

# Research Statement

Zhihao Li

School of Computer Science and Technology

Xidian University

## 1. My Research Interests

Over the past five months, my research interests have focused on *multi-modal learning* and *large language models (LLMs)*. With the rise of large language models, it is impossible to train models with billions of parameters directly on downstream datasets, but well-trained pre-trained models have the potential to perform downstream tasks. Given the diverse range of downstream tasks, their datasets are often collected from the internet, which can contain noise that affects performance. My research aims to explore a robust prompt learning mechanism to ensure that pre-trained models, after fine-tuning, can adapt to downstream noisy datasets. I am very interested in this research because it has the potential to greatly promote the widespread application of large models to real-world tasks in an efficient manner.

## 2. My Research Preliminaries

Through my undergraduate studies, I have gained enough capability to pursue research work in this domain, specifically manifested in the following aspects:

### I. Mathematics Level

I have a deep passion for mathematics. I've been studying math for nearly 12 years from elementary school through high school, consistently achieving excellent grades. I scored 145 out of 150 on math subject of Nationwide Unified Examination for Admissions to General Universities and Colleges. In university, I've also excelled in my foundational math courses, scoring above 95 in all of them, including 95 in Advanced Mathematics (I), 98 in Advanced Mathematics (II), 97 in Linear Algebra, 95 in Probability and Statistics, and 97 in Discrete Mathematics. I've also received *the Grand Prize* of Shaanxi Province Higher Mathematics Competition for College Students and *the Provincial First Prize* in National Mathematics Competition for University Students.

### II. Programming Ability

I place great emphasis on improving my programming skills. I am proficient in Python, MATLAB, and C programming languages. I'm also familiar with C++, Shell scripting, and experienced in using frameworks like Pytorch and Daxl. Additionally, I'm well-versed in Git, GitHub, and Linux tools, actively building and expanding my technical stack.

In practice, I have always served as the programmer in mathematical modeling competitions, achieving outstanding results. In the 2023 National College Mathematical Modeling Competition, I wrote a genetic algorithm to solve a complex large-scale data problem involving product pricing, ultimately earning *the National Second Prize*. In the 2024 American College Mathematical Modeling Competition, I designed an improved Markov chain algorithm to predict football match outcomes, allowing for dynamically updated winning probabilities. This achievement led me to win *the International First Prize*.

### III. Research Projects

During my undergraduate studies, I completed two research projects, from competition practice to scientific training, which greatly enhanced my research skills and expanded my research foundation.

#### *III.1 2D Virtual Try-On based on Deep Learning*

**The problem** Virtual Try-on refers to trying on new clothing based on images of the person and new clothing. This project requires participants to design a technical solution process, as well as complete tasks such as virtual try-on model *training, optimization, and engineering deployment*.

**Method** We first optimized the top-tier model by pre-pruning it to address its large parameter count, determining the optimal sparsity level for the model. Subsequently, we conducted actual model pruning and retraining. To further compress and optimize the model, we applied quantization compression, focusing on post-training static quantization with perceptual training fine-tuning. Additionally, for critical computation modules within the model, we designed an Img2Col algorithm to replace time-consuming convolution operations with matrix multiplication, accelerating inference for those modules. Finally, we deployed and optimized the executable files using the OpenVINO inference framework, resulting in faster computation and smaller model size.

**Implementation results** Our optimized model achieved an FID score of 8.906, which is very close to the results of top-tier publications. After deployment and engineering, the model's size was reduced from 279MB to 34.95MB, with the FID score maintaining at a level of 11.503. The compression ratio reached 87.47%, and the speedup ratio was 10.896 times. This is an outstanding result, which supported us in achieving the prestigious honor of winning the National First Prize.

#### *III.2 Interactive Learning Mechanism for Visual Language Models*

**The problem** Research on how to enable pre-trained multi-modal models to robustly transfer to downstream noisy datasets through prompt learning, facilitating adaptation for downstream tasks.

**Method** We propose an interactive learning mechanism for VLMs by constructing two sets of prompts that guide each other. A Gaussian Mixture Model (GMM) is used to identify reliable labels and noisy labels. Reliable labels undergo smooth augmentation, while noisy labels undergo ensemble augmentation. Finally, both sets of prompts are fine-tuned simultaneously using soft labels.

**Implementation results** We observe that our proposed mutual learning mechanism enables vision-language models to achieve higher accuracy across various noise scenarios. This enhanced mutual learning paradigm ensures more efficient model adaptation to downstream tasks.

### IV. Research Habits

*What is scientific research? To me, scientific research is about exploring new knowledge and venturing into uncharted territories.*

**My research philosophy** What I pursue is to always uphold my research beliefs, knowing that they can be realized and can change the world. I believe that outstanding researchers must stay true to their theories, maintaining their focus for decades, eventually achieving their greatest research outcomes. After understanding the true essence of scientific research, I am fully prepared to immerse myself in it. The greatest

joy I find in scientific research is the thrill and excitement of discovering the unknown. I've also experienced frustration and disappointment from repeated trials and setbacks, but I always understand that discovering the truth is a rare event, and it is the mission of researchers to find that rare event and make it a reality.

**My project logging habit** When I work on a research project, I will always start a project log that accompanies the project throughout its lifecycle, documenting the daily challenges encountered, the exploration of solutions, and the acquisition of new knowledge. This not only helps to better manage the project's development progress but also aids in building problem-solving skills.

**My paper-reading habit** In an academic career, reading other people's papers often helps to understand their research progress, thought processes, discoveries, and challenges. To make the most of these insights, I conduct a thorough analysis while reading a paper and create a brief reading report. This approach significantly aids in quickly identifying my own research topics.

**My research innovation** In my field of expertise and research direction, I have accumulated numerous research theories and the necessary mathematical foundations. For every paper I read, I extract the core contributions and innovations, and I consider how the authors developed these ideas and how they could inspire my research. This process allows me to continuously build new research momentum and spark new innovation throughout my academic career.

## **V. Personal Characteristics**

My greatest characteristic is that I never give up and have intrinsic motivation. Throughout my journey, I've experienced some success but also faced significant failures. However, every time, I stood firm and made it through the toughest moments. My second major characteristic is optimism. When faced with adversity, I always adapt and make the best choices based on the current situation, maintaining a positive attitude. Another of my traits is persistence. Whether mentally or behaviorally, I have a strong willpower, which has been a critical factor in my journey thus far.

With these precious characteristics, I hold a firm belief that I will be able to successfully pursue my path in scientific research.

## **3. My Research Proposals**

According to my personal research interests, I will conduct in-depth research in the direction of multi-modal learning and explore its generation and robustness for downstream tasks. My research plan is divided into the following three stages:

**First stage of research** Reserve research work in the direction of multi-modal learning and transfer learning, explore the innovations and shortcomings of recent research work, and under the guidance of the mentor, do one thing to the extreme and step into my own research field.

**Second stage of research**, based on my own research preparations and proposed innovative ideas, do some research work on promoting the adaptability of the pre-trained multi-modal model to downstream tasks, and make the best of the feasible parts.

**Third stage of research**, start to put forward an idea independently, conduct experimental verification, analyze results, improve ideas, continuously improve and do my best, and gradually realize independent scientific research.