A MAJOR PROJECT REPORT

ON

REAL TIME SIGN LANGUAGE RECOGNITION USING TRANSFER LEARNING

Submitted in partial fulfillment of the requirement

for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

BY

G Harsha Vardhan	**19P61A0566**
G Harsha Vardhan Rao	**19P61A0575**
S Harshith	**19P61A0581**

Under the esteemed guidance of

Mr. G. Anil kumar

Associate Professor

Dept. of CSE

Counselling Code : VBIT

VIGNANA BHARATHI Institute of Technology

(A UGC Autonomous Institution, Approved by AICTE, Accredited by NBA & NAAC-A Grade, Affiliated to JNTUH)

December - 2022

Counselling Code : VBIT

VIGNANA BHARATHI Institute of Technology

(A UGC Autonomous Institution, Approved by AICTE, Accredited by NBA & NAAC-A Grade, Affiliated to JNTUH)

Aushapur (V), Ghatkesar (M), Hyderabad, Medchal – Dist, Telangana – 501 301.

DEPARTMENT
OF
COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

*This is to certify that the major project titled "**REAL TIME SIGN LANGUAGE RECOGNITION USING TRANSFER LEARNING** "submitted by **G Harsha Vardhan(19P61A0566), G Harsha Vardhan Rao(19P61A0575), S Harshith(19P61A0581)** in B.tech IV-I semester Computer Science & Engineering is a record of the bonafide work carried out by them*

The Design embodied in this report have not been submitted to any other University for the award of any degree

INTERNAL GUIDE
Mr. G. Anil Kumar
(Associate Professor)

HEAD OF THE DEPARTMENT
Dr. M. Venkateswarao Rao

EXTERNAL EXAMINER

DECLARATION

We, **G Harsha vardhan, G Harsha vardhan rao**, **S Harshith** bearing hall ticket numbers **(19P61A0566, 19P61A0575, 19P61A0581)** hereby declare that the mini project report entitled "**REAL TIME SIGN LANGUAGE RECOGNITION USING TRANSFER LEARNING**" under the guidance of **Mr. G. Anil kumar,** Associate Professor, Department of Computer Science and Engineering**, Vignana Bharathi Institute of Technology, Hyderabad**, have submitted to Jawaharlal Nehru Technological University Hyderabad, Kukatpally, in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Computer Science And Engineering.

This is a record of bonafide work carried out by us and the results embodied in this project have not been reproduced or copied from any source. The results embodied in this project report have not been submitted to any other university or institute for the award of any other degree or diploma.

G HARSHA VARDHAN (19P61A0566)

G HARSHA VARDHAN RAO(19P61A0575)

S HARSHITH (19P61A0581)

ACKNOWLEDGEMENT

We are extremely thankful to our beloved Chairman **Dr. N. Goutham Rao** who took keen interest to provide us the infrastructural facilities for carrying out the project work.

Self-confidence, hard work, commitment and planning are essential to carry out any task. Possessing these qualities is sheer waste, if an opportunity does not exist. So, we whole-heartedly thank **Dr. P. V. S. Srinivas, Principal**, and **Dr. G. Sreeram**, Head of the Department, Computer Science and Engineering for their encouragement and support and guidance in carrying out the project.

We would like to express our indebtedness to the project coordinator, **Dr. K. Kalaivani**, Associate Professor, Department of CSE, **Mr. G. Srikanth Reddy**, Associate Professor, Department of CSE, **Mrs. P. Subhadra**, Associate Professor, Department of CSE, for her valuable guidance during the course of project work.

We thank our Project Guide, **Mr. G. Anil kumar,** Associate Professor, for providing us with an excellent project and guiding us in completing our mini project successfully.

We would like to express our sincere thanks to all the staff of Computer Science and Engineering, VBIT, for their kind cooperation and timely help during the course of our project. Finally, we would like to thank our parents and friends who have always stood by us whenever we were in need of them.

ABSTRACT

Communication is very crucial to human beings, as it enables us to express ourselves. We communicate through speech, gestures, body language, reading, writing or through visual aids, speech being one of the most commonly used among them. However, unfortunately, for the speaking and hearing-impaired minority, there is a communication gap. Sign language is learned by deaf and mute and usually it is not known to normal people. It strikes our mind that to bridge the gap between hearing impaired and normal people & make communication easier.

