

ABSTRACT

Communication plays a significant role in making the world a better place. Communication creates bonding and relations among the people, whether persona, social, or political views. Most people communicate efficiently without any issues, but many cannot due to disability. They cannot hear or speak, which makes Earth a problematic place to live for them. Even simple basic tasks become difficult for them. Disability is an emotive human condition. It limits the individual to a certain level of performance. Being deaf and dumb pushes the subject to oblivion, highly introverted. In a world of inequality, this society needs empowerment. Harnessing technology to improve their welfare is necessary. In a tech era, no one should be limited due to his or her inability. The application of technology should create a platform or a world of equality despite the natural state of humans. On the other hand, technology is the most innovative thing on Earth for every time the clock ticks, researchers, software engineers, programmers, and information technology specialists are always coming up with bright ideas to provide convenience to everyone. This paper shows how artificial intelligence is being used to help people who are unable to do what most people do in their everyday lives. Aligned with communication, D-talk is a system that allows people who are unable to talk and hear be fully understood and for them to learn their language easier and also for the people that would interact and communicate with them. This system provides detailed hand gestures that show the interpretation at the bottom so that everyone can understand them. This research allows the readers to learn the system and what it can do to people who are struggling with what they are not capable of and will provide the technical terms on how the system works.

Key words: Machine learning, disability application, sign language, image processing, neural networks, artificial intelligence.

INTRODUCTION

Communication should be universal without any barriers or limitations. This paper establishes a method for providing equality, turning the disabilities of the hearing and, or speech impaired individuals to abilities, creating a base where both the disabled and the able can communicate without any barrier. Our objective is to blend deaf and dumb within society and make them able to use their personal computers more effectively and efficiently. Our idea is to create sign assistance, like many applications which is using voice assistance such as Siri on iOS and Cortana on windows. There is need to develop an application that will create an interactive platform where the sign language can be translated to voice output and writing, and voice and writing input can also be converted to sign language. The bigger picture is creating an interactive model of communication for deaf and dumb people. Developing an app will support this vulnerable society of impaired people and enhance communication among people. The application will allow ease in communication, improving their interaction, and hence better life. This project will be a noble cause and translating the sign language into understandable words is the goal. Microsoft Windows will come in handy to enhance the actualization of this application.

According to the World Health Organization, the world population experiencing hearing and speech challenges approximates over 466 million people globally. With such disability, instead unequally distributed resources, these people are vulnerable to discrimination. The fact that every human being, abled or disabled, is entitled to a good life with equal opportunities calls for affirmative action. This society requires attention from all quarters, especially on technological enhancement, to ensure the disabled get a comfortable life.

1.1 Research Question

The process of this application can be daunting, but the value is priceless. Being able to create something to serve people in need is uncountable. The focus of this research is to answer questions related to sign recognition. Therefore, we narrow the research question to align with this application that will be developed by machine learning and artificial intelligence to make the right decisions.

Several questions considered as research questions, which are:

- What will happen when the application recognizes the image?
- What is the process of recognition of the image by the application?
- How the image recognized by the application?

2. THEORETICAL BACKGROUND

This section provides a theoretical background for D-talk to have a better understanding of the process to be used in the application. This application can provide a helpful tool for communication between the deaf and the external world. The sign language recognition program, which is required to understand sign languages, has been studied extensively for years. The studies are focused on various input sensors, gesture segmentation, feature extraction, and classification methods. This paper is aimed at evaluating and comparing the methods used in the sign recognition systems, classification methods used and identifies the most promising approach for this project. Despite recent advances in classification methods, many of the recent works proposed to apply primarily to classification methods, such as deep learning. This paper focuses on the classification methods used in the prior recognition scheme for sign Recognition.

2.1 Machine Learning

It is essential to choose the right strategy; Machine-learning techniques are often used to do this. Machine learning is part of artificial intelligence (AI). It can be defined as an algorithm that focuses on computer program development. Machine learning considered an application that use to increase computer ability to learn from previous experiences.

