Kaiwen Zhang

zkw21@mails.tsinghua.edu.cn · https://kevin-thu.github.io/homepage · @sze68zkw

Education

Tsinghua University, Beijing, China Sep. 2021 - Present

Bachelor student in Computer Science and Technology (**GPA: 3.96/4.0, Top 5 / 180**)
Selected courses of A+/A: Fundamentals of Programming, Introduction to Artificial Intelligence, Artificial Neural Networks, Fundamentals of Computer Graphics, Principles of Signal Processing, Introduction to Computer Systems, Software Engineering, Data Structures, Calculus, Linear Algebra, Probability and Statistics

Publications

Research interests: Generative Models, World Models, Computer Vision and 3D Vision.
Kaiwen Zhang, Yifan Zhou, Xudong Xu, Xingang Pan* and Bo Dai.
DiffMorpher: Unleashing the Capability of Diffusion Models for Image Morphing *CVPR 2024*
Project: https://kevin-thu.github.io/DiffMorpher_page
Paper: https://arxiv.org/abs/2312.07409
Code: https://github.com/Kevin-thu/DiffMorpher (***430+ stars***)

Honors and Awards

China Computer Federation (CCF) Outstanding Student Sept. 2024
SenseTime Scholarship (25 undergraduates all over China) Jun. 2024
National Scholarship (Highest honor for undergraduates in China, $\sim$ Top 0.2%) Oct. 2023
Second Prize (3 / 109), The Jittor AI Challenge (Image Generation Track) Sept. 2023
Tsinghua University Comprehensive Excellence Scholarship (Top 5% undergraduates) Oct. 2022
Meritorious Winner (Top 10%), American Mathematical Contest in Modeling Mar. 2022

Experience

Horizon Robotics Aug. 2024 - Present

Research Intern Mentors: Dr. Xiaoxiao Long and Dr. Wei Yin

- Led research on large-scale video generation world model for autonomous driving combining autoregressive and flow matching modeling for pose control and long-horizon high-quality generation.
- Proposed multi-forward technique to address the autoregressive drift issue in long video generation.

Intelligent Digital Creation Group, Shanghai AI Lab Jul. 2023 - Jan. 2024

Research Intern Primary Advisor: Prof. Xingang Pan @ NTU MMLab

- Developed image morphing methods based on diffusion models, enhancing image quality and video smoothness in real image interpolations, achieving performance that significantly surpasses previous methods.
- First-authored a paper accepted at CVPR 2024 (Rating: 5/5/4), receiving high praise for its novelty, simplicity, elegance, and impressive results.

Knowledge Engineering Group, THU Sep. 2022 - Jun. 2023

Research Assistant Advisor: Prof. Jie Tang

- Designed and implemented a VideoVAE to enhance temporal consistency in video generation.
- Implemented *Pix2Video: Video Editing using Image Diffusion* based on the Stable Diffusion.
- Led summer training courses in computer vision and generative models, handout available at https://summer23.net9.org/pdfs/CV_Kevin.pdf.

Skills & Misc

- Programming: Python (including PyTorch), C/C++, LaTeX, TypeScript (React), R
- Language: English (fluent, with TOEFL score: 105, L30 / R29 / S23 / W23), Chinese (native)