To overcome this barrier, we propose a method where we collect sign language gestures using webcam and by using Transfer learning, we train a Tensor Flow model to create a Real-time Sign Language Recognition system. This will help a lot of people in communicating with deaf and mute people.

CONTENTS

CHAPTER **PAGE NO.**

LIST OF FIGURES

1. Introduction

1.1. Introduction to the system 1

1.2. Problem Definition 1

1.3. Objectives 2

1.4. Aim of the project 2

2. Literature Survey

2.1. Existing System 5

2.2. Proposed System 5

2.3. Scope of the project 6

3. Analysis

3.1. Technical Feasibility 8

3.2. Operational Feasibility 8

3.3. Economical Feasibility 9

4. Hardware and Software requirements

4.1. Software requirements 11

4.2. Hardware requirements 11

5. System Design

5.1. Software Design 13

5.2. Input Design 14

5.3. Output Design 14

5.4. Architecture 15

5.5. UML Diagrams 16-22

List of Figures

S.NO.	Figure Name	Page No.
1.	Use case Diagram	17
2.	Class Diagram	19
3.	Activity Diagram	21
4.	Sequence Diagram	22

CHAPTER - 1

1. INTRODUCTION

1.1. INTRODUCTION TO THE SYSTEM

Sign languages are one of the means of communication through body movements especially of the hands and arms used when spoken type of communication is not possible. It has become the core form of communication for the communities of deaf and mute people. Hence to overcome this barrier of communication among spoken people and sign language users. We have proposed our model which uses a popular deep learning technique called “Transfer Learning” .

Dumb and deaf persons experience difficulties connecting with computers in the workplace because they cannot hear them. It is also risky to travel places alone since they cannot hear cars, bikes, or other people approaching. They can't immediately adapt to their surroundings or respond to other people, and expressing oneself is difficult. Sign language has a long history in western societies as a visual language or technique of communication, dating back to the 17th century. Traditional gestures, mimics, hand signs, and figure spelling, as well as the use of hand position to represent letters of the alphabet, make up sign language. A sign can also represent an entire thought or statement. The major goal is to deliver speech and text output for deaf persons utilising hand gesture sign language without the use of any sensors in a smart method.

1.2. PROBLEM STATEMENT

The problem statement centres around the concept of a camera-based sign language recognition system for the deaf, which would transform sign language gestures to text and subsequently text to speech. Our goal is to create a user-friendly and straightforward solution.

Dumb individuals communicate via hand signs, thus normal folks have a hard time understanding what they're saying. As a result, systems that recognise various signs and deliver information to ordinary people are required.

1.3. OBJECTIVE

The main Objective of Sign Language Recognition (SLR) systems is to provide an efficient and accurate way to convert sign language into text or voice has aids for the hearing impaired for example, or enabling very young children to interact with computers (recognizing sign language), among others. Goal of a sign language detecting system is to provide a practical mechanism for normal and deaf

individuals to communicate through hand gestures. The proposed system will be used with a webcam or any other in-built camera that detects and processes indicators for recognition.

1.4. AIM OF THE PROJECT

Main Aim of Sign Language Recognition (SLR) systems is to provide an efficient and accurate way to convert sign language into text or voice has aids for the hearing impaired.

CHAPTER - 2

2. LITERATURE SURVEY

Literature review of our proposed system shows that there have been many explorations done to tackle the sign recognition in videos and images using several methods and algorithms.

Siming proposed a system having a dataset of 40 common words and 10,000 sign language images. To locate the hand regions in the video frame, Faster R-CNN with an embedded RPN module is used. It improves performance in terms of accuracy. Detection and template classification can be done at a higher speed as compared to single stage target detection algorithm such as YOLO.On the problem of RGB sign language image or video recognition in practical problems , the paper merges the hand locating network, 3D CNN feature extraction network and LSTM encoding and decoding to construct the algorithm for extraction. This paper has achieved a recognition of 99% in common vo- cabulary dataset.

Let's approach the research done by Rekha. which made use of YCbCr skin model to detect and fragment the skin region of the hand gestures. Using Principal Curvature based Region Detector, the image features are extracted and classified with Multi class SVM, DTW and non-linear KNN. The experimental result obtained were 94.4%for static and 86.4% for dynamic.

