Yilin Zhao

Email: zhaoyilin424@163.com/ermuzzz2001@gmail.com
Linkedin: linkedin.com/in/ermu2001
Mobile: +1-732-333-7430/+86-186-1300-8800
Git Remotes: github.com/ermu2001 and gitee.com/ermu2001

EDUCATION

New York University   New York, United States
Master of Science in Computer Engineering with current GPA 4.0/4.0 (expected)   *Sep. 2023-May 2025*
  Courses: *Machine Learning, Probability and Stochastic Processes, Computing Systems Architecture*

Nankai University   Tianjin, China
Bachelor of Engineering in Computer Science with GPA 87.79/100   *Sep. 2019-June 2023*

PROFESSIONAL EXPERIENCES

Baidu Inc.–Intern Algorithm Research Developer (Full-time)   Shenzhen, China
Mentor: Xintong yu   *Dec 2022-Aug. 2023*

Research and develop in the fields of modern Generative Artificial Intelligence. Analyse problems related to Natural Language Processing, Computer Vision, Generative Machine Learning. Design solutions during cross modality AI-system constructing. Contributed projects are listed:

- **Ernie4/Baidu-Wenxin Yiyan(Mar.-Aug. 2023)**: **Research** in the fields of Large Language Models and Cross Modality Models. Develop and experiment on Multi-modal Large Language Model Baidu-Wenxin Yiyan.
  * **Cross Modality Data Construct**: Research, design and construct data for cross modality training. Contributed 4 million samples of data for Multimodel-Erniebot training originated from Web coarse.
  * **Cross Modality Model Experiment**: Research, design, train and evaluate the cross modal language models. Especially foucus Enabling the Multimodal Model with grounding ability. Contributed in curriculum evaluation of Large Language Models and Multimodal Models
- **Ernie-Vilg2/Baidu-Wenxin Yige(Dec. 2022-Feb. 2023)**: **Research** and develop on Ernie-Vilg2, a Chinese Diffusion image generative model. Improve the algorithm for better performance. The encountered problems and contributions are listed below:
  * **Diffusion Process Accelerate**: Research and experiment on acclertating methods for diffusion models.
  * **Extend Controllability Modules Design**: Design extra modules and training process for ernie-vilg2, achieving 20% improvement on performance with better control.
  * **Specialized Image Generative Model Finetune**: Finetune diffusion models with artistic images for better image generation. Achieve 23% performance gain under human evaluation.

ZMO.AI–Intern Algorithm Developer (Full-time)   Shenzhen, China
Mentor: Liqian Ma   *July 2022-Sep. 2022*

Design and Implement ImgCreator's algorithm pipeline to meet the urge for image generation/editing. Evaluate and deploy certain models to assist generative models for better overall performance. Main contribution are listed:
- **Evaluate SOTA Vision Models**: Research and evaluate the State of the Art Vision Models. Decide and deploy models a image generate/edit pipeline.
- **Pipeline design and implement**: Integrate state of art methods to assist a generative model for better image generation. Design the procedure of the integrated image generation/editing pipeline. Implement the pipeline for the algorithm service.

RESEARCH EXPERIENCES

Video Magic LLM: Universal Vision Understanding Language Model   Bytedance Inc.
Advisor: Daquan Zhou   *Dec. 2023-Present*

Construct a SOTA Image-Video Large Multimodal Language Models. Conduct Comprehensive Study of it's Understanding for Vision Modal Information.
- **Video-Language Model Training**: Implement Data and Training Pipeline for Video Large Language Model Training
- **Multi-Vision Modal Evaluation**: Conduct Comprehensive Evaluation for Large Multimodal Language Models

ChatAnything: Facetime Chat with LLM-Enhanced Personas   Open-sourced Project
Advisor: Daquan Zhou   *April 2023-Nov. 2023*

Utilize current foundation models to reinvent the pipeline for arbitrary personas talking head. Achieve using text prompt to run everything.
- **Face Control for initial frame**: Incorporate the diffusion control method to ensure the generated initial frame to fit in the talking head module.
- **Mixture of Diffusers/Voices**: Utilize Large Language Models for zero shot arbitrary module selection with user input. Ensure the consistency between different modal generation.
- **Evaluation Dataset**: Construct evaluation dataset for the specificial pipeline. Enable the automatic evaluation for the desiged methods.

Style Image generation with Text Prompt Diffusion Models   Media Computing Lab, Nankai University
Thesis Advisor: Qibin Hou   *Sep. 2022-June 2023*

Propose novel strategy for style image generation with diffusion models. Narrow the gap between the research of image style transfer and image generation. Main contributions are listed below:
- **Leverage Style Transfer for Diffusion Models**: Design latent style loss and optimization strategy for latent diffusion models. Overcome the difficulty of rare style generation and transferring with diffusion models.
- **Propose Pipeline for Style Image Generation and Style Transfer with Latent Diffusion Models**: Propose a novel latent style image generation and transfer pipeline. Risen the CLIP-score of generated style images by large compared to baselines.