An Cao

📞 +1-4162549939 ✉ an.cao.link@gmail.com 🔗 ancaoai ⌨ CA2528357431 G An Cao 🌐 MyWebsite

EDUCATION

University of Toronto   **Sep 2024 – Jan 2026**
Master of Science in Applied Computing (AI concentration) - GPA: 3.93/4.0   Toronto, ON, Canada
- Recipient of Vector Scholarship in Artificial Intelligence for top 100 AI master's students in Canada

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

TECHNICAL SKILLS

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

EXPERIENCE

Modiface (L'Oréal AI)   **Toronto, ON, Canada**
Machine Learning Intern   May 2025 – Dec 2025
Project: *Digital Dermatologist: Foundational Explainable Vision-Language Model for Skin Health*
- Employed LLMs for data augmentation, expanding dataset by 15 times and imputing absent text modality.
- Applied semi-supervised learning to synthesize missing segmentation masks in partially annotated skin datasets.
- Integrated SAM with LLaMa for multimodal responses of medical text insights and skin concern segmentation masks.
- Aligned LLaMA embeddings and SAM to ground text insights in segmentation masks for visual explainability.
- Boosted diagnostic accuracy and coverage by 31% over SOTA via parallel quantized LoRA finetuning on 4 A100s.
- Achieved a 23% increase in IoU for skin concern segmentation, outperforming existing specialized models.
- Utilized contrastive learning to finetune an embedding model, enabling chat-based product recommendations.

Vector Institute   **Toronto, ON, Canada**
AI Technical Specialist   June 2025 – Present
- Integrated Qwen multi-modal LLM into audio-text RAG system for real-time voice-based grounded health insights.
- Led a RAG system optimized for live tabular stock data, matching human performance with 85% retrieval recall.
- Designed a multi-agent system for Anti-Money Laundering, attaining 76% info inconsistency detection accuracy.

Vector Institute   **Toronto, ON, Canada**
Machine Learning Associate   Sep 2024 – May 2025
Project: *DiligenceGPT: AI for Due Diligence*
- Applied LLMs to extract finance data from unstructured documents and real-time online info with 89% coverage.
- Built a RAG agent on vector database of minute-level live data, delivering business analysis with traceable datapoints.
- Designed a LLM+ML quantitative company evaluator with 97% consistency, exceeding human expert performance.
- 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 conversational RAG agent that suggests relevant database values, accelerating complex query formulation.
- Leveraged LLMs to decompose complex queries into atomic sub-queries, ensuring retrieval covers all specified objects.
- Solved semantic retrieval bottlenecks by fine-tuning the embedding model, achieving a 49% improvement in recall.
- Implemented an async streaming backend with FastAPI to maintain low-latency responses under concurrent user loads.

Huazhong University of Science and Technology   **Wuhan, Hubei, China**
Deep Learning Research Assistant   Sep 2021 – Aug 2024
- Built Diff-STAR, a Student-Teacher model combining Diffusion and ViT, achieving SOTA in image harmonization.
- Proposed LisaCLIP, a zero-shot text-driven adaptive model, enabling image manipulation using text without training.
- 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**)
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**)
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.