In Pigou L, a low cost approach has been used for image processing. The capture of images was done with a green background so that during processing, the green colour can be easily subtracted from the RGB colourspace and the image gets converted to black and white. The prototype has correctly recognised 92% of the sign gestures.

The paper by M. Geetha and U. C. Manjusha[7], make use of 50 specimens of every alphabets and digits in a vision based recognition of Indian Sign Language characters and numerals using B-Spine approximations. The region of interest of the sign gesture is analysed and the boundary is removed . The boundary obtained is further transformed to a B-spline curve by using the MaximumCurvature Points(MCPs) as the Control points. The B-spline curve undergoes a series of smoothening process so fea- tures can be extracted. Support vector machine is used to classify the images and the accuracy is 90.00%.

A similar work was done by J Huang [10].He created his own dataset using Kinect and got a total a total of 25 vocabularies which are used in everyday lives. He then applied a 3D CNN in which all kernels are also in 3D. The input of his model consisted of 5 important channels which are colour-r, colour-b, colour-g, depth and body skeleton. He got an average accuracy of 94.2%.

2.1. EXISTING SYSTEM

Sign Language Recognition (SLR) system, which is required to recognize sign languages, has been widely studied for years.The studies are based on various input sensors, gesture segmentation, extraction of features and classifcation methods.This paper aims to analyze and compare the methods employed in the SLR systems, classi cations methods that have been used, and suggests the most promising method for future research. Due to recent advancement in classifcation methods, many of the recent proposed works mainly contribute on the classifcation methods, such as hybrid method and Deep Learning. This paper focuses on the classifcation methodsused in prior Sign Language Recognition system. Based on our review, HMM based approaches have been explored extensively in prior research, including its modifcations.

This study is based on various input sensors, gesture segmentation, extraction of features and classification methods. This paper aims to analyze and compare the methods employed in the SLR systems, classifications methods that have been used, and suggests the most reliable method for future research. Due to recent advancement in classification methods, many of the recently proposed works mainly contribute to the classification methods, such as hybrid method and Deep Learning. Based on our review, HMM-based approaches have been explored extensively in prior research, including its modifications.Hybrid CNN-HMM and fully Deep Learning approaches have shown promising results and offer opportunities for further exploration.

The Disadvantages are

- Highly Expensive.
- Even if performed by any other ML techniques and With online, they Require tons of images to train for a single gesture

2.2. PROPOSED SYSTEM

The proposed system can recognize the static word sign and represent its label for better communication. The deaf person should submit a gesture or sign image to the system in the proposed system. The system uses a mat lab image processing technique to analyse the sign input and classifies it for recognised identification. When the input image matches the specifi ed dataset, it then starts the voice media through the system. In addition, the output will be displayed in text format. This is a working prototype for the conversion of sign language to speech and text.

Hence, We are implementing using a popular Deep Learning technique known as “Transfer Learning” which help us fasten the training process and also require very less number of images for training the model.

2.3. SCOPE OF THE PROJECT

Sign languages are developed primarily to aid deaf and dumb people. They use a concurrent and specific combination of hand movements, hand shapes and orientation in order to convey particular information.

CHAPTER - 3

3. ANALYSIS

The major step in analysis is to verify the feasibility of the proposed system. “All projects are feasible given unlimited resources and infinite time“. But in reality, both resources and time are scarce. Project should confirm to be time effective and should be optimal in their consumption of resources. This plays a constant role in approval of any project.

Three key considerations involved in the feasibility analysis are

- Technical Feasibility
- Operational Feasibility
- Economical Feasibility

3.1. Technical Feasibility

To determine whether the proposed system is technically feasible, we should take into consideration the technical issues involved behind the system. Android project uses the android-based technologies, which is rampantly employed these days worldwide. The world without the internet is incomprehensible today. That goes to render that the proposed system istechnically feasible.

3.2. Operational Feasibility

To determine the operational feasibility of the system we should take into consideration the awareness level of the users. This system is operationally feasible since the users are familiar with the android technologies and hence there is no need to gear up or train the personnel to use the cell phones. Also, the system (android phones) is very friendly and easy to use.

