

Kai XU
3896 Haven Ave, Fremont, CA 94538
(480) 758-0625
kaix.personal@gmail.com
 [linkedin.com/in/kai-xu-92807471/](https://www.linkedin.com/in/kai-xu-92807471/)
 Kai Xu

RESEARCH INTEREST

Computer vision, image processing, compressive sensing, machine perception, hardware-friendly deep learning, information-aware sensing, pose estimation and tracking, digital avatar, video understanding, voice clone/multi-speaker multilingual TTS.

WORK EXPERIENCE

Staff Machine Learning Engineer

Jan 2022 -
Present

CoCoPIE.ai
Palo Alto, California

1. Developed an automatic model compression and deployment framework for CV/NLP applications such as classification, object detection, segmentation, video action recognition, pose estimation and voice clone/TTS. Please click to watch our demos: <https://www.cocopie.ai/#/services>
2. Lead and collaborate with student interns on multiple projects, including face detection, face swap, and 3D talking head generation.

Senior Computer Vision Software Engineer

April 2021 -
Jan 2022

Innopeak Technology Inc. (OPPO Research US)
Palo Alto, California

1. Developed and deployed a hand pose estimation method for OPPO AR glasses using built-in bird-eye and ToF cameras. Our method is similar to the "MEgATrack" project at Facebook Reality Lab (FRL). I propose several techniques to improve pose estimation and tracking accuracy. The invention is successfully deployed on OPPO AR glasses and showcased on OPPO INNO DAY. Please click to watch our products: <https://www.youtube.com/watch?v=fm0bzCU9ZR4>
2. Developed a hand tracking and gesture recognition method for OPPO SmartTV and filed one patent. The invention is successfully deployed on OPPO SmartTV and showcased on OPPO INNO DAY.

PhD Research Intern

2020

Kwai Inc.
Palo Alto, California

1. Developed a music generation system using deep learning for short videos.
2. Define a new research problem of generating soundtracks with appropriate mood, genre, and tempo synchronized with the video contents.
3. Build a new dataset for the video-based music generation task by crawling data from the backend database at Kwai.

PhD Research Intern

2019

Alibaba DAMO Academy
Sunnyvale, California

1. Developed a frequency-domain data pre-processing pipeline for accelerating computer vision systems.
2. Both communication bandwidth between CPU-GPU/AI accelerator and data pre-processing time cost can be reduced for a typical DL-based CV system.

3. Published one paper at CVPR 2020 and filed five patents. Please click to view my presentation: <https://www.youtube.com/watch?v=dy7M1y1xE5k>

PhD Research Intern

2017

Samsung Research America Inc.
Mountain View, California

1. Developed an end-to-end lipreading system for Samsung mobile phones. Published one paper at FG 2018, which was cited by Prof. Nando de Freitas at Oxford University and DeepMind.
2. Please click to view my presentation: <https://www.youtube.com/watch?v=i-gFMLnFfqE>

PhD Research Assistant

2015-2019

Arizona State University
Tempe, Arizona

Computer vision, computational imaging, compressive sensing.

EDUCATION

Ph.D. in Computer Engineering (Computer Science)

