Memo

To: Professor Pisano

From: Hassan Hijazi, Nicholas Hardy, Riya Deokar, Marybel Boujaoude, Jazmyn Walker

Team: EchoView.AI

Date: 11/14/23

Subject: First Prototype Testing Report

1.0 Required Materials

Hardware:

- OLED Display
- ESP32-C3 Module
    - Power Supply: 3.0V to 3.6V

Software:

- Espressif IoT Development Framework (ESP-IDF)
- Xcode
    - Swift Programming Language

2.0 Setup

Our testing setup involves a structured plan focusing on both hardware integration and software development. The hardware module consists of the OLED display, which will show the converted speech-to-text, and the ESP32-C3, which is necessary for its Wi-Fi and Bluetooth capabilities. Regarding the software aspect, we will be using Xcode and the Swift programming language to develop an iOS compatible mobile application, tailored for user interaction and data visualization. The app will include a feature where the phone's microphone will listen and take in audio input and then convert it into written text directly on the screen. This is a backup option for our user in scenarios where the device might run out of battery and lose power. After initial testing, we will focus on combining both aspects of the hardware and software pieces to seamlessly work together and create a cohesive system for the user.

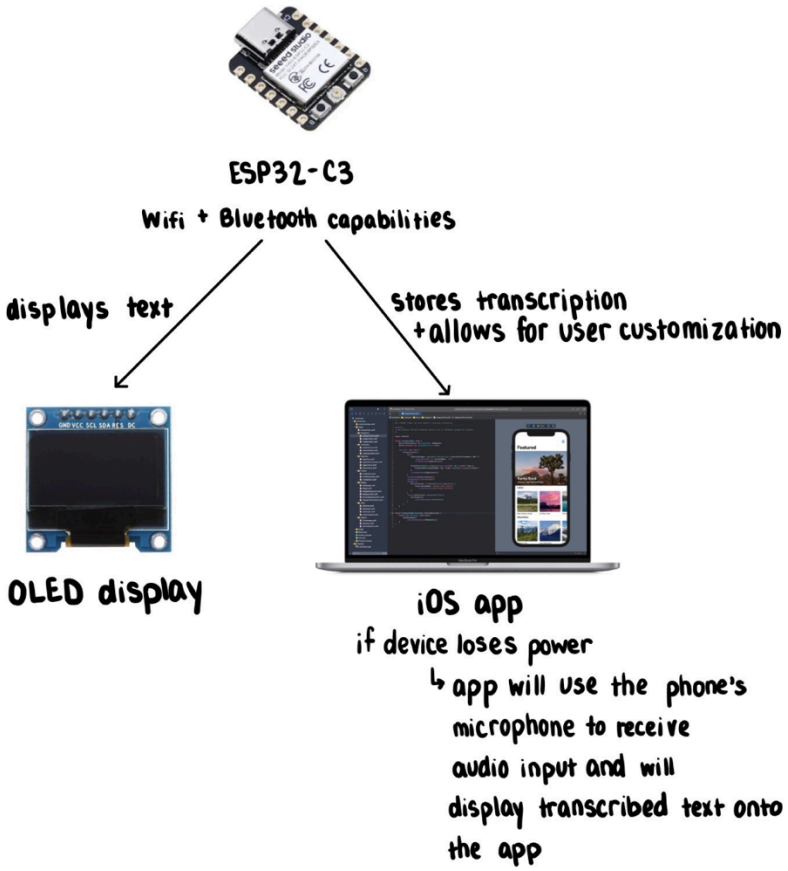

Figure 1: Illustration of Setup and Process Flow

2.1 Pre-testing Setup Procedure

ESP-32-C3 side:

1. Install ESP-IDF and Python 3 using HomeBrew
2. Set up the tools used by ESP-IDF (compiler, debugger, Python packages, etc.)
3. Set up the environment variables
    a. Run in terminal: .$HOME/esp/esp-idf/export.sh

Xcode side:

1. Download the Xcode app

2.2 Testing Procedure

1. Test the OLED display by running the code onto the device and ensuring the text displays properly
2. Check the audio capture and processing capabilities on the iOS mobile application through Xcode on the iPhone simulator

2.3 Measurable Criteria

The criteria for successful running and output is as follows:

I. The app successfully takes in audio input and accurately displays the text onto the screen in real-time with minimal latency
II. The OLED displays text that is easily readable

3.0 iOS Mobile Application Testing Results

To test the accuracy of the iOS mobile application, we created a short paragraph of sixty-six words and read it out loud as the app was working. We then reviewed the transcribed text and compared it to the written text. We found how many words were incorrect to determine the accuracy of our algorithm.

