Sign Language Translator

A Project Report Submitted to

Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal
Towards Partial Fulfillment for the Award of

Bachelor of Technology
(Computer Science and Engineering)

Under the Supervision of

Prof. Priyanka Jangde
Prof. Narendra Pal Singh

Submitted By

Aditya Sharma (0827CS201016)
Aanya Chourasiya (0827CS201003)
Amit Kumar Yadav (0827CS201031)
Akshay Keswani (0827CS201022)

Department of Computer Science and Engineering
Acropolis Institute of Technology & Research, Indore
July-Dec 2022

EXAMINER APPROVAL

The Project entitled "**Sign Language Translator**" submitted by Aditya Sharma (0827CS201016), Aanya Chourasiya (0827CS201003), Amit Kumar Yadav (0827CS201031), Akshay Keswani (0827CS201022) has been examined and is hereby approved towards partial fulfillment for the award of Bachelor of Technology degree in Computer Science & Engineering; Engineering discipline, for which it has been submitted. It understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein, but approve the project only for the purpose for which it has been submitted.

(Internal Examiner) **(External Examiner)**
Date: **Date:**

GUIDE RECOMMENDATION

This is to certify that the work embodied in this project entitled "**Sign Language Translator**" submitted by Aditya Sharma (0827CS201016), Aanya Chourasiya (0827CS201003), Amit Kumar Yadav (0827CS201031), Akshay Keswani (0827CS201022) is a satisfactory account of the bonafide work done under the supervision of Prof. Priyanka Jangde and Prof. Narendra Pal Singh are recommended towards partial fulfillment for the award of the Bachelor of Technology (Computer Science & Engineering) degree by Rajiv Gandhi Proudyogiki Vishwavidhyalaya, Bhopal.

(Project Guide) **(Project Coordinator)**

STUDENTS UNDERTAKING

This is to certify that a project entitled "**Sign Language Translator**" has been developed by us under the supervision of Prof. Priyanka Jangde and Prof. Narendra Pal Singh Rathore. The whole responsibility of work done in this project is ours. The sole intention of this work is only for practical learning and research. We further declare that to the best of our knowledge, this report does not contain any part of any work which has been submitted for the award of any degree either in this University or in any other University / Deemed University without proper citation and if the same work is found then we are liable for explanation to this.

Aditya Sharma (0827CS201016)

Aanya Chourasiya (0827CS201003)

Amit Kumar Yadav (0827CS201031)

Akshay Keswani (0827CS201022)

Acknowledgement

We thank the almighty Lord for giving me the strength and courage to sail out through the tough and reach on shore safely. There are a number of people without whom this project's work would not have been feasible. Their high academic standards and personal integrity provided me with continuous guidance and support. We owe a debt of sincere gratitude, deep sense of reverence and respect to our guide and mentors **Prof. Priyanka Jangde and Prof. Narendra Pal Singh Rathore**, Associate Professor, AITR, for their motivation, sagacious guidance, constant encouragement, vigilant supervision and valuable critical appreciation throughout this project work, which helped us to successfully complete the project on time. We express profound gratitude and heartfelt thanks to **Dr Kamal Kumar Sethi**, HOD CSE, AITR Indore for his support, suggestion and inspiration for carrying out this project. I am very much thankful to other faculty and staff members of CSE Dept, AITR Indore for providing me all support, help and advice during the project. We would be failing in our duty if we do not acknowledge the support and guidance received from **Dr S C Sharma**, Director, AITR, Indore whenever needed. We take the opportunity to convey my regards to the management of Acropolis Institute, Indore for extending academic and administrative support and providing me with all necessary facilities for the project to achieve our objectives. We are grateful to our parents and family members who have always loved and supported us unconditionally. To all of them, we want to say, “Thank you”, for being the best family that one could ever have and without whom none of this would have been possible.

Aditya Sharma (0827CS201016)

Aanya Chourasiya (0827CS201003)

Amit Kumar Yadav (0827CS201031)

Akshay Keswani (0827CS201022)

Executive Summary

This project is submitted to Rajiv Gandhi Proudyogiki Vishwavidhyalaya, Bhopal(MP), India for partial fulfillment of Bachelor of Technology in Computer Science & Engineering; Engineering branch under the sagacious guidance and vigilant supervision of **Prof. Priyanka Jangde and Prof. Narendra Pal Singh Rathore**. This Project focuses on a review of the literature on hand gesture techniques and introduces their merits and limitations under different circumstances. The theories of hand segmentation and the hand detection system, may be used to construct hand gesture recognition using Python and OpenCV. The use of hand gestures as a natural interface motivates research in gesture taxonomies, representations, and recognition algorithms, as well as software platforms and frameworks, all of which are briefly covered in this Project.

List of figures

Table of contents

Sign Language is the Noblest Gift God has Given to Deaf People.

-George Veditz

Chapter 1. Introduction

Introduction

1.1 Overview

The increased public acceptance and funding for international projects emphasizes the necessity of sign language. The desire for a computer-based solution is significant in these age of technology for the deaf people. However, re- searchers have been attacking the problem for quite sometimes now and the results are showing some promises. Although interesting technologies become available for voice recognition, but there is currently no commercial solution for sign recognition on the market.