There are three types of ML:

- Supervised: The application has previous experience of new data using labeled examples to predict future events.
- Semi-supervised: Data can be labeled and unlabeled for training.
- Unsupervised: Data that unlabeled from experience. The system cannot figure out the right input so that the result will be wrong.

2.2 Deep Learning

Compared to standard algorithms, neural networks can solve somewhat complicated issues at a much easier level about the complexity of algorithms. Neural networks can solve somewhat complicated issues at a much easier level concerning the complexity of algorithms. The neural network builds to mimic human brain neural function but with the mathematical functions. One of the neural networks is the multi-layer network as shown in figure. It includes three layers, the input layer, many hidden layers, and the output layer. The input layer passes data without modification. Hidden layers process the data, and the output layer converts hidden layers to output as a classification. Collecting datasets for training takes time to process. As the number of configuration increases, training samples increase. Most data in the world not uniformly distributed.

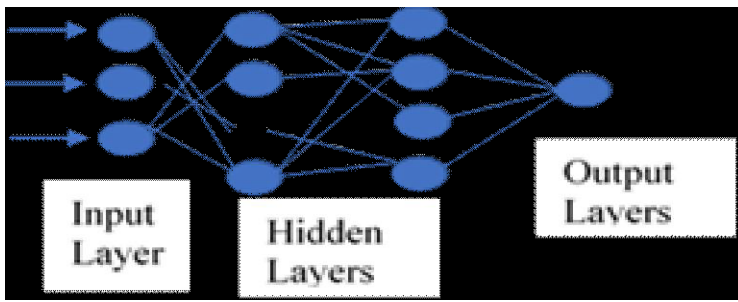


Figure 1: Multilayer Network.

3. METHODOLOGY

This paper presents an understanding of complex hand movements. A framework based on Hidden Markov Models (HMMs) is provided for modeling and recognition of complex gesture trajectories. A detection algorithm is used to detect the hand of the user, and a contour-based hand tracker is developed, which combines condensation and partitioned sampling. The proposed approach can attain automated online identification of hand movements and can effectively reject atypical movements. The experimental results indicate that the proposed algorithm can produce better results for recognition than the conventional method of hand recognition. The hand gesture recognition system consists of three major parts: palm detection, hand tracking, and trajectory recognition. Figure 2 provides an overview of the hand gesture recognition process. The hand tracking function is enabled when the device senses an open hand in front of the camera; when the user finishes the gesture, the hand gesture classification based on HMM is disabled. The basic algorithmic structure for our process of recognition is the following: -

- Detect the palm from the video and initialize the tracker with the template of hand shape.
- Track the hand motion using a contour-based tracker and record the trajectory of the palm center.
- Classify the gesture using HMM, which gives the maximum probability of occurrence of observation sequence.

