MAST:

Myo Armband Sign-Language Sign Translator

Project Proposal

(Capstone Design Project (SWE3028-41) Sungkyunkwan University)

- Team 1 (College of Computing – Software):

So Soonhyuk

Muhammad Shakeel Dahab

Go Seokhwan

- **Contents:**

1) Abstract

Over the last couple of years, there has been an increase in hearing-impaired people. According to a research done by J. Fellinger, D. Holzinger and R. Pollard around 15-26% of people around the world suffer from hearing disabilities, with about 7 per 10,000 of the general population become deaf before learning any language [1]. Mute individuals have a communication problem dealing with other people. It is hard for such individuals to express what they want to say since sign language is not comprehensible for everyone. Especially in urgent situations like visiting a hospital, a lack of correctness and speed of communication is crucial to a patient. Concerning this situation, this project aim is to develop a MYO Armband Sign-language Translator (MAST) system that can be used for the hospitalized mute patients. MAST uses a set of eight sensors to detect the muscle movement and along with our trained model we aim to generate efficient results.

2) Introduction:

Sign language translation is a popular and a growing field in computer science research. The sign language is a language that uses the hand movement to express words. It is estimated that the people who are using sign language exceed 80 million.

This section will provide a introduction to our project which includes trends in the field of this project, motivation behind choosing this topic, key idea of this project and the expected results.

To translate sign language, tracking a person's movement and determining what gesture a person is perfoming needs to be decided by a computer or a complex machine. Thus, there were several attempts to resolve this problem through abundant amount of research. Currently, there are four major tools or trends to implement the sign language translation [2,3,4].

A. Wired Gloves:

These can provide input to the computer about the hand/fingers' position and the rotation of the hands using magnetic or inertial tracking devices. The first commercially available hand-tracking glovetype device was the Data Glove, a glove-type device whichcould detect hand position, movement and finger bending. This uses fiber optic cables running down the back of the hand.Light pulses are created and when the fingers are bent, lightleaks through small cracks and the loss is registered, giving an approximation of the hand position. However, those are economically expensive.

B. Depth-Aware Cameras:

Using specialized cameras such as structured light or time of-flight cameras, one can generate a depth map of what is being seen through the camera at a short range, and use this data to approximate a 3d representation of what is being seen.

C. Stereo Camera:

Using two cameras whose relations to one another are known, a 3d representation can be approximated by the output of the cameras. To get the cameras' relations, one can use a positioning reference such as a lexia-stripe or infra-red emitters in combination with direct motion measurement (6D Vision) so that the gestures can directly be detected.

D. Single Camera:

A standard 2D camera can be used for gesture recognition where the resources/environment would not be convenient for other forms of image-based recognition. A Software-based gesture recognition technology using a standard 2D camera that can detect robust hand gestures, hand signs, as well as track hands or fingertips [5].

Currently, many hospitals changed their system to so-called 'Smart-Hospital'. But in case of the disabled, when they are hospitalized, sign-language translator must attend with them for communication. Our team thought that we can develope a tool or a machine that can effectively translate words between those disabled people and a non-sign language user. The Key idea of this project is making a full system (both hardware and software) for mute people to detect any emergency and to translate their issues in hospitals. After that we can train the model to specific gestures. We found several tools that we can use to implement our goal and chose the most appropriate tool for our project which is MYO arm band and an artificial 3d printed arm.

With this project, we are expecting to make a more-efficient way of communication between hospitized patient with communication problem. Moreover, we want to offer another tool which is myo armband and artificial arms for the sign language translation that might be a good replacement from current tools.

3. Related Work

This section is more detailed version of the “Project Survey”. In this part we will focus on describing the related work that has been done in this area in a more systematic way. It is really essential to mention that most of the research that has targeted this field is either dependent on Statistical Machine Translation (SMT) or Neural Network based Computer Vision. Thus, it is really important to describe those mechanisms before diving into describing their applications.