1.2 Problem Statement

This method takes pictures on camera as touch data. The vision-based approach focuses heavily on touch captured images and brings out the main and recognizable feature. Colour belts were used at the beginning of the vision-based approach. The main disadvantage of this method was the standard color to be applied to the fingers. Then use bare hands instead of colored ribbons. This creates a challenging problem as these systems require background, uninterrupted lighting, personal frames and a camera to achieve real-time performance.

In addition, such systems must be developed to meet the requirements, including accuracy and robustness. The Model will be trained by several set of Gestures of sign language using OpenCV and Python API. We will Expand several gestures of hands by our own to train the model. The gestures will be taken by the camera from multiple directions. Several different methods have been tested so far. The first is to build a 3D human hand model.

The model is compared to hand images with one or more cameras, and the parameters corresponding to the shape of the palm and the combined angles are estimated. These parameters are then used to create the touch phase. The second is to take a picture using the camera and extract certain features and those features are used as input in the partition algorithm to separate. To use

the Model, the user shows some hand gestures in front of the camera. Upon encountering similar gestures in database, the model will show text/sentence as output representing the translation of the sign language. Upon encountering any gesture that is not present in the database, the model will display an Error message.

1.3 Objectives

The goal is to make computers to understand human language and develop a user-friendly human computer interface (HCI). Some steps toward this goal include teaching a computer to recognize speech, facial emotions, and human gestures. Gestures are the nonverbally exchanged information. Gesture recognition is an aspect of human-computer interaction that demonstrates an academic treatise and is a vital to popularizing the notion of a human-to-human connection, open dialogue, that must imply the correlation between the user and the machine.

Gesture analysis is a scientific field that can recognize gestures such as hand, arm, head, and even structural motions that usually entail a certain posture and/or motion. Using hand gestures, the individual may send out more information in a shorter amount of time. Several approaches were explored to apply computer-vision ideas to the real time processing of gesture outputs. The Computer Vision study concentrates on gesture recognition in the open CV framework using the Python language.

Language is a huge part in communication. Languages are useless to a person with a disability. Gesture is a vital and meaningful mode of communication for the visually impaired person.

1.4 Survey of Existing Systems

Based on our review, HMM-based approaches have been extensively explored in prior research, including its modifications. Deep Learning, such as Convolutional Neural Networks, has been popular in the past five years. Hybrid CNN-HMM and completely deep learning systems have yielded encouraging results and provide avenues for additional research. Clustering and high computational needs, however, continue to stymie their adoption.

The output of the sign language will be displayed in the text form in real time. This makes the system more efficient and hence communication of the hearing and speech impaired people easier. The images captured through web cam are compared and the result of comparison is displayed.

It is required to make a proper database of the gestures of the sign language so that the images captured while communicating using this system can be compared. For making the database, we would be capturing each gesture from more than 2 angles so that the accuracy of the system will be increase significantly. The more angles you take, the better is the accuracy and the more amount of memory is required.

Chapter 2 . Model Architecture

2.1 Use of sensory devices

It uses electromechanical devices to provide exact hand configuration, and position. Different glove based approaches can be used to extract in formation. But it is expensive and not user friendly.

2.2. Vision based approach

In vision based methods computer camera is the input device for observing the information of hands or fingers. The Vision Based methods require only a camera, thus realizing a natural interaction between humans and computers without the use of any extra devices. These systems tend to complement biological vision by describing

2.3. Technical Requirements

- Python 3.7
- Tensorflow
- Keras
- OpenCV
- NumPy
- Matplotlib

2.3.1 Python :

Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation. Python is dynamically-typed and garbage-collected. It supports multiple programming paradigms, including structured, object-oriented and functional programming.

2.3.2 TensorFlow :

Tensorflow is an open source software library for numerical computation. First we define the nodes of the computation graph, then inside a session, the actual computation takes place. TensorFlow is widely used in Machine Learning.

2.3.3 Keras :

Keras is a high-level neural networks library written in python that works as a wrapper to TensorFlow. It is used in cases where we want to quickly build and test the neural network with minimal lines of code. It contains implementations of commonly used neural network elements like layers, objective, activation functions, optimizers, and tools to make working with images and text data easier.

2.3.4 OpenCV :

OpenCV(Open Source Computer Vision) is an open source library of programming functions used for real-time computer-vision. It is mainly used for image processing, video capture and analysis for features like face and object recognition. It is written in C++ which is its primary interface, however bindings are available for Python, Java, MATLAB/OCTAVE.

2.4 Comparative Study of Existing Systems

Based on our review, HMM-based approaches have been extensively explored in prior research, including its modifications. Deep Learning, such as Convolutional Neural Networks, has been popular in the past five years. Hybrid CNN-HMM and completely deep learning systems have yielded encouraging results and provide avenues for additional research. Clustering and high computational needs, however, continue to stymie their adoption.

The output of the sign language will be displayed in the text form in real time. This makes the system more efficient and hence communication of the hearing and speech impaired people easier. The images captured through web cam are compared and the result of comparison is displayed.

It is required to make a proper database of the gestures of the sign language so that the images captured while communicating using this system can be compared. For making the database, we would be capturing each gesture from more than 2 angles so that the accuracy of the system will be increase significantly. The more angles you take, the better is the accuracy and the more amount of memory is required.

2.5 Requirement Identification

- Windows computer or Linux
- Python Version 3.6 or above
- CMOS sensor (Webcam)
- Python Installed Windows Operating system or Linux Os Machine.

