EchoSign

Team D7: Ria Balli, Ricky Gao, and Somya Pathak
18-500 Capstone Design, Spring 2024
Electrical and Computer Engineering Department
Carnegie Mellon University

Product Pitch

ASL is the default mode of communication for many deaf speakers across the United States. However, most non-deaf speakers still do not have a working knowledge of the language. This severely limits the abilities for these groups to communicate with each other.

Our project proposes a glove sensor-based design that allows for quick translation of ASL words to audible English output. The hardware is composed of a two-glove system fitted with flex sensors and IMU for data collection, and an ARDUINO for Bluetooth data transmission. The software composes of an ML model for classification. This design has allowed us to achieve the following metrics: **Accuracy: 89%, Latency: 0.65s, Vocabulary: 11 signs**

System Architecture

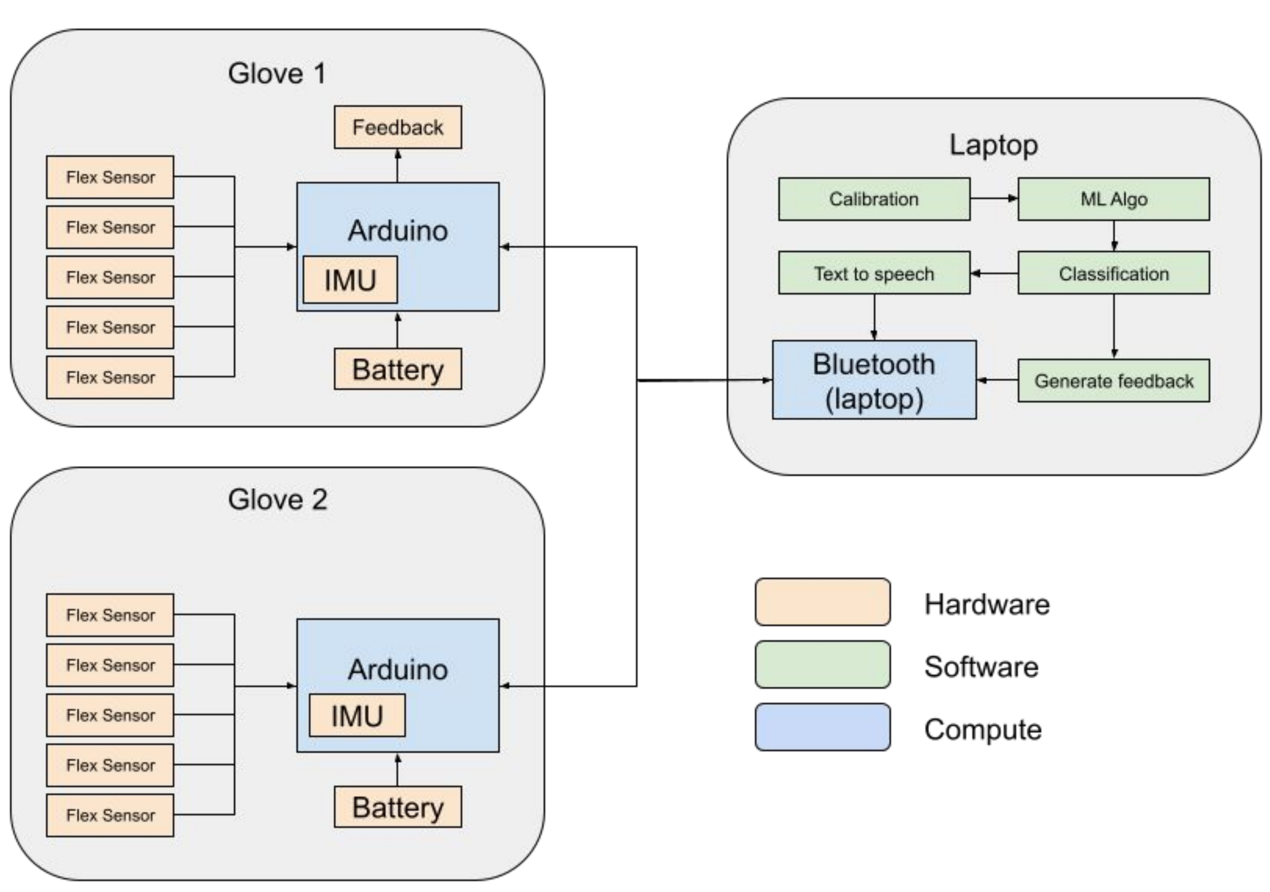

Figure 1: Hardware Architecture

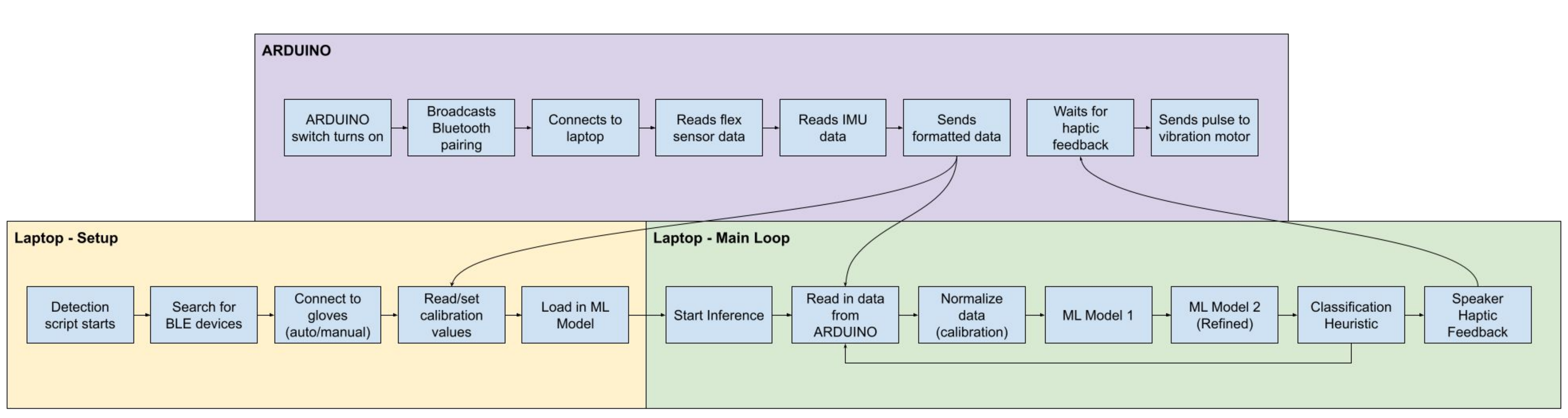

Figure 2: Software Architecture

Conclusions & Additional Information

http://www.eceS24-teamxx.cmu.edu/~ece500/projects/S24-teamd7

We are proud of the product that we eventually developed. We were able to create a system that can be wireless and successfully detect a variety of ASL words. We did have to limit the size of our vocabulary due to sensor limitations and our latency was a bit higher than our requirements.

However, we are confident that our project can serve as a proof-of-concept for future work into using double-handed glove designs to help the deaf community. We also believe that there are many aspects to continue innovation on. These include the inclusion of touch sensors, Wi-Fi capabilities for faster transmission, RNN architectures for error correction, and eliminating the need for a laptop.

System Description

Our main hardware system primarily leverages the use of flex sensors that detect the bending of the fingers through variable resistance with the help of an op amp gain circuit. In addition, we used the ARDUINO Nano 33 BLE Sense Rev2 which has an onboard IMU that assists us in identifying hand orientation. This compact board also allows for wireless capabilities and sends the data to our laptop through that channel. On the laptop we handle classification using a neural network. The neural network features a 4-layer, fully-connected architecture with 128 nodes per hidden layer, using a sigmoid activation function. Data to train the NN was collected from the three members of the group and subsequently processed. We also leverage the laptop's speaker to voice the predicted word as well as a vibrating motor on the glove to indicate the successful prediction

