

Sign Language Recognition using Open CV

Saiteja Juluru¹, Bhavya Empati², Deepak Kallepalli³, Anitha V⁴

Department of Computer Science and Engineering, Dr B V Raju Institute of Technology, Hyderabad, India
julurusaitaja93[at]gmail.com

Department of Computer Science and Engineering, Dr B V Raju Institute of Technology, Hyderabad, India
bhavyaempati[at]gmail.com

Department of Computer Science and Engineering, Dr B V Raju Institute of Technology, Hyderabad, India
deepakbabu825[at]gmail.com

Department of Computer Science and Engineering, Dr B V Raju Institute of Technology, Hyderabad, India
anitha.vippadapu[at]gmail.com

Abstract: Sign language recognition is a project which utilizes advancements in technology to assist especially aided individuals. We planned to build a sign detector for this purpose. Through this paper, we are trying to recognize and display messages according to hand movements. The growth in the existing technologies, as well as extensive research, is used to assist the deaf and dumb. This can be extremely useful for deaf and a dumb person in interacting with others, as understanding sign language is not something that many possess. The modules such as OpenCV and Keras of python are being used to accomplish this task.

Keywords: Sign language, Keras, OpenCV, Python

1. Introduction

A. Motivation

Conversing with people with a hearing disability is a major challenge. Deaf and Mute people use a hand gestured sign language to communicate, which leads to others, outside their community, often facing difficulty in recognizing their language by the signs they make. Hence, there is a need for systems that recognize the different signs and conveys the information to common folks.

B. Problem Statement

A sign detector is used to detect hand and other gestures where these gestures are a part of the sign language. The detector then displays messages for the specially aided people based on these specific movements. The concepts of Computer vision and Machine learning algorithms can be used in this regard.

C. Sign Language and Hand Gesture Recognition

Sign language recognition is a technique of translating a user's gestures into text. It helps individuals, who are unable to interact, with others thereby reducing the visible gap between the special aided individuals to the general public. Raw images/videos are turned into text that can be read and comprehended using image processing techniques and neural networks which links the gesture to its corresponding text in the training data. Dumb people are frequently cut off from normal social interaction and it has been found that they struggle to engage with regular folks through their gestures often, as only handful of them are understood by the majority of people. This means, that people with hearing impairment or deaf people cannot talk like normal people, so they have to depend on some sort of visual communication most of the time.

In the specially aided community, Sign Language has been a primary means of communication. As with any other

language, it includes grammar and vocabulary but uses visual modality for exchanging information. The problem arises when dumb or deaf people try to express themselves to other people with the help of this sign language grammar. As it is very distinctive for other individuals outside their community, this method of communication is limited to the family of the said specially aided person.

In this age of technology, the demand for a computer-based system is highly demanding for the dumb community. However, researchers have been attacking the issue for quite some time now and the results are showing some promise. Although impressive speech recognition systems were created earlier, there are yet any commercially available character recognition devices.

The goal is to familiarize the computer to recognize human speech and create a user-friendly human-computer interface (HCI). A few steps in this technique include training your computer to recognize voices, expressions, and human gestures. Gestures are nonverbal exchanges of messages, and a person can make a variety of gestures at once. Since human signals are perceived through vision, this is a fascinating subject for computer vision researchers. Complex programming strategies are required to code these gestures in machine language and help in creating an HCI that can detect human gestures. For better output creation, we are relying on Image Processing and Template Matching in our research.

D. Objective

- Create a huge number of relevant datasets.
- To reduce the noise and acquire the ROI, use the suitable picture pre-processing techniques.
- Develop a model and framework for CNN that can be used in terms of training pre-processed data and achieve optimal efficiency.
- Create a real-time gesture prediction algorithm.

E. Organization of the document

Chapter 1 has introduction. Introduction consists of motivation, problem definition, sign language and gesture recognition, and the objectives of project.

