# Kai Huang

CONTACT

3700 O'Hara St, Pittsburgh, PA 15213

Information Dept. of Electrical and Computer Engineering

University of Pittsburgh

Pittsburgh, PA, 15213 Homepage: https://hellokevin07.github.io/

RESEARCH INTERESTS On-device AI, AI for systems, AI-assisted wireless communication systems

EDUCATION

University of Pittsburgh, Pittsburgh, PA

Ph.D. student, Electrical and Computer Engineering

Advisor: Prof. Wei Gao

University of Science and Technology of China (USTC), Hefei, Anhui, July 2019

**B.E.**, Electronic Information Engineering

RESEARCH EXPERIENCE Research Assistant

2019-present

On-going

Telephone: (412)277-5047

Email: k.huang@pitt.edu

Dept. of Electrical and Computer Engineering, University of Pittsburgh

- Designing a selective training scheme that can accelerate on-device neural network training.
- Developed and implemented an offloading scheme that allows extremely weak devices (e.g., MCUs with <1MB memory) to achieve real-time (<20ms) neural network inference. It is the first work that leverages Explainable AI to speed up neural network inference on weak devices.
- Developed and implemented a backscatter system that leverages neural network inference to improve its RF energy efficiency by up to 3.5x. The neural network is tailored based on the domain knowledge of backscatter communication, and hence is very lightweight and can be effectively trained even with a limited amount of data.

#### **PUBLICATIONS**

## **Conference Papers**

- \* indicates equal contributions
- 1. Xiangyu Yin, <u>Kai Huang</u>, Erick Forno, Wei Chen, Heng Huang, Wei Gao. "PTEase: Objective Airway Examination for Pulmonary Telemedicine using Commodity Smartphones." In Proceedings of the 21st International Conference on Mobile Systems, Applications, and Services (**MobiSys'23**) (Accepted)
- 2. <u>Kai Huang</u>, Boyuan Yang, Wei Gao. "ElasticTrainer: Speeding Up On-Device Training with Runtime Elastic Tensor Selection." In Proceedings of the 21st International Conference on Mobile Systems, Applications, and Services (**MobiSys'23**) (Accepted)
- 3. Chen Ruirong, Kai Huang, Wei Gao. "AiFi: AI-Enabled Interference Cancellation in WiFi Networks with Commodity PHY-Layer Information." Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems (SenSys'22), pp. 134-148. 2022.
- 4. <u>Kai Huang</u>, Wei Gao. "Real-time Neural Network Inference on Extremely Weak Devices: Agile Offloading with Explainable AI." In Proceedings of the 28th Annual International Conference on Mobile Computing and Networking (**MobiCom'22**), pp. 200-213. 2022.
- 5. <u>Kai Huang</u>, Ruirong Chen, Wei Gao. "RAScatter: Achieving Energy-Efficient Backscatter Readers via AI-Assisted Power Adaptation." In 2022 IEEE/ACM Seventh International Conference on Internet-of-Things Design and Implementation (**IoTDI'22**), pp. 1-13. IEEE, 2022.
- Song, Xingzhe, <u>Kai Huang</u>, Wei Gao. "FaceListener: Recognizing Human Facial Expressions via Acoustic Sensing on Commodity Headphones." In 2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN'22), pp. 145-157. IEEE, 2022.
- 7. Boyuan Yang, Ruirong Chen, <u>Kai Huang</u>, Jun Yang, Wei Gao. "Eavesdropping user credentials via GPU side channels on <u>smartphones</u>." In Proceedings of the 27th ACM International

- Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS'22), pp. 285-299. 2022.
- 8. Yihao Liu\*, <u>Kai Huang</u>\*, Xingzhe Song, Boyuan Yang, Wei Gao. "MagHacker: eavesdropping on stylus pen writing via magnetic sensing from commodity mobile devices." In Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services (**MobiSys'20**), pp. 148-160. 2020.

## Public Speaking

## Presentations

- 1. "AiFi: AI-Enabled WiFi Interference Cancellation with Commodity PHY-Layer Information." In Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems (Sen-Sys), Boston, USA, Nov 2022.
- 2. "Real-time neural network inference on extremely weak devices: agile offloading with explainable AI." In Proceedings of the 28th Annual International Conference on Mobile Computing And Networking (MobiCom), InterContinental Sydney, Australia, Oct 2022
- 3. "RAScatter: Achieving Energy-Efficient Backscatter Readers via AI-Assisted Power Adaptation." In 2022 IEEE/ACM Seventh International Conference on Internet-of-Things Design and Implementation (IoTDI), Virtual, May 2022
- 4. "Towards Real-time Neural Network Inference on Extremely Weak Devices", Elijah Group Meeting, Dept. of Computer Science, Carnegie Mellon University, November 2021
- "Tailoring Neural Network Designs to Computing System Domains", Elijah Group Meeting, Dept. of Computer Science, Carnegie Mellon University, March 2021

# TEACHING AND MENTORING EXPERIENCE

## Teaching:

• Teaching Assistant, ECE1175 - Embedded Systems Design Dept. of Electrical and Computer Engineering, University of Pittsburgh	Spring 2021
• Teaching Assistant, ECE1175 - Embedded Systems Design Dept. of Electrical and Computer Engineering, University of Pittsburgh	Fall 2020
• <b>Teaching Assistant</b> , ECE0202 - Embedded Processors and Interfacing Dept. of Electrical and Computer Engineering, University of Pittsburgh	Spring 2020

## Professional Activities

#### Journal Reviewer

• IEEE Transactions on Mobile Computing

## Conference Reviewer

- IEEE International Conference on Mobile Ad-Hoc and Smart Systems (MASS) 2022,
- IEEE Conference on Computer Communications (INFOCOM), 2022, 2023