Project No
TEYDEB 3205018

Speech Processing at the Edge for Smart Building Automation

Funding Programme
TUBITAK (Turkish NCP)

Project Start Date
01.01.2021

Project End Date
30.06.2022

Coordinator
Interra

Partners
SESTEK

Objectives

The project originated from the need to activate smart building systems offline, allowing the use of touch panel devices without an internet connection, as well as mobile devices and applications. Developed collaboratively by Interra and SESTEK as a custom R&D project, the system enables the setting and usage of peripherals by processing voice commands at the edge.

Once the building automation system is triggered with the wake word, SESTEK's speech recognition component takes the stage. Interra enables users to command the system just after the "beep," for instance, turning on the lights of the room or vice versa. The edge device in the project is responsible for collecting and processing speech data locally, which reduces latency, enhances privacy, and minimizes the need for constant connectivity to the cloud. Given the device's limited computing power, efficiently processing data with a minimized model while optimizing latency posed an additional challenge addressed during the project.

An example use case scenario unfolds as follows:

1. The user initiates the building automation system with the trigger word "Hey Interra."

2. The system transitions to command listening mode following the "beep" sound.

3. The user issues the command "turn on the lights."

4. The embedded Speech Recognition module is engaged to detect the command immediately after the voice activation and noise filtering steps.

5. The pertinent automation process is executed.

SESTEK's Ambition to be a Part of This Project

SESTEK promotes expanding the use cases of its solutions within a collaborative partnership to create a rapidly implementable product with high commercialization potential.

In this project, SESTEK leverages voice control for smart automation in home and office environments, implementing embedded speech processing on low-computing edge devices without cloud transfers, thereby addressing privacy concerns for users.

Project Rationale

Smart automation in home and office environments using speech recognition lies in enhancing convenience, efficiency, and privacy.
By enabling voice-controlled automation without the need for constant internet connectivity or cloud-based processing, this approach ensures real-time, secure, and immediate control over smart devices.

SESTEK's Role and Achievements

- Embedded wake word detection on edge: The user is enabled to activate the device's "command listening mode" by the wake word "Hey Interra." The process is entirely executed on the device. Due to the device's low CPU capacity, a model that consumes minimal CPU while maintaining high recognition accuracy was required, without compromising recognition success. Thereby it has been optimized to ensure CPU consumption does not exceed 35%, achieving a recognition accuracy of 98%.

- Embedded speech processing on edge: The system switches into the command listening mode right after the wake word, followed by a request for a command from the user regarding the action she/he wants to be performed by the automated system. As is the case with wake word trigger low CPU capacity requirement, high recognition accuracy maintenance its importance. While optimizing CPU consumption not to exceed 63%, the recognition accuracy has been maintained at 92% and above up to a distance of 3 meters.

- Dynamic Command List: To provide enhanced and elevated user experience, dynamic user commands were built so that more than one command could be executed. For example, the user can freely define a command such as "light up the environment" instead of the "turn on the light" command. In order not to reduce the recognition success, lists of voice recognition (such as grammar, and garbage) are provided to be dynamic.

- Sound classification: As part of the study to distinguish sounds containing human voice from those that do not, the MAP (Mean Average Precision) score within 527 classes is 0.41, within 137 classes is 0.55, and within 7 classes is 0.64.

Interra User Interface