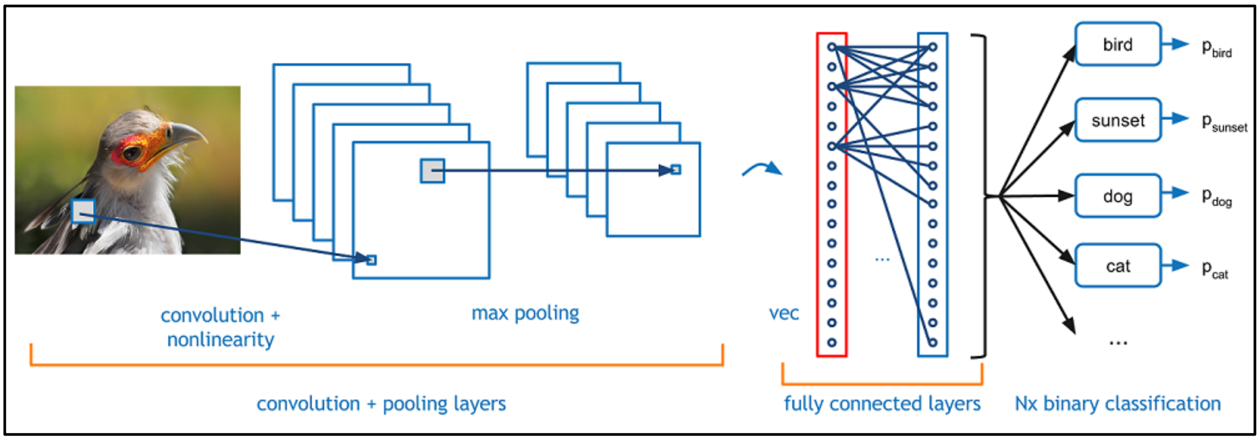
1) Statistical Machine Translation (SMT): It is regarded as a subfield of natural language processing which investigates how to automatically translate text or speech across human languages which uses probabilistic estimated from parallel corpora which needs large quantities of training data. In simple probabilities the formula can be describes as follows [3]:

$$\begin{aligned}\hat{T} &= arg\,max_T\; P(T|S)\\ &= arg\,max_T\; \frac{P(S|T)P(T)}{P(S)}\\ &= arg\,max_T\; P(S|T)\,P(T)\end{aligned}$$

Translation model
Encodes the faithfulness of T as a translation of S

Language model
Encodes the fluency of T in the target language

2) Computer Vision: It is a field of artificial intelligence that trains computers to interpret and understand the visual world. Using digital images from cameras and videos and deep learning models, machines can accurately identify and classify objects — and then react to what they "see." It is essential to mention that machine are usually trained with a Convolutional Neural Network described as follows [4]:



Although, at a first glance those solutions might seem to be accurate and efficient however we will discuss the details of why our suggested solution has an upper hand and how novel our approach is compared to the previously suggested approaches by drawing a stark comparison.

One of the most notable works that has been done in this area was published by the University of Guelph in an IEEE conference under the title of " Sign Language Translator and Gesture Recognition" in which they proposed a custom sign language translation system which uses a specialized glove programmed with an SMT algorithm to translate certain gestures into written text with an accuracy of 96% [5]. On the other hand, the practicality of such glove is debatable as it is only able to translate 20 gestures to 20 English letters rather than phrases or sentences which in fact is inconvenient. Another similar SMT based system was introduced in the paper "Statistical Sign Language Translation" which had the same limitations but with a slight improvement of targeting simple word in the German Language [6]. To compare those two papers to our approach, we are inferring the fact that those projects that focus on the use of sign language translator in normal, casual situations but aren't concerned with more real-life use cases such as targeting the medical field.

Another Interesting paper that was suggested in the previously mentioned conference by the University of Wisconsin; "Real-time Sign Language Translation based on Neural Network Architecture". In this paper, the researchers suggested a very similar approach but instead of using SMT and a glove they built their mechanism based on computer vision technology where the mute people have to do some gestures Infront of a camera so that it can be translated [7]. On the other hand, this approach has three main flaws:

