# Research proposal:Development of an Embedded Speech-to-Text and Emotion Recognition Device with Haptic Feedback for Individuals with Hearing Impairments

Student Name: Mrityunjay Gupta

December 14, 2024



## Project Overview

The proposed research focuses on developing an embedded, wearable device designed to assist individuals with hearing impairments by converting spoken language into text and classifying the speaker's emotional state through audio processing. The device will also provide personalized haptic feedback to represent different family members, using distinct vibration patterns to convey both identity and emotion. By integrating speech recognition and emotion detection algorithms optimized for resource-constrained embedded systems, this innovative solution aims to enhance communication and social interaction for users, offering not only textual information but also emotional and identity cues through alternative sensory modalities.

### Subject

Embedded Systems, Computer Science, Machine Learning.

### Ref ID

OMO/RE/170

# Problem

Individuals with hearing disabilities often face significant challenges in verbal communication, particularly in understanding spoken language and perceiving the emotional context of conversations. Traditional assistive devices primarily focus on amplifying sound but do not provide insights into the speaker's mood or identity. For example, a person with profound hearing loss may rely on lip-reading or sign language but still miss critical emotional cues, leading to misunderstandings and social isolation. There is a need for a comprehensive solution that not only transcribes speech but also conveys emotional and identity information through alternative sensory modalities.

# Importance of Solving the Problem

Addressing this problem is crucial for enhancing the quality of life and social integration of individuals with hearing impairments. By providing speech-to-text conversion coupled with emotion classification and personalized haptic feedback, the proposed device can bridge communication gaps. It empowers users to engage more fully in conversations, recognize emotional nuances, and identify speakers based on haptic signals. This innovation has the potential to reduce social barriers, improve interpersonal relationships, and contribute to inclusivity in various social settings.

# Research Description

The research focuses on designing and developing an embedded device that integrates three core functionalities: speech-to-text conversion, emotion classification through audio processing, and personalized haptic feedback. The device will utilize advanced digital signal processing techniques and machine learning algorithms optimized for real-time execution on embedded hardware. Speech input captured via microphones will be processed to generate textual transcripts using state-of-the-art speech recognition models like Deep Neural Networks (DNNs) or Recurrent Neural Networks (RNNs). Simultaneously, the device will analyze acoustic features such as pitch, timbre, and prosody to classify the speaker's emotional state using algorithms like Support Vector Machines (SVM) or Convolutional Neural Networks (CNNs). The haptic feedback system will employ actuators capable of producing distinct vibration patterns corresponding to different family members or emotional states, leveraging techniques from haptic communication and sensory substitution research.

# Research Hypothesis

- **H1:** It is possible to achieve accurate real-time speech-to-text conversion on an embedded platform suitable for wearable devices.
- **H2:** Emotional states of a speaker can be reliably classified through audio signal processing techniques implemented on embedded hardware.
- **H3:** Customized haptic feedback can effectively convey information about the speaker's identity and emotional state to individuals with hearing impairments.

# Research Questions

- **RQ1:** What are the optimal speech recognition algorithms for real-time transcription on resource-constrained embedded systems?
- **RQ2:** Which audio features are most indicative of emotional states, and how can they be efficiently extracted and classified in real-time?
- **RQ3:** How can haptic feedback patterns be designed to be easily distinguishable and convey information like identity?

# Approach

- **Literature Review:** Conduct a comprehensive survey of existing technologies in speech-to-text conversion, emotion recognition from audio signals, and haptic feedback mechanisms, focusing on embedded implementations.
- Adapt lightweight speech recognition algorithms suitable for embedded systems.
- Identify and extract relevant audio features (e.g., MFCCs, pitch contours) and develop classification models optimized for real-time processing.
- Select appropriate embedded hardware platforms (e.g., Rpi) capable of handling the computational requirements.
- Design the haptic feedback system using actuators (e.g., vibrotactile motors) and develop a mapping between feedback patterns and specific speakers or emotions.
- Integrate all modules into a single embedded system, ensuring efficient communication between components and overall system optimization.
- Implement power management strategies to make the device suitable for wearable applications.
- Conduct functional testing to validate the accuracy of speech-to-text conversion and emotion classification.

# Prerequisites

- In-depth understanding of digital signal processing and machine learning techniques for speech and audio analysis.
- Basic understanding in deep learning frameworks like PyTorch or TensorFlow.
- Familiarity with previous studies on haptic communication and sensory substitution devices.
- Proficiency in embedded system programming (C/C++), real-time operating systems, and hardware-software co-design.

# References


1. Marwa, Tharwat., Yasmin, Wardak., Shumokh, Balbaid., Ejlal, Radin. (2024). 1. Wearable Device With Speech and Voice Recognition for Hearing-Impaired People. doi: 10.1109/lt60077.2024.10469567
2. Seong, Jae, Lee., Sunmee, Kang., David, K., Han., Hanseok, Ko. (2016). 2. Dialogue enabling speech-to-text user assistive agent system for hearing-impaired person.. Medical & Biological Engineering & Computing, doi: 10.1007/S11517-015-1447-8


### General Note:

1. The device will be a prototype only and not a final minimum viable product (MVP) or production-grade device, but it will be functional enough to demonstrate the project idea.
2. The software component will be implemented in Python and will not be implemented for iOS or Android devices.

# FAQ's

### What is a research description?

- A research description outlines the main objectives, methods, and scope of a study. It serves as a comprehensive overview of the research project, helping you to understand what the study aims to achieve and how it plans to do so.

## What is a research hypothesis?

- A research hypothesis is a specific, testable prediction about what you expect to happen in your study. It is based on knowledge, theory, and research related to the topic and directs the focus of the study.

## What is a research question?

- Research questions are the questions around which you center your research. They should be clear, focused, and researchable within the constraints of your project.

## What if I don't fulfil all the prerequisites?, Can I still Continue?

- Yes, you can continue with the idea; just let your mentor know beforehand, and they will guide you with the required skills.

## What are the time commitments I will need to give for research?

- The program requires a commitment of 30 hours over a period of 4 months. This allows for a well-paced and immersive research experience without overwhelming your academic schedules.
- In case there is no experience of TECH / Analysis, then the student needs to allocate 20 hours for skilling as well.
- All sessions for Research and Skilling are aligned one-on-one with mentor
- Sessions are one-on-one with the mentor assigned to you. Ideally, the sessions should be twice a week, and this will ensure you complete the research project in 4 months.
- In case of an exam, health, or personal reasons, you can take a pre-informed break, which will carry forward the time to complete accordingly.

## Who will be my mentor?

- Each student will be assigned a dedicated mentor, who is full-time with OMOTEC, and will guide you through every step of the process. All our mentors are qualified engineers. Your mentor will be your partner in exploring and taking your idea to execution, overcoming challenges, and achieving success.