



Zhengze Yu

 [zhengze-yu.github.io](https://github.com/zhengze-yu) |  zhengze347@outlook.com

EDUCATION

University of Chinese Academy of Sciences
B.E. in Artificial Intelligence

Beijing, China
Sep 2021 – Jun 2025

- **Advisors:** Prof. Changsheng Xu (Academic),  [Google Scholar](#).
Prof. Junyu Gao (Research),  [Google Scholar](#).
- **Relevant Coursework:** Fundamentals of Programming and Experiments in C, Data Structure and Algorithms, Pattern Recognition and Machine Learning, Principles of Artificial Intelligence, Discrete Mathematics, Cognitive Neuroscience, Knowledge Engineering, Text Data Mining, Game Theory.


***Admission to UCAS's undergraduate program requires a **top 0.1%** ranking in China's national college entrance examination (Gaokao), representing the most competitive and selective academic pathway in China. UCAS is affiliated with the Chinese Academy of Sciences (CAS), which ranks **No.1 globally in the Nature Index** and **No.5 worldwide in AI (CSRankings, AI all areas)**.*

RESEARCH EXPERIENCE

Multi-modal Video Editing with Attention Injection 
Institute of Automation, Chinese Academy of Sciences

Advised by Prof. Junyu Gao
Sep 2024 - Apr 2025

- Led an undergraduate thesis within a National Natural Science Foundation project, targeting controllable and multimodal video editing systems.
- Architected an attention-injection-based video editing framework integrating text, image, and video-level transformations for seamless multimodal editing.
- Constructed a comprehensive evaluation benchmark for video editing, introducing eight metrics across semantic consistency, temporal consistency, and generation quality to address the lack of standardized protocols.
- Curated a dataset derived from V2VBench and VideoP2P to ensure compatibility with high-fidelity video synthesis research.
- Optimized VideoP2P performance by leveraging large language models (LLMs) for intelligent prompt rewriting, achieving more contextually coherent edits and deeper semantic alignment between text and visual content.

Video Generation Models and Optimization Techniques 
Institute of Automation, Chinese Academy of Sciences

Advised by Prof. Junyu Gao
Mar 2024 - Jul 2024

- Conducted a systematic literature review on video generation, covering diffusion-based, autoregressive, GAN-based, and transformer-based paradigms, with an emphasis on long-video synthesis and temporal consistency.
- Analyzed core architectural designs in modern video diffusion models (e.g., 3D U-Net, latent diffusion, spatiotemporal attention) and compared divide-and-conquer versus temporal autoregressive generation strategies.
- Studied and summarized model optimization techniques for large generative models, including quantization, knowledge distillation, pruning, and low-rank approximation, with a focus on computational efficiency.
- Developed proficiency in technical academic writing, including mathematical formulation, figure preparation, structured comparison of models, and citation management for survey-style research.

PROJECTS

Cross-Domain Multilingual Text Clustering Analysis

University of Chinese Academy of Sciences

Advised by Prof. Jiajun Zhang

Mar 2024 - Sep 2024

- Applied pretrained BERT models to extract sentence-level semantic representations from multilingual corpora spanning five specialized domains (IT, Koran, Law, Medical, Subtitles).
- Conducted cross-domain and multilingual clustering experiments using K-means to analyze the semantic separability of contextual embeddings across English and German text data.
- Evaluated clustering performance with supervised accuracy, achieving an overall accuracy of 85.6%, and validated results through confusion matrix analysis.
- Performed visualization and error analysis to identify systematic misclassifications, revealing semantic overlap patterns between Medical and IT domains.

Multilingual Text Compression via Subword Tokenization

University of Chinese Academy of Sciences

Advised by Prof. Jiajun Zhang

Feb 2024 - Mar 2024

- Built a multilingual text compression pipeline using the training-monolingual-news-commentary corpus across five languages (cs, de, en, es, fr), ensuring cross-language comparability through aligned news samples.
- Implemented Byte Pair Encoding (BPE) from character-level initialization, iteratively constructing subword vocabularies via frequent pair merging.
- Trained BPE vocabularies with sizes ranging from 500 to 3000 and conducted controlled ablation studies to analyze vocabulary size effects.
- Defined and evaluated compression rate (encoded length / original byte size) as the core metric, reporting results with tables and curves.
- Analyzed cross-lingual compression trends, identifying consistent behaviors across languages and differences in compression efficiency among cs, es, and others.

ACADEMIC SERVICE & LEADERSHIP

19th “Light of Automation” Public Science Day

Institute of Automation, Chinese Academy of Sciences

May 2023

- **Organizer, Presenter:** Involved in all stages of the exhibition, from coordinating displays on humanoid robots, multimodal AI, and brain-inspired systems to presenting generative video synthesis technologies that exemplify advances in large-scale AI models.

RoboRAV Asia

Jiuhua Resort & Convention Center

Sep 2021

- **Judge:** Advised participants on optimizing robot design and programming; Judged robot combat matches, ensuring fair play.

SKILLS

Machine Learning & AI: Pytorch, Tensorflow, Pandas, SQL, R.

Programming Languages: C, C++, Python, Java.

Specialized Knowledge: Diffusion Models, Large Language Models, Multimodal Learning, Video Synthesis.

Languages: English (Full Professional Proficiency), Mandarin Chinese (Native).