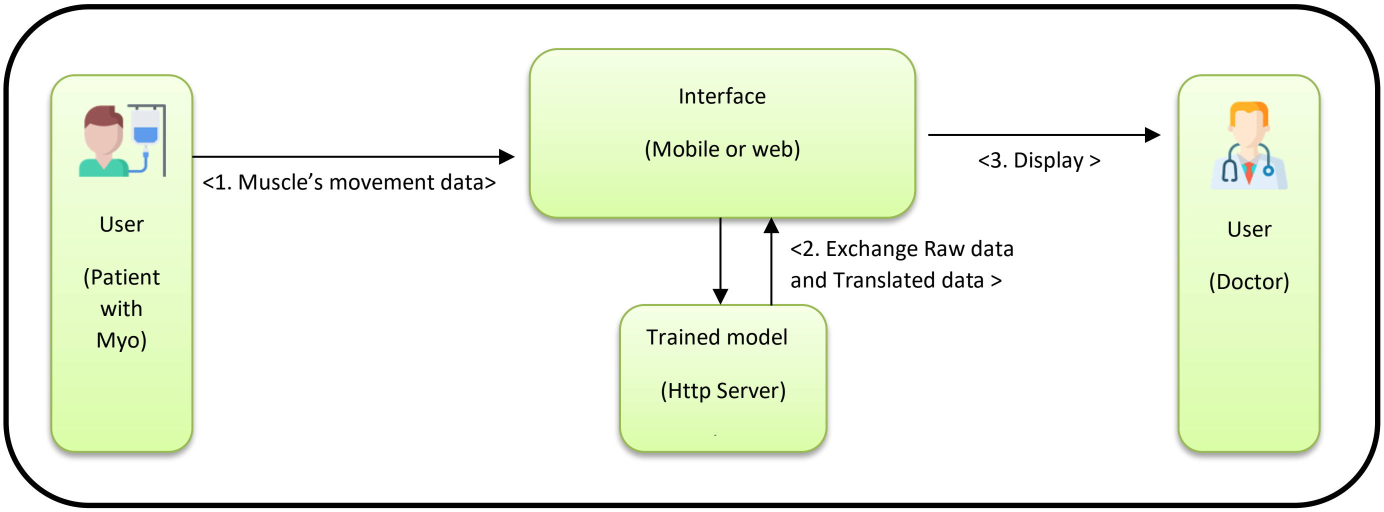
1) The limited ability of translating English letters rather than more useful sentences or phrases.
2) The high latency that is due to the workflow of the suggested architecture where pictures are first captured then stored and after that a CNN (Convolutional Neural Network) is applied.
3) The limitation due to the requirement of having a very clear picture under perfect lighting condition.

As far for our pros as compared to the previously paper, since we are using Myo band and artificial hands, our project will have a higher accuracy than other computer vision-based projects. Moreover, the machine learning model that we will use is extremely easy to train as it doesn’t involve complex CNN techniques used in the afore-mentioned method although we will still need large sets of data to generate more accurate results. Currently, research works have focused mainly on the recognition signs from images or video sequences that have been recorded under controlled conditions. Those conditions include resolution, light and shadows. But with MAST, we do not take image or video as an input thus our results are independent of such factors.

That being said, newer papers have suggested better approaches using advanced CNNs. On the other side of the spectrum, the usability of such techniques is questionable as there are more than 240 different sign languages in the world and all the suggested papers focus on one of these sign languages. So even if the approach and accuracy is 100% the usability case of these mechanisms is very limited. For example, in the journal of computational and theoretical nanoscience, a paper was suggested to translate simple phrases from Portuguese sign language to English written text [8]. While another paper under the title “Vision-based sign language translation device” generated similar results but for the Indian sign language [9]. From a practical viewpoint, such implementations have very limited applications in real world as it is only helpful for people who are familiar with the Portuguese sign language. In our approach, instead of focusing on a single limited language we are trying to build a more comprehensive general translation system which focuses on the medical field and the smart hospitals equipped with MYO bands programmed with pre-trained custom gestures that will help any mute disabled person to easily use the gestures’ manual to convey their issue to the doctors without the need of a third part interpreter.

