

# Shitian Zhao

Homepage:  
<https://zhaoshitian.github.io/>

East China Normal University, Shanghai, China  
☎ (+86) 13569820135, 18817598970  
✉ [zhaosh5t5an@gmail.com](mailto:zhaosh5t5an@gmail.com)

## Interests

Multi-modal Language Model, Large Language Model, Diffusion Model, Text Image Synthesis and Autoregressive Image Modeling

## Education

2020-2024 **Major in Electronic Engineering**, *East China Normal University(ECNU)*

## Experience and Background

July 2022 - **Numerical Inference in NLP**, *Internship in NJUNLP(Nanjing University)*

Aug. 2022 ○ Supervised by Jiangzhou Ju.

○ Study the numerical inference problem under the NER paradigm, implement a Bert-BiLSTM-CRF model.

July 2023 - **Multi-modal Large Language Model**, *Internship in CCVL(Johns Hopkins University)*

Nov. 2023

○ Supervised by Zhuowan Li, Yan Wang and Alan Yuille.

○ Design a prompt engineering framework, named Causal-CoG for multi-modal language models.

○ One paper accepted by CVPR2024(poster highlight top 2.8%).

Oct. 2022 - **Vision and Language Model**, *Internship in East China Normal University*

July. 2024 ○ Supervised by Yan Wang.

○ Prompt tuning for vision and language pretrained models.

○ Two papers submitted.

Dec. 2023 - **Generative Model**, *Internship in Shanghai AI Lab, Alpha-VLLM team*

Present ○ Supervised by Peng Gao.

○ Be engaged in SPHINX-X, LUMINA-T2X.

○ One paper accepted by ICML2024.

## Publication

**Causal-CoG: A Causal-Effect Look at Context Generation for Boosting Multi-modal Language Models (CVPR24, poster highlight top 2.8%)**

○ Shitian Zhao, Zhuowan Li, Yadong Lu, Alan Yuille, Yan Wang

**SPHINX-X: Scaling Data and Parameters for a Family of Multi-modal Large Language Models (ICML24)**

○ Peng Gao\*, Renrui Zhang\*, Chris Liu\*, Longtian Qiu\*, Siyuan Huang\*, Weifeng Lin\*, Shitian Zhao, Shijie Geng, Ziyi Lin, Peng Jin, Kaipeng Zhang, Wenqi Shao, Chao Xu, Conghui He, Junjun He, Hao Shao, Pan Lu, Hongsheng Li, Yu Qiao

### **Unleashing the Potentials of Likelihood Composition for Multi-modal Language Models**

- **Shitian Zhao**, Renrui Zhang, Dongzhi Jiang, Xu Luo, Yan Wang, Shanghang Zhang, Peng Gao (Submitted)

### **CLIP with Vector-Quantized Prompts for Continual Learning**

- **Shitian Zhao**, Yadong Lu, Xingran Xie, Qingli Li, Alan Yuille, Yan Wang (Submitted)

### **Prompting Universal Model for CT Image Segmentation from a Causal Perspective**

- Boxiang Yun, **Shitian Zhao**, Qingli Li, Alex Kot, Yan Wang (Submitted)

---

## Honours and Awards

- Outstanding Graduate Thesis
- Excellent student scholarship
- The third prize, province level, China Undergraduate Mathematical Contest in Modeling

---

## Technical Skills

Programming Languages C, C++, PYTHON, SHELL, LATEX

DeepLearning Framework PYTORCH, TRANSFORMERS, DIFFUSERS