An Cao

+1-4162549939 an.cao.link@gmail.com ancaoai CA2528357431 An Cao MyWebsite

EDUCATION

University of Toronto **Sep 2024 – Dec 2025**
Master of Science in Applied Computing (AI concentration) - GPA: 3.93/4.0 Toronto, Canada
- Recipient of Vector Scholarship in Artificial Intelligence for top 100 AI master's students in Ontario

Huazhong University of Science and Technology **Sep 2020 – Jun 2024**
Bachelor of Engineering in Software Engineering (AI track) - GPA: 3.96/4.0 Wuhan, China
- Excellent Undergraduate Graduates, Merit Students, Outstanding Undergraduates in Academic Performance

TECHNICAL SKILLS

- **Language & ML Frameworks**: Python, Pytorch, HuggingFace, DeepSpeed, Langchain, OpenAI, Vertex AI, MLflow
- **ML Techniques**: Multi-Modal Learning, Vision-Language Models, Audio-Language Models, LLM Finetuning, Segmentation, Explainable AI, Data Augmentation, Computer Vision, RAG, Embedding Models, Vector Database
- **Models**: LLaMa, LLaVa, Qwen, Transformer, BERT, LoRA, SAM, Unet, CLIP, Diffusion, ViT, Audio Flamingo
- **Tools, Infra & Deployment**: GCP, Azure, AWS, SQL, MongoDB, PostgreSQL, Docker, FastAPI, Flask

EXPERIENCE

Modiface **May 2025 – Present**
Machine Learning Intern Toronto, Canada
Project: *Digital Dermatologist: Foundational Explainable Vision-Language Model for Skin Health*
- Incorporated SAM to LLaMa to deliver segmentation masks with text insights with to enhance explainability
- Performed alignment on VLM and Segmentation model, enabling the model to deliver medical text and visual proof
- Applied LoRA to finetune the model for medical tasks, reaching 52% segmentation IoU and 0.15 text Cross Entropy
- Leveraged Semi-Supervised learning to finetune the model on partially annotated and modal-incomplete datasets
- Employed LLMs and Unet for data augmentation, expending dataset by 3 times and imputing 2 absent modals
- Utilized contrastive learning to finetune an embedding model, enabling product recommendations from skin analysis

Vector Institute **July 2025 – Present**
AI Technical Specialist Toronto, Canada
- Integrated Qwen multi-modal LLM into audio-text RAG system for second-level voice-based grounded health insights
- Led a RAG system optimized for live tabular stock market data, achieving 85% recall in relevant stock retrieval
- Designed a multi-agent system to detect inconsistent datapoints, attaining 76% accuracy in Anti-Money Laundering

Vector Institute **Sep 2024 – May 2025**
Machine Learning Associate Toronto, Canada
Project: *DiligenceGPT: AI for Due Diligence*
- Applied multi-modal LLMs to fuse images and texts, structuring data from uncurated documents with 89% coverage
- Built a RAG agent with a vector database and minute-level live data sources, providing real-time traceable insights
- Combined Neural Networks and LLMs to quantitatively evaluate companies with analysis, reaching 97% consistency
- Orchestrated async model deployment to parallelize inference, reducing response latency by 70% on average

Project: *Audience Builder: Conversational Database Agent for Synthetic Society*
- Created a RAG agent that analyzes database schema to suggest attributes in dialogues, speeding up customer defining
- Finetuned the embedding model using contrastive learning, improving retrieval recall to 89% in the vector database
- Utilized LLMs to decouple inputs into atomic queries, achieving 82% question-wise attribute recommendation coverage
- Implemented a async streaming backend with FastAPI to maintain low-latency responses under concurrent user loads

Huazhong University of Science and Technology **Sep 2021 – Aug 2024**
Deep Learning Research Assistant Wuhan, China
- Built Diff-STAR, a Student-Teacher model combining Diffusion and ViT, achieving SOTA in image harmonization
- Proposed LisaCLIP, a zero-shot text-driven model, enhancing image manipulation precision via adaptive optimization
- Collaborated on Virtual Try-On and Street Semantic Segmentation projects, refining models and doing ablation studies

PUBLICATIONS

Diff-STAR **(First Author, Published by IMAVIS, SCI Q1)** **Sep 2023 – Aug 2024**
A. Cao and G. Shen, "Diff-STAR: Exploring student-teacher adaptive reconstruction through diffusion-based generation for image harmonization," Image Vis. Comput., vol. 151, p. 105254, Nov. 2024, doi: 10.1016/j.imavis.2024.105254.

LisaCLIP **(First Author, Accepted by IJCNN as ORAL)** **Sep 2022 – Feb 2023**
A. Cao, Y. Zhou, and G. Shen, "LisaCLIP: Locally Incremental Semantics Adaptation towards Zero-shot Text-driven Image Synthesis," in 2023 International Joint Conference on Neural Networks (IJCNN), Jun. 2023, pp. 1–10. doi: 10.1109/IJCNN54540.2023.10191516.