

Zixin (Stan) Wan

781-547-9050 | zixinwan04@gmail.com | [linkedin.com/in/zixinwan](https://www.linkedin.com/in/zixinwan) | stanzixinwan.github.io

EDUCATION

Columbia University <i>Master of Science in Computer Science</i>	Aug. 2026 – Dec. 2027 (Expected) New York, NY
Brandeis University <i>Bachelor of Science in Computer Science and Biology</i>	Aug. 2022 – May 2026 Waltham, MA
<ul style="list-style-type: none">• GPA: 3.76/4.00 <i>Magna cum laude, Dean's List</i>• Relevant Coursework: Machine Learning, Natural Language Processing, Linear Algebra, Statistics, Data Structures & Algorithms, Computer Systems, Operating Systems, Theory of Computation.	

EXPERIENCE

Visiting Student Researcher <i>Shanghai Jiao Tong University</i>	June 2026 – Sept. 2026 Shanghai, China
<ul style="list-style-type: none">• Engineered a high-throughput multimodal data pipeline synchronizing dual-arm joint states and 3-channel RGB camera feeds at 30 FPS, reducing I/O latency by 35% across 100+ manipulation episodes.• Scaled SFT and advantage-conditioned policy training for an InternVL-1B Vision-Language-Action model.• Improved cross-task policy success rates by 18%, utilizing PyTorch FSDP on an 8×A100 GPU cluster.• Built a low-latency human-in-the-loop teleoperation and intervention framework for closed-loop evaluation.	
Machine Learning Research Assistant <i>Brandeis University (Van Hooser Lab)</i>	Sept. 2024 – May 2026 Waltham, MA
<ul style="list-style-type: none">• Authored and defended an Honors Thesis on functional decoding and neural tuning dynamics in visual cortex.• Architected an automated scientific compute pipeline in Python/MATLAB for 2-photon calcium imaging data, transforming terabytes of raw fluorescence streams into memory-mapped binary representations.• Developed 50+ modular feature extraction scripts, cutting data processing time by 40%.• Maintained CI/CD workflows for open-source scientific tools (<code>vhlab-toolbox</code> and <code>NDI-matlab</code>).	
Software Engineer Intern <i>Noah AI</i>	May 2024 – Aug. 2024 Remote
<ul style="list-style-type: none">• Designed and deployed an enterprise hybrid RAG pipeline combining dense vector embeddings (FAISS) with BM25 sparse keyword search over medical literature, improving retrieval and reducing hallucinations by 15%.• Implemented Cross-Encoder re-ranking and structured JSON Schema validation layers, enforcing output format constraints and robust exception handling across high-concurrency LLM agent workflows.• Developed backend microservices using FastAPI and Next.js modules, optimizing API response times.	

PROJECTS

Cross-Domain Text-to-SQL System <i>PyTorch, Transformers, LoRA/QLoRA, FAISS, Docker, FastAPI</i>
<ul style="list-style-type: none">• Fine-tuned encoder-decoder models via QLoRA parameter-efficient tuning on the Spider benchmark, achieving an 82% execution accuracy across unseen, complex multi-table relational schemas.• Constructed a schema-aware RAG pipeline utilizing FAISS vector indexing.• Containerized the inference service with Docker and built asynchronous FastAPI endpoints.
LLM-Powered Nutrition Tracking Platform <i>Python, TypeScript, PostgreSQL, Prisma, Vitest</i>
<ul style="list-style-type: none">• Engineered an AI agent orchestration system that parses unstructured text and dietary inputs into validated macronutrient data structures using structured output parsers and strict JSON schemas.• Designed a high-concurrency PostgreSQL relational database schema managed by Prisma ORM.• Built automated CI/CD pipelines using Vitest and GitHub Actions, ensuring 90%+ code coverage.

TECHNICAL SKILLS

Programming	Python, Java, C/C++/C#, JavaScript, SQL, Go, HTML/CSS
Frameworks	React, Next.js, FastAPI, Node.js, Prisma ORM
Dev & Infra	Git/Github Actions, Linux, CI/CD, Vitest, PostgreSQL, Docker
Libraries	PyTorch, FAISS, Scikit-learn, Pandas, NumPy