Chapter 2 has literature survey. It consists of existing system, disadvantages of existing system and proposed system.

Chapter 3 has Analysis. It includes software requirements specifications and architecture design. Architecture represents the workflow of the project and also explains about the major algorithms involved in the project.

Chapter 4 has Module Design. It consists of all the modules that are there in the process. Here, palm detection model and hand landmark detection are elaborated.

Chapter 5 has implementation and results of the project. Here, key functions and methods of implementation are elaborated.

Chapter 6 is about testing and validation. Testing represents about the accuracy of the project. Validation gives us idea about the results of the project and variation between proposed system and existing systems.

Chapter 7 is conclusions and future enhancements. Extensions of the projects can be made in the future are discussed here.

2. Literature Survey**a) Introduction**

A review of the literature for the proposed framework indicates that numerous methods and algorithms have been used to tackle sign detection in videos and photos. Understanding neural networks was a major part of the domain analysis we conducted for the project.

b) Existing System

Glove-based method:

- In this method the signer has to wear a hardware glove, while the hand movements are getting captured.
- Based on the movement of the glove a message will be generated.

Vision-based method:

- It is further classified into static and dynamic recognition.
- Statics is concerned with the diagnosis of static gestures (2d-images), whereas dynamic is concerned with the live record of gestures on a real-time basis. This entails the employment of a camera to record movement.

c) Disadvantages of Existing System

- Gloves are typically expensive, and not everyone can afford them.
- It's a hassle to trace a defect in a glove when isn't functioning.
- The parts that are being used to make these gloves are pricy. These materials can only be obtained in specific locations.

d) Proposed System

- Creating a system that can identify Fingerspelling-based hand movements and combine them to form a whole word.
- This can be implemented with Python's OpenCV and Keras modules, which identify digits and characters to display data in a variety of signals and body gestures.

3. Analysis**a) Introduction**

Sign language (SL) is a visual-gestural communication technique used by specially aided individuals. Computer recognition of sign language begins with the recording of sign gestures and culminates with creation of text or speech. The generated language is shown, allowing individuals to comprehend the message being delivered by others.

b) Software Requirements Specifications**1) Software Requirements**

The software requirements for this design are listed below:

- Python (3.7.4)
- IDE (Jupyter)
- Numpy (version 1.16.5)
- cv2 (openCV) (version 3.4.2)
- Keras (version 2.3.1)
- Tensorflow (version 2.0.0)
- Spyder IDE

2) Hardware Requirements

The hardware requirements for this design are listed below:

- Processor: Intel Core i3
- RAM: 4GB
- Hard Disk: 256 GB Minimum

c) Architecture Diagram

In the sign language recognition model, the architecture provides a blueprint and ideal techniques to follow so as to developed a well-structured application as per our requirement. This architecture mainly involves three phases:

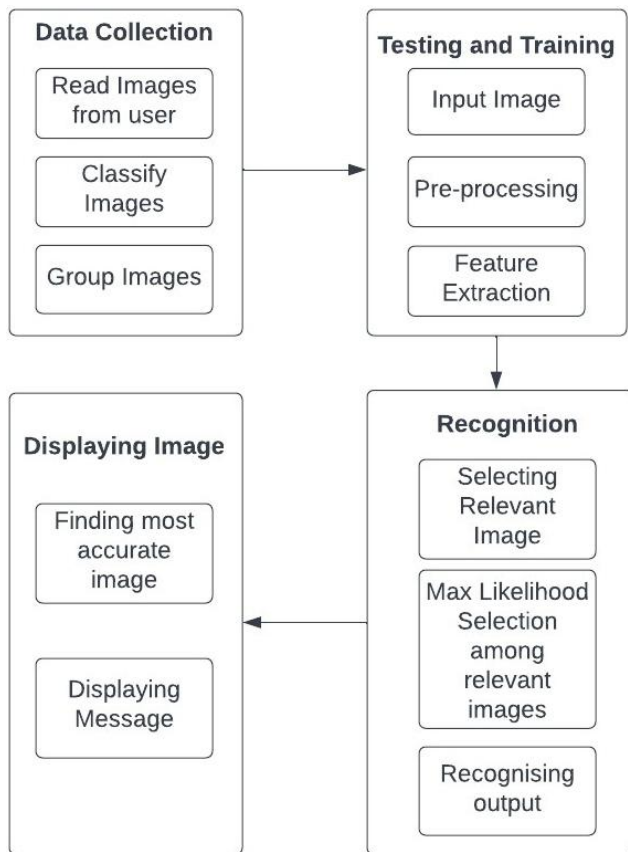


Figure 1: Architecture Model

Phase 1- Data Collection Phase: A model is being built and a sequence of images is being fed to the model. This phase is useful to further train the model based on the type of symbol.

Phase 2- Training and Testing Phase: This phase involves a set of inputs to the model and a particular output is being expected. Based on the outputs, the accuracy of the model is being identified.

Phase 3- Recognition of output: In this phase, an image is being given as input to the model. Based on the images being trained to the model, it matches the image with a particular output and generates a message. The sign being identified in the above phase has a particular meaning and a message is being assigned in the training phase. The meaning is displayed to the user.

d) Flowcharts

Each one of the phases in the architecture are further explained and sketched graphically in the form of flowcharts. These flowcharts further explain how each of the phases described in the architecture diagram functions.

This helps us in deeply understanding it's core features, usability and its relation with the other respective phases present.

Data Collection Phase:

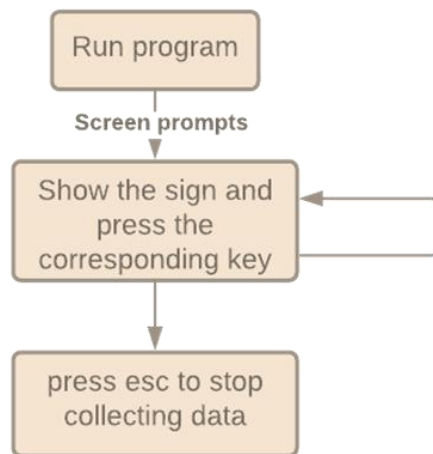


Figure 2: Flow of Data Collection Phase

Training and Testing Phase:

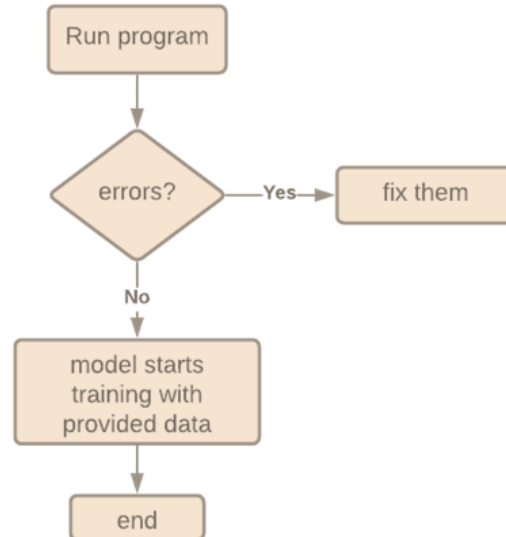


Figure 3: Flow of Training and Testing Phase

Recognition Phase:

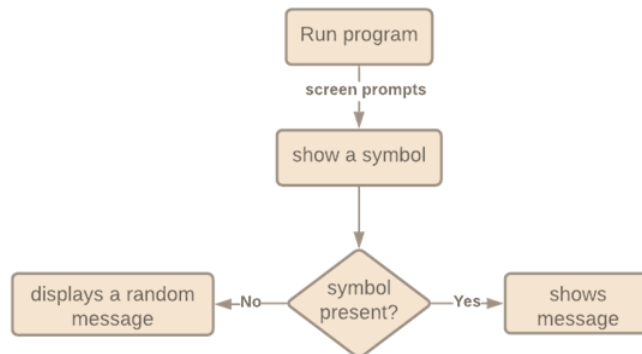


Figure 4: Flow of Recognition Phase

4. Design

F. Introduction

One of the primary issues that a person who is specially aided goes through is that they are unable to express their emotions as freely as others in this world. The developments such as Google Assistance or Apple's Siri are based on voice control and such features, although brought about an extreme change in the world of technology, are of no use to

them as it works based on voice recognition. So, the resort to other ways, such as, the sign language.

While sign language plays a huge role in the means of communication for the deaf-mute people, it is often observed how it gains less attention as most people are unfamiliar with it. Hence, one of the solutions to talk with these individuals is the hand gesture mechanism. Designing an application which recognizes hand gestures and help in the non-verbal communication has been developed over the years by many manufacturers around the world, but they are neither budget-friendly or flexible for the end-users.

G. Module Design and Organisation

Module 1- Data Collection: In the data collection phase, different inputs are received from the user and these inputs are saved for future reference. After the data collection is done, the model is being trained to obtain accurate results.

Module 2- Training and Testing: In the training and testing phase, each class is trained with a multi-class support vector machine (MSVM). Hu Invariant moment and structural shape descriptors are combined to make a combinational feature vector that is to be extracted from the input image in the testing phase after applying pre-processing.

Module 3- Recognition: In the recognition phase, different classes are used for testing an input gesture. The outcome with the most probable group is identified to recognize the gesture. Finally, after the recognition of the input message, its meaning is displayed on the screen.

Testing and Validation

H. Introduction

This paper mainly aims to predict a message based on the sign it has read. This is done by training the model into displaying a specific message when it captures an image. Each image has its respective message and the accuracy of the exact message to be displayed depends on how well the model has been trained. Usually, the model is fed with huge number of images and based on the user input, an accurate message is being displayed.

I. Output Screenshots

The set of screenshots listed below displays a set of messages based on the image it has captured. It is to be noted that each of its message differs from one another.

In the below screenshot, the model is trained in such a way that it displays the message "THANK YOU" when it captures the hand gesture of keeping your palms together.



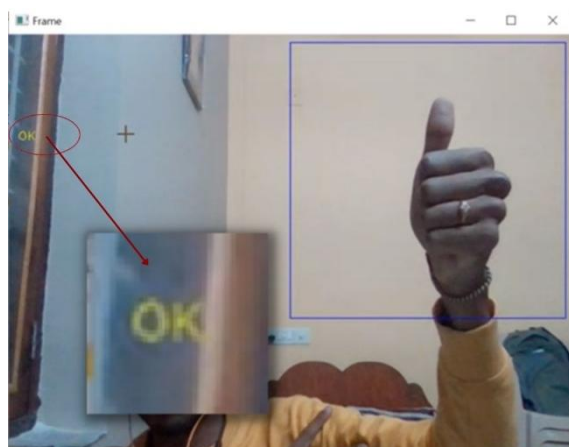
Screen 1: Screen displaying the message "THANK YOU"

In the below screenshot, the model is trained in such a way that it displays the message "LOVE YOU" when it captures the hand gesture of holding a fist.



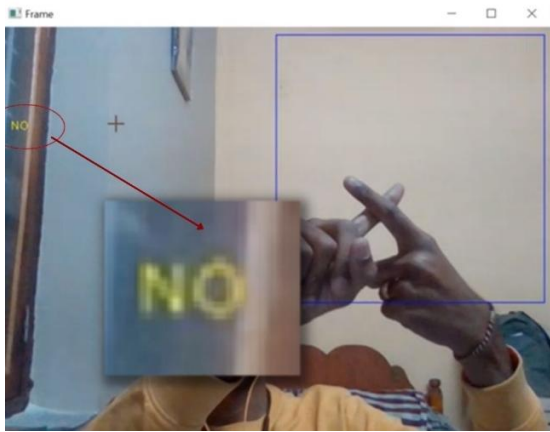
Screen 2: Screen displaying the message "LOVE YOU"

In the below screenshot, the model is trained in such a way that it displays the message "OK" when it captures the hand gesture of showing your thumbs up.



