A

PROJECT REPORT

ON

"SIGN LANGUAGE INTERPRETATION SYSTEM"

SUBMITTED TO

SHIVAJI UNIVERSITY, KOLHAPUR

IN THE PARTIAL FULFILLMENT OF REQUIREMENT
FOR THE AWARD OF DEGREE BACHELOR OF TECHNOLOGY IN
COMPUTER SCIENCE AND ENGINEERING

SUBMITTED BY

18UCS087 Miss. Pise Uma Jagdish

18UCS110 Miss. Shingade Amruta Satish

18UCS114 Miss. Sidnale Prajakta Chandrakant

18UCS115 Miss. Sutar Priyanka Sanjay

18UCS121 Miss. Wadkar Aakanksha Deepak

UNDER THE GUIDANCE OF

PROF. V. V. Kheradkar

DKTE
Promoting Excellence in
Teaching, Learning & Research

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DKTE SOCIETY'S TEXTILE AND ENGINEERING INSTITUTE,

ICHALKARANJI

(An Autonomous Institute)

ACCREDITED WITH 'A+' GRADE BY NAAC

An ISO 9001-2015 Certified

2021-22

D.K.T.E. SOCIETY'S

TEXTILE AND ENGINEERING INSTITUTE, ICHALKARANJI
(AN AUTONOUMOUS INSTITUTE)

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that, project work entitled

"SIGN LANGUAGE INTERPRETATION SYSTEM"

is a bonafide record of project work carried out in this college by

18UCS087	Miss. Pise Uma Jagdish
18UCS110	Miss. Shingade Amruta Satish
18UCS114	Miss. Sidnale Prajakta Chandrakant
18UCS115	Miss. Sutar Priyanka Sanjay
18UCS121	Miss. Wadkar Aakanksha Deepak

is in the partial fulfillment of award of degree Bachelor in Engineering in Computer Science & Engineering prescribed by Shivaji University, Kolhapur for the academic year 2021-2022.

PROF. V. V. KHERADKAR
(Project Guide)

PROF. (DR.) D.V. KODAVADE
(HOD CSE DEPT.)

EXAMINER: _____

PROF.(DR.) P.V. KADOLE
(DIRECTOR)

_____

DECLARATION

We hereby declare that, the project work report entitled "**Sign Language Interpretation System**" which is being submitted to D.K.T.E. Society's Textile and Engineering Institute Ichalkaranji, affiliated to Shivaji University, Kolhapur is in partial fulfillment of degree Bachelor of Engineering (Computer Science and Engineering). It is a bonafide report of the work carried out by us. The material contained in this report has not been submitted to any university or institution for the award of any degree. Further, we declare that we have not violated any of the provisions under Copyright and Piracy / Cyber / IPR Act amended from time to time.

Student Name	PRN No	Signature
Miss. Pise Uma Jagdish	18UCS087	
Miss. Shingade Amruta Satish	18UCS110	
Miss. Sidnale Prajakta Chandrakant	18UCS114	
Miss. Sutar Priyanka Sanjay	18UCS115	
Miss. Wadkar Aakanksha Deepak	18UCS121	

ACKNOWLEDGEMENT

With great pleasure we wish to express our deep sense of gratitude to Prof. V. V. Kheradkar sir for his valuable guidance, support and encouragement in completion of this project report.

Also, we would like to take opportunity to thank our head of department Dr. D. V. Kodavade sir for his co-operation in preparing this project report.

We feel gratified to record our cordial thanks to other staff members of Computer Science and Engineering Department for their support, help and assistance which they extended as and when required.

Thank you.

18UCS087	Miss. Pise Uma Jagdish
18UCS110	Miss. Shingade Amruta Satish
18UCS114	Miss. Sidnale Prajakta Chandrakant
18UCS115	Miss. Sutar Priyanka Sanjay
18UCS121	Miss. Wadkar Aakanksha Deepak

ABSTRACT

In this project, we proposed to build a complete sign language interpretation system which will be a platform to learn Indian Sign Language (ISL). The system is divided into two parts that is conversion of sign video to audio format and conversion of speech to sign video. The system will be useful for normal, deaf and mute people which provide efficient and accurate way to convert sign language to text and audio as well as vice-versa

The models will be based on convert sign language to audio and video to speech. The CNN model will be used to extract frames from the video and to predict hand gestures. It is a multi-layered feed forward neural network mostly used in image recognition. They have applications in image and video recognition, recommender systems, image classification, image segmentation, frame extraction, natural language processing. CNNs are regularized versions of multilayer perceptron.

In this project GUI used as a platform. Which makes it easier to manage multiple tasks. At one single time users can work and view two or more programs. GUI uses a combination of technologies and devices to provide a platform that users can interact with, for the tasks of gathering and producing information.

INDEX

1. Introduction

Communication is a way of sharing or expressing information, feelings with someone. Wherever life exists, communication also exists. It is an unavoidable system. For successful and effective communication, it is very important to know about the communication process. Communication process involves sender, message, receiver, channel, feedback.

Overall, the world is having around 15 -20 % of the deaf and dumb population. A World Health Organization (WHO) report says around 63 million people in India suffer from either complete or partial deafness, and of these, at least 50 lakhs are children. India holds the world ‘s 20% of the Deaf and Dumb Population. They are facing many problems, because of lack of communication as deaf people are unable to listen and dumb people are unable to speak. Considering deaf and dumb people, it becomes tougher for them to communicate using natural language. So, they use sign language to communicate with themselves and with the entire world. Sign language is the daily language of communication between deaf and dumb people, which is the most comfortable, effective and natural way of communication between deaf and dumb people.

Sign language is defined as a way to communicate using hand gestures and symbols for words or letters of the alphabet, often used by deaf and dumb people. Sign language is composed of continuous gestures, which are used by deaf and dumb for their talking. It is a well-arranged gesture where every sign has a specific meaning assigned to it. It’s the main form of communication for the Deaf and Dumb community, but sign language can be useful for other groups of people as well. There is not any specific sign language used around the world. Similar to spoken language, sign languages expanded naturally through different groups of people interacting with each other, thus there are many varieties. There are somewhere around 300 different types of sign language used around the globe today. Most countries that distribute the same spoken language do not essentially have the same sign language as each other. English for example, has three varieties: American Sign Language (ASL), British Sign Language (BSL) and Australian Sign Language (Aslan).

There are different types of sign languages that are used throughout the world. Indian Sign Language (ISL), American Sign Language (ASL**),** British, Australian and New Zealand Sign Language (BANZSL), Chinese Sign Language (CSL), French Sign Language (LSF), Japanese Sign Language (JSL) Syllabary. Indian sign language (ISL)is the language which uses both hands to make signs while British and American sign language uses single handed sign language.

Sign language is a highly important communication tool for many deaf and dumb people. Sign languages are the adopted languages of the deaf people provides full access to communication. Even if sign languages are used basically by people who are deaf, they are also used by others people, similarly people who can hear but can’t speak i.e., deaf and dumb. Some child with autism spectrum disorder (ASD) conflict growing verbal communication. Learning a sign language can be an accessible communication tool for some child with ASD. It can be advice when the spoken word is actually impractical, such as talking subaquatic, talking through glass, from a distance, at a loud music show, and talking with your chunk. Sign languages can also let you talk to someone without interference others with noise. Accordingly, there is a demand of the systems which make the different signs and transmit the data to the normal people.

As the project deals with alphabets and numeric, greetings in Indian Sign Language, it can be extended to common expressions and also words which can be more effectual for deaf and dumb and normal people in communication and understanding. Now, India is developing at a rapid pace in

terms of digital and technological advances, this project could be one of the gradation stones where technology meets humanity and helps the deaf and dumb community.

As the sign language recognition systems provide a channel for communication between deaf and dumb people and normal people. This project proposing a system for transforming voice commands or text commands to signs of Indian sign language using hand gestures. Indian sign language does not have a standard database like American Sign Language database (ASL) which can be accessed from the internet. Hence, here created database contains Indian sign gestures. In this project the input to the system is either English speech or text, which is then processed and anyone can learn that sign gestured. Further, testing will be done by capturing gestures performed by people in front of camera and avatar will also be provided.

The Greek philosophers Plato, Socrates, and Aristotle remain the first person in history to create around sign language and deaf and dumb person of their association. The history of deaf and dumb people and their perception causes up deaf and dumb history. The Deaf and dumb perception is an ability that is intent on sign language and communication between one another. dissimilar other cultures the Deaf and dumb ability is not identical with any native land as it is a global culture. By a few, deafness. And dumbness may be viewed as a defect, but the Deaf and Dumb world sees itself as a language minority.

There are many applications of this sign language recognition system, as communication is the very important part of our life. Wherever, life exists, communication also exists. To make communication easy between deaf -dumb and normal people as well as normal people can also learn this sign gestures and he/she can train other deaf and dumb people, in such cases people can use this system to learn the sign language.

a. Problem Definition

A Complete Sign Language interpretation system which will be a platform to learn Indian Sign Language as well as to make communication easy between deaf-dumb and normal people.

This project proposing a system for transforming voice commands to video, contains Indian sign language using hand gestures as well as for converting video into the audio format.

b. Aim and Objectives

Aim:

The proposed system aims to provide a learning platform that helps deaf and dumb as well as normal people for learning sign language that will make communication easy.

Objectives:

- To build a machine learning model that will be able to classify various hand gestures for sign language recognition.
- To convert Indian Sign Language to Audio for normal people.
- To convert speech to Indian Sign Language for deaf and dumb people.
- To make the communication experience as complete as possible for both hearing and deaf people.

c. Scope and Limitations

Scope:

- By using this system, anyone can learn Indian Sign Language.
- It makes communication easy between Deaf -Dumb and Normal people.

Limitation:

- Limited to English sentences only. This system is limited to English sentences. It can take input in English language and will provide output in English language. Furthermore languages can be added.
- Standalone Project. This project is a standalone project. As this is the early-stage project, here only provided the video input directly, not recorded at real time.
- Limited Database. At this stage, this system is considering limited database containing some words and some sentences. Further, more sentences can be added into it.

d. Timeline of the project

This project was started in the month of August, 2021. The project approval of this project was done on 14 August, 2021. After 7 days of that gathered all the requirements which are necessary for the project. After that, analysis and design were done till 14th oct, 2021. The next part that is coding was completed. And then testing up to 20th April, 2022.