2015-2021

Arizona State University
Tempe, AZ, USA
Advisor: Fengbo Ren

M.S. in Electrical Engineering

2011-2014

University of Electronic Science and Technology of China
Chengdu, Sichuan, China

B.S. in Electrical and Electronics Engineering

2007-2011

Shandong University
Jinan, Shandong, China

PUBLICATION

1. [Kai XU](#), Minghai Qin, Fei Sun, Yuhao Wang, Yen-Kuang Chen, Fengbo Ren, "Learning in the Frequency Domain" *In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
2. [Kai XU](#), Zhikang Zhang and Fengbo Ren, "LAPRAN: A Scalable Laplacian Pyramid Reconstructive Adversarial Network for Flexible Compressive Sensing Reconstruction" *In European Conference on Computer Vision (ECCV)*, 2018.
3. [Kai XU](#), Dawei Li, Nick Cassimatis and Xiaolong Wang, "LCANet: End-to-End Lipreading with Cascaded Attention-CTC" *In IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, 2018, poster spotlight.
4. [Kai XU](#) and Fengbo Ren, "CSVideoNet: A Real-time End-to-end Learning Framework for High-frame-rate Video Compressive Sensing" *In IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2018.
5. [Kai XU](#), Yixing Li, and Fengbo Ren, "A Data-Driven Compressive Sensing Framework Tailored for Energy-efficient Wearable Sensing" *In International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2017, oral.
6. [Kai XU](#), Yixing Li, and Fengbo Ren, "An Energy-Efficient Compressive Sensing Framework Incorporating Online Dictionary Learning for Long-Term Wireless Health Monitoring" *In International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2016.
7. Zhikang Zhang, [Kai XU](#), and Fengbo Ren, "CRA: A Generic Compression Ratio Adapter for End-To-End Data-Driven Image Compressive Sensing Reconstruction Frameworks" *In International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2020, oral.

8. Yixing Li, Zichuan Liu, Kai XU, Hao Yu, and Fengbo Ren, "A GPU-Outperforming FPGA Accelerator Architecture for Binary Convolutional Neural Networks" *In ACM Journal on Emerging Technologies in Computing Systems (JETC)*, 2018.
9. Duo Lv, Kai XU, and Dijiang Huang, "A Data Driven In-Air-Handwriting Biometric Authentication System" *In International Joint Conference on Biometrics (IJCB)*, 2017.
10. Yixing Li, Zichuan Liu, Kai XU, and Fengbo Ren, "A 7.663-TOPS 8.2-W Energy-efficient FPGA Accelerator for Binary Convolutional Neural Networks (Abstract Only)" *In ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA)*, 2017.
11. Yuhao Wang, Xin Li, Kai XU, Fengbo Ren, H. Yu, "Data-Driven Sampling Matrix Boolean Optimization for Energy-Efficient Biomedical Signal Acquisition by Compressive Sensing" *In IEEE Transactions on Biomedical Circuits and Systems (TBioCAS)*, 2016.

PATENTS

1. Fengbo Ren, Kai XU. Real-Time End-to-End Learning System for High-frame-Rate Video Compressive Sensing Network.
2. Fengbo Ren, Kai XU, Zhikang Zhang. LAPRAN: A Scalable Laplacian Pyramid Reconstructive Adversarial Network for Flexible Compressive Sensing Reconstruction.
3. Minghai Qin, Yen-Kuang Chen, Kai XU, Yuhao Wang, Fei Sun, Yuan Xie. Reconstructing Transformed Domain Information in Encoded Video Streams.
4. Kai XU, Fei Sun, Minghai Qin, Yen-Kuang Chen. Static Channel Filtering in Frequency Domain.
5. Kai XU, Fei Sun, Minghai Qin, Yen-Kuang Chen. Data Preprocessing and Data Augmentation in Frequency Domain.
6. Kai XU, Minghai Qin, Yuhao Wang, Fei Sun, Yen-Kuang Chen, Yuan Xie. Techniques for Determining Importance of Encoded Image Components for Artificial Intelligence Tasks.
7. Kai XU, Minghai Qin, Yuhao Wang, Fei Sun, Yen-Kuang Chen, Yuan Xie. Techniques to Dynamically Gate Encoded Image Components for Artificial Intelligence Tasks.

CONFERENCE PRESENTATION

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	2020
IEEE International Conference on Automatic Face & Gesture Recognition (FG)	2018
European Conference on Computer Vision (ECCV)	2018
IEEE Winter Conference on Applications of Computer Vision (WACV)	2018
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)	2016, 2017

HONORS

ASU Outstanding Computer Engineering PhD Student (top 1)	2021
ASU Engineering Graduate Fellowship	2020
ASU CIDSE Doctoral Fellowship	2019
ASU Completion Fellowship (top 3)	2019

PROPOSAL

1. Neural Network Based Video Compression: A Real-time End-to-end Learning Framework for High-Frame-Rate Camera, co-authored with Prof. Fengbo Ren, **Google faculty research award**, 2017.