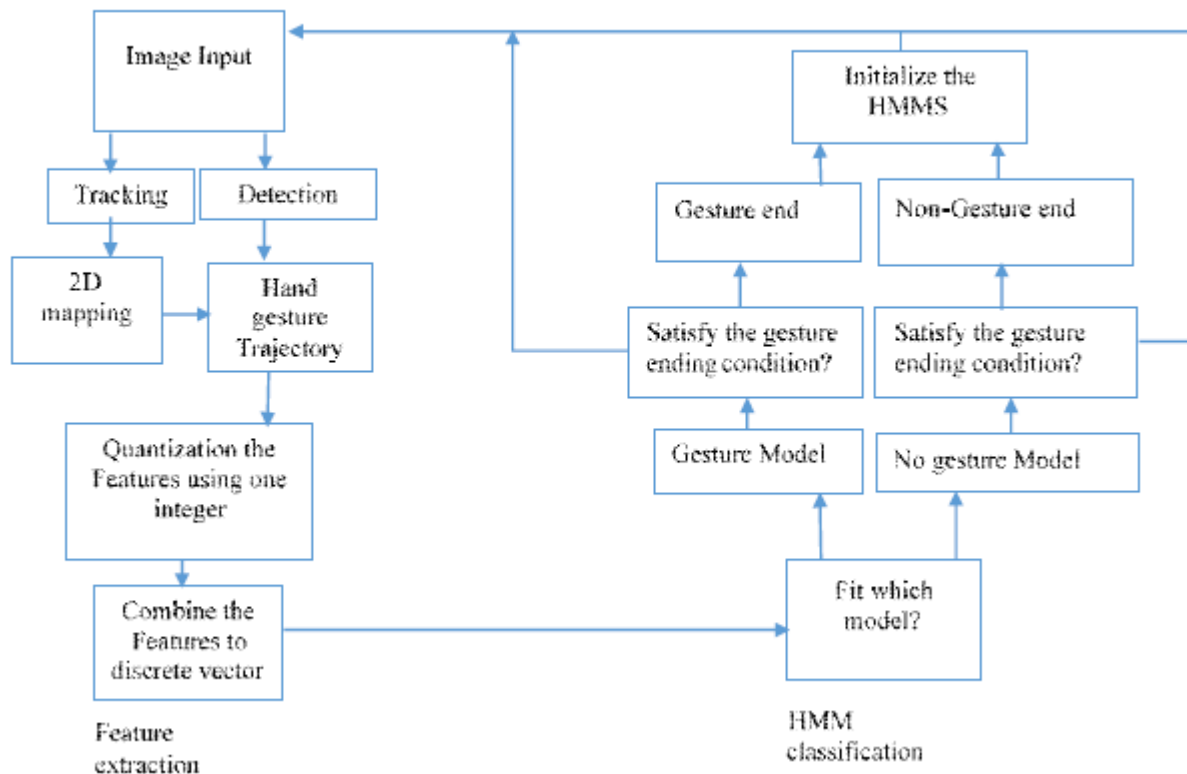


Figure 2: Overview of Hand Gesture Recognition Process.

3.1 Gesture Recognition

We explored one way to identify simple hand gestures and implement two basic gesture controls: movement of the cursor and mouse click. The figure 3 describes the basic process of hand gesture recognition.

- By using vision-based recognition, the computer captures the sign to find the gesture acquisition.
- Hand tracking can be done by using clustering algorithms that able to treat each finger as a cluster and delete the empty spaces between them or multi-scale color feature hierarchies that provide users' hand and the different background shades of colors to identify and remove the background. Hand tracking is the computer's ability to track the user's hand and split it from the background or any other objects.

- Feature extraction depends on the application. On D-talk, finger status, skin color, alignments of the finger, and the palm position are taken into consideration.
- After features extracted, they sent to training and testing classification algorithms to reach the output.

We explored one way to identify simple hand gestures and implement two basic gesture controls: movement of the cursor and

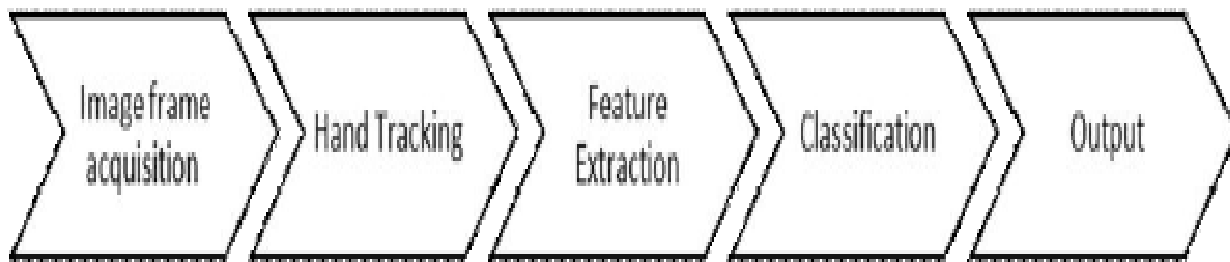


Figure 3: Basic steps of hand recognition process.