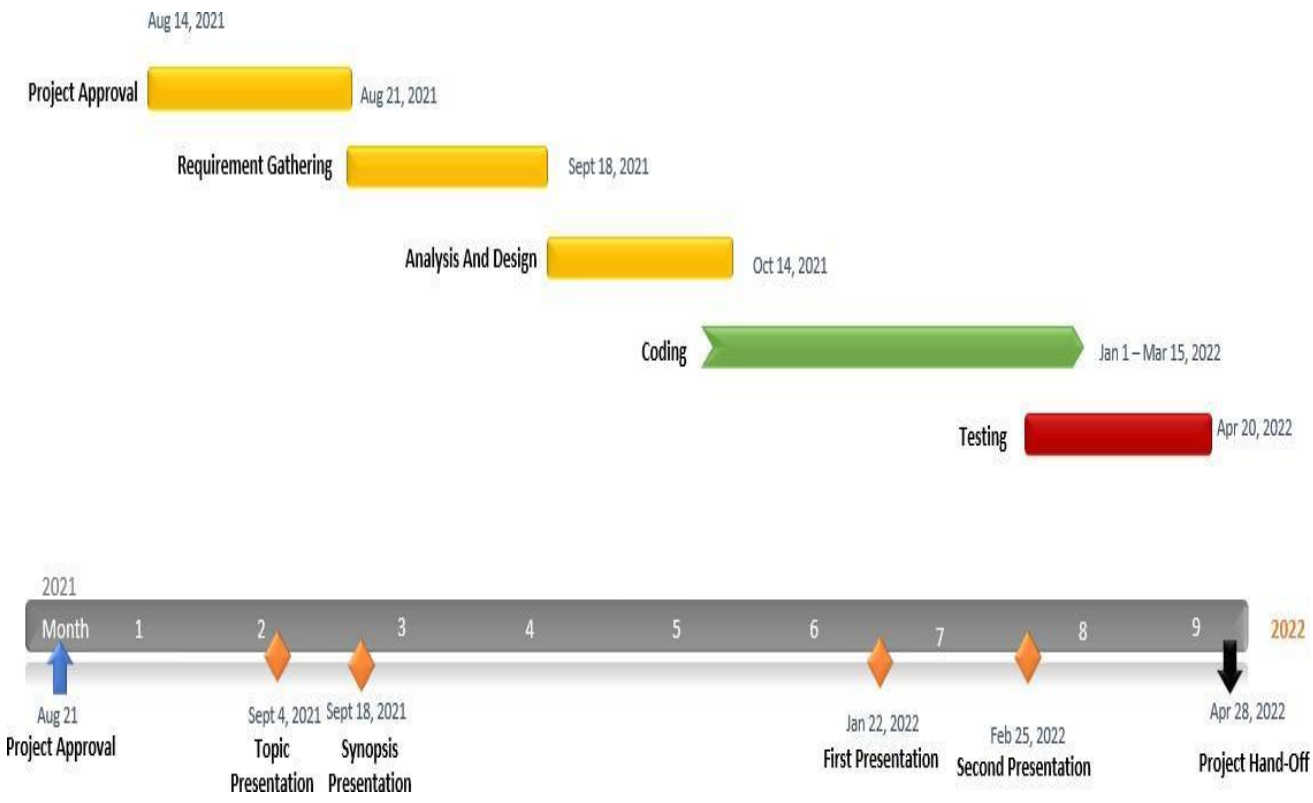


Figure. 1.1 Timeline

e. Project Management Plan

Sr. no	Work	Start Date	End Date	Duration	Status
1	Project Domain Finalization	11/08/2021	12/08/2021	7 Days	Completed
2	Problem Statement Finalization	17/08/2021	18/08/2021	7 Days	Completed
3	Previous Work Study	20/8/2021	16/09/2021	7 Days	Completed
4	Requirement Analysis, SRS	19/09/2021	29/10/2021	14 Days	Completed
5	Technology Related Study	2/11/2021	9/11/2021	7 Days	Completed
6	Model Identification	22/01/2022	1/02/2022	7 Days	Completed
7	Model Building (CNN)	4/02/2022	7/02/2022	7 Days	Completed
8	Testing Of the Model	8/02/2022	15/02/2022	7 Days	Completed
9	Model building	20/02/2022	25/02/2022	5 Days	Completed
10	Testing and validation	15/02/2022	28/02/2022	13 Days	Completed
11	Graphical User Interface	14/02/2022	24/02/2022	10 Days	Completed
12	Report	03/04/2022	23/04/2022	20 Days	Completed

f. Project Cost

The software is built using open-source software and no any extra hardware is used other than those present in the system. Hence the cost of the project includes only the cost of the machine. The cost depends on amount of hours you have spent on the project and the amount per hours.

Line of code: To develop the system 4585 lines of codes are required.

KLOC: KLOC is the estimated size of the software product indicates in Kilo Lines of Code.

KLOC = LOC / 1000

= 4585 / 1000

= 4.585

Effort: The effort is only a function of the number of lines of code and some constants evaluated according to the different software systems.

$E = a (KLOC)^{b}$

$= 2.4 (4.585)^{1.05}$

$= 2.4 * 4.9477$

$= 11.8745$

Time: The amount of time required for the completion of the job, which is, of course, proportional to the effort put in. It is measured in the units of time such as weeks, months.

Time = c (Efforts)^{d}

$= 2.5 (11.8745)^{0.38}$

$= 2.5 * 2.5606$

$= 6.4015$

Persons Required: Persons required is nothing but effort divide by time.

Persons Required = Efforts / Time

$= 11.8745 / 6.4015$

~ 1.858

Other Cost factors:

Hardware/Software	**Cost (Approximately) in Rs.**
Computer System with i7 7th generation or above	70000
Python IDE(Spyder Environment) to run machine learning models	0
16GB or above RAM and graphics card	6000
Electricity	560
Internet	2400
Mobile	15000
Total Cost	93960

2. Background & Literature Overview:

a. Literature overview

P. V. V. Kishore and P. Rajesh Kumar [1] proposed Video Based Indian Sign Language Recognition System. In this paper, A video segmentation technique is proposed which detects shapes of various hand signs and head movement. Fourier Descriptors is used to extract shape outline with minimum number of pixels for an image frame without losing shape information. They chose Takagi-surgeon-kang (TSK) or simply surgeon type Fuzzy inference system for pattern classification. The total 80 signs are used for testing with 10 different signers and the recognition rate is up to 96%.

Pratik Lehar, Rathan G N[2] proposed Indian Sign Language Translation using Deep Learning. In this paper, they present three deep architectures to translate a given video sequence containing the Indian Sign Language sentence to English Language sentence. They used three approaches. First using an LSTM based Sequence to Sequence model (Seq2Seq), second using an LSTM based Seq2Seq model utilizing attention, third using an Indian Sign Language Transformer.

Kusunoki Kirori Dutta, Satheesh Kumar Raju K, Anil Kumar G S, Sunny Arkia Swarna B[3] from Department of Electrical and Electronics Engineering, M.S.Ramaiah Institute of Technology, Bengaluru proposed Double Handed Indian Sign Language to Speech and Text. In this paper they proposed system to provide speech to speechless, the double handed Indian Sign Language is captured as a series of images and it's processed with the help of MATLAB and then it's converted to speech and text.

Siming He [4] from Ridley College, St. Catharine's, Canada proposed Research of a Sign Language Translation System Based on Deep Learning, in which he proposed hand locating and sign language recognition of common sign language based on neural network. To solve the problem of RGB sign language image or video recognition and to build the recognition algorithm, this paper combines hand locating network, 3D CNN feature extraction network and LSTM encoding and decoding.

Shravani K, Shree Lakshmi A, Sri Geethika M , Dr. Sapna B Kulkarni [5] from Computer Science, RYM Engineering College, VTU University, Belagavi, India proposed Indian Sign Language Character Recognition. The system consists of several steps which are Image collection, Image pre-processing (segmentation), Feature extraction. They used alphabets and numeric data for sign language. Image pre-processing and Bag of words model is used to extract the features from the collected segmented data. To map the alphabets with images histograms are generated. In the last step, these features will be fed to Support Vector Machine (SVM) model for the classification of images.

Debashish Das Chakradhar, Pradeep Kumar, Shubham Mandal, Paratha Pratima Roy, Masakazu Imamura and Byung-Gyu Kim [6] from Department of Computer Science and Engineering, Indian Institute of Technology Roorkee proposed 3D Avatar Approach for Continuous Sign Movement Using Speech/Text. In this paper there are three modules. Firstly, the

input speech is converted into an English sentence using the IBM-Watson service. Then these English texts are converted into the corresponding ISL sentence using regular expression and script generator. Finally, each word of the ISL sentence is translated into an equivalent sign movement, which is represented by a 3D avatar.

López-Ludeña, V.; San-Segundo, R.; Morcillo, C.G.; López, J.C.; Muñoz, J.M.P. [7] proposed Increasing adaptability of a speech into sign language translation system. In this paper, system is made up of a speech recognizer (for decoding the spoken utterance into a word sequence), a natural language translator (for converting a word sequence into a sequence of signs belonging to the sign language), and a 3D avatar animation module (for playing back the signs. For rapidly design of the required signs, the avatar animation module includes a new editor. the effort for adapting the system to a new domain has been reduced more than 50% while the whole translation presents a SER (Sign Error Rate) lower than 10% and a BLEU higher than 90%.

b. Critical appraisal of another people work

I. P. V. V. Kishore and P. Rajesh Kumar [1] proposed Video Based Indian Sign Language Recognition System. In this paper, A video segmentation technique is proposed which detects shapes of various hand signs and head movement. The work was accomplished in four stages. In video pre-processing, to reduce the complexity, RGB indexed video frames are converted to gray scale indexed video frames. to remove the high frequency noise and smoothen the image frames, Gaussian low pass filter is applied to the gray scale image. By fusing discrete wavelet transform (DWT) and canny operated images segmentation is done. Fourier Descriptors is used to extract shape outline with minimum number of pixels for an image frame without losing shape information. They chose Takagi-sugeno-kang (TSK) or simply sugeno type Fuzzy inference system for pattern classification. The total 80 signs are used for testing with 10 different signers and the recognition rate is up to 96%.

II. Siming He [2] from Ridley College, St. Catharine's, Canada proposed Research of a Sign Language Translation System Based on Deep Learning, in which he proposed hand locating and sign language recognition of common sign language based on neural network, it includes: 1. Faster R-CNN is established for hand locating network to recognize the sign language video or the part of the hand in the picture, and the result of recognition is handed over to subsequent processing; 2. A 3D CNN feature extraction network and a sign-language recognition framework are constructed for the sign language images of sequences based on long- and short-time memory (LSTM) coding and decoding network. 3. To solve the problem of RGB sign language image or video recognition and to build the recognition algorithm, this paper combines hand locating network, 3D CNN feature extraction network and LSTM encoding and decoding.

III. Shravani K, Shree Lakshmi A, Sri Geethika M, Dr. Sapna B Kulkarni [3] from Computer Science, RYM Engineering College, VTU University, Belagavi, India proposed Indian Sign Language Character Recognition. In this paper they used Indian Sign Language gestures taken from static pictures. The system consists of several steps which are Image collection, Image pre-processing (segmentation), Feature extraction. They used alphabets and numeric data for sign language. Image pre-processing and Bag of words model is used to extract the features from the

collected segmented data. To map the alphabets with images histograms are generated. In the last step, these features will be fed to Support Vector Machine (SVM) model for the classification of images.

IV. Debashish Das Chakradhar, Pradeep Kumar, Shubham Mandal, Paratha Pratima Roy, Masakazu Imamura and Byung-Gyu Kim [4] from Department of Computer Science and Engineering, Indian Institute of Technology Roorkee proposed 3D Avatar Approach for Continuous Sign Movement Using Speech/Text. In this paper there are three modules. Firstly, the input speech is converted into an English sentence using the IBM-Watson service. Then these English texts are converted into the corresponding ISL sentence using regular expression and script generator. Finally, each word of the ISL sentence is translated into an equivalent sign movement, which is represented by a 3D avatar.

c. Investigation of current project and related work

After reviewing found that complete Sign Language interpretation system is not yet built. In this project, the system proposed to build a complete sign language interpretation system which will be a platform to learn Indian Sign Language (ISL). It consists of datasets of sentences.

The system is divided into two parts i.e., conversion of sign videos to audio format and conversion of speech to sign video. This system will be useful for normal, deaf and mute people which provide efficient and accurate way to convert sign language to text and audio and vice versa.

3. Requirement analysis

a. Requirement Gathering

Module-1 (Sign Video to Audio)

1. As a User, user needs to create the sign videos first, so that system can perform further conversions like sampling of video, extracting the frames from the videos and segmentation etc. and get the correct output of the video in audio format.
2. As a system, system needs to divide the videos into frames, so that system can easily perform the further operations.
3. As a system, system needs to perform segmentation, so that system will resize and enhance the frames for better quality of image.
4. As a system, system needs to train the model, so that system will give proper output as per the user's requirement either audio or video.
5. As a system, system needs to save the model, so that system will not need to train model again and again in future use.
6. As a system, system need to predict the probability of input video by comparing the frames of that video to get the output as a speech.

Module-2 (Speech to Sign Video)

1. As a User, user needs to give input as a speech, so that system will give the output in video format.
2. As a system, system first convert speech into text format, text must be in lowercase and also without any punctuation mark.
3. As a system, system needs to search video from the training folder, so that the system can get the file from the folder.
4. As a system, system then compare the text with index file, if the text matches with index file, then video will play. And if text does not match then it will not play video.

b. User Stories

Module-1 (Sign Video to Audio)

1. User first needs to create the sign videos with mobile phone.
2. System can divide the videos into frames with frame rate of 0.25.
3. System can resize the frames with size 300 * 350 pixel and enhance the quality of image.
4. System then train the model and save the model for further use.
5. After the training the model it will predict the output for the input video.
6. Then system will give the output in the audio format.