Test paragraph:

"Testing a speech-to-text system requires a well-structured paragraph that includes a variety of sounds and word combinations. This test paragraph includes common and uncommon words, challenging the system's ability to accurately transcribe speech. It contains plosives, fricatives, and nasals, ensuring a comprehensive assessment of the system's capabilities. The paragraph should be read aloud at a natural pace, with clear enunciation. The time is 4:30 pm."

While performing the test, we kept in mind a few possibilities that might lead to error for example:

- Clarity and Accuracy
  - Question: Will the app be able to accurately transcribe clear speech?
  - Testing Procedure: One person read the test paragraph out loud in a clear, natural tone
  - Result: The app was able to accurately transcribe spoken words

- Background Noise Handling
  - Question: Is the app able to accurately transcribe speech in noisy environments?
  - Testing Procedure: One person read the paragraph while in a noisy lab environment with other people talking
  - Result: The system was able to filter out background noise and still accurately transcribed the speech

- Speech Variability
  - Question: Can the system handle different speech patterns?
  - Testing Procedure: We had multiple different people with different accents, speech speeds, and tones read the paragraph
  - Result: The system was accurately able to transcribe various speech patterns

Table of Results:

# of Correct Words	Percentage Accuracy
65	98%
62	93%
66	100%
61	92%
60	90%
64	96%
65	98%
63	95%
64	96%
65	98%
Result	96%

4.0 Discussion of test results

The average accuracy of our app according to our table of test results is 96%. This indicates high performance and proves that our app is able to accurately transcribe spoken words in different scenarios. Our lowest percent accuracy was 90% while our highest was 100%. This demonstrates the consistency that our app has as it is a small range of recorded percentages. While the results are predominantly positive, there is still room for improvement as there are slight variations in the accuracy. Overall, the system gives high accuracy results which seems ready for real-world deployment and further testing, however continuous monitoring and user feedback will be essential for improvement.

5.0 OLED Testing Results

We initially had the OLED fully running successfully with the ESP32-C3, where it was able to display text onto the screen, however during in-lab testing we received an error saying “OLED configuration failed”. After further examination of the device, we realized the OLED was cracked and therefore needed to be replaced. Once this issue was resolved and a new replacement OLED was obtained, we were able to successfully run the code onto the device and display written text.

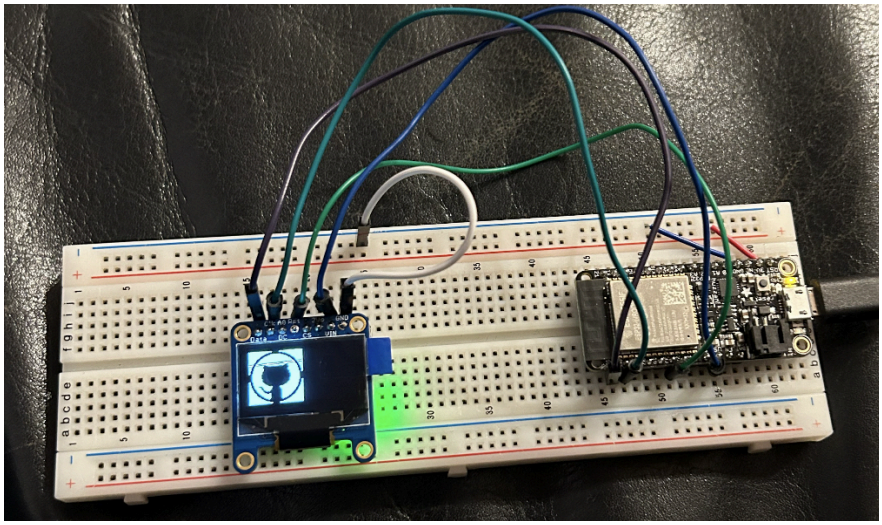

Figure 2: OLED display running with ESP32

6.0 Conclusion

Overall, the first prototype testing of our EchoView.AI project has shown immense progress and promising results. In our iOS mobile application testing, we achieved an average accuracy rate of 96%. This high level of performance shows the reliability and success of our speech-to-text algorithm among various environments and speech patterns. The challenges we faced with the OLED display, specifically with the cracked screen, were valuable learning experiences as they highlighted the importance of durability in our hardware choices, especially for a wearable device intended for daily use. As we move forward, our focus will be on refining the system based on the insights gained from this initial testing phase.