Hardware & Software Requirement - Windows computer or Linux, Python installed and Libraries, CMOS sensor (Webcam) , Python Installed Windows Os or Linux Os Machine. Data Collection – The first step is the data-set collection. We collect the full frame (RGB) image and the corresponding depth map from the camera. It actually understands the gesture and finger positions and update it in the Database where all relevant Collection of images are stored. Collection of Gestures in Database – The set of all Images conveying the same message will be grouped together as an entity in the database, thus the Database will serve as a collection of multiple number of entities, each of which represents a set of actions / gestures of hand conveying a message

Chapter 3 . Proposed System

Proposed System

3.1 Proposal

- Data Collection – The first step is the data-set collection. We collect the full frame (RGB) image and the corresponding depth map from the camera. It actually understands the gesture and finger positions and update it in the Database where all relevant Collection of images are stored.
- Extraction of Hand-Gesture – The Set of relevant images will be processed and assigned a Message that the particular sign language will be conveying. All the images from multiple angles having similar Messages will be grouped together and will be assigned the Text message to be displayed in the Model.
- Collection of Gestures in Database – The set of all Images conveying the same message will be grouped together as an entity in the database, thus the Database will serve as a collection of multiple number of entities, each of which represents a set of actions / gestures of hand conveying a message.
- Hand-Gesture Recognition using Deep Learning – Upon User encounter with the model, the user will show hand gestures in front of the camera. Using OpenCV, those images will be processed and compared with pretrained and pre-processed set of images.
- Upon any image in database getting matched with the Gesture, the Mapped Text message linked with the gesture will be displayed on the screen in the form of Text/sentence.
- If the gesture performed by the user does not matches any linked image in the database, in that scenario, and Error message will be displayed on the screen of the user.

3.2 Benefits of Proposed System

The current project had a lot of challenges that are overcome by this system :

- **Easy to Use :** The user interface is so basic that it is very easy to understand and operate.

- **Hand Gesture recognition :** Users, especially deaf can use it because of its Sign language detection feature.
- **2 way communication :** Using sign language translator, deaf people can communicate with normal people and in this way the communication no longer remain 1 way only but rather be interactive and two way communication.

3.3 Diagrams

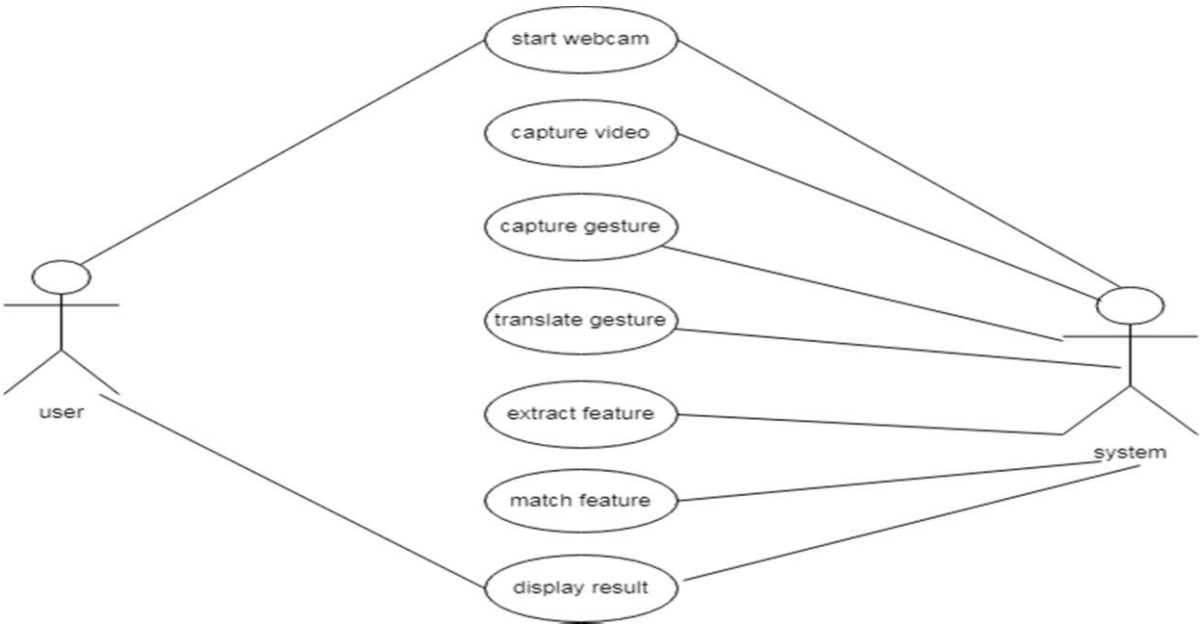

Figure 1: Use Case Diagram

A use case is a methodology used in system analysis to identify, clarify and organize system requirements. The use case is made up of a set of possible sequences of interactions between systems and users in a particular environment and related to a particular goal. The method creates a document that describes all the steps taken by a user to complete an activity. Here The set of states mentioned in Use case diagram represents the set of actions that are performed during the operational phase of project.