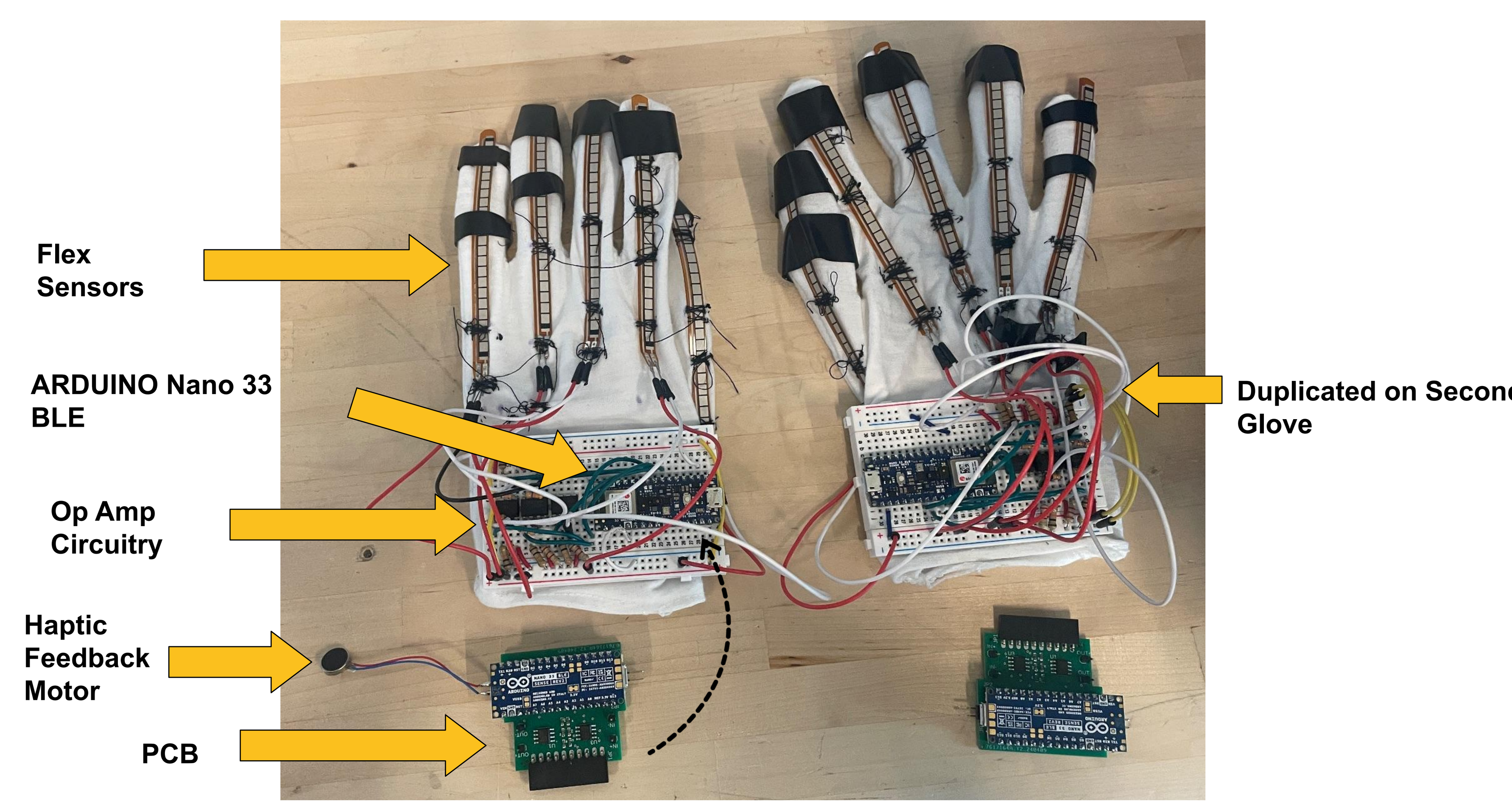

Figure 3: Glove Design

System Evaluation

First, we made sure to benchmark the performance of the ML model based on a separate test set. We also performed additional real-time tests to ensure performance. The figures below shows the real-time and test-time test results. The real-time tests were done by having a person sign each word in a random order for at least 12 times per word. We were also able to benchmark latency performance by measuring the time it takes from the initial signing to the speaker output. This was determined to be around **0.6-0.7 seconds**.