3.3. Economical Feasibility

To decide whether a project is economically feasible, we have to consider various factors as:

- Cost benefit analysis
- Long-term returns
- Maintenance costs

The proposed system is android based. It requires average computing capabilities and access to the internet, which are very basic requirements hence it doesn't incur any **additional economic overheads, which renders the system to be economically feasible.**

CHAPTER - 4

4. HARDWARE AND SOFTWARE REQUIREMENTS

4.1 Hardware Requirements

- The Hardware Interfaces Required are:
- Camera: Good quality,3MP
- Ram: Minimum 8GB or higher
- GPU: 4GB dedicated
- Processor: Intel Pentium 4 or higher
- HDD: 10GB or higher 7

4.2 Software Requirements

- Software requirements Operating System : Windows, Mac, Linux
- SDK: OpenCV ,TensorFlow, Keros, Numpy

CHAPTER - 5

5 . SYSTEM DESIGN

System design is the transition from a user-oriented document to programmers or database personnel. The design is a solution, specifying how to approach to the creation of a new system. This is composed of several steps. It provides the understanding and procedural details necessary for implementing the system recommended in the feasibility study. Designinggoes through logical and physical stages of development. Logical design reviews the present physical system, prepare input and output specification, details of implementation plan and prepare a logical design walkthrough.

The database tables are designed by analysing functions involved in the system and format of the fields is also designed. The fields in the database tables should define their role in the system. The unnecessary fields should be avoided because it affects the storage areas of the system. Then, in the input and output screen design, the design should be made user friendly. The menu should be precise and compact.

5.1 SOFTWARE DESIGN

In designing the software, the following principles are followed:

- Modularity and partitioning: software is designed in such a way that each system should consist of hierarchy of modules and serve to partition into separate function.
- Coupling: modules should have little dependency on the other modules of a system.
- Cohesion: modules should carry out the operations in a single processing function.
- Shared use: avoid duplication by allowing a single module which is called by other, that needsthe function it provides.

5.2 INPUT DESIGN

Considering the requirements, procedures are adopted to collect the necessary input data in most efficiently designed format. The input design has to be done keeping in view that, the interaction of the user with the system should be in the most effective and simplified way.Also, the necessary measures are taken for the following

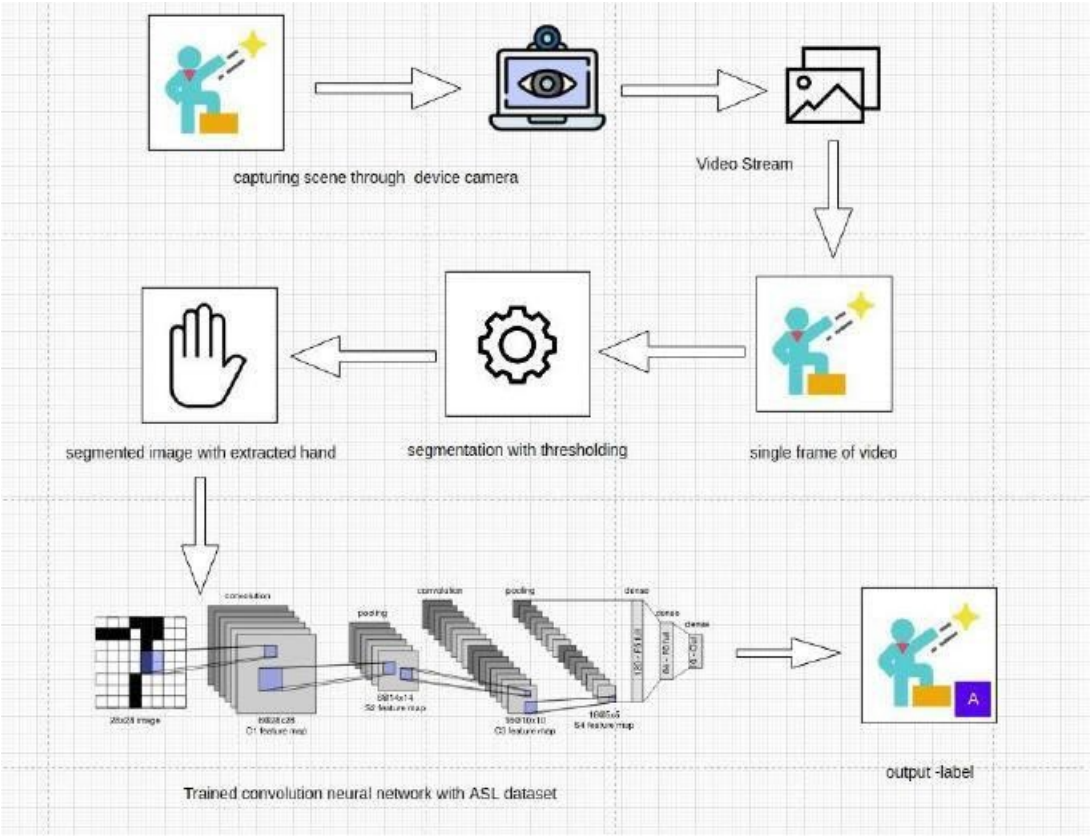
- Controlling the amount of input
- Avoid unauthorized access to the users
- Eliminating the extra steps
- Keeping the process simple
- At this stage the input forms and screens are designed.