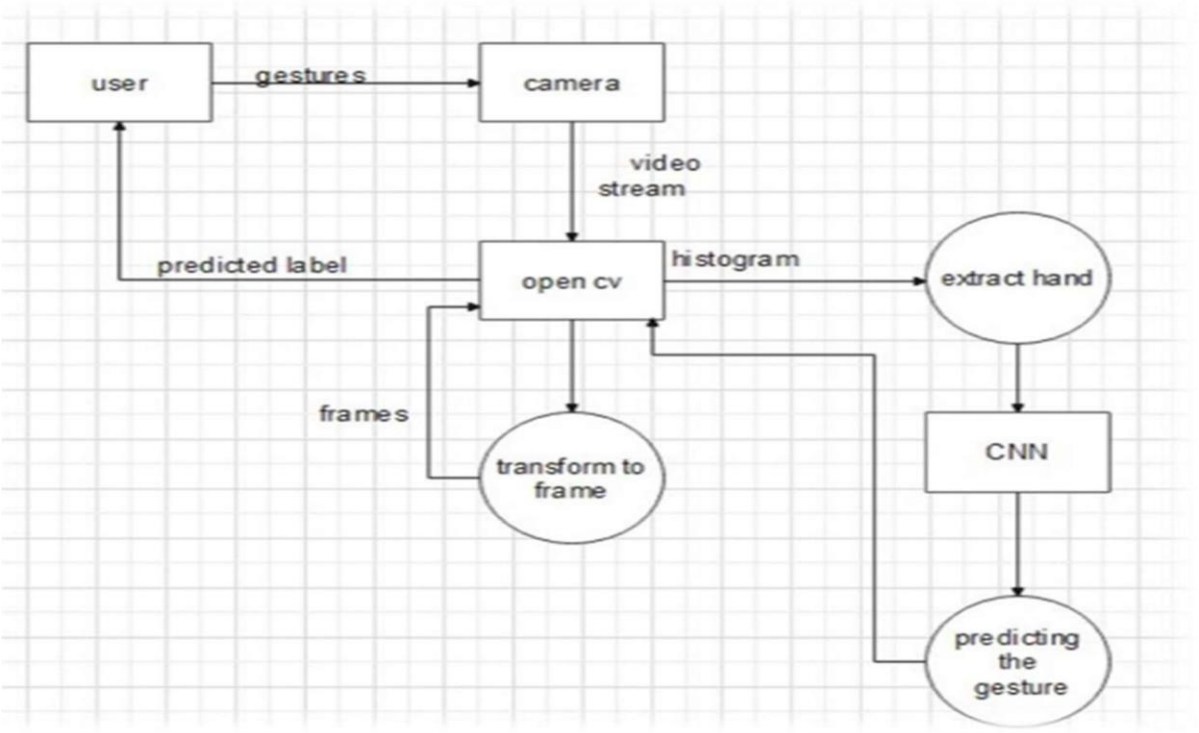

Figure 2. Data Flow Diagram

A data flow diagram (DFD) maps out the flow of information for any process or system. It uses defined symbols like rectangles, circles and arrows, plus short text labels, to show data inputs, outputs, storage points and the routes between each destination According to the DFD diagram, the user's gestures are tracked by the camera, out of which hand gestures are extracted. Later, using openCV, the extracted hand gestures are transformed in skeleton like dotted patterns which is detected by the model upon being compared with existing gestures.

