Zhengze Yu

 zhengze-yu.github.io | ✉ yuzhengze21@mails.ucas.edu.cn

EDUCATION

University of Chinese Academy of Sciences   Beijing, China
B.E. in Artificial Intelligence   *Sep 2021 – Jun 2025*

- **Advisors:** Prof. Changsheng Xu (Academic), G Google Scholar.
  Prof. Junyu Gao (Research), G 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.

*****Admitted to UCAS through the National College Entrance Examination (Gaokao),
ranking within the top 1‰ *nationwide — the most competitive and selective pathway in China.*

RESEARCH EXPERIENCE

Multi-modal Video Editing with Attention Injection 🌐   Advised by Prof. Junyu Gao
Institute of Automation, Chinese Academy of Sciences   *Sep 2024 - Apr 2025*

- Architected a multi-stage cross cultural video transformation framework by integrating text, image, and video-level adaptations, enabling seamless multimodal content generation.
- Curated a specialized evaluation dataset derived from V2VBench and Video-P2P, tailored for high-quality video-to-video synthesis research.
- Conducted rigorous quantitative benchmarking on GLP and Video-P2P models, analyzing semantic/temporal consistency and generation quality to identify optimal editing boundary.
- Optimized VideoP2P performance by leveraging LLMs for intelligent prompt rewriting, significantly enhancing the contextual relevance of edited content.
- Investigated hyperparameter sensitivity in P2P editing, specifically mapping the impact of attention-map values on semantic consistency.

Video Generation Models and Optimization Techniques 🌐   Advised by Prof. Junyu Gao
Institute of Automation, Chinese Academy of Sciences   *Mar 2024 - Jul 2024*

- Conducted an in-depth survey of state-of-the-art video generation models like Sora, mastering the core mechanics of Diffusion Models (DDPM) and their evolution into 3D U-Net architectures for spatiotemporal consistency.
- Systematically analyzed long-video generation frameworks, comparing Divide-and-Conquer (keyframe interpolation) and Temporal Autoregressive strategies (sequential fragment generation) to address the limitations of finite computational resources.
- Model Compression & Optimization: Investigated efficiency enhancement techniques for Vision Transformers (ViT), including Uniform Quantization , Knowledge Distillation (soft/hard distillation) , Pruning based on importance scoring , and Low-Rank Approximation.
- Spatiotemporal Modeling Expertise: Explored advanced cross-modal integration methods using Cross-Attention mechanisms and latent space utilization (LVDM) to maintain high fidelity and temporal coherence in text-to-video (T2V) synthesis.

PROJECTS

Cross-Domain Multilingual Text Clustering Analysis 🌐   Advised by Prof. Jiajun Zhang
University of Chinese Academy of Sciences   *Mar 2024 - Sep 2024*

- Leveraged BERT pre-trained models to perform deep semantic encoding on 10,000 text samples across five specialized domains (IT, Koran, Law, Medical, and Subtitles) in both English and German.
- Developed and executed K-means and Hierarchical Clustering algorithms to categorize high-dimensional text data. Optimized the Within-Cluster Sum of Squares (WCSS) to achieve minimal intra cluster variance.
- Established a robust evaluation framework using Confusion Matrices and clustering coefficients to assess model accuracy. Successfully achieved high precision in IT and Subtitles categories, correctly identifying 1,896 IT-related samples.
- Conducted a granular diagnostic of misclassifications, identifying a specific semantic overlap where approximately 33% of medical texts were confused with IT terminology. Proposed targeted optimization strategies to enhance domain-specific recognition.
- Formulated the iterative clustering process by defining distance functions $d(x, m_k^{(t)})$ and centroid update rules, ensuring mathematical rigor in the unsupervised learning pipeline.

Multilingual Text Compression via Subword Tokenization 🌐   Advised by Prof. Jiajun Zhang
University of Chinese Academy of Sciences   *Feb 2024 - Mar 2024*

- Implemented strategies to enhance token representation for LLM, utilizing state of the art NLP techniques and data driven optimization.
- Designed subword tokenization pipelines for comprehensive multilingual news corpora (cs, de, en, es, fr), demonstrating proficiency in cross-lingual text processing.
- Trained BPE vocabularies of varying sizes (500–3000), optimizing the trade-off between token fragmentation and compression efficiency.
- Conducted systematic evaluation on test sets, calculating compression rates and identifying vocabulary configurations that maximize text representation efficiency.
- Extracted insights on tokenization behavior across languages, showing ability to analyze multilingual patterns and adapt methods to diverse linguistic structures.

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.

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