Chenyi Zhuang

(+86) 19883840507 chenyi.zhuang@nuaa.edu.cn Homepage

RESEARCH INTERESTS

Computer Vision, Computer Graphics, Generative Models (particularly, GANs and Diffusion Models), Explainable AI, Multimodal Learning, Representation Learning for text, image, video, and 3D.

EDUCATION

Nanjing University of Aeronautics and Astronautics September 2022 - April 2025
Master of Electronic Information in Computer Technology *Nanjing, China*

- GPA: 91/100
- Core Modules: Python For Data Science (100), Computer Vision and Artificial Intelligence (94), Mathematical Foundations in Information Security (93), Advanced Engineering Mathematics (92).
- Thesis: Research on Consistency of Target and Controllability of Synthesis for Deep Generative Models

Zhejiang Sci-Tech University September 2018 - June 2022
Bachelor of Engineering in Digital Media Technology *Hangzhou, China*

- GPA: 90/100 (ranked #2/83)
- Core Modules: Linear Algebra (99), Audio and Video Signal Processing (95), Digital Image Analysis and Artistic Processing (95), Computer Graphics (94), Discrete Mathematics (92).
- Thesis: Design and Implementation of Form Recognition and Reconstruction Algorithm Based on Image Processing (awarded as outstanding graduation thesis)

PUBLICATIONS

** indicates an equal contribution.*

Chenyi Zhuang, Ying Hu, Pan Gao. Magnet: We Never Know How Text-to-Image Diffusion Models Work, Until We Learn How Vision-Language Models Function. Advances in Neural Information Processing Systems (2024). *[arXiv] [code]*

Qingguo Liu, **Chenyi Zhuang**, Pan Gao, Jie Qin. CDFormer: When Degradation Prediction Embraces Diffusion Model for Blind Image Super-Resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2024). *[arXiv] [code]*

Ying Hu*, **Chenyi Zhuang***, Pan Gao. DiffuseST: Unleashing the Capability of the Diffusion Model for Style Transfer. In Proceedings of the 6th ACM International Conference on Multimedia in Asia (2024). *[arXiv] [code]*

Chenyi Zhuang, Pan Gao, Aljosa Smolic. StylePrompter: All Styles Need Is Attention. In Proceedings of the 31st ACM International Conference on Multimedia (2023). *[arXiv] [code]*

Ying Hu, **Chenyi Zhuang**, Pan Gao. StyTips: Towards High-Quality, Efficient and Controllable Style Transfer via Transformer Filtering Prompts. (accepted to Computational Visual Media)

Zhi Zuo*, **Chenyi Zhuang***, Pan Gao, Jie Qin, Hao Feng, Nicu Sebe. Uni4D: A Unified Self-Supervised Learning Framework for Point Cloud Videos. (under review) *[arXiv]*

Chang Xie*, **Chenyi Zhuang***, Pan Gao. PiCo: Enhancing Text-Image Alignment with Improved Noise Selection and Precise Mask Control in Diffusion Models. *[arXiv]*

PATENTS

Training-free method and system for text-image generation based on diffusion model
Chenyi Zhuang*, Ying Hu, Pan Gao.* Chinese Patent (CN118485074A).

RESEARCH EXPERIENCE

Research Assistant September 2022 - Present
Generative Models for Image Synthesis, Inversion, and Editing

- Explored generative models (e.g., GANs and diffusion models) and identified their challenges in unconditional or conditional synthesis, and editing on real images, particularly for human faces.
- Investigated the compositional understanding of vision-language models and analyzed how the CLIP text encoder with inaccurate concept representations can affect text alignment.
- Designed a novel framework to address the attribute binding issue that manipulates the text embedding of each object to enhance disentanglement between concepts in a training-free manner.

Project Leader September 2023 - November 2024
Postgraduate Research & Practice Innovation Program of NUAA

- Developed a diffusion-based multi-modal visual-guided style transfer approach, combining textual and spatial features of images, and separating the injection in different denoising steps.
- Visualized the intermediate representations of two injection modules in both feature space and Fourier space to verify the enhancement of high-frequency information after injection.
- Built a web application with Gardio that integrates the above style transfer pipeline with a well-designed user interface and straightforward controllable UI components for non-technical users.

AWARDS & SCHOLARSHIPS

Outstanding Graduate Student 2025
Nanjing University of Aeronautics and Astronautics, Nanjing, China

Hua Wei Scholarship 2024
Nanjing University of Aeronautics and Astronautics, Nanjing, China

Innovation Research Advanced Individual 2023
Nanjing University of Aeronautics and Astronautics, Nanjing, China

First Price of Academic Scholarship 2022 - 2024
Nanjing University of Aeronautics and Astronautics, Nanjing, China

Provincial Outstanding Undergraduate Student June, 2022
Zhejiang Provincial Government, Hangzhou, China

First Prize in Provincial Undergraduate Competition May 2022
9th Haikang Cup Student Service Outsourcing Innovation and Application Competition, China

Third Prize in National Undergraduate Competition August 2021
12th Student Service Outsourcing Innovation and Entrepreneurship Competition, China

Second Prize in Provincial Undergraduate Competition May 2021
17th Challenge Cup Extracurricular Academic and Technological Competition, China

SKILLS

Technical skills: Programming languages (Python, C++), Python libraries for machine and deep learning (PyTorch, Diffusers, NumPy), LaTex, Microsoft Office, Linux.
Language skills: Mandarin - mother tongue; English - fluent (IELTS 7/6).