Module-2 (Speech to Sign Video)

1. User needs to give input in speech format.
2. System then convert speech into text format which used for system compatibility.

3. Then text will search for the index file in the folder and then it will compare with the index file with the text to get output in video format.
4. And then it will give the output in sign video format.

c. Requirement Specification

Module-1 (Sign Video to Audio)

No.	Requirement	Essential /Desirable	Description of requirement
RS1	The user needs to create the sign videos.	Essential	Sign videos are the first requirement. These videos are further used converted into frames to get the correct output.
RS2	The System needs to divide the frames.	Essential	Sign videos divided into frames with frame rate of 0.25.
RS3	The system needs to perform segmentation.	Essential	The frames need to be resized and enhance the quality of image/frames. After the segmentation process image size will become 300 * 350 pixel.
RS4	The system needs to train the CNN model.	Essential	After segmentation of images the quality of the image get increased then this is useful for the training the model. This model is mainly used to process pixel data and used in image recognition and processing.
RS5	The system saves the model.	Essential	After training the model, it needs to save the model for further use.
RS6	The system predicts the output.	Essential	After segmentation of the frames the system then predict the output for the input.
RS7	The system converts the video into speech.	Essential	After the segmentation and training of model, the will converts the sign videos into speech.

Module-2 (Speech to Sign Video)

No.	Requirement	Essential /Desirable	Description of requirement
RS8	The user gives the input as a speech.	Essential	The next part of the project is to convert speech to sign video. So, it is important to convert speech to video. If user wants to give input as a speech, then this method can be used.
RS9	The system needs to convert speech to text.	Essential	The conversion of speech to text is important step because it makes documentation flexible.
RS10	The searching and comparing of index file.	Essential	After getting the text format of the speech, it is easy to do further processes. The text then searches for index file and compare with index file.
RS11	The system converts speech into video format.	Essential	After that it will give the output in speech format.

d. Use Case Diagram

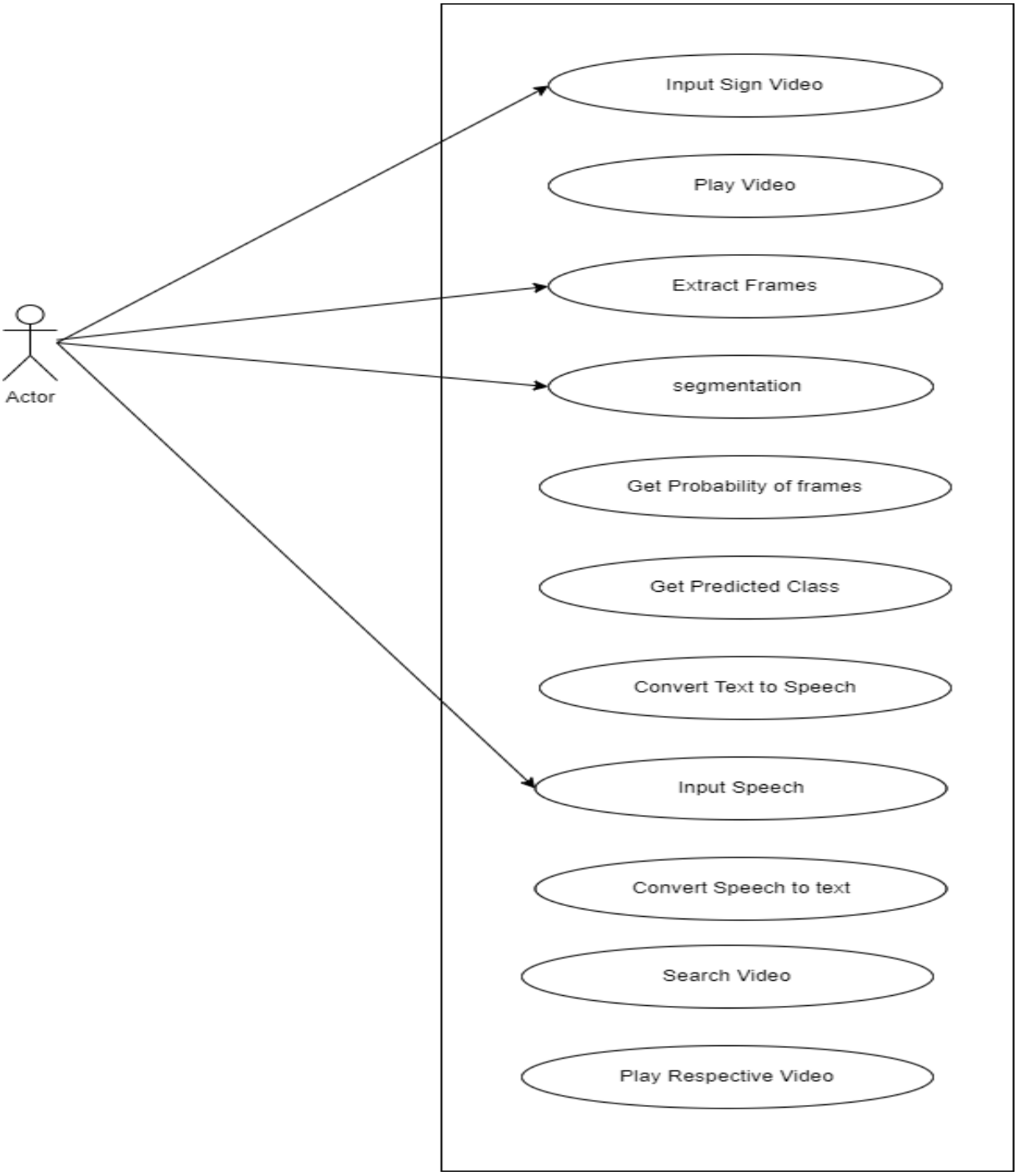


Figure 3.1: Use case Diagram

The above use case diagram consists of:
1.Extract frames:
User input video and extract the frames with framerate of 0.25.
2. Segmentation:
The frames need to be resized and enhance the quality of image/frames. After the segmentation process image size will become 300 * 350 pixel.

3. Get probability of frames:
Probability calculated for each of the sign and class also predicted.
4. Convert text to speech:
After prediction of class system give output in text format and this can be converted to speech.
5. Search video:
For converting speech to video, system first convert speech to text and then it search for video.

4. System design

a. Architectural Design

The user creates the sign videos and store it as dataset. After that the frames get extracted from the videos. After that the frames get extracted from the video with framerate of 0.25. Then the frames get resized and enhanced. Then model get trained and saved for further use. The output of video can be given in speech format.

Then second way is that user give input in speech format. That speech gets converted into text format. Then the video can be searched from index file and after searching it compared with index file. And then the output came in video format.

According to overall working of code, the architectural diagram can be given as follows:

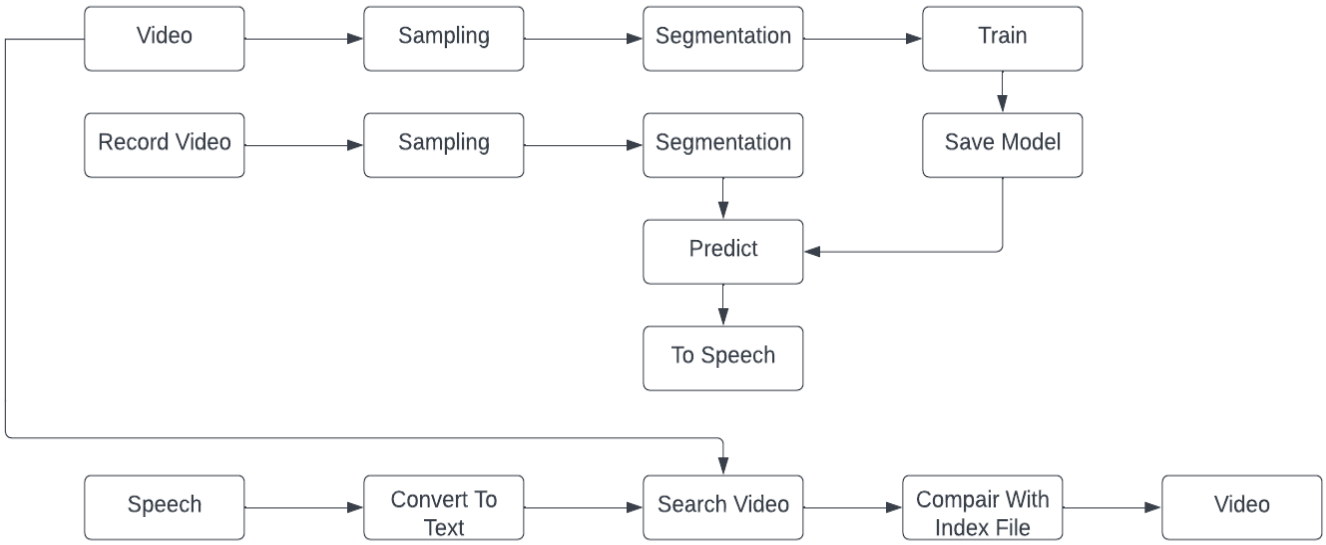


Figure 4.1: Architecture Diagram

The above architecture diagram consists of:

1. Sampling: User give input video and extract the frames with framerate of 0.25.

2. Segmentation: The frames need to be resized and enhance the quality of image/frames. After the segmentation process image size will become 300 * 350 pixel.

3. Train: After segmentation of images the quality of the image get increased then this is useful for the training the model. This model is mainly used to process pixel data and used in image recognition and processing.

4 Save model: After training the model, it needs to save the model for further use.

5. Predict model: After segmentation of the frames the system then predict the output for the input.

6. Convert to speech: After the segmentation and training of model, the will converts the sign videos into speech.

7. Convert speech to text: The next part of the project is to convert speech to sign video. So, it is important to convert speech to video

8. Search video: For converting speech to video, system first convert speech to text and then it search for the video.

b. User Interface Design

The User Interface diagram is an extended diagram type that provides a set of wire framing toolboxes with a rich palette of user interface elements for Android and Apple devices, as well as for web pages and dialogs.

This diagram indicates the main activities performed by the user.

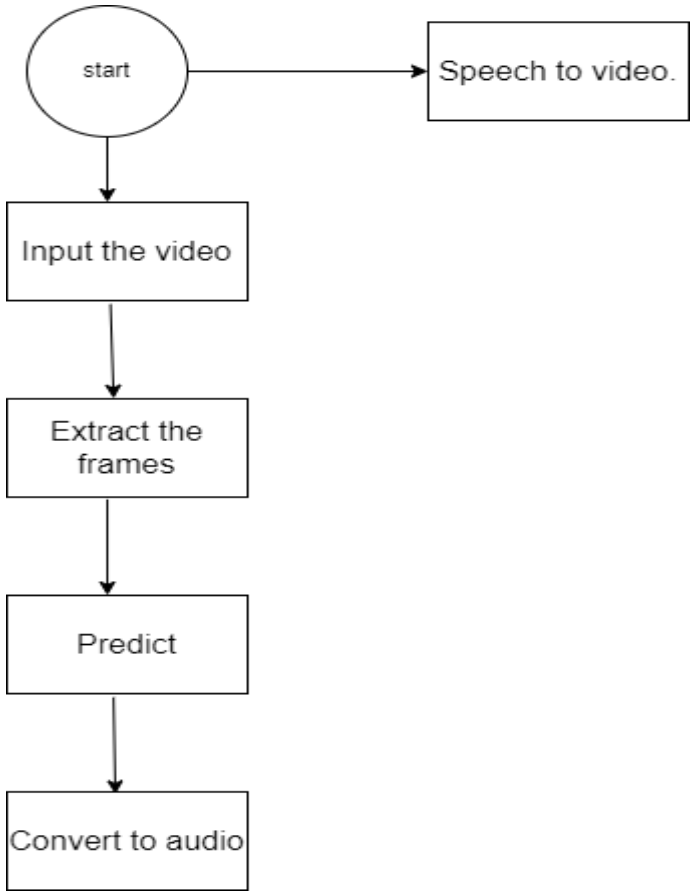


Figure. 4.2: User Interface Design

User interface diagram consists of following:

User first give the input in speech or video form and accordingly the system provides output. If system got input as video, then it will extract the frames, predict the class for the frames and convert to audio. And if it is providing speech as input then it will convert to text first and search for video and give output.

c. Algorithmic description of each module

1. Module (Sign video to audio)

Step 1: Start
Step 2: Browse Input as a Sign video from user.
    If valid video
        Go to Step 3.
    End
    Else
        Go to Step 2.
    End
Step 3: Extract video frames.
    Resize the frames.
    Display number of frames extracted.
Step 4: Predict video.
    Read model from json file.
    Print maximum predicted column.
    Print action corresponding to maximum predicted column.
Step 5: Convert predicted video to audio format.
    Save the converted audio in a mp3 file
    Play mp3 file.
Step 6: End.

