

Personalized Automatic Speech Recognition on Mobile Computing Devices

As artificial intelligence (AI) algorithms enable transformative new user experiences in mobile computing devices, data security and privacy has become increasingly important. In a typical deployment scenario, AI models are trained in the cloud with massive amounts of data and deployed to mobile devices "as-is". While such an approach is generalizable to a majority of the population, certain user groups experience subpar performance (e.g. as a result of training bias introduced by imbalanced training datasets). Edge-based machine intelligence, where user personalization and incremental training is performed natively on the mobile device itself, offers a potential solution to these challenges. Such design goals impose challenging constraints on system-level performance and energy efficiency and necessitates hardware/software co-design at the system level.

The Wang lab at Stanford developed an adaptive automatic speech recognition (ASR) system for edge devices designed to address these challenges and pave a path towards enabling personalized ASR experiences for a multitude of users (e.g. privacy conscious users, non-native speakers, etc.). The technology employs a multi-faceted approach which leverages the triplet loss function and acoustic embeddings at the software level, and associative memories at the hardware level to rapidly identify words in constant time. The ASR technology employs an energy efficient architecture, provides per-user personalization and is highly adaptable.

Applications

- Personalized Automatic Speech Recognition
- Speaker Identification
- Acoustic Signal Identification

- Multilingual Translation

Advantages

- Per-user Personalization
- Energy Efficient Architecture
- Highly Adaptable

Patents

- Published Application: [WO2023159072](#)

Innovators

- William San-Hsi Hwang
- Shan Wang

Licensing Contact

Luis Mejia

Senior Licensing Manager, Physical Sciences

[Email](#)