

Research Statement

[Ziyang Zhou]

Research Interests

My research interests are primarily focused on the fields of computer science, especially in algorithms, artificial intelligence (AI), Audio, NLP, and the development and optimization of large language models (LLMs). I have conducted in-depth studies of various deep learning architectures, including **LLama**, **GPT**, **Qwen**, BERT, and Transformer, and applied model optimization techniques such as distillation in real-world projects. Additionally, I am deeply interested in **AI+** and **human-computer interaction (HCI)**, exploring how AI technologies can enhance user experiences and lives through intelligent system design.

I believe that with the increasing computational power and the rapid iteration of models, AI can be leveraged to solve increasingly complex computational tasks, such as information retrieval, speech generation and enhancement, multimodal content generation, and intelligent human-computer interactions. My goal is to contribute to the development of AI technologies and to drive the field toward a more intelligent and efficient future. I also aim to explore how AI can be applied to various domains, benefiting the greater society.

Research Background and Contributions

Currently, my research work focuses primarily on audio processing tasks, including **Acoustic Scene Classification** and **Accent Recognition**. During this process, I have developed a deep understanding of the feature extraction mechanisms in HuBERT, BEATs, CIAP, HTSAT and have combined them with contrastive learning methods to create an efficient audio classification pipeline. I am also learning some methods about prompt and agent in LLM for some classification tasks.

To further enhance my research skills, I have actively participated in multiple academic replication projects, gaining experience in the full process of reading papers, replicating code, and optimizing models. These projects have provided me with valuable hands-on experience in data processing and model tuning. I am proficient in using the **Linux system** and the **VS-Code** development environment and have hands-on experience with **PyTorch**, allowing me to efficiently develop and debug deep learning models. In terms of technical innovation, I have proposed a **loss function optimization strategy** in a multi-task learning project, which successfully balanced the performance of both **speech classification** and **speaker verification** tasks.

Future Research Directions

In my future research, I plan to expand into the following areas:

- **Optimization of Large Language Models and Multimodal Applications:** I aim to explore the potential of large language models in multimodal tasks involving both speech and text. For instance, I am interested in developing **NLP** that combine speech and text for real-time translation, emotion recognition, and educational assistance. I am also interested about the application with prompt and agent.
- **Cross-disciplinary Research in Audio Processing:** I have passion in the intersection of audio signal processing and other disciplines such as education and healthcare. Specifically, I want to develop an on-device AI assistant that integrates speech recognition, computer vision, and user behavior modeling to provide personalized, context-aware assistance.
- **Human-Computer Interaction and Intelligent System Design:** I intend to expand my research into the field of **human-computer interaction (HCI)**, exploring how intelligent system design can improve user experience and efficiency. This could involve designing systems for education, healthcare, or other fields where users can interact with AI-powered systems seamlessly.
- **AI+:** I hope to leverage current AI technologies to develop smart applications that address specific problems in areas such as education or healthcare. Furthermore, I hope to use AI to establish specialized networks for collaboration, empowering professionals and academics to connect and innovate more effectively.

I am eager to participate in these breakthrough research projects and contribute to the development of efficient and scalable solutions through active discussion, experimentation, and innovative design.

Personal Strengths

I believe that my greatest strength in research lies in the combination of **technical depth** and **innovation ability**. I have a strong background in computation and algorithms, with deep understanding and practical experience, particularly in **Transformer architectures** and **speech processing applications**. One of my key strengths lies in my ability to maintain a forward-thinking perspective and keen insight into emerging technologies. I am deeply attuned to the rapid advancements in AI and machine learning, and I actively explore how these innovations can be applied across different domains. This forward-looking mindset, combined with my continuous learning from cutting-edge research, allows me to identify opportunities for breakthroughs and contribute unique ideas that can drive meaningful progress in the field.

Moreover, my international educational and collaborative experiences provide me with a diverse perspective and allow me to work efficiently in global environments. In team projects, I excel at summarizing meeting content, identifying the core issues quickly, and proposing constructive solutions. This openness to ideas and my passion for continuous learning allow me to break new ground in my research and provide real value to the teams I work with.