Kaihua Chen

https://kaihua-chen.github.io/
Pittsburgh, PA, United States
kaihuac5@gmail.com or +1 412-320-0373

Education

Carnegie Mellon University, Pittsburgh, PA — Aug 2023-Dec 2024
Master of Science in Computer Vision
Advisor: Prof. Deva Ramanan
GPA: **4.17/4.00**
Courses: Machine Learning, 3D Vision, Generative Models, Math for Robotics

China Agricultural University, Beijing, China — Sept 2019-Jul 2023
Bachelor of Engineering in Computer Science (Minor: Mathematics)
GPA: **3.91/4.00** (rank: 1/39)

Selected Publications

Kaihua Chen[*], Tarasha Khurana[*], Deva Ramanan. *Reconstruct, Inpaint, Test-Time Finetune: Dynamic Novel-view Synthesis from Monocular Videos.* Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS), 2025.

Kaihua Chen, Deva Ramanan, Tarasha Khurana. *Using Diffusion Priors for Video Amodal Segmentation.* Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2025.

Yizhou Zhao, Hengwei Bian, **Kaihua Chen**, et al. *Metric from human: Zero-shot monocular metric depth estimation via test-time adaptation.* Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS), 2024.

Cheng-Yen Hsieh, **Kaihua Chen**, Achal Dave, Tarasha Khurana, Deva Ramanan. *TAO-Amodal: A Benchmark for Tracking Any Object Amodally.* arXiv:2312.12433, 2024.

Other Publications

Kaihua Chen, Tingting Zhu, Shaofeng Li, Yinxue Shi. *Facial keypoint-based segment-level driver yawning detection by graph-temporal convolutional neural network modeling.* International Conference on Intelligent Computing (ICIC), 2025 (released on arXiv in 2023).

Kaihua Chen, Tingting Zhu, Shaofeng Li, Yinxue Shi. *Real-time Yawning Detection Based on Machine Learning Algorithm and Time Series Classification using Facial Feature Points.* IEEE International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), 2021.

Research Experience

Carnegie Mellon University, Pittsburgh, PA — Jan 2024-Present
Research Associate at Deva's Lab, advised by Prof. Deva Ramanan

Camera-conditioned Long Video Generation as World Model (Ongoing)

- Investigate autoregressive video diffusion framework for camera-conditioned long-horizon video generation with large view transitions, focusing on mitigating error accumulation, maintaining long-term spatial memory, and enabling accurate camera control

Dynamic 4D Novel-view Synthesis from Monocular Videos

- Designed CogNVS, a video diffusion based on CogVideoX, to synthesize free-viewpoint novel views from monocular videos, by inpainting and fixing MegaSAM point cloud renders
- Devised a self-supervised data generation strategy that repurposes any 2D videos into geometry-structured training pairs, enabling training on web-scale datasets and test-time finetuning during inference for improved generalization
- The work is accepted to NeurIPS '25: https://cog-nvs.github.io/

Diffusion Priors for Video Amodal Segmentation

- Developed a two-stage video diffusion model based on Stable Video Diffusion with depth conditioning to synthesize wholes of partially or fully occluded objects, enabling user-interactive object-level amodal segmentation and content completion in videos
- Established TAO-Amodal Masks benchmark (https://tao-amodal.github.io/) by evaluating baselines, including fine-tuned amodal segmentors, self-supervised completion algorithms, and ViT regression models
- The work is accepted to CVPR '25: https://diffusion-vas.github.io/

University of Toronto, Toronto, ON, Canada Jul 2022-Oct 2022
ML Research Intern at Lin Brain Lab, advised by Prof. Fa-Hsuan Lin
Correlating Human Neural Activity and Behaviors by Machine Learning

- Investigated correlations between fMRI brain imaging, EEG signals, natural stimuli, and subjective feelings, employing a spherical convolution-based UNet and LSTM autoencoder pretrained on Human Connectome Project (HCP) emotion tasks
- The work is accepted as a poster at ISMRM '23

China Agricultural University, Beijing, China Apr 2021-Mar 2022
Undergraduate Researcher at Department of Computer Science
AI-Powered Facial Behavior Analysis

- Innovated a Graph-Temporal Convolutional Network (GTCN) for real-time drowsiness detection based on facial keypoints recognized by OpenPose, achieving a 2.6% accuracy improvement on YawDD compared with SOTA and deploying the system on Raspberry Pi 4B
- The work is accepted to HPBD&IS '21 and ICIC '25 (initially released in 2023)

Course Projects

Zero-shot Monocular Metric Depth Estimation (CMU 16-825) Jan 2024-May 2024

- Developed a zero-shot adaptation method that converts relative depth to metric depth using generative painting and human mesh recovery, achieving strong generalization in unseen scenes; the work is accepted to NeurIPS '24

Robust 3D Surface Reconstruction (CMU 16-811) Oct 2023-Dec 2023

- Implemented Poisson Surface Reconstruction in Python, introducing a random adaptive resampling technique to address highly non-uniform point clouds

Intelligent Bitcoin and Gold Trading Strategy (2022 MCM Finalist) Feb 2022

- Devised a trading strategy based on temporal deep learning prediction and greedy Markov decision process, achieving 30x return in simulated trading

IoT System for Greenhouse Environment Monitoring Apr 2020-Mar 2021

- Constructed an Arduino UNO-based IoT system featuring a PyQt GUI, enabling data collection, wireless transmission, and remote control, and obtained software copyright

Awards

2022 International Mathematical Contest In Modeling (MCM), Finalist (Top 1%)
2021 National Scholarship (Top 1%)
2020 Arawana Scholarship (Top 2%)
2020, 2021 Outstanding Student Scholarship, First Prize (Top 5%)
12^{th} Chinese Mathematics Competitions (CMC), Second Prize
31^{st} Beijing Mathematics Competition (Group A), Second Prize
2020 Huawei Cup Mathematics Competition, Third Prize

Personal

Programming Languages: Python (Proficient); C/C++, Java, MATLAB, SQL (Familiar)
Libraries: PyTorch, Diffusers, NumPy, Matplotlib, Scikit-learn, OpenCV, Open3D, Gsplat, Git
Hobbies: Movies, Soccer