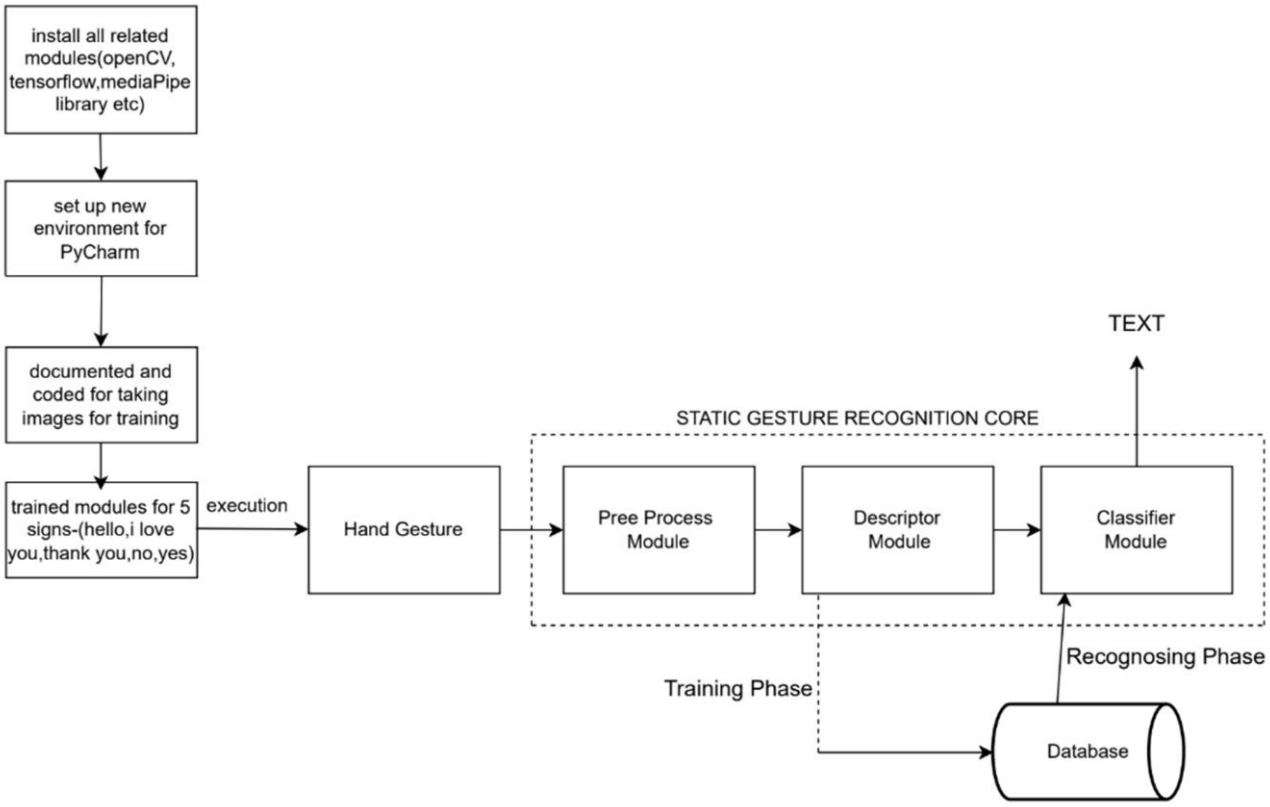

Figure 3 : Process flow Diagram

A Process Flow Diagram is a type of flowchart that illustrates the relationships between major components at an industrial plant. It's most often used in chemical engineering and process engineering, though its concepts are sometimes applied to other processes as well. It’s used to document a process, improve a process or model a new one.

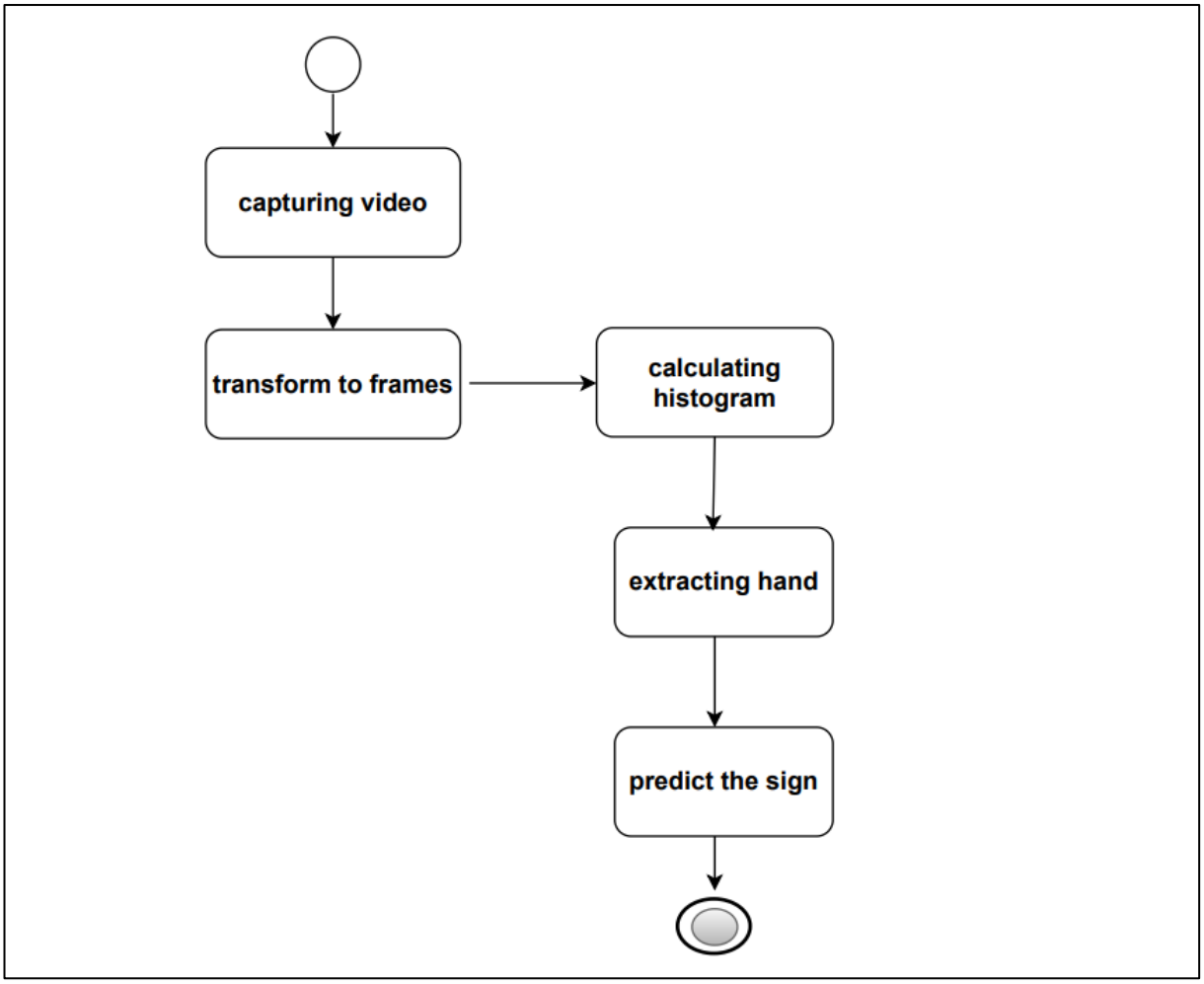

Figure 4 : State Chart Diagram

Statechart diagram is one of the five UML diagrams used to model the dynamic nature of a system. They define different states of an object during its lifetime and these states are changed by events. Statechart diagrams are useful to model the reactive systems. Statechart diagram describes the flow of control from one state to another state. States are defined as a condition in which an object exists and it changes when some event is triggered.

