Open Thesis / Project / Paid Student Job

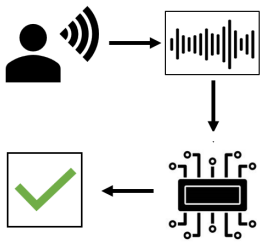
Robust Speaker Recognition for Deeply Embedded IoT Devices

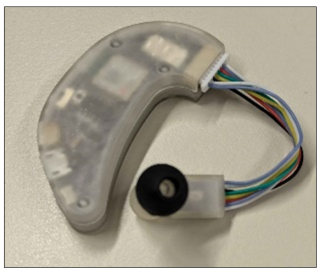
Note: *this work is suitable for a MSc project/thesis, and can also be performed as a* ***paid student job****.*

Motivation

With the emergence of powerful speaker recognition systems like Alexa and Siri, people are becoming fond of using voice activation instead of physical input. Together with the rapid growth of the Internet of Things (IoT), an increasing number of embedded devices are being deployed in various settings that could benefit from voice activation. However, embedded IoT devices are often not as powerful as Alexa or Siri, and only have a fraction of their memory and processing power.

Our goal is to let constrained embedded IoT devices (i.e., deeply embedded devices with limited memory and computational resources) perform speaker recognition autonomously: this could be used to detect people's presence in a room (e.g., to easily clock in and out of work), or to eliminate the need for bulky keyboard or pin code inputs. Instead, only a small and cheap microphone needs to be present on the IoT device. Towards this goal, we have shrunk a state of the art model to fit the resource constraints of low-power IoT devices while retaining on par accuracy, and deployed it on a tiny custom-made earpiece. However, deploying a model in the real world often leads to worse performance (e.g., in noisy environments) and we hence want to study how to increase the robustness of our approach.





Goals and Tasks

Within this context, students can explore several directions and perform different tasks, such as:

- Understand how state-of-the-art speaker recognition systems work, and how they can be shrunk to fit on embedded devices;
- Implement and potentially shrink a noise removal model on an embedded device that can run before the speaker recognition model;
- Leverage a second microphone pointed into the ear to distinguish external noise from speech, or leverage an in-ear pressure sensor to detect if the person is speaking;
- Develop a prototype of a robust speaker recognition system running on a constrained IoT device (e.g., our earpiece embedding the `nRF5340`).

Target Group

- Students of ICE/Telematics;
- Students of Computer Science;
- Students of Electrical Engineering.

Required Prior Knowledge

- Basic knowledge of machine learning;
- Solid skills in Python and C programming;
- Experience with microcontrollers.

Contact Person

- Dipl.-Ing. Markus Gallacher
  markus.gallacher@tugraz.at
- Assoc.Prof. Carlo Alberto Boano
  cboano@tugraz.at

4480 – Institute of Technical Informatics (ITI)

Low-Power Embedded Networked Systems (LENS) Group

Group leader: Assoc.Prof. Carlo Alberto Boano