# YANG Zhenhao

**Tel:** +86-13272651988

**Email:** uestc.yangzhenhao@gmail.com

**Address:** Uestc Shahe Campus, No.251, 1st Ring Rd East
1 Section, Chenghua District, Chengdu, China 610056

## EDUCATION BACKGROUND

**University of Electronic Science and Technology of China (UESTC)**          09/2021-06/2025

Bachelor of Engineering

- Major: Software Engineering ("Internet +" Cultivation Experimental Class)
- GPA: 4.0/4.0    |    Average Score: 93.8/100
- Ranking: 2/83

## INTERESTS & SKILLS

**Interests**:

- Multimodal Understanding (video and image understanding based on large vision-language model)
- Multimodal Generation (multimodal generation techniques covering image, video and 3D generation)

**Skills**:

- Programming Languages: C/C++, Python, Matlab
- Web Technologies: HTML5, CSS, JavaScript

## RESEARCH EXPERIENCE

**Towards Open-vocabulary HOI Detection with Calibrated Vision-language Models          09/2023-03/2024
and Locality-aware Queries**

*Project Leader*

- Developed a two-stage approach with a visual language pre-trained model (VLP) and locality-aware human object interaction decoder
- Observed the intrinsic noise in VLP when it is applied to the HOI domain and enhanced model performance via calibrating V&L teacher and integrating spatial priors in HOI detector
- **Outcome**: Completed a paper accepted by ACM MM 2024 (CCF-A) as the first author

**Scene Graph Generation Based on Language Modeling and Discrete Feature          04/2023-08/2023
Learning**

*Core Member*

- Utilized the CLIP model for encoding scene graph (SG) features, constructing a codebook for knowledge extraction and vector quantization inspired by VQ-VAE
- Inspired by Pix2Seq, used Transformer to encode and decode visual features into discrete tokens which were then reconstructed as embeddings for the classifier to output the SG
- **Outcome:** Developed a novel framework for Scene Graph Generation (SGG) and achieved competitive performance in zero-shot SGG scenarios

**Customized Car Image Generation Based on Diffusion Model**          03/2024-06/2024

*Project Leader*

- Constructed a customized car image generation model based on InstantID for tuning-free image generation
- Used BLIP2 Caption capability to create an Image-Caption dataset in the automotive domain. Combine CLIP Embedding, Canny image, and car key points as guiding conditions for model training
- **Outcome**: Developed a customized car image generation model and implemented it at ByteDance's 2C business, enabling functions like car replication, posture adjustments, and style modifications

**An Improved Multimodal Large Model for Multitasking Collaboration Based on          01/2024-03/2024
Vertical Application**

*Project Leader*

- Collected a multimodal VQA dataset and a Chinese dialogue dataset and constructed training and evaluation sets for the automotive domain
- Applied Kmeans to filter out highly similar images when constructing test sets and trained with the Yi-6B base model and LLaVA framework
- **Outcome:** Achieved accuracy surpassing that of similarly sized open-source models in both general domains (CMMMU) and the automotive domain

**A Security Detection System for Wire Fraud Web Pages Based on Large Language Models and Graph Neural Networks**     02/2023-06/2023

*Project Leader*
- Developed a webpage security detection system, amalgamating webpage content from crawlers with IP location results using graph neural network IP location technology
- Deployed large language models within the LangChain framework, including ChatGLM, to provide security assessment hints through the Prompt project's pre-training
- **Outcome:** Won the National Grand Prize in the 8th C4-Network Technology Challenge and declared one software copyright as the second author

**Virtual Simulation Video Data Understanding Technology for Intelligent Driving Vehicles**     11/2023-12/2023

*Project Leader*
- Integrated BLIP pre-trained model for graphical matching to predict weather, time of day, and other labels
- Expanded the dataset with custom driving data and self-trained a CNN to predict road conditions
- Utilized pre-trained YOLO v8 model to predict nearby traffic participants and their behavior
- **Outcome:** Achieved a rank of 28/1843 in the Alibaba Cloud Tianchi Big Data Competition

**Financial Information Collection and Analysis System for Listed Companies**     09/2022-01/2024

*Project Leader*
- Employed convolutional neural networks to extract features from time series financial data of listed companies
- Utilized transformer models for detailed feature extraction, facilitating stock price and trend prediction
- **Outcome:** Deployed the system on the server side and completed several software design proposals

## INTERNSHIP EXPERIENCE

**Beijing ByteDance Technology Company Limited**     01/2024-Present

*Graphic Image Algorithm Engineer Intern*
- Developed automotive multimodal lagre language models (MLLM) and established evaluation metrics for content understanding, improving the Chinese language comprehension capability of MLLM and optimizing them for automotive applications
- Participated in the development of the AIGC graphic model and AIGC video model in the automotive field, aiming at personalized content customization and application to eco-creation, and committed to building an automotive multimodal atomic capability library

## HONOURS & AWARDS
- The National Scholarships for the Academic Years 2022-2023 & 2021-2022
- The National Grand Prize of the 8th C4-Network Technology Challenge
- The National First Prize of the 16th Computer Design Competition of Chinese College Student
- The Outstanding Student Scholarship (Grade 1) for the Academic Years 2022-2023 & 2021-2022
- The 2023 Outstanding Undergraduate Award of Association for Computing Machinery Chengdu Chapter

## EXTRACURRICULAR ACTIVITIES
- Vice Secretary, Undergraduate Party Branch || , UESTC     09/2023-Present
- Center Manager of Competition Centre, School of Information and Software Engineering, UESTC     03/2022-02/2024
- Visiting Scholar, Nanyang Technological University (NTU), Singapore     08/2023
- Volunteer Teacher, Weiyang Volunteer Teacher Team, UESTC     07/2022

## OTHER INFORMATION
- **Language:** Mandarin (Native), English (CET-6: 580/710)
- **Hobbies:** Guitar, Cycling, Music