2. Module (Speech to Sign video)

Step 1: Start
Step 2: Get Input as speech.
Step 3: Convert speech to text.
Step 4: Convert text into lowercase.
Step 5: Separate text using commas.
Step 6: Print Separated text.
Step 7: Search separated text in the directory.
    If text found
        Play video
    End
    Else
        Video not found.
    End

d. System Modeling

i. Dataflow Diagram:

A data flow diagram (DFD) maps out the flow of information for any process or system.

User first create the video then it goes into system for the extraction, segmentation, prediction, conversion into speech and gives output in audio format.

In another way, the system get input in speech format and then searching, comparing and converting to text to video format can be done in dataflow diagram.

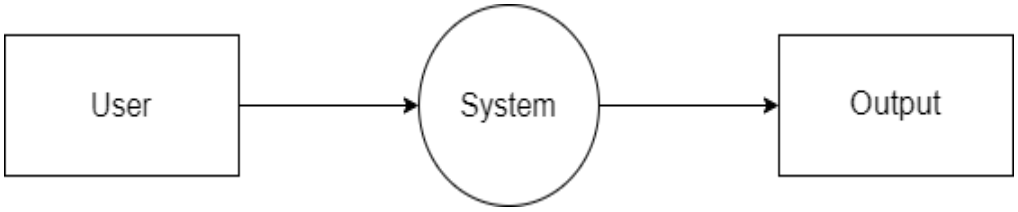


Figure 4.3: DFD0

Dataflow diagram gives information about input and output to the system. First, user give input then system work according to that and gives output.

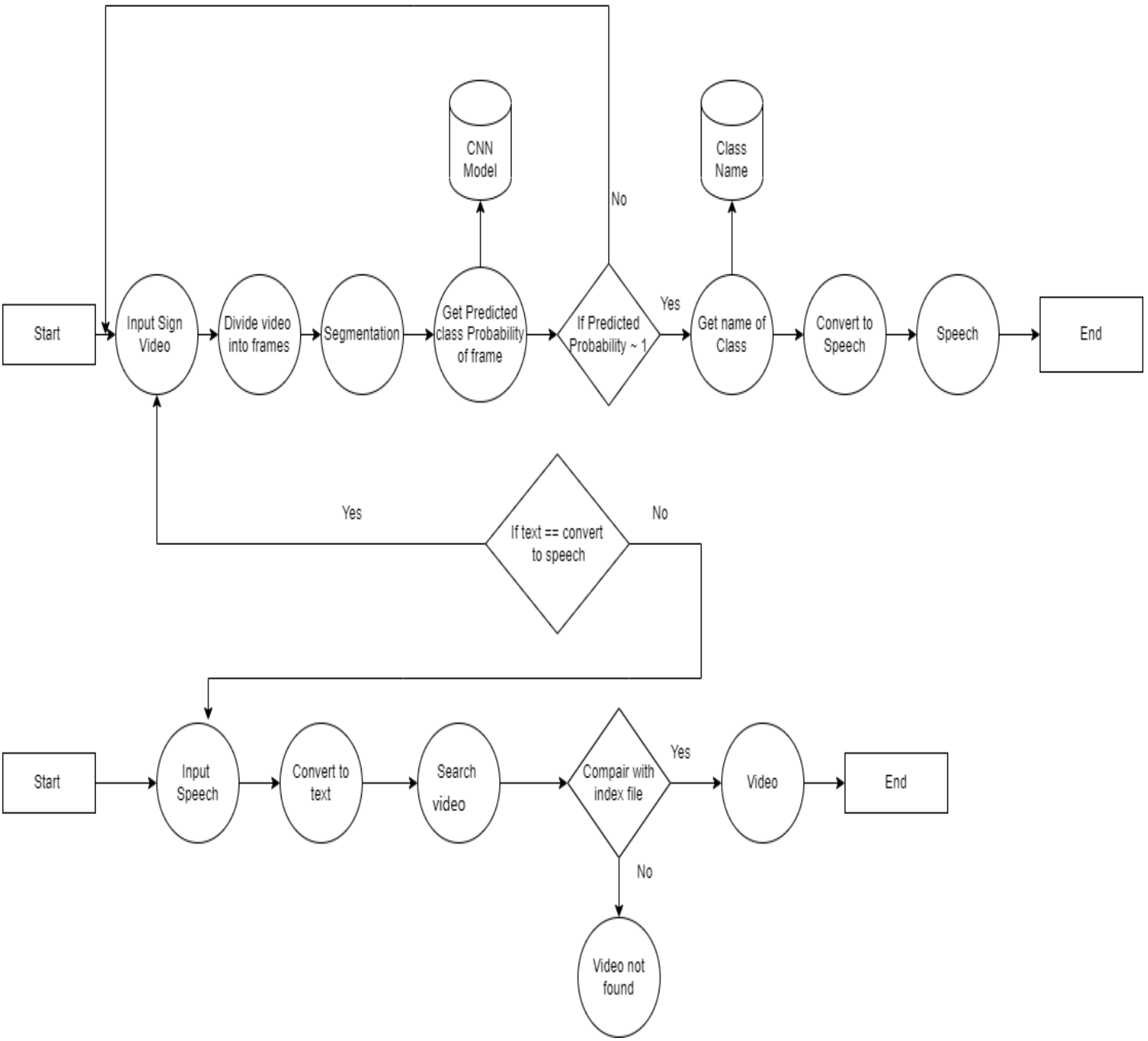


Figure 4.4: Dataflow Diagram

The above dataflow diagram consists of: Sampling, Segmentation, Train, Save model, Predict model, Convert to speech, Convert speech to text, Search video.

ii. Sequence Diagram

The Sequence Diagram shows the object interaction arranged in time sequence**.** The solid bars show the active time of the module**.**

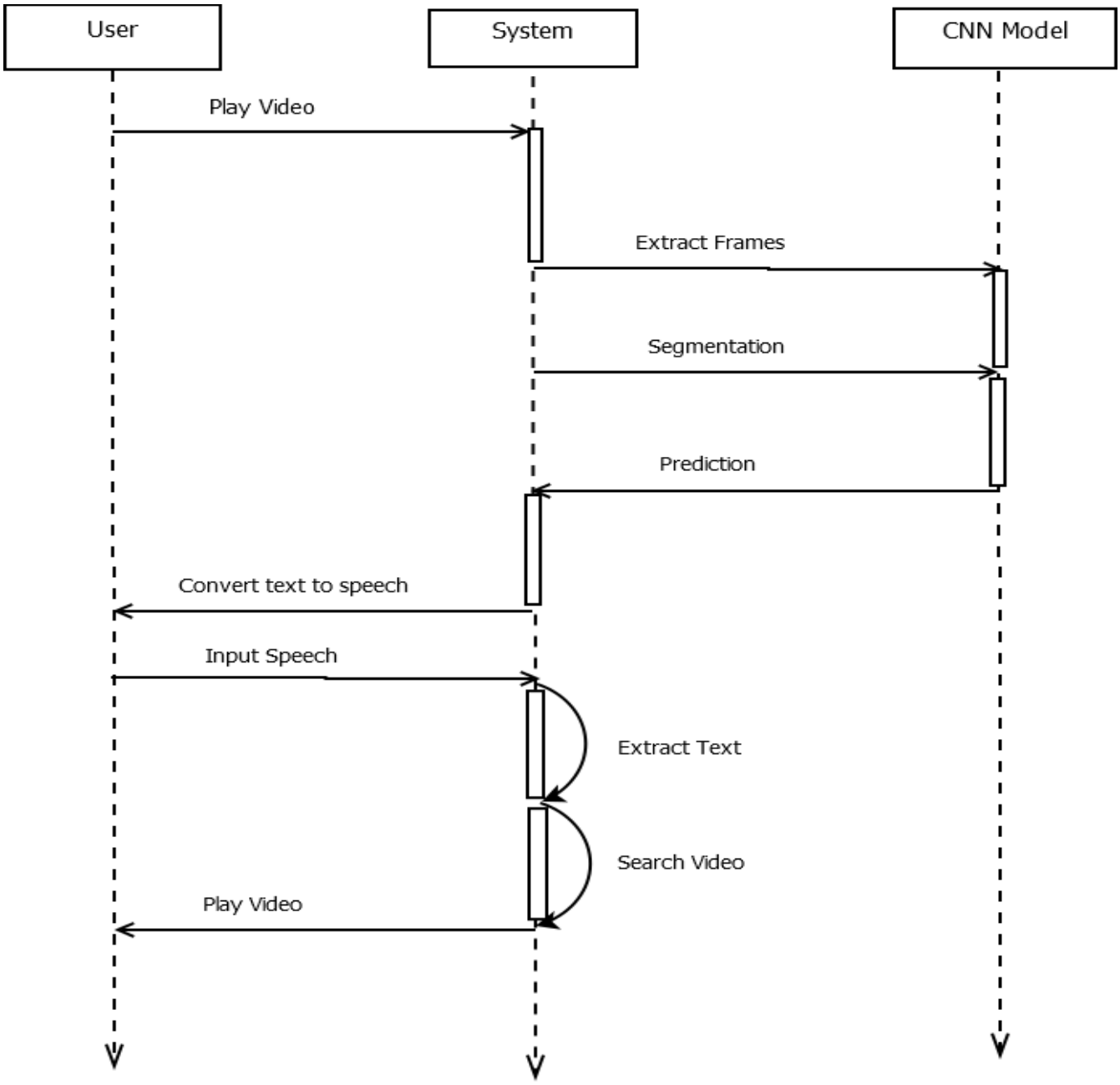


Figure 4.5: Sequence Diagram

According above sequence diagram, there are three main parts of the system. UI, Model of sign language recognition & Database. As data i.e., videos are stored in database.

For video to audio, the frames get extracted then the prediction of class done and then the video get converted to audio. Similarly, the speech to video also done by comparing the video with stored video.

iii. Activity Diagram

Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system.

