Project Summary

Overview: The outcome of this project is to use a gesture recognition device to convert American Sign Language (ASL) to a text sentence. This project is meant to help deaf and hearing disabled individuals convert signs, in ASL, to sentences that can be read by anyone. The goal of this project is to provide a device that can be trained to learn new gestures and also recognize what gesture is being signed by the user. In order to achieve this goal, a gesture recognition device, which in this case is the Myo Armband, will be manipulated to take the electromyographic data from the device and use machine learning techniques to be able to train and create a new gesture. After this data has been gathered and analyzed, a program will take this data and be able to recognize which new gesture is being performed.

Keywords: gesture recognition, natural language processing, human-computer interaction, biometrics, sign language

Subtopic Name: Electronic Hardware, Robotics, and Wireless Technologies (EW)

Intellectual Merit:

This Small Business Innovation Research Phase I project demonstrates how gesture recognition devices can be used to gather and analyze gyroscope, spatial, accelerometer, and electromyography data to train the device to recognize new gestures. Some technical hurdles involved in this project will include the details in the data. The project will gather a certain number of samples for each gesture to recognize. The device listener captures data for every millisecond and if every gesture takes about 10 seconds to perform, that's already 10000 lines of information. As the number of new gestures increases, the file and data sizes rapidly increase, posing the issue of how to simplify the data. Another technical difficulty to account for is the confidence in the program's ability to identify gestures. This requires finding a confidence interval for determining a gesture off of the dataset or changing the sensitivity of the armband so that the data points are more precise. The goal of the program is to be able to collect data from the armband, analyze and train this data so that the armband will recognize it, construct a sentence off of the gestures signed by the user, and convert the sentence from American Sign Language syntax to that of English syntax.

Broader/Commercial Impact:

Pending the commercialization of this project, it is likely that American Sign Language translators can be utilized in schools, workplaces, and even homes. The goal of the program is to allow the deaf and hearing disabled to be able to communicate freely with the hearing. Like any other language, American Sign Language should be just as easy to translate as it is for people to translate spoken languages on their phones. The fact that American Sign Language is not a spoken language shouldn't prevent the portability and ease at which American Sign Languages translators should be. Instead, it is an opportunity to look at alternative methods to translation. Attempts at American Sign Language translators utilize cameras or a Kinect, providing neither portability or ease for the user. Using a gesture recognition device such as the Myo armband allows individuals to imagine a different realm of translation, specifically for unspoken languages.

Elevator Pitch:

In the United States alone, it is estimated that there are about 11 million people who are deaf or hearing disabled. Commonly, individuals who are deaf or have hearing disabilities turn to American Sign Language, or ASL, as their main form of communication. As a result of this, American Sign Language has become a widely used method of communication among both the deaf and hearing population alike. Like any other language, there is a need for the ability to be able to translate American Sign Language in order to create communication amongst individuals who may not be able to understand American Sign Language. If American Sign Language were to be compared to other spoken languages in the United States, it would be placed behind Spanish and Chinese. The target market for this device applies to individuals who are deaf or have hearing disabilities who use American Sign Language as their primary way of communication. Additionally, those without these disabilities but still know American Sign Language can find this device to be beneficial. But what sets this device apart from other devices on the market? The largest difference is the ability to easily take the device anywhere. With an armband gesture recognition device, a user no longer has to calibrate a Kinect or some other sort of camera. Additionally, this project aims to allow the user to sign in traditional American Sign Language syntax of "Topic + Comment" and have this returned in traditional English syntax that could be read by anyone who speaks English. These small differences from other American Sign Language translator provide ease of use and portability, setting it apart from competitors.

The Commercial Opportunity:

Currently, the market for American Sign Language (ASL) converters is vastly unexplored. With only two or three ASL converters out there, there is plenty of room to make improvements and develop. Additionally, the need for ASL converters is rising as the number of deaf and hearing impaired individuals starting to use sign language is rising. This product is aimed to be used by the deaf, hearing impaired, and any individual that may use ASL in their day-to-day life.