5.3 OUTPUT DESIGN

All the screens of the system are designed with a view to provide the user with easy operations in a simpler and efficient way, with minimum key strokes possible. Important information is emphasized on the screen. Almost every screen is provided with no error and important messages and option selection facilitates. Emphasis is given for faster processing and speedy transactions between the screens. Each screen assigned to make it as much user friendly as possible by using interactive procedures. In other words, we can say that the user can operate the system without much help from the operating manual.

5.4 ARCHITECTURE

Convolutional neural networks (CNN) is a special architecture of artificial neural networks, proposed by Yann LeCun in 1988. CNN uses some features of the visual cortex. One of the most popular uses of this architecture is image classification. For example Facebook uses CNN for automatic tagging algorithms, Amazon — for generating product recommendations and Google — for search through among users' photos.



5.4.1 Architecture of Sign Language recognition System

5.5 UML DIAGRAMS

Unified Modelling Language

UML is an acronym that stands for **Unified Modeling Language**. Simply put, UML is a modern approach to modeling and documenting software. In fact, it's one of the most popular business process modeling techniques. It is based on **diagrammatic representations** of software components. As the old proverb says: "a picture is worth a thousand words". By using visual representations, we are able to better understand possible flaws or errors insoftware or business processes.

The elements are like components which can be associated in different ways to make a complete UML picture, which is known as diagram. Thus, it is very important to understand the different diagrams to implement the knowledge in real-life systems.

Any complex system is best understood by making some kind of diagrams or pictures. These diagrams have a better impact on our understanding. If we look around, we will realize that the diagrams are not a new concept but it is used widely in different forms in different industries.

We prepare UML diagrams to understand the system in a better and simple way. A single diagram is not enough to cover all the aspects of the system. UML defines various kinds of diagrams to cover most of the aspects of a system.

You can also create your own set of diagrams to meet your requirements. Diagrams are generally made in an incremental and iterative way.

USE CASE DIAGRAM

A use case diagram contains four components:

- The boundary, which defines the system of interest in relation to the world aroundit.
- The actors, usually individuals involved with the system defined according to their roles.
- The use cases, which are the specific roles played by the actors within and aroundthe system.
- The relationships between and among the actors and the use cases.