Chapter 4 . Implementation

Implementation

4.1 GESTURE CLASSIFICATION AND METHODOLOGY

1. Installed all the relevant modules involved in project like Tensorflow, OpenCV, Keras, Numpy.
2. Setted up New Environment for Pycharm Notebook.
3. Documented and Coded for taking inages for training.
4. The Code part upon being executed, made separate folders for Images to be proessed in the notebook and in system.
5. In the next, step, we coded for Camera tracing and Image capturing.
6. Here for initial phases for Training, We used 7 Signs –
   Hello, Thank You, A, B, C, No, Yes and Acropolis.
7. Upon running the camera tracing and image capturing code the camera window opened.
8. We Took 20 Images per Sign from different Directions & different Alignments.
9. Through this, 6 different files for all 6 signs got Imported in Images Folder each containing bunch of 20 images all representing same sign.
10. Using Label-Img, All 600 Images were Traced and Hand gesture is Cropped for every Image, While cropping, Each cropped gesture is labeled a Title describing its Sign.
11. Now we are left with 300 images (All 600 in paired with their Cropped label image)
12. Out of 300 pairs of Images, 20% Image pairs were sent into Testing folder, rest 80% image pairs were transferred to training folder.
13. Now Through OpenCV and cv2, We code for Application interface and Live detection and Text displaying on screen.
14. Upon running, the camera window opens and user can Use gestures infront of camera.
15. Upon detection and matching of any Gesture with trained images, If gesture gets matched, Corresponding Sign Text get displayed on screen.
16. Since the camera remains active until No external toggle, The user can keep on using gestures and detection continues.
17. The detection process can be stopped simply by closing the window.

4.2 Libraries

- Python 3.7
- Tensorflow
- Keras
- OpenCV
- NumPy
- Matplotlib

4.3 Testing

Testing is the process of evaluation of a system to detect differences between given input and expected output and also to assess the features of the system. Testing assesses the quality of the product. It is a process that is done during the development process.

Tests can be conducted based on two approaches – Functionality testing, Implementation testing

The testing method used here is Black Box Testing. It is carried out to test functionality of the program. It is also called ‘Behavioral’ testing. The tester in this case, has a set of input values and respective desired results. On providing input, if the output matches with the desired results, the program is tested ‘ok’, and problematic otherwise.

Everything in this project is tested, testing was also done while making this project simultaneously and after making the project.

4.4 Test Results

We were able to achieve training accuracy of 89.30% and testing accuracy of 98.5%. Our model may show good accuracy while predicting results both offline and online.