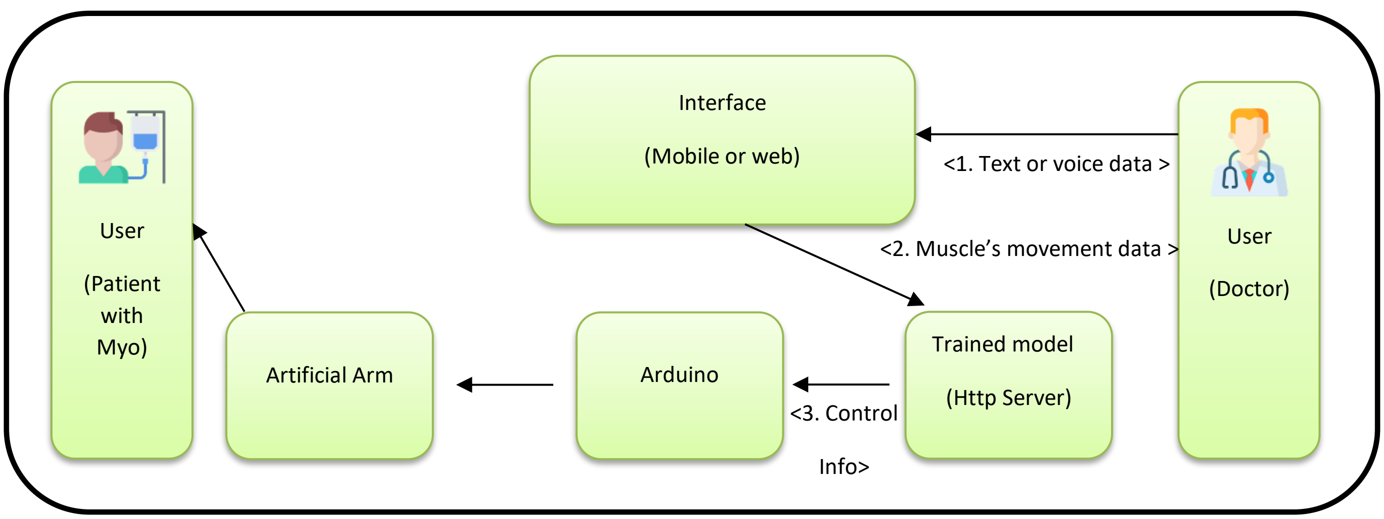
4. Problem Formulation

This section is focused on what this project's objective and constraints are. Our goal is the implementation of sign language translator in the medical field. To be more specific, building model which is trained and a mobile or web environment that interacts between model and users. Below diagram shows how system works and its flow:



<Architecture diagram (from patient to doctor)>

As we shown in above diagram, Myo band attached to a patient sends raw myograph data to Interface. And the Interface exchanges raw data with translated data through REST API Server. Doctor will see an understandable language displayed in interface.



<Architecture diagram (from doctor to patient)>

In case of a doctor, they can send a text or voice messages to the interface. Next, trained model request Arduino to control artificial arm according to muscle data which is received. Arduino will control motors to display sign language which go well with the input data from doctor.

5. Project Milestone

Phase	Expected Deadline
Conceptualization &Survey & Proposal	4/6
Project Proposal Presentation & User Requirements Documentation	4/13
Environmental Setting & Design 1	4/20
Design 2	4/27
Development (Frontend and Backend)	5/31
System Integration & Deployment &Testing	6/9

6. Reference List

[1] J. Fellinger, D. Holzinger, and R. Pollard, “Mental health of deaf people”, Lancet, vol. 379, no. 9820,pp. 1037–1044, 2012.

[2] Oliveira, T. Escudeiro, P. Escudeiro, N. Rocha, and E.Barbosa. “Automatic Sign Language Translation to Improve Communication”, 2019.

[3] P. Lison, “Introduction to Statistical Machine Translation,” *Maskinlæring meetup at Language Technology Group of University of Oslo*, May 2016.

[4] “Parallel architectures and computer vision,” *Computer Vision, Graphics, and Image Processing*, vol. 46, no. 1, p. 136, 1989.

[5] M. Elmahgiubi, M. Ennajar, N. Drawil, and M. S. Elbuni, “Sign language translator and gesture recognition,” *2015 Global Summit on Computer & Information Technology (GSCIT)*, 2015.

[6] J. Bungeroth and H. Ney, “Statistical Sign Language Translation,” *Lehrstuhl fur Informatik VI*.

[7] P. Mekala, Y. Gao, J. Fan, and A. Davari, “Real-time sign language recognition based on neural network architecture,” *2011 IEEE 43rd Southeastern Symposium on System Theory*, 2011.

[8] P. Escudeiro, N. Escudeiro, R. Reis, M. Barbosa, J. Bidarra, and B. Gouveia, “Automatic Sign Language Translator Model,” *Advanced Science Letters*, vol. 20, no. 2, pp. 531–533, Jan. 2014.

[9] Y. Madhuri, G. Anitha, and M. Anburajan, “Vision-based sign language translation

device," *2013 International Conference on Information Communication and Embedded Systems (ICICES)*, 2013.