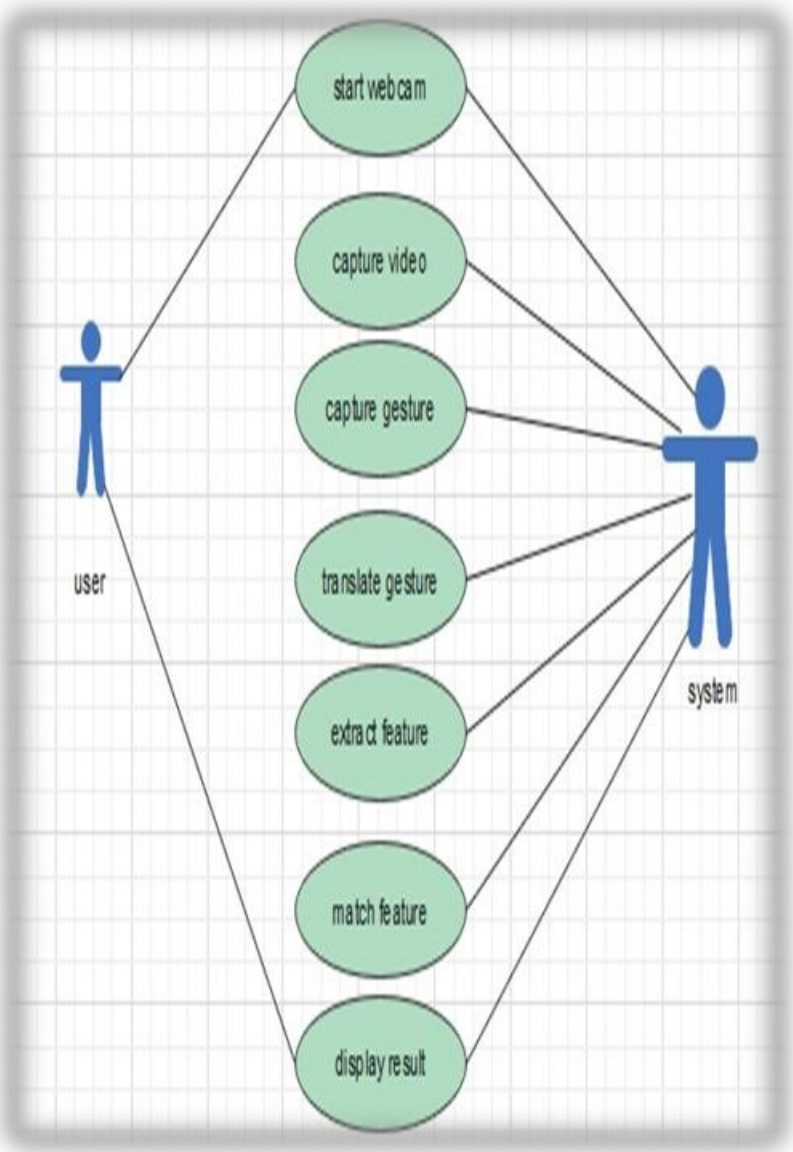


Fig:5.5.1 Use case Diagram for Real Time Sign Language Recognition

CLASS DIAGRAM

The purpose of the class diagram is to model the static view of an application. The class diagrams are the only diagrams which can be directly mapped with object oriented languages and thus widely used at the time of construction.

PURPOSE OF CLASS DIAGRAMS

- Analysis and design of the static view of an application.
- Describe responsibilities of a system.
- Base for component and deployment diagrams.
- Forward and reverse engineering

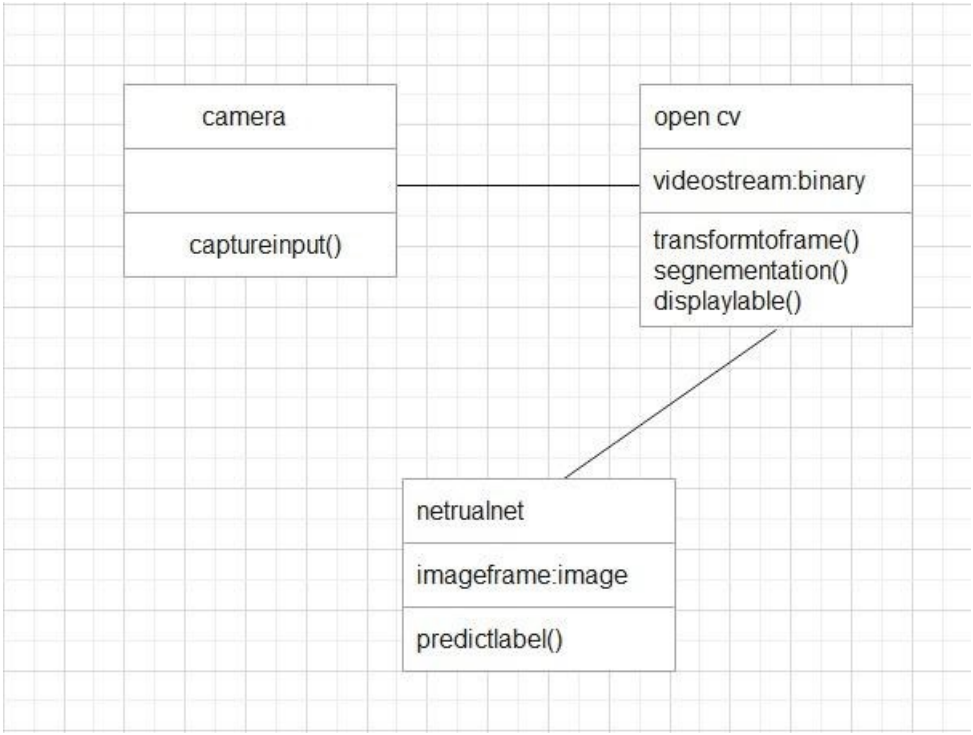


Fig:5.5.2 Class Diagram for Real Time Sign Language Recognition

ACTIVITY DIAGRAM

- It shows the flow of the various activities that are undergone from the beginning till the end.
- It consists of the activities that are held and carried out throughout the session from starting till the ending stage.

