

Yingyi Luo

Chicago | yingyiluoneu@gmail.com | 773-456-4688

EDUCATION

University of Illinois Chicago

PhD candidate

Master of Science in Electrical and Computer Engineering

Chicago, IL

01/2024 - Now

08/2021 - 12/2022

Northeastern University

Bachelor's Degree in Engineering

Shenyang, China

08/2017 - 07/2021

- Awards: Merit-Based University Scholarship (Top 15%; Awarded in 06/2021, 09/2020, and 09/2019)

RESEARCH & TECHNICAL PROJECTS

TradingAgents: Multi-Agent LLM Platform for Financial Analysis, *AI Engineer / Researcher*

from 07/2025

- Designed and implemented a robust **multi-agent framework** to decompose complex trading workflows into 8 specialized AI roles (e.g., Market Analyst, Manager Agent), enabling modular, scalable decision-making. Used StateGraph and ToolNode abstractions to orchestrate agent interactions; implemented asynchronous agent communication for real-time strategy updates.
- Built and published a **fully local LLM-powered system** using Qwen, LLaMA3, and other Ollama-compatible models, ensuring agent functionality without reliance on cloud services. <https://github.com/LuoYingyi/TradingAgents/releases/tag/ollama>
- Diagnosed and resolved multi-model orchestration issues, including API incompatibilities, function/tool constraints, and model-specific quirks (OpenAI, Ollama, Claude, Gemini), enabling seamless cross-agent collaboration.
- Engineered a custom **in-memory embedding and retrieval system**. Achieved **end-to-end data privacy compliance** by eliminating cloud dependencies, API key usage, and external embedding calls, making the system secure for financial institutions.

FASTML and Signal Processing Research Group, *University Of Illinois Chicago*

from 09/2023

- Developed an AIN sputtering **model prediction** using DFT and multiple dense layers within a machine learning framework, with findings submitted for presentation at the upcoming IEEE VTS conference.
- Utilized autoencoders and Discrete Cosine Transform to compress data and reached an accuracy of 82% in predicting orientation.

Research Assistant, *Chinese Academy of Sciences*

08/2020 - 09/2020

- **Lead an Image Machine Learning project:** Performance metrics (Normalized Difference Vegetation Index) based on 80+ qualified regions of interest selected by a multi-scale **Convolutional Neural Network** with an accuracy of 74.2% using **Python**.
- Built a filtering algorithm in Python to select 50 qualified images from millions of satellite images, with an accuracy of 87.7%.

WORK EXPERIENCE

Machine Learning Researcher, *Fermilab*

02/2024– 03/2025

Batavia, Illinois

- Participated in the **Fast ML** project, specifically focused on Fast Machine Learning (hls4ml) in **neuroscience**.
- **Machine Learning Model development and Optimization for FPGA:** Designed and refined neural network models using Discrete Cosine Transform (DCT), optimizing the model structure to enable efficient deployment on micro-hardware devices.
- Collaborative **Neural Signal Research:** Led the development of data reconstruction models, contributing to the Fermi workshop.

Software Engineer in Neuroscience Company, *NeuroLux Inc.*

02/2023– 08/2023

Northfield, Illinois

- Implemented **recording functionality** for multiple devices, enabling simultaneous data capture and individual device control while optimizing storage and retrieval of plot settings using **C#**.
- Redesign and layout of the **front-end pages** using **XAML**, resulting in a remarkable **30%** increase in customer.
- Created an **SQL** database for devices and stored device-related information to enable **CRUD** operations and data analysis.
- Actively participated in the software packaging process, ensuring the proper handling and accessibility of local **JSON files**.

PUBLICATIONS

- Yingyi Luo, Talha M. Khan, Emadeldeen Hamdan, Xin Zhu, Hongyi Pan, Didem Ozevin, A. Enis Cetin. "AIN Sputtering Parameter Estimation Using a Multichannel Parallel DCT Neural Network." **IEEE 42nd VLSI Test Symposium (VTS)**, 2024.
- Emadeldeen Hamdan*, Yingyi Luo* (co-first authors), et al. "Real-time Instantaneous Phase Estimation Using a Deep Dual-Branch Complex Neural Network." **IEEE Transactions on Biomedical Engineering**, under review, 2025.

SKILLS

AI Agents & LLMs: LangChain, LangGraph, AutoGen-style Frameworks, Ollama, OpenAI API, Function Calling

ML & Deep Learning: Python including PyTorch, scikit-learn, Transformers, DCT-based Neural Networks

Others: JIRA, Professional in Sizchuan cuisine, Pokémon Trainer