Kimi Chen

hello@kimichen.net | github.com/kimichenn | linkedin.com/in/kimichenn | kimichen.net

EDUCATION

University of California, San Diego (GPA: 3.96) **Expected Graduation: June 2028**
B.S. in Data Science and B.S. in Applied Mathematics, Minor in Cognitive Science *San Diego, CA*

EXPERIENCE

Research Assistant **March 2025 – Present**
Professor Zhiting Hu's Research Lab *San Diego, CA*
- Engineered an end-to-end World Model pipeline using SOTA VLMs to simulate 2D/3D gameplay; **benchmarked agentic behavior** to assess consistency, controllability, and interaction fidelity.
- Orchestrated multi-node Slurm jobs on H100 and H200 clusters, using Python multiprocessing to cut data generation time by ~**90%** (3M frame-action pairs in 12h).

Computer Vision Lead **September 2024 – Present**
Triton Unmanned Aerial Systems *San Diego, CA*
- Refactored the codebase (~100 files) into modular stages, reducing new model integration time by **94% (2 days to <3 hours)** and enabling rapid field testing.
- Developed a synthetic dataset generation pipeline to benchmark CV models and finetuned YOLO v11/OWLv2 to ONNX, achieving **realtime detection** on NVIDIA Jetson Nano during flights.
- Led the development and deployment of an OpenCV based orthomosaic stitcher for high resolution panoramic maps.

Data Science Intern **May 2023 – August 2023**
Lanner Inc. *Taipei, Taiwan*
- Performed data analysis and designed gradient-boosted and time-series models that flagged procurement cost anomalies with **96% precision** and generated optimal reorder dates, accelerating supply-chain decisions.
- Introduced a Robotic Process Automation (RPA) pipeline for data organization tasks through Power Automate, reducing processing times by **70%**.

Web Developer Intern **April 2023 – November 2023**
Silverline Educational Advisory Services *Boston, MA (Remote)*
- Integrated Strapi CMS and built guardrails for content rendering; cut manual update time ~**90%** and reduced production errors.
- Improved performance/SEO by **30 Lighthouse points** via asset optimization; reduced CLS and TBT.
- Led a CSS refactor into modular, maintainable patterns and ensured responsive design across breakpoints.

PROJECTS

File Order Randomizer – Adobe Premiere Pro Extension [[Link](#)]
- Identified a critical UX gap in creative workflows and engineered a native solution using Adobe CEP and ExtendScript, implementing recursive algorithms to flatten nested bin structures.
- Solved complex timeline logic by calculating frame-precise "tick" parameters to ensure non-overlapping clip insertion, scaling the tool to 8,**000+ downloads with a 4.9 rating** and validation from Adobe engineers.

Self-Fed Memory – Personal AI Assistant [[Link](#)]
- Built a retrieval-augmented generation (**RAG**) system that ingests Markdown notes, chunks and embeds them, and answers questions with user-specific context via OpenAI embeddings and Pinecone.
- Engineered a multi-query retrieval system with time-decay scoring and semantic routing to optimize context relevance and long-term memory personalization.

Disinformation Detection via Various Deep Learning Models [[Link](#)]
- Benchmarked LSTM with and without GloVe, as well as large language models (LLMs) such as DistilBERT, Llama 2, and GPT 4 Turbo in PyTorch, exposing a 40 percentage point gap between in domain (**99% accuracy**) and out of distribution (OOD) splits.
- Designed and ran a human benchmark with ~**100 participants**, then led a detailed error analysis that exposed a false-negative bias in all models and informed recommendations for more diverse, multimodal training data to boost real-world generalization.

TECHNICAL SKILLS

Languages & Scripts: Python, C++, SQL databases, Java, TypeScript, React
Machine Learning: PyTorch, TensorFlow, Scikit-Learn, Pandas, NumPy, OpenCV
Other Tools: Git, GitHub, Docker, Slurm, Linux (Ubuntu, Fedora), Power Automate, Matplotlib