4. IMPLEMENTATION AND RESULTS

In the implementation phase, developers change several tasks that they were planned to do. They notice that they can build the system without preparing any training and testing images as they were plan. The code is depending on skin color and contour to find the right sign. Moreover, developers narrow the tasks to only one task which is browse websites only. Moreover, the result was precise and accurate aligned with the methodology and testing that was used. This signifies that developing modern technology assists disabled individuals specifically deaf-dumb on interacting among people. The measurement variables along with the supporting evidence from the methodology concluded that the measures taken to evaluate this study were supported all throughout. Meanwhile, the efficiency and effectiveness of the system provide the utmost benefit of disabled individuals by offering convenience and being able to make their lives easier and better for there are no required training or specificities for

them to use the system. Thus, as a result, D-talk allows everyone to determine the hand gestures that are being projected and be able to come up with interpretations on enabled individuals. Hence, communications between deaf-dumb and enabled individuals are way easier and lacks misunderstandings are being prevented this time. This application can catch finger shapes by using the code for Extract skin color and draw lines around the hand. As a result, the system will recognize any element in the frame. The application main screen is shown in figure 4. Thus, users must be careful about what is inside the frame to avoid any other unwanted requests. This system will recognize any element in the box, and the brightness does not matter. D-talk is a dynamic system that includes three gestures in total to browse websites. All that users need to implement this system is WiFi connection and webcam to capture user gestures. The following figures 5 and 6 show the hand gestures that are used for orders that the system can recognize to browse websites. It could be used as a guide for users.