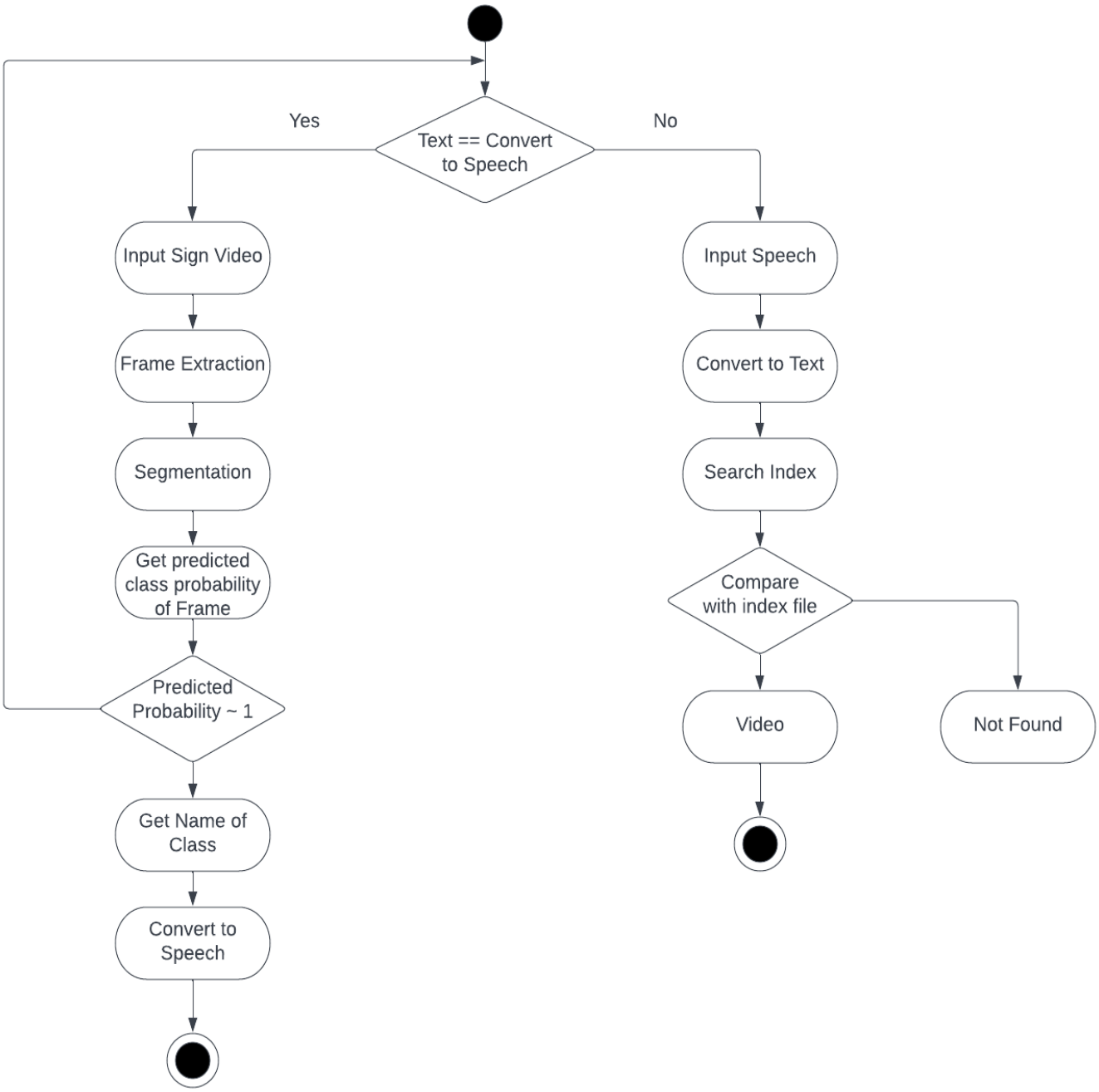


Figure. 4.6: Activity Diagram

According to above activity diagram, videos are created by user and stored in database. For video to audio, the frames get extracted then the prediction of class done and then the video get converted to audio. Similarly, the speech to video also done by comparing the video with stored video

iv. Component Diagram

A component diagram breaks down the actual system under development into various high levels of functionality. Each component is responsible for one clear aim within the entire system and only interacts with other essential elements on a need-to-know basis.

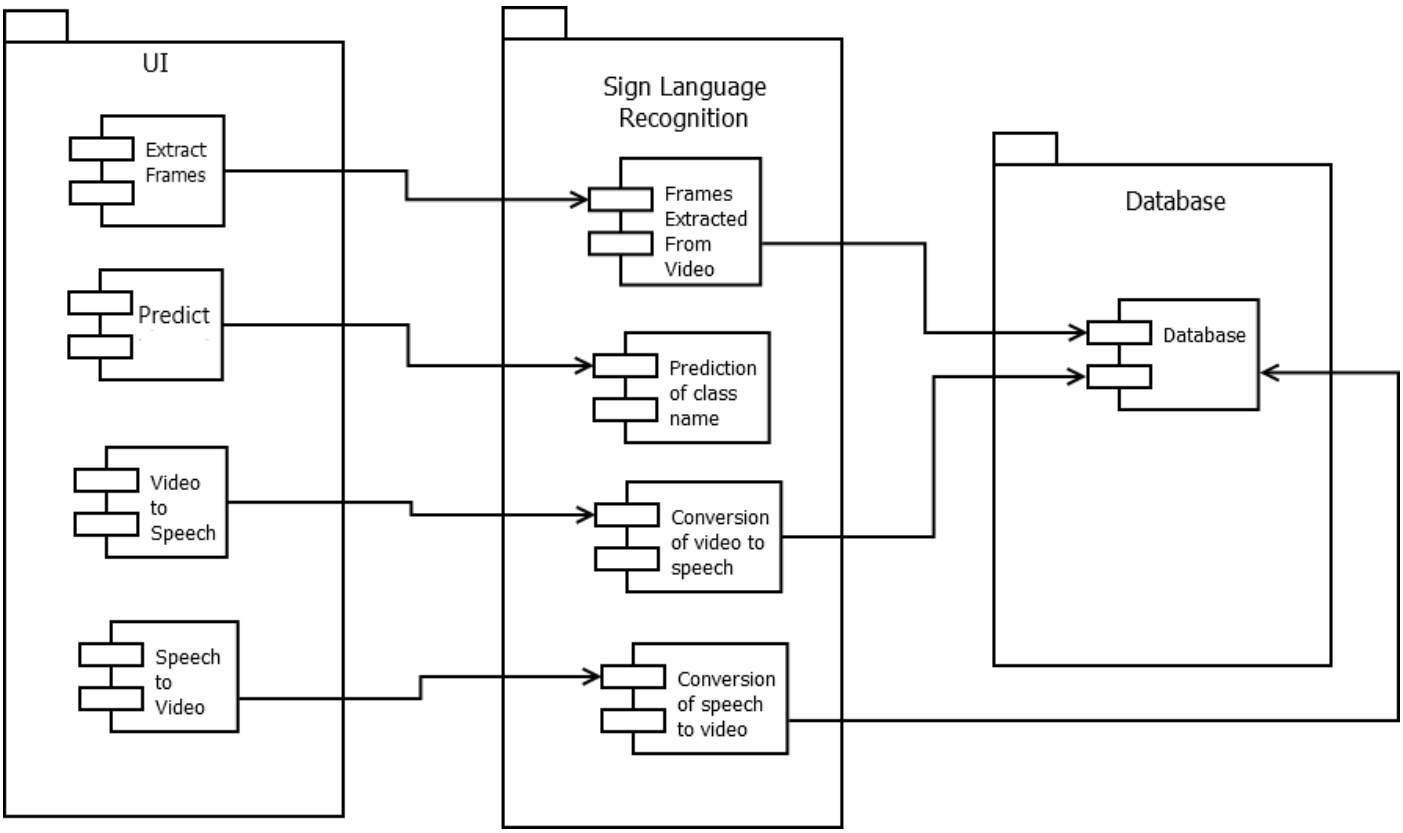


Figure 4.7: Component Diagram

According above component diagram, there are three main parts of the system. UI, Model of sign language recognition & Database. As data i.e., videos are stored in database. For video to audio, the frames get extracted then the prediction of class done and then the video get converted to audio. Similarly, the speech to video also done by comparing the video with stored video.

5. Implementation

a. Environmental Setting for Running the Project

1. Install python 3.8 on your system.

2. You can also check whether the installation was successful by typing python – V in Command Prompt. The output should display your installed version of Python.

3. Install following packages:

1. **TensorFlow**: It is used to build and train CNN models. Use command to install TensorFlow- pip install tensorflow.
2. **Gtts**: It is googled text to speech API which reads text aloud. Use command to install gtts- pip install gtts.
3. **Os**: The os module in python provides functions like identifying the current directory, fetching its contents, removing a directory, etc. Use command to install os- pip install os.
4. **Opencv**: It is used for all sorts of image and video analysis like recognition and detection. Use command to install opencv- pip install opencv.
5. **Numpy**: This library used to convert lists into numpy arrays. Use command to install - pip install numpy.
6. **Pandas**: It is used for exploring and cleaning data. Use command to install- pip install pandas.
7. **Pyttsx3**: It is used to convert Text into speech. Use command to install- pip install pyttsx3.
8. **Tkinter**: Tkinter is the standard GUI library for Python. Use command to install Tkinter- pip install tkinter.
9. **Matplitlib**: Matplotlib is used for plotting graphs of accuracy and for data visualization. Use command to install matplotlib-pip install matplotlib.
10. **Datasets**: Datasets is a library for easily accessing and sharing datasets. Use command to install datasets- pip install datasets.
11. **Sklearn**: It consists of lots of tools for machine learning. Used to build ML models. Use command to install sklearn- pip install sklearn.
12. **PIL**: PIL is library is used to open image content into an array. Use command to install PIL- pip install pillow.
13. **Speech Recognition**: It offers easy audio processing and microphone accessibility. Use command to install Speech Recognition- pip install Speech Recognition.
14. **Pyaudio**: Pyaudio is used to play and record audio. Use two commands to install Pyaudio- pip install pipwin and pipwin install pyaudio.

4. Record the video through mobile as a good quality webcam.

b. Detailed Description of Methods

1. **browse_button ()**: This method is used to select a directory to browse the video. It will print the name of the selected video file from the directory.

2. **play_video ()**: This method is used to select the video which will be played. With the help of os module it will play the browsed video.

3.**extract button ()**: This method extracts frames. Initialize variable frame rate to 0.25. cvtColor() method used to convert image frames into grayscale. Resize the frames using resize() method. Initialize variables as y = 300, h = 350, x = 50, w = 300. Frames are resized to 350*650 pixels. Declare variable count to 0 to count the number of frames extracted. Declare variable success to True which tells whether frames are extracted or not.

4.**most_frequent (List)**: Input to this method is a list of probabilities of features. This method returns a column index which has maximum probability compared to others.

5.**predict_button ()**: This method takes output from most_frequent method to print maximum predicted column. Also prints action corresponding to the maximum predicted column.

6.**videoToSpeech_button ()**: This method converts predicted video to audio format. Declare variable mytext which is the text that you want to convert to audio i.e. value of action variable. Pass the text and language to the gTTS() method to convert into audio. The converted audio is saved in a mp3 file named welcome. With the help of the os module play the welcome.mp3 file.

7.**speechToVideo_button ()**: In this, the SpeakText() method takes the input as speech. Microphone () is used as a source for input. The user's input is listened and using recognize_google() audio is recognized. Recognized audio is converted into text and that will be displayed. The text is compared with the index file using the casefold() method. If both matches then with the help of os module search video corresponding to the text in the index file.

c. Implementation Details

1. Module -1 (Sign video to audio):

    i. Browse Video:
        1. Select a directory to browse the video using method askopenfilename() method.
    ii. Extraction of Frames:
        1. Initialize variable frame rate to 0.25.
        2. Convert frames to grayscale using cvtColor() method.
        3. Resize the frames into 300 * 350 pixel using resize() method.
        4. Declare variable count to 0 to count the number of frames extracted.
        5. Declare variable success to True which tells whether frames are extracted or not.
        6. Save frames as JPEG files.
        7. Display number of frames extracted.
    iii. Prediction of frames:
        1. Read model from json file.
        2. Declare variable num which takes value of the most predicted column.
        3. Print maximum predicted column.
        4. Print action corresponding to the maximum predicted column.
    iv. Conversion of video to audio:
        1. Declare variable mytext which is value of action variable.

2. Pass the text and language to the gTTS() method to convert into audio.
3. Save the converted audio in mp3 file.
4. Play mp3 file.

2. Module -2 (Speech to Sign video):

1. Initialize the recognizer to listen to recognize the voice.
2. Use the microphone as a source for input.
3. Wait for the second to let recognizer adjust the threshold based on the surrounding noise level.
4. Declare a variable which takes listen() method to listen to the user's input.
5. Declare variable myText to Recognize audio using google speech API and audio file is converted into text using recognize_google() method.
6. Convert text to lowercase using lower() method and remove the whitespaces.
7. Compare text with line in index file having class name.
8. Search video with respect to text from the database.
9. Play the respective video.