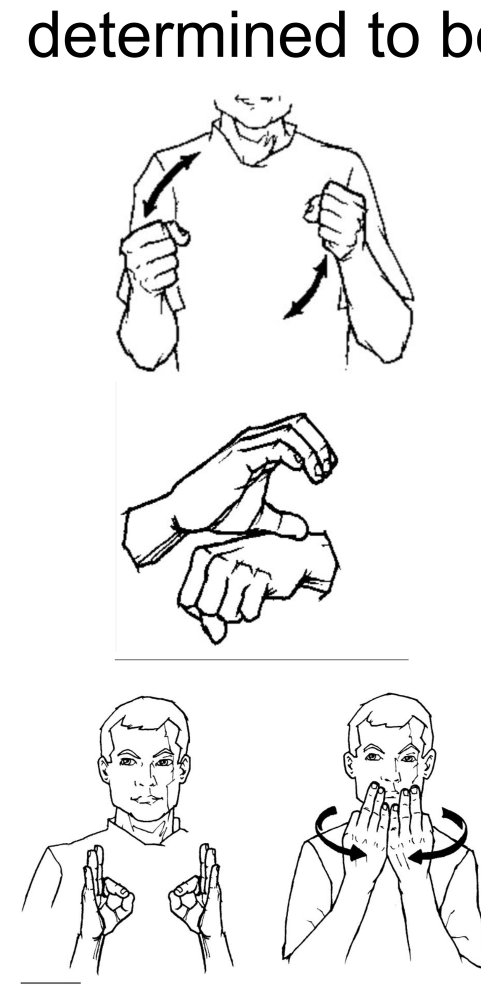

Figure 4: Example Vocabulary (Top-Down: Car, Church, Family)

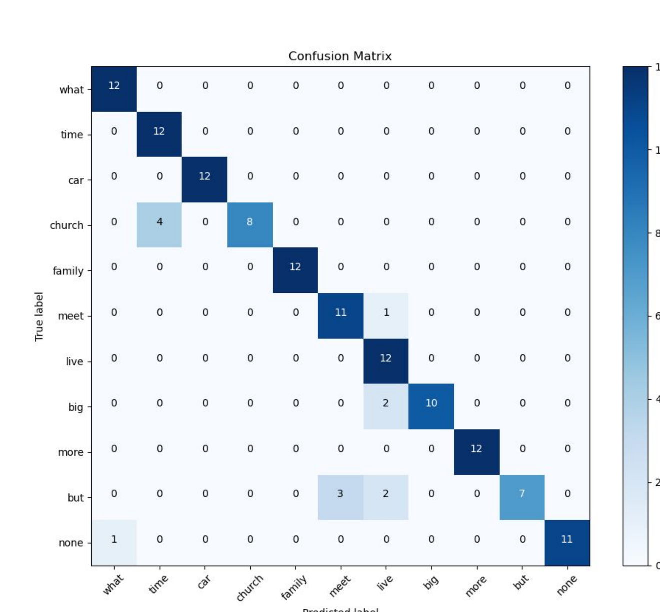

Figure 5: Confusion Matrix (Real-Time)

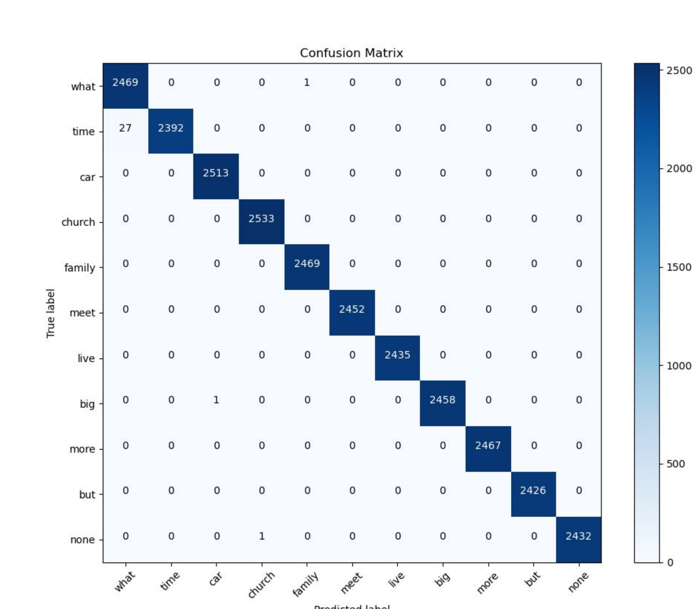

Figure 6: Confusion Matrix (Test-Time)

Electrical & Computer ENGINEERING
Carnegie Mellon