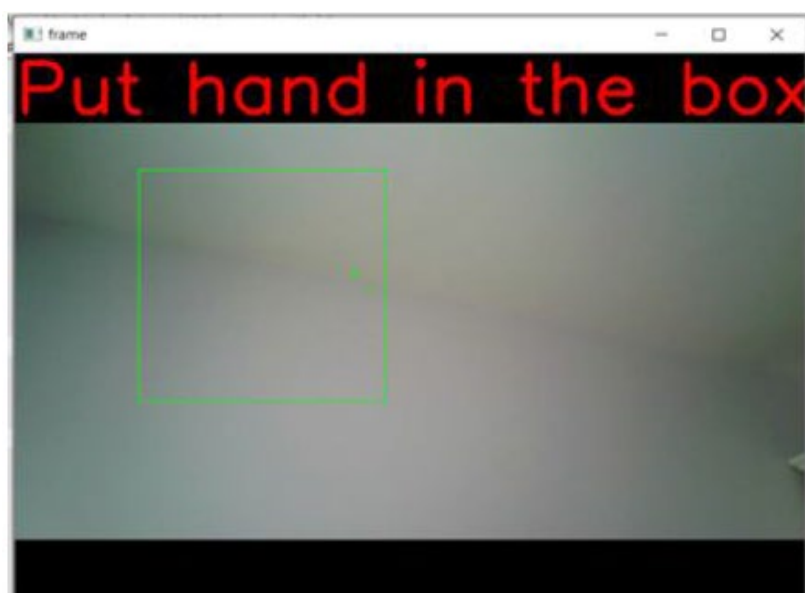


Figure 4: Application main screen

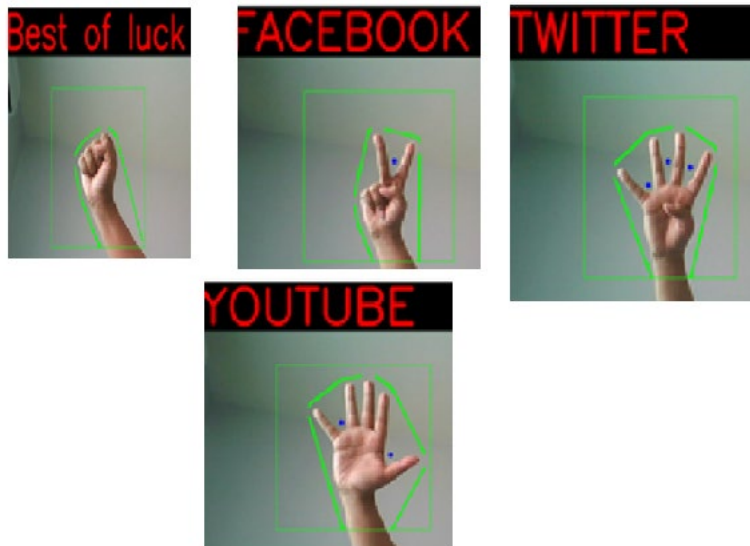


Figure 5: Signs being used by System

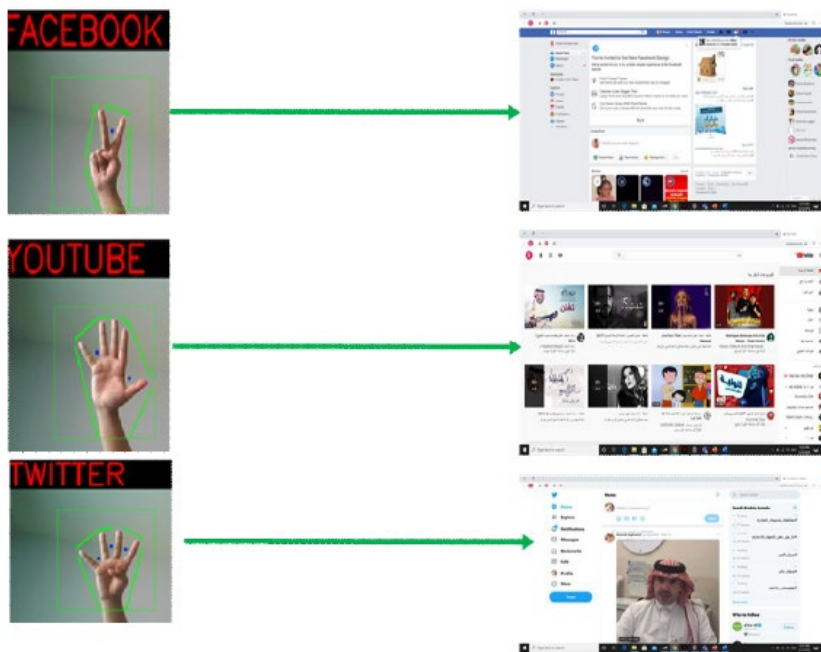


Figure 6: Results of Sign Detection.

5. CONCLUSION

The main objective of this research has been achieved successfully. Gesture interpretation works best in case users who understand sign language may interact with people who are unfamiliar with sign language. Speech interpretation is helpful for sign language non-speakers who want the accompanying

hand sign to be understood. Room conditions such as lighting can play a role in predicting the outcome of poor lighting. The light that is either too bright or too dim will result in inaccurate hand segmentation, resulting in inaccurate gesture prediction. The type of inaccuracy can emerge from the user's peripherals, such as poor web camera performance or poor microphone quality. In a nutshell, the development of technology is essential, and its deployment in sign language is highly critical. It will serve to bring efficiency in communication, not only to the deaf and dumb but those with the ability to hear and speak as well. In addition to creating opportunities for their career growth, it will enhance their social life through effective communication. Making an impact and changing the lives of the deaf and dumb through technology will be an innovation of the year worth the time and resources.

At the beginning of the D-Talk idea, the developers think to have more than one task for this application, but in the end, they narrow the task to have only one. They thought to have an open calendar, lunch Microsoft office word, and browse the website. The final task was to lunch three websites, Facebook, Twitter, and YouTube.