6. Integration & Testing

a. Description of the Integration Module

1.Sign video to audio:

System should browse video from database. For that a select a directory to browse the video. Next, Extract the Frames. Extract frames having frame rate 0.25. Convert frames to grayscale. Resize the frames into 300*350 pixels. Save frames as JPEG files. Display Number of frames extracted. Then, Predict the frames. Read model from json file. Print maximum predicted column. Print action corresponding to the maximum predicted column. Finally, Convert predicted video to audio format. Save the converted audio in a mp3 file. Play mp3 file.

2. Speech to Sign video:

Initialize the recognizer to listen to spoken words and identify them. Use the microphone as a source for input. Listens for the user's input. Extract Text from speech. Compare text with indexfile.txt having class name. Search video with respect to text from the database. Play the respective video.

b. Testing

1.Unit Testing

1.Video to Audio**:**

Sr. No	Test Case Description	Expected Output	Actual Output	Status
1	Check whether video is browsed from system and displayed on the screen	System needs to browse video from database	System is able to browse video from database	Pass
2	Check whether browsed video is played after clicking on Play / Run Video button	System should be in position to play appropriate video	System is able to play correct video	Pass
3	Check whether proper frames are extracted	System needs to do extraction of frames from given video	System is able to extract the frames from video	Pass
4	Check whether system can predict class name	System should be in position to predict correct class name of given video	System is able to predict correct class name of given video	Pass
5	Conversion from video to audio	System should be in position to convert video to audio format	System is able to convert video to audio format	Pass

2.Speech to Video**:**

Sr. No	Test Case Description	Expected Output	Actual Output	Status
1	Check whether system is accepted speech	System should be in position to accept speech	System is able to accept speech	Pass
2	Recognition of Speech	System needs to correctly recognize speech comes from user	System is able to recognize speech from user	Pass
3	Extraction of text from speech	System should in position to extract the text from speech	System is able to extract the text from video	Pass
4	Searching of video from database	System needs to find respective video from database	System is able to find respective video from database	Pass
5	Run or play the respective video	System needs to run correctly founded video	System is able to run correct video	Pass

2. System testing

Sr. No	Test Case Description	Expected Output	Actual Output	Status
1	Check whether browsed video displayed on screen	System needs to browse video from database	System is able to browse video from database	Pass
2	Check whether the particular video is played	System should be in position to play appropriate video	System is able to play correct video	Pass
3	Check how many frames are extracted	System needs to do extraction of frames from given video	System is able to extract the frames from video	Pass
4	Check whether proper Class Name is predicted	System should be in position to predict correct class name of given video	System is able to predict correct class name of given video	Pass
5	Video To Audio	System should be in position to convert video to audio format	System is able to convert video to audio format	Pass
6	Speech To Video	System should be in position to convert speech to Video format	System is able to convert speech to video format	Pass
7	Recognition of Speech	System needs to correctly recognize speech comes from user	System is able to recognize speech from user	Pass

8	Extraction Of Text from Speech	System should in position to extract the text from speech	System is able to extract the text from video	Pass
9	Searching Of Video from Database	System needs to find respective video from database	System is able to find respective video from database	Pass
10	Run Or Play the Respective Video	System needs to run correctly founded video	System is able to run correct video	Pass

7. Performance Analysis

This project is based on image processing which is part of a convolutional neural network concept. In this project, first created the video which are used for the conversion purpose and many more. The frame first gets extracted from videos with framerate of 0.25 per second as the video timing is more the number of frames extracted is more. After extraction of frames, the segmentation is done. After the segmentation, get the probability frame of that class and the it will predict the text. According to the performance of system, as number of frames extracted are more the accuracy of prediction of gestures is more.

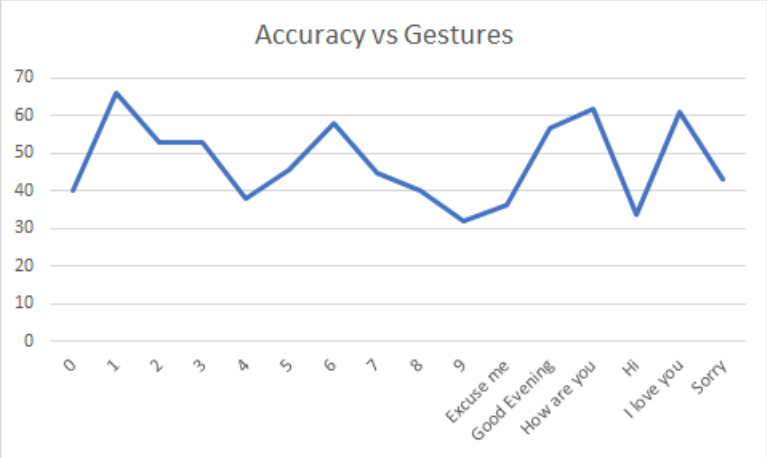


Figure 7.1: Accuracy vs Gesture
(x-axis: Gestures & y-axis: Accuracy)

The accuracy for each gesture varies as the changes appears in the dataset. The above graph shows the information about accuracy vs gestures. Here, the x-axis shows the gestures and y-axis shows accuracy.

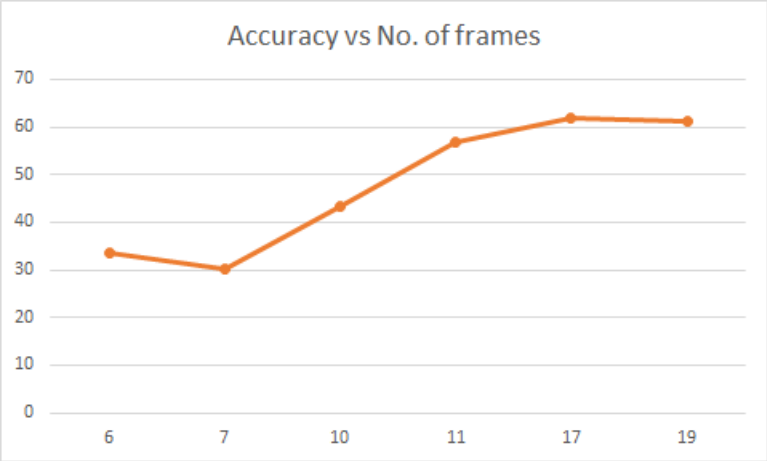


Figure 7.2: Accuracy vs No. of frames
(x-axis: No. of Frames & y-axis: Accuracy)

The above graph shows the information about accuracy vs gestures. Here, the x-axis shows the number of frames extracted and y-axis shows accuracy.

Training and validation accuracy:

The main advantage of CNN is its high Accuracy. Image classification involves the extraction of features from images to observe some pattern in the dataset. For classification of frames, the training accuracy of the model on training dataset is between 95-99% and validation accuracy is between 94-99%. Training and validation loss is very low up to 0.06%.

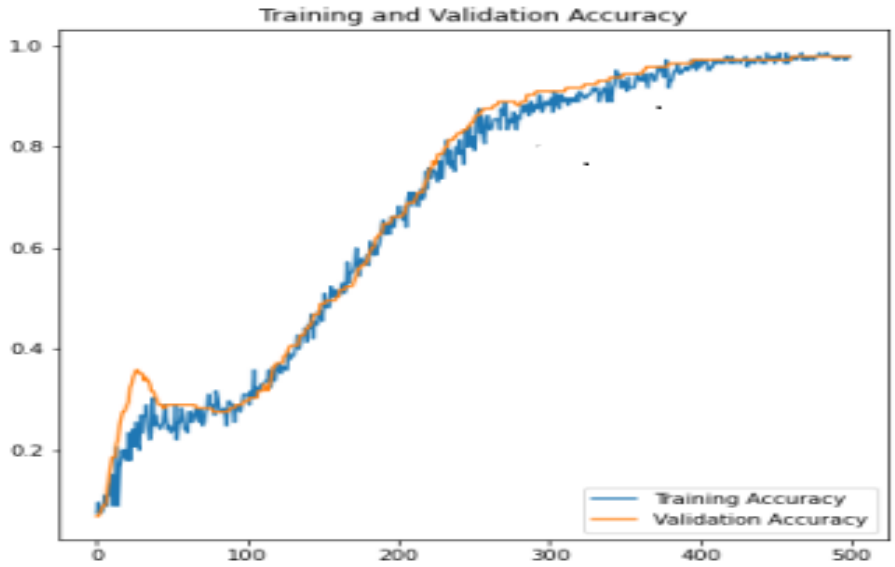


Figure 7.3: Training and validation accuracy

Figure 7.4: Training and validation loss

8. Future Scope

The main aim of this project is to provide a platform which makes communication easy between deaf-dumb community and normal people. This system will take input in the video format it will convert into text and then Audio as well as it takes input as Video and convert into speech. This system will also help in learning Indian Sign Language.

This System have some limitations like it is limited to English Language only. It is a standalone project. Here provided limited database.

So, as stated earlier it can further be improved with Realtime implementation of project with live webcam. And database can be further extended. Further many languages can be added.

Although this project has achieved high accuracy, does not include all the sign language words so the data set is limited in scope. As a research direction with broad application and development space, sign language recognition still has much room for improvement. At the same time, most methods of sign language recognition now only consider the accuracy of the algorithm. However, for the application of sign language recognition in real scene, real-time performance is another important index. Therefore, it is also a direction worthy of breakthrough about how to improve the speed of hand locating and recognition of sign language words.

9. Applications

Following are the applications where this system can be used:

Learning Indian Sign Language:

Studying Indian Sign Language boosts better awareness of and sensitivity to the deaf and dumb community. There is an exponent need to increase the availability of qualified ISL interpreters in the community and mainstream programs in schools and colleges/universities. Some places interpreters are in large demand include hospitals, courts, governmental agencies, community activities. Deaf and hard of hearing people expert in ISL may be interested in becoming ISL teachers

Communication:

Communication is an important aspect while sharing or exchanging information. It plays a very important role in our life. It brings people closer to each other with better understanding. Wherever life exists, communication also exists. It is an unavoidable system. For successful and effective communication, it is very important to know about the communication process. Communication process involves sender, message, receiver, channel, feedback. Communication between normal people is easy but communication between deaf-dumb and normal people is difficult. Thus, this system provides platform so communication becomes easy between themselves.

10. Installation and User Manual

Installation Guide:

1. Check configuration of your laptop.
2. Install Python 3.8 on your machine.
3. Import the project on your machine.
4. Run the project.
5. Run UI.
6. Install following packages:
    1. **TensorFlow**: It is used to build and train CNN models. Use command to install tensorflow- pip install tensorflow.
    2. **Gtts**: It is googled text to speech API which reads text aloud. Use command to install gtts- pip install gtts.
    3. **Os**: The os module in python provides functions like identifying the current directory, fetching its contents, removing a directory, etc. Use command to install os- pip install os.
    4. **Opencv**: It is used for all sorts of image and video analysis like recognition and detection. Use command to install opencv- pip install opencv.
    5. **Numpy**: This library used to convert lists into numpy arrays. Use command to install - pip install numpy.
    6. **Pandas**: It is used for exploring and cleaning data. Use command to install- pip install pandas.
    7. **Pyttsx3**: It is used to convert Text into speech. Use command to install- pip install pyttsx3.
    8. **Tkinter**: Tkinter is the standard GUI library for Python. Use command to install Tkinter- pip install tkinter.
    9. **Matplitlib**: Matplotlib is used for plotting graphs of accuracy and for data visualization. Use command to install matplotlib-pip install matplotlib.
    10. **Datasets**: Datasets is a library for easily accessing and sharing datasets. Use command to install datasets- pip install datasets.
    11. **Sklearn**: It consists of lots of tools for machine learning. Used to build ML models. Use command to install sklearn- pip install sklearn.
    12. **PIL**: PIL is library is used to open image content into an array. Use command to install PIL- pip install pillow.
    13. **Speech Recognition**: It offers easy audio processing and microphone accessibility. Use command to install Speech Recognition- pip install Speech Recognition.
    14. **Pyaudio**: Pyaudio is used to play and record audio. Use two commands to install Pyaudio- pip install pipwin and pipwin install pyaudio.