4.5 Demonstration

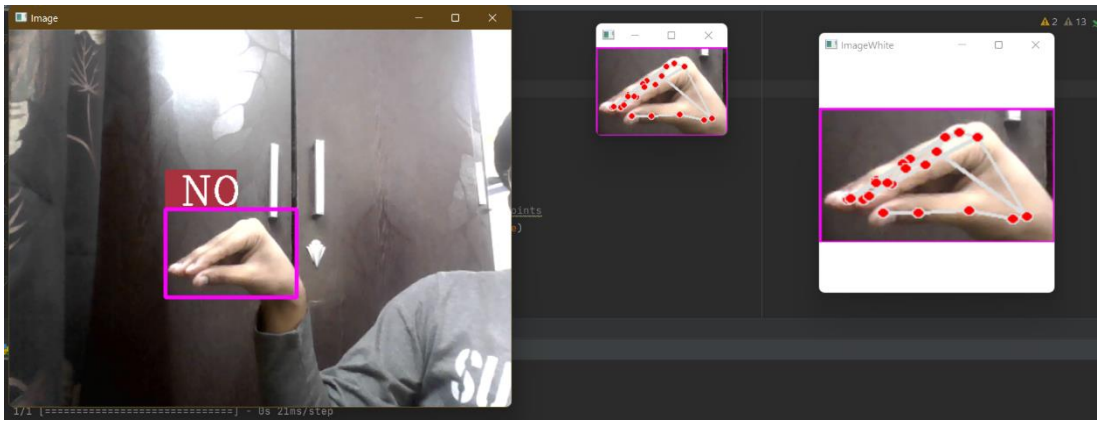

Figure 5: Demonstration of NO

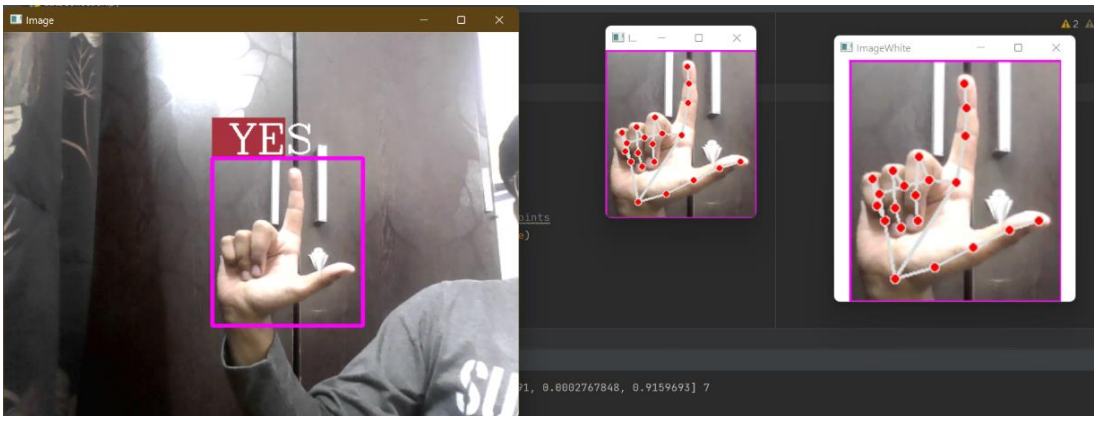

Figure 6: Demonstration of YES

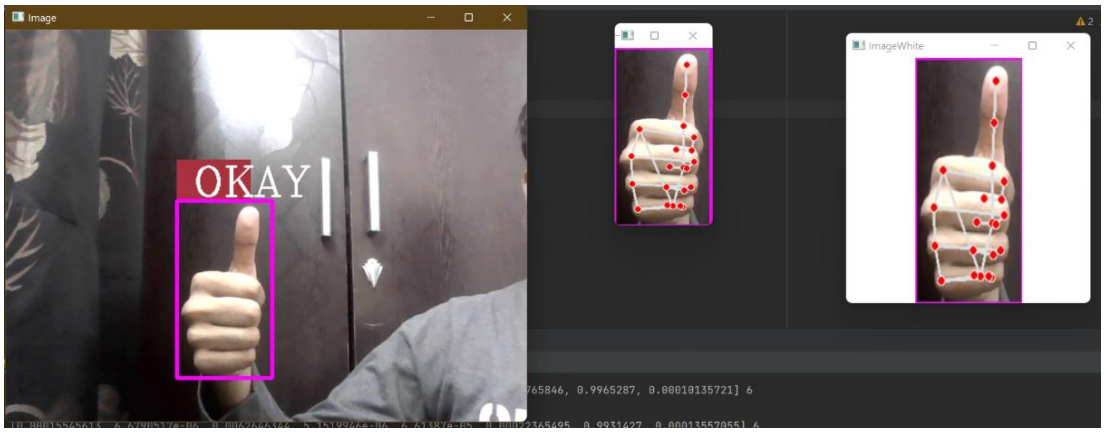

Figure 7: Demonstration of OKAY

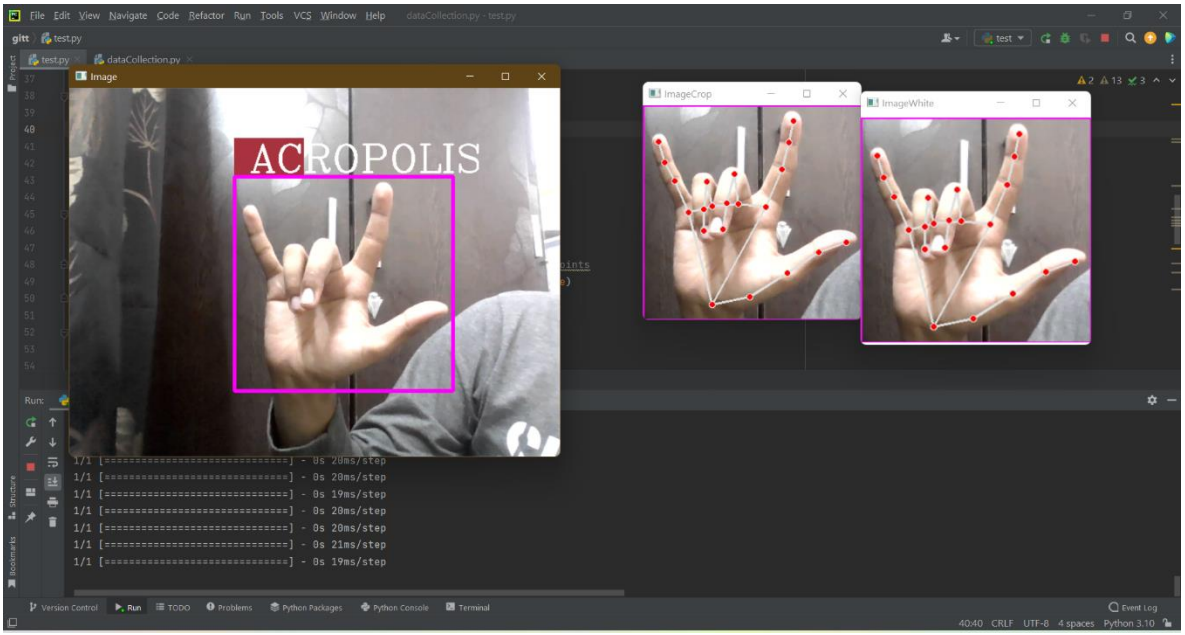

Figure 8: Demonstration of ACROPOLIS

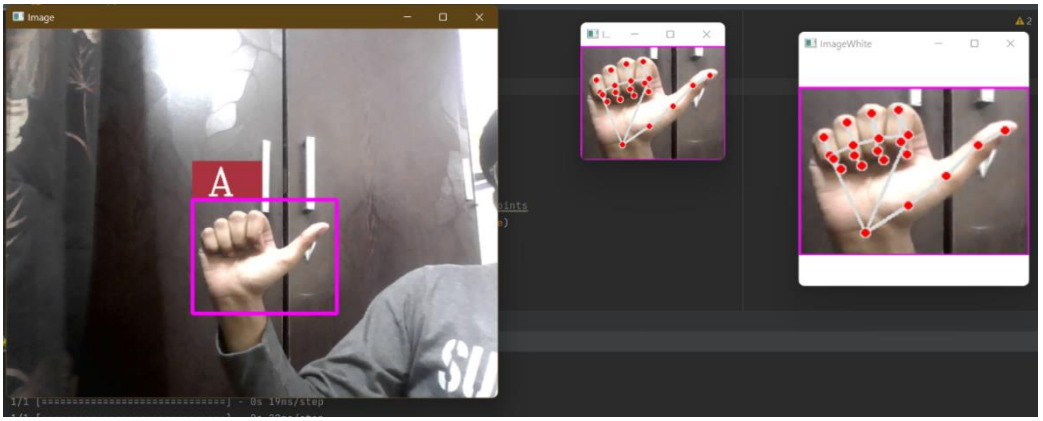

Figure 9: Demonstration of A

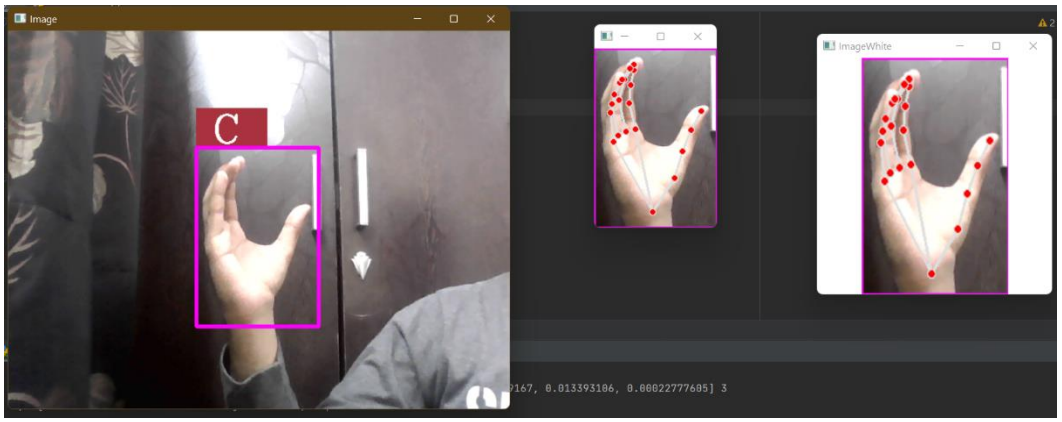

Figure 10: Demonstration of C

Chapter 5. Conclusion

Conclusion

The aim of this project is to predict the ISL alphanumeric hand-gestures in real time. The above work shows that it can be solved with better accuracy when we actually consider the segmented RGB hand-gestures. By applying depth-based segmentation we remove the overheads of dynamic background. The segmented RGB hand-gestures were fed to 3 layered CNN for training and testing in real time.

We were able to achieve training accuracy of 89.30% and testing accuracy of 98.5%. Our model may show good accuracy while predicting results both offline and online. Our project aims to make communication simpler between deaf and dumb people by introducing Computer in communication path so that sign language can be automatically captured, recognized, translated to text and displayed it on Screen. There are various methods for sign language conversion.

5.1 Limitations & Challenges

The main challenge of vision-based hand detection is to cope with the large variability of human hand's appearance due to a huge number of hand movements, to different skin-colour possibilities as well as to the variations in view points, scales, and speed of the camera capturing the scene.

