor, paper accepted at **SemEval 2025** — contributed to methods design, experiments evaluation, and literature analysis.*
- *Coauthor, paper under review at **EMNLP 2025** — contributed to scripting inference logic, case studies, and related work analysis.*

Autonomy Software Engineer Sept 2024 – Apr 2025
Waterloo Aerial Robotics Group (WARG) *Waterloo, ON*

- Retuned CNN-based object detection model (YOLOv5) for UAV imagery and landmark detection, **achieving 90% mAP**.
- Ported telemetry from ground station to RPi over MAVLink FTP and TCP protocol, **improving telemetry transfer rate by 43%**.
- Designed unit and integration tests in PyTest to validate mission accuracy, reducing navigation and landing errors.

Software Engineer Jan 2025 – Mar 2025
H-O-M-E AI *Calgary, Alberta*

- Led a two-dev sprint to build and deploy a **voice-ordering MVP in 7 days**, containerized and tuned Llama on EC2 G5.
- Developed and shipped dynamic floor-plan PDF generation/export pipeline in React and Tailwind front-end with Node backend, converting live design data into downloadable PDFs under 5s.
- Refactored **40% of the project monorepo** and decreased initial page load latency through microservice architecture.
- Performed iterative A/B testing cycles to refine functionality, optimize performance, and ensure stable version rollout.

PROJECTS

SONA | *Flask, FastMCP, PyTorch, SocketIO, Whisper, Kokoro, Gemini, Picovoice, Next* *Jun 2025*

- An acronym for **Semantic Operational & Natural Assistant**, a real-time assistant with wake word of ”Hey SONA”, developed for desktop productivity hands free.
- Transcribed with Whipser, thinks via Gemini, and replies with kokoro TTS model, acts like an intelligent personalized assistant.
- Designed and wired our own **Model Context Protocol** for Gemini to control native tools and be integrated end-to-end.

Gesture Recognition System | *Python, TensorFlow, OpenCV, NumPy, Keras, MediaPipe* *Apr 2025*

- Trained **2 CNN models** in TensorFlow to recognize spatial gestures and motion sequences for real-time input classification.
- Used OpenCV + MediaPipe to build a low-latency gesture control system with **80% accuracy** on custom hand landmark data.
- Captured training data via CSV logging and evaluated performance with confusion matrices and live system tests.

State Machine Developer Assistant | *Flask, Gemini, Monaco Editor, React Flow, Next.js, TypeScript, CSS* *Mar 2025*

- Built a **node-based visual IDE** for state machine design with React Flow, generating exportable OOP logic structures.
- Connected Gemini LLM to a Flask backend to convert Abstract Syntax Trees into Python class hierarchies and dynamic code.
- Added syntax validation and live feedback using Monaco Editor to streamline UX

AI Image Summarizer | *OpenCV, PyTorch, LangChain, Python, Next.js, TypeScript* *Feb 2025*

- Built a real-time image captioning system that turns visual content into natural-language descriptions using LangChain, PyTorch and Flask.
- Used OpenCV for image preprocessing and connected to a responsive frontend built with Next and TypeScript.

EDUCATION

University of Waterloo Waterloo, ON
B.A.Sc. in Computer Engineering | GPA: 3.8/4.0 | President's Scholarship of Distinction ($2000) *Expected Apr 2028*