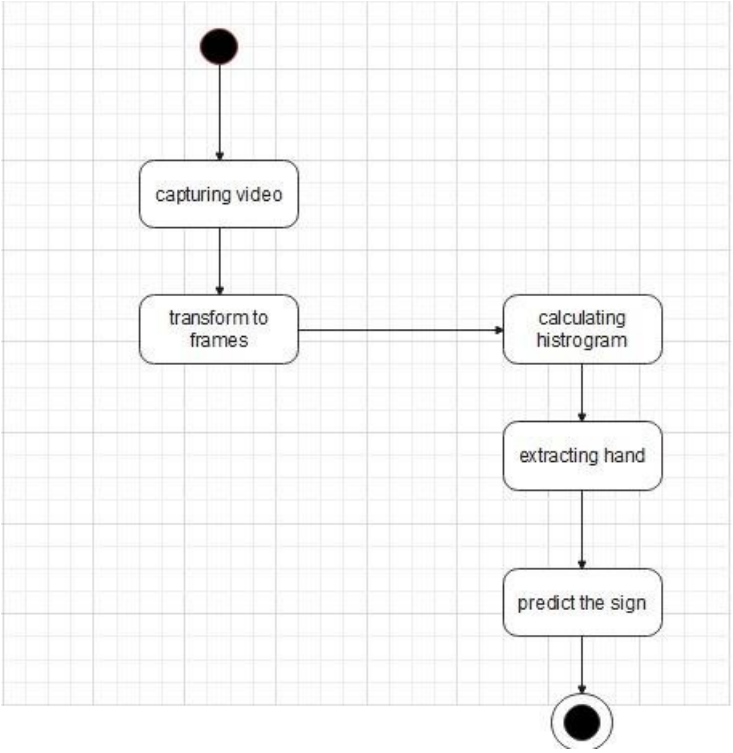


Fig:5.5.3 Activity Diagram for Real Time Sign Language Recognition

SEQUENCE DIAGRAM

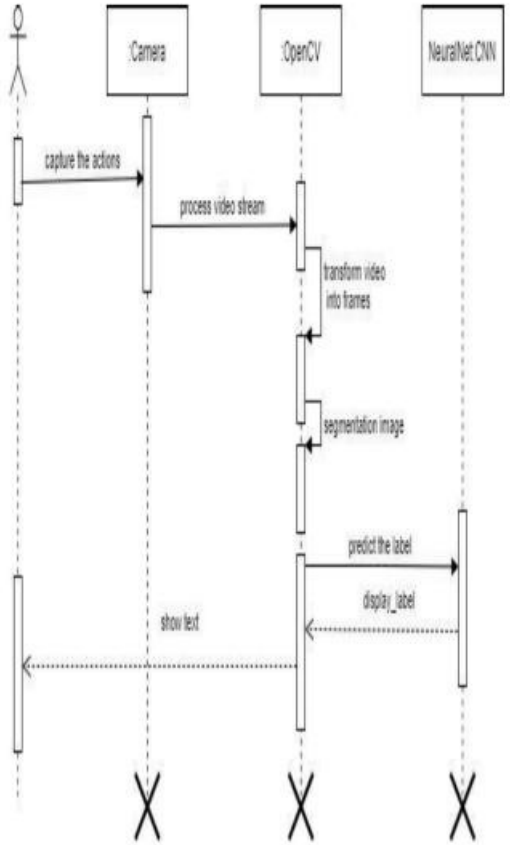


Fig:5.5.4 Sequence Diagram for Real Time Sign Language Recognition

- It shows the sequence of the steps that are carried out throughout the process of execution.
- It involves lifelines or life time of a process that shows the duration for which the process is alive while the steps are taking place in the sequential manner.
- Sequence diagram specifies the order in which the various steps are executed.