There were many challenges faced by us during the project. The very first issue we faced was of dataset. We wanted to deal with raw images and that too square images as CNN in Keras as it was a lot more convenient working with only square images. We couldn't find any existing dataset for that hence we decided to make our own dataset. Second issue was to select a filter which we could apply on our images so that proper features of the images could be obtained and hence then we could provided that image as input for CNN model. We tried various filter including binary threshold, canny edge detection, gaussian blur etc. but finally we settled with gaussian blur filter. More issues were faced relating to the accuracy of the model we trained in earlier phases which we eventually improved by increasing the input image size and also by improving the dataset.

Bibliography

1. Vinayak S Kunder, Aakash A Bhardwaj, and Vipul D Tank. Sign language recognition system with speech output.

2. Malavika Suresh, Avagyan Sinha, and RP Aneesh. Real-time hand gesture recognition using deep learning. International Journal of Innovations and Implementations in Engineering.

3. GRS Murthy and RS Jadon. A review of vision-based hand gestures recognition. International Journal of Information Technology and Knowledge Management.

4. Mukesh Kumar Makwana, " Sign Language Recognition", MTech thesis, Indian Institute of Science, Bangalore.

5. S. Hussain, R. Saxena, X. Han, J. A. Khan, and H. Shin, “Hand gesture recognition using deep learning”, 2017 International SoC DesignConference (ISOCC).

6. J. -. Lementec and P. Bajcsy, "Recognition of arm gestures using multiple orientation sensors: gesture classification", Proceedings.

7. Ms. Rashmi D. Kyatanavar, Prof. P. R. Futane,Comparative Study of Sign Language Recognition Systems, International Journal of Scientific and Research Publications.


SLT
Sign Language Translator