

# KELING YAO

+1 (878)-256-9379 | [kenny@andrew.cmu.edu](mailto:kenny@andrew.cmu.edu) | [Home Page](#) | [Linkedin](#) | [github](#) | Pittsburgh, PA

## EDUCATION

### Carnegie Mellon University

*Master of Science in Computer Vision*

**Dec 2025**

*Pittsburgh, PA*

### The Chinese University of Hong Kong

*Bachelor of Science in Data Science and Big Data Technology (CS track)*

**Jul 2024**

*Shenzhen, China*

- Major GPA: **3.87/4.00 (top 6%)**, Cumulative GPA: 3.72/4.00

### University of California, Berkeley

*Exchange Student in Computer Science*

**May 2023**

*Berkeley, CA*

- Major GPA: **4.00/4.00 (top 1%)**, Cumulative GPA: **4.00/4.00**

## SKILLS

**Programming Languages:** Python, C, C++, C#, Java, JavaScript, SQL, MATLAB, PowerShell, RISC-V, LaTeX

**Developer Tools:** Docker, AWS, Unity, Blender, Git, Azure, OpenGL

**Frameworks:** Numpy, PyTorch, TensorFlow, OpenCV, ROS, Sklearn, pandas

**Other:** Machine Learning, Computer Vision, Image Segmentation, VR/AR/XR, Generative AI, Large Language Models, Multimodal Models, Diffusion Models, Data-driven, Data Processing, Robot Learning, Reinforcement Learning, Graphics, DNN, R&D

## WORK EXPERIENCE

### DeepMirror

**May 2024 – Aug 2024**

*Machine Learning Engineer Intern*

*Guangzhou, China*

- Implemented a real-time 3D bounding box generator by training a high-resolution saliency detection model, significantly improving the surface detail of meshes and resolving floater issues in NeRF 3D reconstruction.
- Implemented and experimented on distortion loss in NeRF 3D reconstruction, leading to a significant reduction in air floaters and mesh surface discontinuities, resulting in a 0.3 PSNR improvement.
- Developed a construction-specific blurry image detection model with an F1 score of 0.99, deployed into mobile devices at 130 FPS, enhancing data robustness for Structure from Motion (SfM) 3D reconstruction pipelines.

## RESEARCH PROJECTS

### University of California, Berkeley

**Dec 2022 – May 2024**

*Research Assistant, FHL Vive Center for Enhanced Reality*

*Berkeley, CA*

- **Published** “Robust 6DoF Pose Estimation Against Depth Noise and a Comprehensive Evaluation on a Mobile Dataset” tailed for **AR/VR mobile applications at DMLR @ ICML as co-first author.**
- Contributed to a Transformer-based 6DoF pose estimator designed to withstand the low-quality depth data from mobile LiDAR sensors, surpassing all existing baselines with AUC results of ADD and ADD-S as 73.99 and 88.10.
- Led and established the **first** mobile RGBD 6DoF pose estimation dataset (DTTDv2), by developing an efficient data collection paradigm (**open-source**), utilizing the OptiTrack system to automatic labelling and improve 80% time.
- Conducted LiDAR depth analysis on 47668 frames and introduced a novel depth-ADD metric for quantitative measurement.

### Microsoft Research Asia

**Jun 2023 – Sep 2023**

*Research Assistant, Multimedia Search and Mining*

*Beijing, China*

- **Published** “Transferring Foundation Models for Generalizable Robotic Manipulation” at WACV (**review**) as **co-author.**
- Collaborated in designing an end-to-end policy model by leveraging Vision Foundation Models and multi-modal multi-view fusion of semantic, geometric, and temporal observations using self-attention mechanism, achieving sample-efficient generalization learning with 81.25% accuracy on real world robot pick and place problem.
- Engineered real-time, precise, and rich human-like data collection pipeline by developing a 30Hz robotic manipulation framework that uses a VR headset to teleport a Franka robot and dexterous hands via a Unity app and UDP connections.

## PROJECTS

### Zi2Calli: a GAN-Based Model for Chinese Fonts Transfer | *The Chinese University of Hong Kong*

**Sep 2023 – May 2024**

- Implemented and trained Zi2Calli (open-source), a GAN-based model with category embeddings and category loss for enhanced one-to-many style transfer, achieving diverse and high-quality calligraphy/font generation in a single framework.
- Curated the first comprehensive dataset of President Xu’s calligraphy, consisting of 484-character images from 95 works using Tesseract OCR.

### General Computational Machine based on GPT-2 | *UC Berkeley*

**Apr 2023 – Aug 2023**

- Constructed and trained a General Computational Machine by incorporating a pre-trained GPT-2 model.
- Finetuned the model to generalize in diverse domain tasks beyond traditional language tasks (Bit-wise operation, Bit Memory, and Image Classification) and visualized the generalization capabilities of transformers by attention map.

### Facial Recognition and Interaction Robotics | *South China University of Technology*

**May 2022 – Sep 2022**

- Deployed face recognition algorithms (yolov5, dlib, arcface) on robot Linux system and was used in commercial robot.
- Developed robot software enabling robots to interact with users, memorize unfamiliar faces, and recognize previously encountered faces.