The two other leading competitor converters provide similar functionality as this project but have setbacks that this project will address and correct. The first major drawback of competitor converters is their lack of mobility. One converter uses an Xbox Kinect to recognize gestures. This requires having a Kinect and some sort of computer or Xbox, meaning very low mobility. Additionally, Kinect's require frequent recalibrations whenever the device is moved or a user is switched. This may serve well in a setting where a user is giving a presentation or video chatting, but offers very little accessibility on an everyday basis

Another important drawback is the quality and range of the camera that will interpret the user's gestures. One converter uses a phone or tablet camera to recognize gestures. Phone and tablet cameras are not always known for their high quality and large depth of field. A lack of a depth of field makes interpreting when a user is doing a sign towards or away from the camera more difficult. In addition to camera quality, if a user is doing larger signs that require being able to view below or above their face, the user must either extend their arm to an uncomfortable length or ask someone else to hold the camera for them. In reality, these are not impossible tasks to perform but ones that a user shouldn't have to face.

The intention of this project is to completely move away from cameras and Kinect's and focus on devices that are mobile and extremely efficient. This project will feature Myo armbands that a user will wear on their forearms. The Myo armband doesn't require being plugged into any computer or phone; the armband does all the work wirelessly. The armband features electromyographic (EMG) sensors that are placed on the inside of the band and read the electrical pulses from muscle contraction and tension to determine what gesture the user is signing. In addition to EMG sensors, the armband also factors in gyroscope, accelerometer, and magnetometer readings to provide a state of the art piece of technology. The Myo allows a user to slip on the armband wherever they are, increasing mobility and making it easy for users of all ages to use.

The other part of the project that sets it apart from its competitors is the translation function. This project differs from others in that it will create sentences based on the syntax that ASL uses. ASL typically follows a subject - predicate sentence structure that can make it difficult to read for someone who may not know this particular syntax; most converters print out in this syntax. Other converters require that a user sign a sentence in typical English syntax, meaning that they have to sign more words and letters than they typically do. This issue is on the smaller scale when it comes to accessibility, but this project provides a solution that can make interacting with a deaf of hearing impaired individual easier than it currently is.

Because the market for ASL converters is rather small, there is little competition and low barriers to entry. This project is meant to be an aid and tool that can be used in schools, professional settings, and even for a day-to-day use. In order to bring the concept to the marketplace, the product should begin in a research and development laboratory. Then the product should be tested in the different settings mentioned above. To make the product beneficial to the user's needs, the product should be given to schools and small businesses for testing. Finally, the product should be launched once it has done successfully in testing and in the laboratories.

The overall purpose of this project is to convert ASL to fully formatted sentences, but there is potential to use different components of this project for different purposes. For example, one of the biggest parts of the project is teaching the Myo armband how to recognize and learn new gestures outside of the five "out of the box" gestures it comes with. This could be adapted to learn different sign languages like French Sign Language, Spanish Sign Language, and more. Additionally, the Myo armband was intended to be able to control devices with gestures. Having the ability to teach the armband new gestures to different commands is a small but helpful ability.

The main function of the text conversion is to take keywords and turn them into sentences. An example of this is converting, "I teacher" to "I am a teacher." There are plenty of algorithms and programs to pull keywords from full sentences, but very few that can make full sentences from keywords. This could be used in predictive typing and for search engine optimization. Regardless of your reason for using this project, there is potential for both its intended use and for those that aren't in its initial target market.

Broader Impact:

Many people are unaware of the tangible barrier that separates the hearing and the deaf. There are about 360 million people in the world that are deaf or have a hearing impairment. About seven million people of this large community reside in the United States alone. Of the 360 million individuals who are part of the deaf community, about 70 million of these people use American Sign Language as their primary language. American Sign Language, or ASL, is widely used by the deaf community and considered a commonly used foreign language. Like any other language, it would be helpful to have converters so that those who do not know American Sign Language can be able to communicate with those who use it. Translators are available to act as a real-time converter but that requires having a translator on hand and takes away from the intimacy of a one-on-one conversation. Other alternatives include using pen and paper to communicate between two people, making it both painstaking and inefficient. My project aims to move away from translators and transcribing and provide an easier solution to bridge the American Sign Language and hearing community.

