Pengfei Zhang   Applying for 2026 Internships (Winter, Spring, Summer, Fall) with F1 CPT (Parttime/Fulltime)
Email : **pengfz5@uci.edu**   Github   LinkedIn   Personal Website

EDUCATION

Year	Degree	Institute	GPA
Fall 2022 -	PHD stu in Computer Science	University of California, Irvine	3.93/4.0
2017 - 2021	B.Eng in Computer Science	University of Science and Technology of China	3.45/4.3

TECHNICAL SKILLS

- Languages. Python, C++, Java, Web Frontend Languages (Vue Framework), Verilog.

- Frameworks. Pytorch, Kafka, Spring Boot, Vue (Ant Design), Langchain, Flask, MySQL, Redis.

- Research Fields: Multimodal Large Language Models (Speech LLM, VLM), LLMs on Low-Resource Settings, Personalized Conversational Agents, Human Video Understanding and Generation

PUBLICATIONS

I. **Multimodal LLMs for Human-Centric Video Understanding and Generation**

This project addresses limitations of multimodal LLMs in understanding human-centric videos with complex motions—a low-resource setting due to the difficulty of motion capture. To tackle this, we propose a semantic alignment method and a coordinated generation approach that produces high-quality motion sequences and human-centric videos to augment training data.

(c) **KinMo: Kinematic-aware Human Motion Understanding and Generation**   *ICCV 2025*
  - Designed an automatic VLM-based dataset collection pipeline that enhances the existing text-motion benchmark by incorporating proposed novel motion representation that decomposes into distinct body movements.
  - Introduced a hierarchical motion semantics approach that progressively fuses details into the global action-level semantics for modality alignment.

(b) **Contextual Gesture: Co-Speech Gesture Video Generation through Context-aware Gesture Representation**   *ACM MM 2025*
  - A framework that improves co-speech gesture video generation through chronological speech-gesture alignment that temporally connects two modalities, and a contextualized gesture tokenization that incorporate speech context into motion pattern representation through distillation

(a) **Handformer2T: A Lightweight Regression-based model for Interacting Hands Pose Estimation from a single RGB Image**   *WACV 2024*
  - Designed a lightweight but high performance model which proposed hand-level tokenization in the transformer based model for interacting hand pose estimation, where only one token was used for each hand.

II. **Speech LLMs on Speech Recognition and Reasoning on Low-Resource Settings**

This project explores limitations of Speech LLMs in understanding speech in low-resource settings, such as low-resource languages and domain-specific terminology in healthcare.

(b) The Hidden Failure Modes of Speech LLMs on Extremely Low-Resource Languages
  - Provide the first systemic analysis of the Speech LLM's mechanism on low-resource languages, highlighting differences from high-resource languages: 1) Audio Encoder Collapse on co-distribution between speech embeddings and text embeddings; and 2) Decoder Hallucination which generates phonetically plausible but linguistically invalid words.

(a) **MedSpeak: Knowledge Enhanced ASR Error Correction framework for Spoken Medical Question Answering**   *submitted to ICCASP 2026*
  - a novel knowledge-enhanced retrieval augmented generation framework through a medical knowledge graph (KG) capturing semantic relationships and phonetic similarities between medical terms

III. **Personalized LLM Agents on Low-Resource Settings**

This project aims to create a LLM-based conversational healthcare agent, tailored to individual patients. To implement this, personalized retrieval-augmented generation and mix-of-expert architectures are proposed.

(c) **Adaptive Constraint Relaxation in Personalized Nutrition Recommendations: An LLM-Driven Knowledge Graph Retrieval Approach** *AMIA 2025*
- Designed an LLM-Driven Knowledge Graph Retrieval System for Personalized Food Recommendation, with LLM-driven constraint analysis and structured relaxation strategies, to resolve "all-or-nothing" limitation for traditional constraint-based recommendation system.

(b) **FoodAgent: A Multi-modal Mixture of Experts Reasoning Agent for Divide-and-Conquer Food Nutrition Estimation** *BSN 2025*
- Designed a multi-modal Mixture-of-Experts (MoE) reasoning framework that improves nutrition estimation through a divide-and-conquer strategy, with extra feature on personalized meal preparation and recommendation based on provided information and requested conversations.

(a) **DEMENTIA-PLAN: An Agent-Based Framework for Multi-Knowledge Graph Retrieval Augmented Generation in Dementia Care.** *AAAI W 2024*
- Introduced an innovative retrieval-augmented generation framework that enhances conversations with mild-stage dementia patients by intelligently integrating multiple knowledge graphs with Large Language Models.

SOFTWARE AND PROGRAMMING PROJECTS

- **Distributed Chatroom with LLaMa-Powered Summarization** *Jan. 2024 - Mar. 2024*
  - Implemented A Multi-topic Web Chatroom which can provide backup on previous conversations and LLaMa Powered summarizations after each refresh
  - Developed the frontend using React, the backend using Spring Boot (Java) and Kafka, and the database using Redis.
  - Implemented Cache Feature on each summarization, which largely decreases latency and improves fault tolerance

- **OpenCHA - an Automatic Conversational Health Agent** *Oct. 2023 - Present* (Collaborative Lab Proj)
  - Implemented a Conversational Health Agent framework leveraging LLMs-based agents (ReAct) as problem solvers, which can address health tasks like stress estimation. Built the frontend using Vue and backend using flask.

Working Experience

- Applied Scientist Intern at Amazon Web Services in San Jose *Jun. 2025 - Sep. 2025*
  - Speech LLMs recognition, understanding, and reasoning in low-resource languages (Paper II.a, II.b).
- Research Science Intern at Flawless. AI. Inc in Los Angeles *Jun. 2024 - Sep. 2024*
  - Multimodal Alignment of LLMs for Human Motion Understanding and Text/Audio-driven Motion Generation/Editing (Paper I.a, I.b, I.c).
- Research Intern in the Chinese University of Hong Kong in Hong Kong *May. 2022 - Aug. 2022*

AWARDS

- Dean's award from UCI *2022-2023*
- National Encouragement Scholarship (top 20%) from USTC *2020*
- National Encouragement Scholarship (top 20%) from USTC *2018*