

# Shikun Sun

Shuangqing apartment, Haidian, Beijing, China (Address)  
100084 (ZIP code)

+86 18811382689 (Tel)  
ssk52839916@gmail.com (Email)  
<https://github.com/skipper17> (Github)

My primary research focus is on diffusion models, with a specific interest in advancing new theoretical frameworks, enhancing controllable generation, exploring RL-based fine-tuning techniques, and improving computational efficiency.

## Education

- **Tsinghua University** Bachelor Degree  
*Computer Science and Technology* 2017.09 - 2021.06  
– Undergraduate Thesis: *Aesthetic Knowledge Guided Image Sentiment Transfer*
- **Tsinghua University** Second Bachelor Degree  
*Math and Applied Math* 2018.09 - 2021.06  
– Undergraduate Thesis: *Application of Invertible Neural Networks in Image Processing*
- **Tsinghua University** Pursuing Doctoral Degree  
*Multimedia Information Processing* 2021.09 - now  
– Expected to graduate in 2026.06

## Grade-Point Average

- **Bachelor of Computer Science and Technology** 3.75
- **Bachelor of Math and Applied Math** 3.26
- **Ph.D. in Multimedia Information Processing (by now)** 3.89

## Social Practice Activities

- **Tsinghua University Bauhinia Volunteer Service Team** *Member*

## Selected Award

- **Tsinghua University Academic Excellence Award** *2018*

## Current Work

- **Minimal Impact ControlNet (MIC):** This approach aims to improve the composition of multiple ControlNets by focusing on three key aspects: data augmentation, feature combination strategies inspired by multi-objective optimization, and ensuring score conservativity. These elements are designed to minimize negative interactions between ControlNets when used together, thereby achieving better overall control and stability in the network's output. *Internship at Alimama.*
- **RL for Diffusion:** We applied DPO to Diffusion Models using the DPO with a novel CLIP method and a simple latent mix. This approach resulted in over a 10% improvement in real-world applications for the Tongyi Wanxiang service, including various scenarios such as layout-to-image generation. *Internship at Alimama.*

## Publications

My research publications since 2023 have been centered on the theory and application of Diffusion Models.

- Ziyi Wang, Xingqi Wang, Zeyu Jin, Xiaohan Li, **Shikun Sun**, Jia Jia. *AI Carpet: Automatic Generation of Aesthetic Carpet Pattern*. In Proceedings of the 30th ACM International Conference on Multimedia (MM'22)

- Zijie Ye, Jia Jia, Haozhe Wu, Shuo Huang, **Shikun Sun**, Junliang Xing. *Salient Co-Speech Gesture Synthesizing with Discrete Motion Representation*. In Proceedings of 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'23)
- **Shikun Sun**, Jia Jia, Haozhe Wu, Zijie Ye, Junliang Xing. *MSNet: A Deep Architecture using Multi-Sentiment Semantics for Sentiment-Aware Image Style Transfer*. In Proceedings of the 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'23)
- **Shikun Sun**, Longhui Wei, Junliang Xing, Jia Jia, Qi Tian. *SDDM: Score-Decomposed Diffusion Models on Manifolds for Unpaired Image-to-Image Translation*. In Proceedings of the Fortieth International Conference on Machine Learning. (ICML'23)
- **Shikun Sun**, Longhui Wei, Zhicai Wang, Zixuan Wang, Junliang Xing, Jia Jia, Qi Tian. *Inner Classifier-Free Guidance and Its Taylor Expansion for Diffusion Models*. In Proceedings of the Twelfth International Conference on Learning Representations. (ICLR'24)
- Zixuan Wang, Jia Jia, **Shikun Sun**, Haozhe Wu, Rong Han, Zhenyu Li, Di Tang, Zhou Jiaqing, Jiebo Luo *DanceCamera3D: 3D Camera Movement Synthesis with Music and Dance*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024. (CVPR'24)
- Shuo Huang, **Shikun Sun**, Zixuan Wang, Xiaoyu Qin, Yanmin Xiong, Yuan Zhang, Pengfei Wan, Di Zhang, Jia Jia *PlacidDreamer: Advancing Harmony in Text-to-3D Generation*. In Proceedings of the 32th ACM International Conference on Multimedia (MM'24)
- Zixuan Wang, Jiayi Li, Xiaoyu Qin, **Shikun Sun**, Songtao Zhou, Jia Jia, Jiebo Luo *DanceCamAnimator: Keyframe-Based Controllable 3D Dance Camera Synthesis*. In Proceedings of the 32th ACM International Conference on Multimedia (MM'24)