As describe by the World Health Organization, hearing loss can have a large impact on both an individual's socialization and ability to function. Many lose their hearing with old age, causing frustration and loneliness at their inability to do what they once could do. In individuals who loses their hearing at a young age, they can find it difficult to interact with their younger peers, as many of them do not know sign language yet. As a result, this difficulty can cause exclusion and lead to many developmental problems down the road. This project will help eliminate some of these problems by allowing a portable and easy to use user interface for users of all ages. The project is intended to be easy to use, meaning a young child should be able to learn how to use the interface just as easily as an elderly adult can. Additionally, someone using this interface will be able to communicate with another individual who does not know sign language. Removing this barrier allows the closing of that bridge between the hearing and the deaf and increased socialization.

Hearing loss is not just a physical disability; instead, it is a handicap in psychological, emotional, and social abilities. Those with hearing loss sometimes lose the ability to function on their own. Some hearing impaired individuals are constrained to either using a translator or just simply being unable to easily communicate with a person who does not understand sign language. At times, people who do not know sign language can be discouraged from interacting with the hearing impaired because of their inability to communicate with them. This disheartenment may translate into academic or professional settings and limit a person's performance and opportunity. As stated earlier, providing a convenient translator that can be worn anywhere and by anyone will break down all previous barriers that deaf and hearing impaired may face. A user can wear the gesture recognition device and hook it up to a phone that can then translate their signs to a full text sentence, all without the use of a translator or pen and paper. A small change like this has the potential to make a great impact. Having the ability to communicate with ease can alleviate the discouragement a person may have once felt or allow an individual to regain the independence they may have once had.

The societal impact for the deaf and hearing impaired population would be substantial and one that is much needed. Other American Sign Language converters don’t provide the ease and mobility that deaf and hearing impaired signers need. the amount of technology that we have nowadays, there is no reason that we should still have to accept that there are barriers between those who can and cannot hear.

Technical Discussion

The project starts off with a standard Myo Armband that recognizes five pre-set gestures. The armband is worn on the forearm and has electromyographic sensors in addition to a gyroscope, accelerometer, and magnetometer that wirelessly connect to a computer. The armband comes with 5 built in gestures, but the goal of this project is to be able to train the armband to recognize new gestures. The armband comes with a standard device listener and a Software Developer Kit (SDK). One of the objectives of this project is to be able to create a new device listener that is adapted to include the new gestures.

Using the SDK, we can capture the electromyographic, accelerometer, gyroscope, and magnetometer data and do some preprocessing in order to be able to distinguish different data sets for each gesture from one another. One technical challenge that this project will face is in deciding how to preprocess the data. The device listener captures data for every millisecond and if every gesture takes about 10 seconds to perform, that's already 10000 lines of information. As the number of new gestures increases, the file and data set sizes rapidly increase, posing the issue of how to simplify the data to the point where it's not too condensed but also not too large.

Once the data has been reduced, the issue of deciding the confidence intervals for the recognized gesture has to be chosen so that if two gestures are similar, how can the program be able to determine whether the user is doing one gesture over another. While most gestures in American Sign Language are very different in movement, gestures such as the alphabet are harder to differentiate. One solution to this problem is to adjust the sensitivity of the armband. If the armband becomes more sensitive to movement, then there will be more precise data to work off of.

Once the data has been captured, analyzed, and preprocessed, the next objective of the program is to be able to train the device listener the new gesture. Using Nick Gillian's Gesture Recognition Toolkit (GRT), the program can load the datasets and uses a custom data structure to apply the Dynamic Time Warping (DTW) algorithm. In this case, the DTW algorithm will be able to compare two gestures despite the amount of time it took to complete them. One of the technical hurdles in capturing the data for each individual gesture is the difference in time. If a user is trying to teach a gesture, it is unlikely that the user will be able to achieve the same gesture in the exact same time frame every single time. Using the DTW algorithm helps create a match between two gestures by converting them to linear sequences.