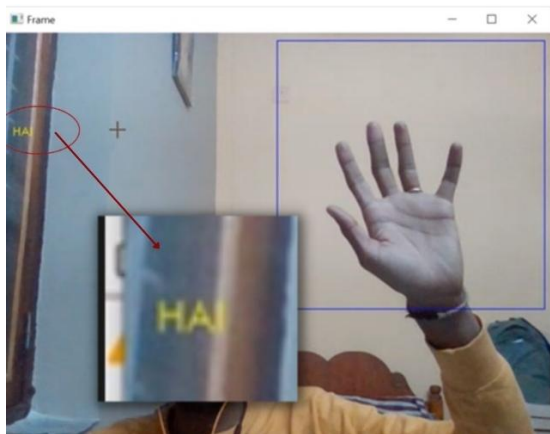
Screen 3: Screen displaying the message "OK"

In the below screenshot, the model is trained in such a way that it displays the message "NO" when it captures the hand gesture of two fingers crossed together.



Screen 4: Screen displaying the message “NO”

In the below screenshot, the model is trained in such a way that it displays the message “HAI” when it captures the hand gesture of your open palm with fingers stretched as if to wave.



Screen 5: Screen displaying the message “HAI”

In the below screenshot, the model is trained in such a way that it displays the message “SORRY” when it captures the hand gesture of your palm shaped in a half-moon or ‘C’.



Screen 6: Screen displaying the message “SORRY”

5. Conclusion and Future work

5.1 Conclusion

By the use of modules such as OpenCV and Keras of Python, a sign language recognition model can be developed which provides aid to disable community. Here, the

designed model is trained by providing a set of inputs, such as images which are then further used to produce a set of messages after capturing a certain hand gesture.

The speciality of this model is that, it can be trained to design various messages with different hand gestures of our choice which helps in bridging the gap between the specially-aided individuals to the common public.

5.2 Future Work

One major enhancement is the increased number of signs and messages. Since our model generates messages only for 6 different signs, we want our model to be able to generate a large number of messages. Another enhancement is that we want to convert this idea to a web or an android application so that it can be used by a wide variety of people.

Acknowledgment

We acknowledge that the success and final outcome of this project required a lot of guidance and assistance from many people and we are extremely fortunate to have got this all along the completion. Whatever we have done is due to such guidance and assistance. We would not forget to thank them. We thank Mrs. V Anitha for guiding us and providing all the support in completing this project.

References

- [1] José Herazo, “Sign language recognition using deep learning”, August 2020.
- [2] Aya Hassonueh, and A.M.Mutawa, “Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG”, March 2020.
- [3] Sanil Jain and K.V.Sameer Raja, “Indian Sign Language Character Recognition” , Indian Institute of Kanpur, April 2019.
- [4] Helen Cooper, Brian Holt, Richard Bowden: “Visual Analysis of Humans”, 2011.
- [5] Aditya Das, Shantanu Gawde, Khyati Suratwala and Dhananjay Kalbande, “Sign Language Recognition Using Deep Learning on Custom Processed Static Gesture Images”, IEEE Conference, 2018.
- [6] Balaji Srinivasan, “Face Mask Detection using Python, Keras, OpenCV and MobileNet”, July 2020
- [7] K Abhijith Bhaskaran, Anoop G. Nair, K Deepak Ram, Krishnan Ananthanarayanan and H.R. Nandi Vardhan, “Smart gloves for hand gesture recognition: Sign language to speech conversion system”, IEEE Conference, 2016.