Pengfei Zhang Applying for Summer 2025 Internship (starting June 2025) with F1 OPT
Email : **pengfz5@uci.edu** Github LinkedIn Personal Website

EDUCATION

Year	Degree	Institute	GPA
Fall 2022 -	PHD student 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: Conversational Agent System, Multimodal Agents, Retrieval Augmented Generation, Speech Generation, Text/Speech Driven Motion Synthesis

PUBLICATIONS

I. **Multimodal Alignment of LLMs for Speech and Visual Understanding and Generation**

(c) **KinMo: Kinematic-aware Human Motion Understanding and Generation** *ICCV 2025*
- Designed an automatic LVM-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 joint-group level movements and interaction information 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 three innovative components: (1) a chronological speech-gesture alignment that temporally connects two modalities, (2) 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. **LLM Agent for Personalized Recommendation and Nutrition Recognition**

(d) **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.

(c) **Self-Reflection Visual Agents for Precise Nutrition Recognition and Personalized Meal Preparation** *Submitted to BSN 2025*
- Designed an self-reflection chatbot which can provide personalized meal preparation and recommendation based on provided information and requested conversations, with extra feature on providing precise nutrition recognition based on user input.

(b) **DEMENTIA-PLAN: An Agent-Based Framework for Multi-Knowledge Graph Retrieval Augmented Generation in Dementia Care.** *AAAI W*
- 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.

(a) **Knowledge-Infused LLM-Powered Conversational Health Agent: A Case Study for Diabetes Patients.** *EMBC 2024*

- Developed a nutrition assessment chatbot powered by a large language model (LLM), incorporating external nutrition knowledge sources, which can analyze user nutritional input and provide a risk assessment.

III. **LLM-enhanced Automatic Speech Recognition**

(a) **MedSpeak: Knowledge Enhanced ASR Error Correction framework for Spoken Medical Question Answering** *submitted to Interspeech2025*

- a novel knowledge-enhanced retrieval augmented generation framework through a medical knowledge graph (KG) capturing semantic relationships and phonetic similarities between medical terms

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 Projects)
  - 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

- 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*