After capturing the data and providing all training, the program will then have to go through the process of predicting the correct gesture. Nick Gillian's GRT provides a function that will use all the analysis from training the datasets to accurately predict the gesture. Machine learning will be implemented in order to maximize the accuracy of the program.

Now the program has accurately predicted a gesture. The next step in the program is to convey this word onto the screen. Each gesture has a name associated with it, which will then be associated to the word that the user is trying to convey. As mentioned in previous parts of this proposal, American Sign Language follows a syntax that varies from English syntax. Typically, American Sign language is signed as "Topic" and then "Comment", where topic is the subject and comment is the action or description. The gestures that the user signs will be passed in the order that they are signed. From here, the program will differentiate the topic from the comment. The program then uses the Natural Language Toolkit to break the sentence down into a parsing tree. The Natural Language Toolkit has a function that can find the parts of speech of each word which will be useful when rearranging the words. A major technical difficulty here is converting from the American Sign Language syntax to English syntax while keeping the same meaning of the sentence. First the program will reorder the sentence so that it follows the traditional English syntax of "Subject" then "Predicate". Then the program will fill in the different parts of speech that are essential to making the sentence comprehensible. But what were to happen if the program puts in a word that makes the sentence unintelligible? In order to correct this problem, there is a database compiled of example sentences that the result sentence will be compared to. The addition of machine learning will also be added to ensure that if a sentence is done incorrectly once, then it will not repeat the mistake again.

R&D Plan

Research:

The first step is to complete the research associated with the Myo armband and find a data set of example sentences. This includes researching the API associated with the Myo armband. Additionally, it will be useful to understand how the Myo device listener interacts with the armband and a program.

*Projected Completion: 2 weeks*

Since the project will also use two different toolkits, Nick Gillian's Gesture Recognition Toolkit (GRT) and the Natural Language Toolkit, it is important to research how to incorporate these libraries into the Myo library. The GRT and the Myo libraries are in C++ and should be easy to implement with one another. The Natural Language Toolkit is in python and will need to interact with both a database and the program.

*Projected Completion: 1 week*

Development:

Once everything will be researched, the development phase will be broken down into two modules. The first part will be the Device Module and the Sentence Module.

Device Module: This part of the development phase will be broken down into three submodules:

1. The Data Module: This module will use the Myo armband to record data. The user will input a name for the gesture and the number of times to be recorded. This will then be put into a .rawmyo format so that the Myo will be able to read from it

*Projected Completion: 1-2 weeks*

2. The Training Module: This module will use the data from the first module and the GRT library to perform preprocessing and algorithms so that it can be passed into the next module. In this module there are a few preprocessing decisions that need to be made. One decision is to decide how to condense the data so that the data doesn't exceed 10000 lines of information. Another decision it to decide what are the confidence intervals for

*Projected Completion: 1-2 weeks*

3. The Prediction Module: This module will create a device listener that calls the data from the Training Module to allow the Myo to access the data for each gesture. The user will then be able to wear the Myo, perform a gesture, and have the interface correctly predict which gesture the user is doing.

*Project Completion: 1 week*

Sentence Module: This part of the development phase will be broken down into two submodules:

1. Armband - Word Module: This module will get a gesture name from the Myo armband which will then be associated with a word. This word will then be appended to a string that will return to be processed.

*Projected Completion: 1-2 days*

2. ASl - English Module: This module will get a sentence of words that were signed by the user. It will then use the Natural Language Toolkit to parse the sentence, perform parts of speech tagging, and place the sentence in the correct syntax. This module will implement machine learning so there will probably be some training involved in here.

*Project Completion: 2-3 weeks*

*Total Completion: 10-12 weeks*