User Manuals:

User first needs to choose the directory for the purpose of storing the dataset. So that it can be used for future use. The figure shown below is showing the directory for storing the data:

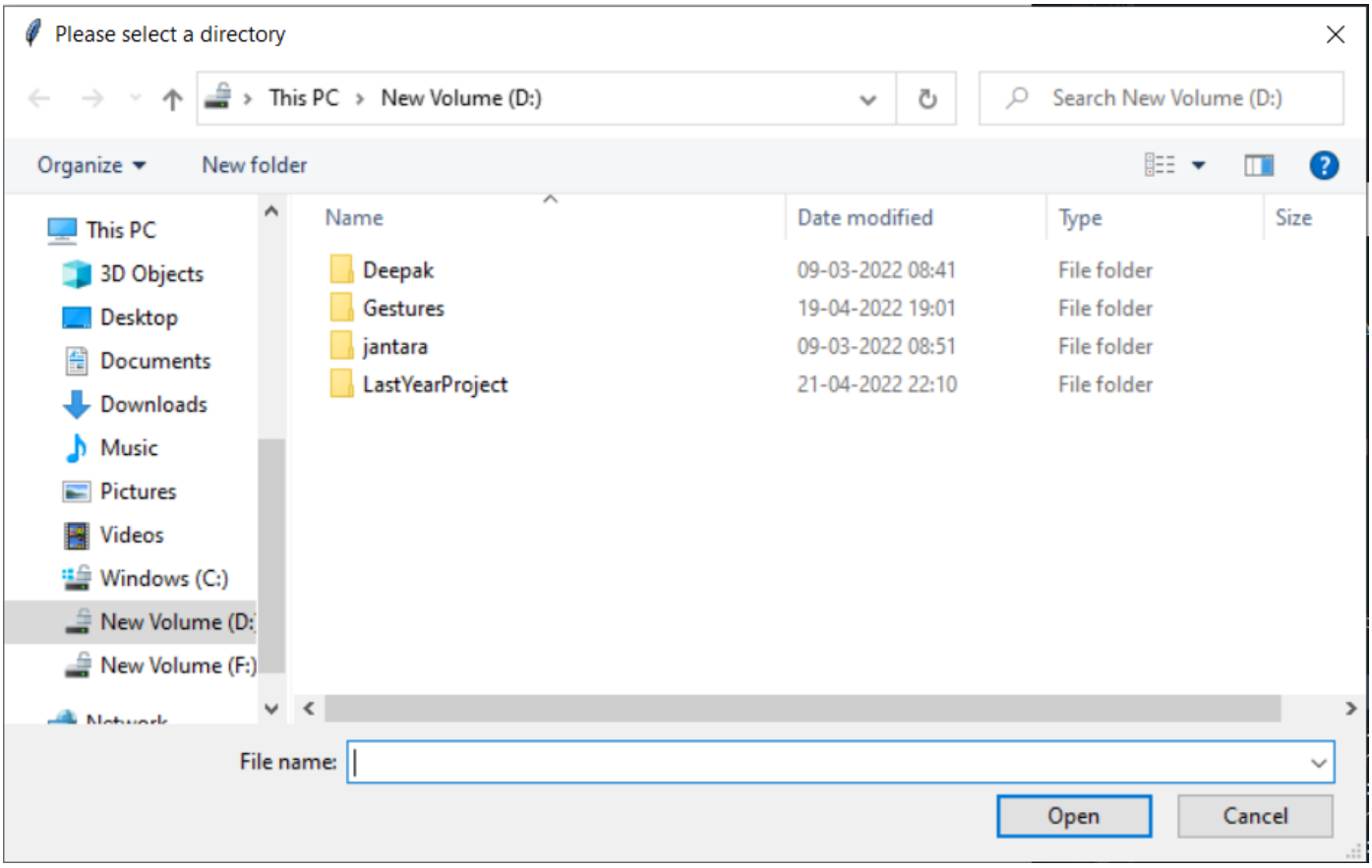


Figure 10.1: showing directory

The videos are created by the user and stored in respective directory. This is shown below:

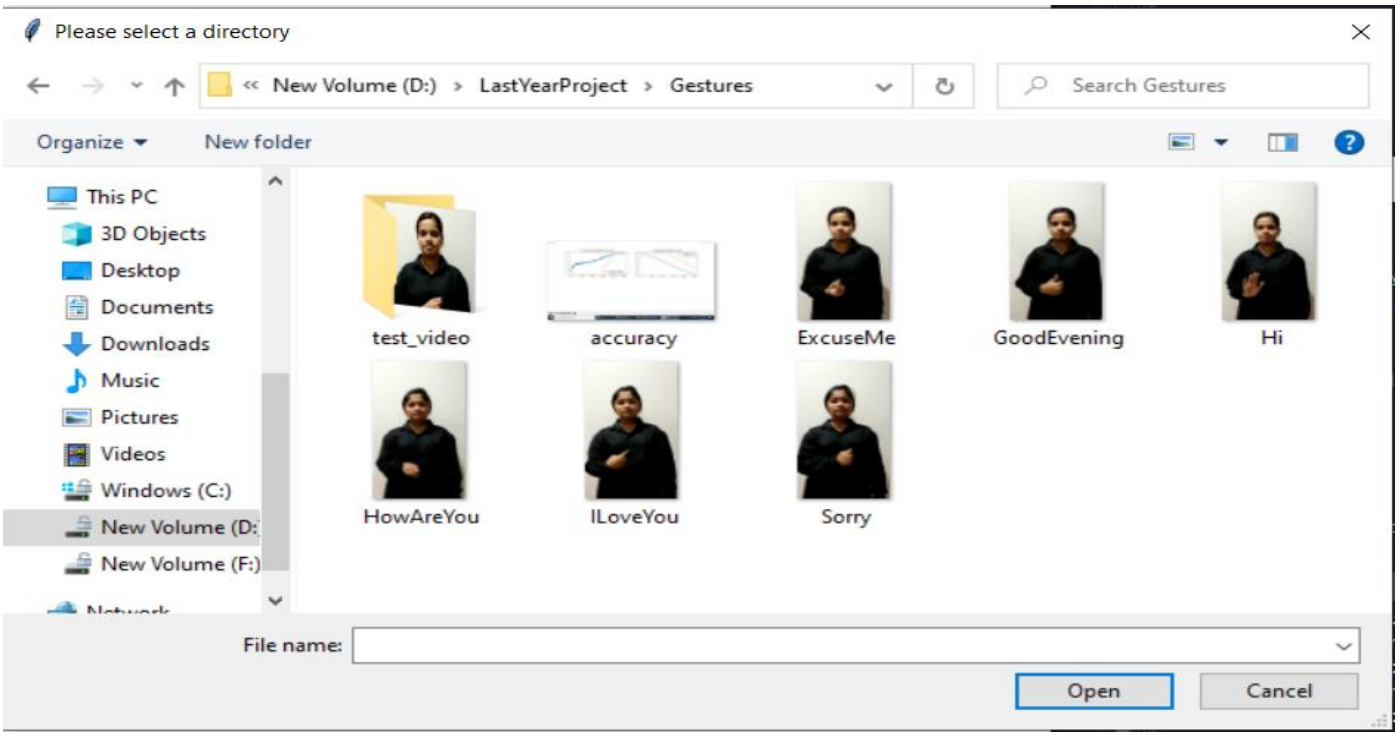


Figure 10.2: videos stored in this directory

Here, system provided user interface which performs different actions that user wants to perform. It includes actions like browse videos, play video, extract frames, predict, convert to speech and speech to video. The user interface design is given as follows:

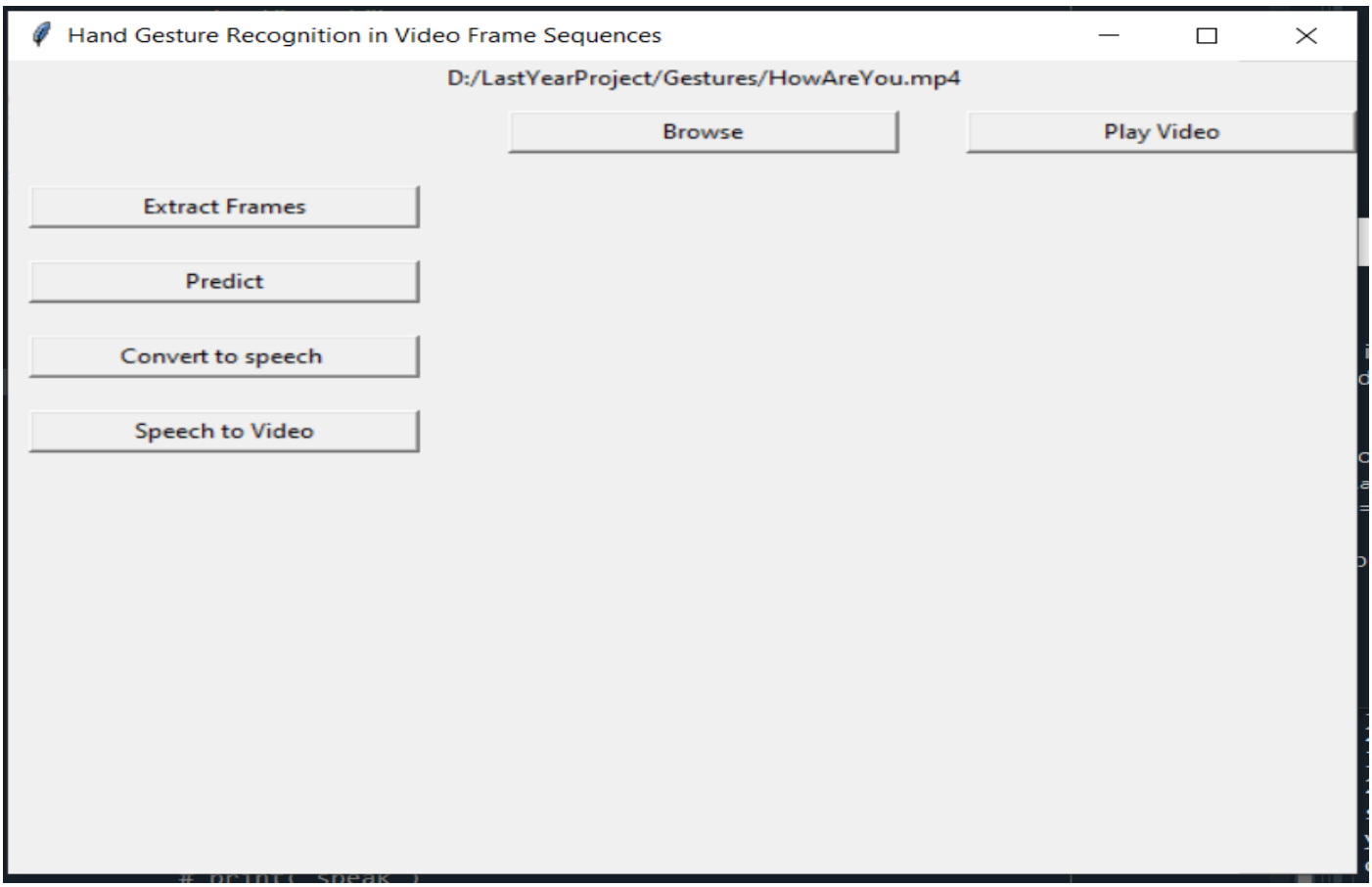


Figure 10.3: User Interface

After clicking on the extract frames button the frames get extracted from the video. Here, user browsed the video of how are you from respected directory and frames get extracted from the video. Number of frames get extracted here are 17. This is shown below:

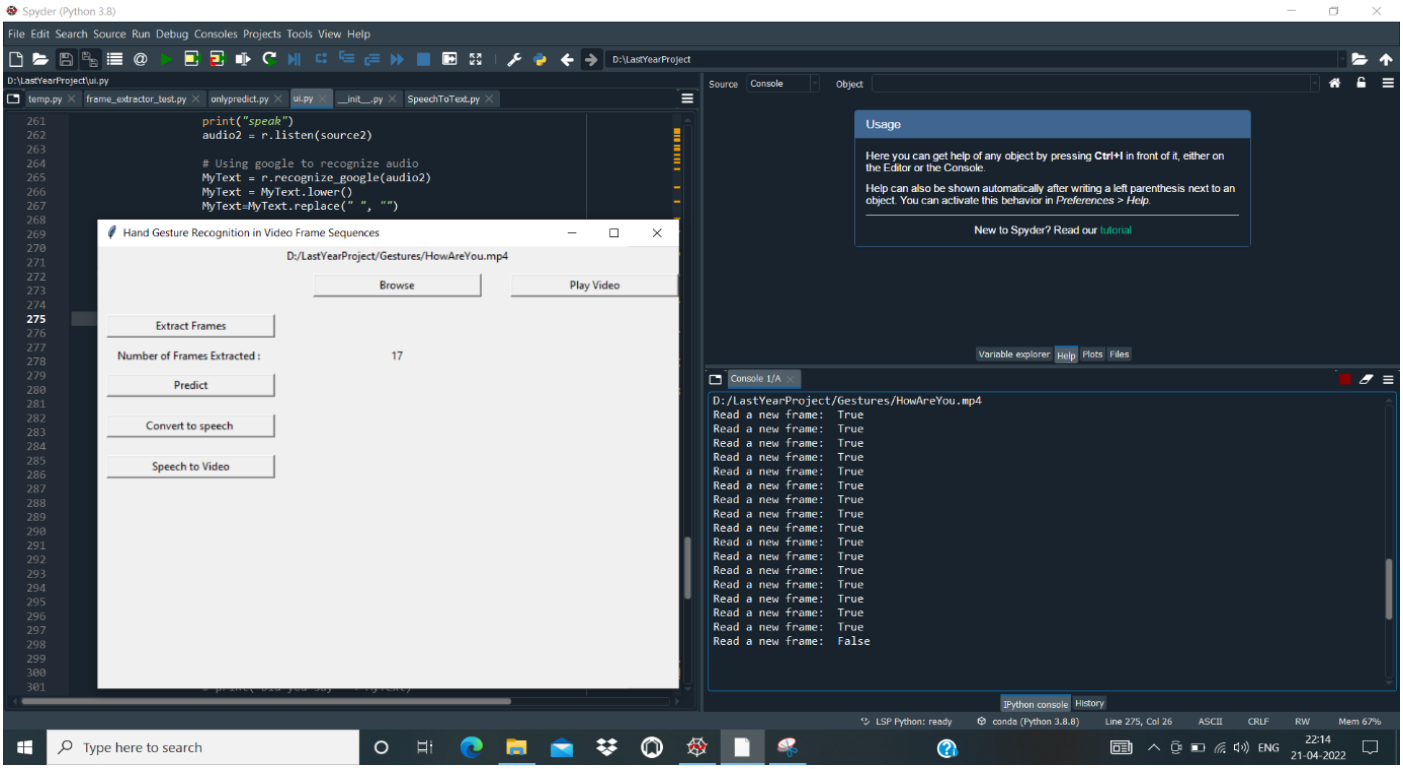


Figure 10.4: number of frames

Extracted frames are shown as follows:

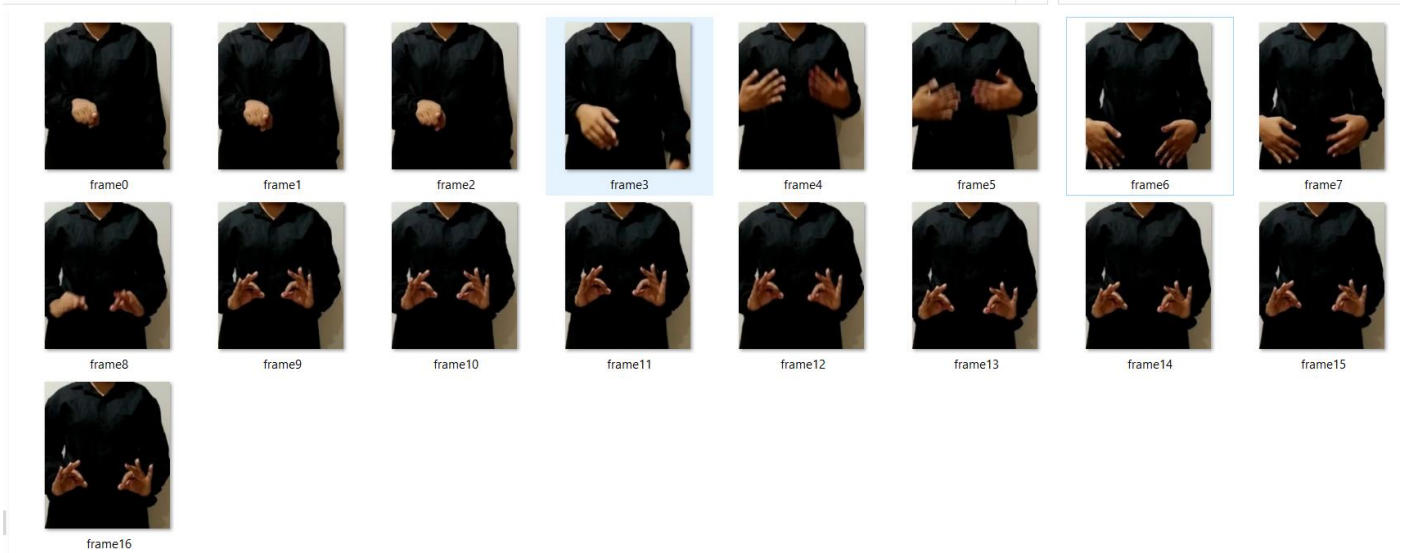


Figure 10.5: Extracted Frames

Here, the frames are predicted according to following factors: 1. Maximum predicted column.
2.Action corresponding to maximum predicted column.
The metrics of probability is shown in the snapshots:

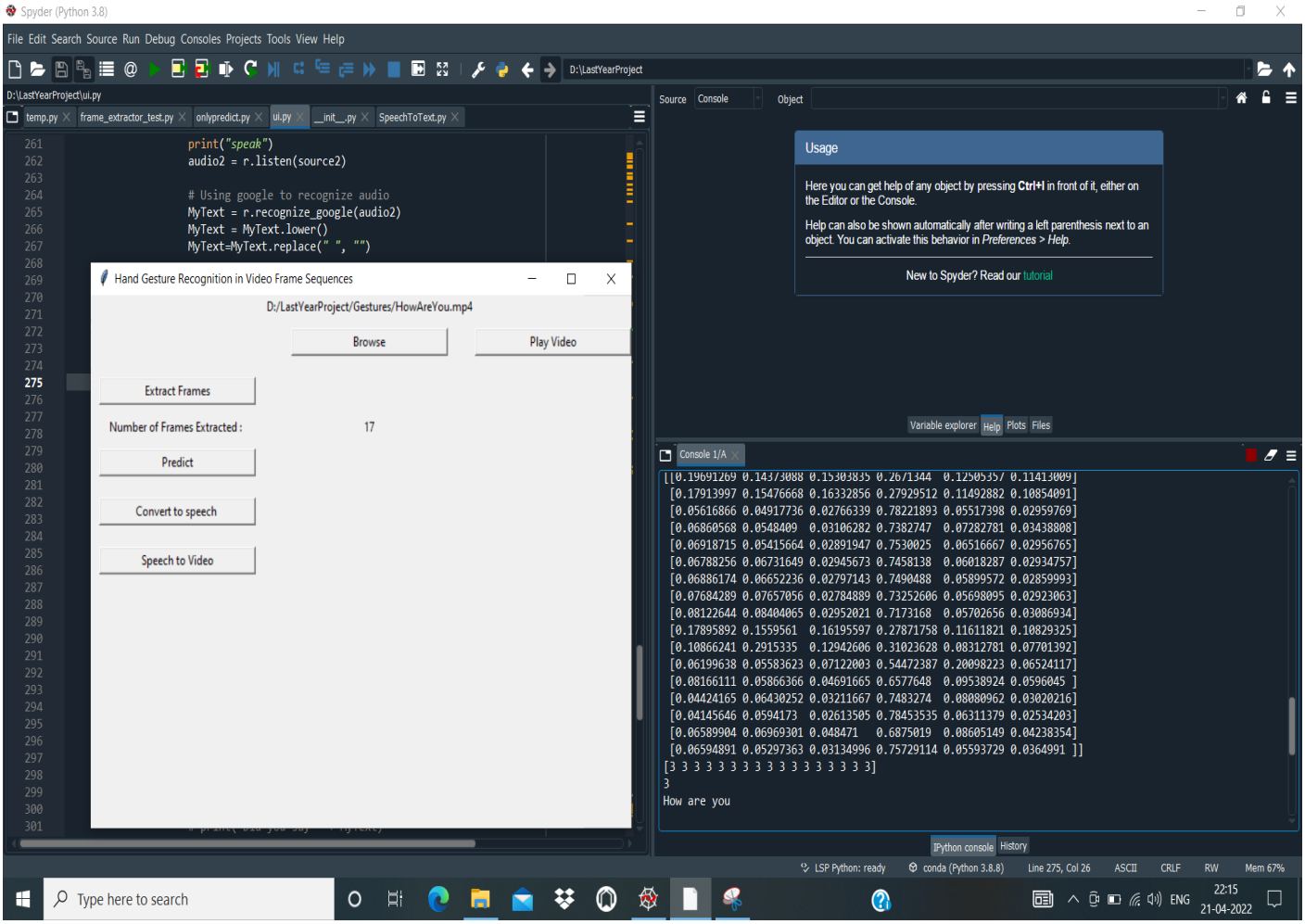


Figure 10.6: Probability Metrics

Convert predicted video to audio format. Save the converted audio in a mp3 file. Play mp3 file:

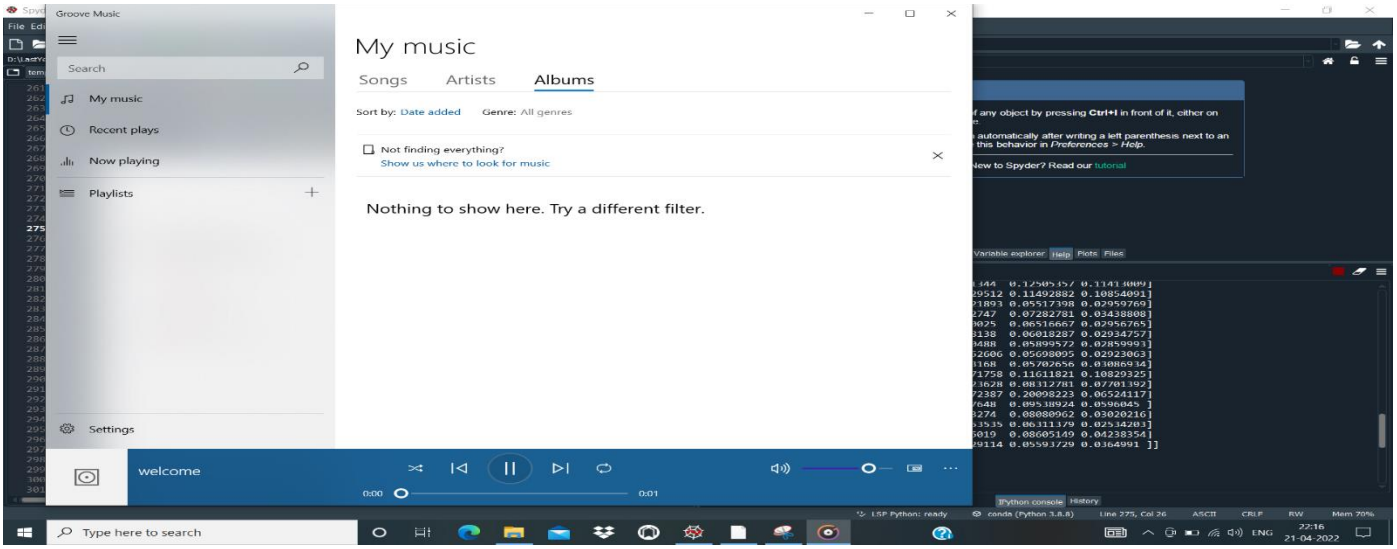


Figure 10.7: convert video to audio

Convert speech to text:

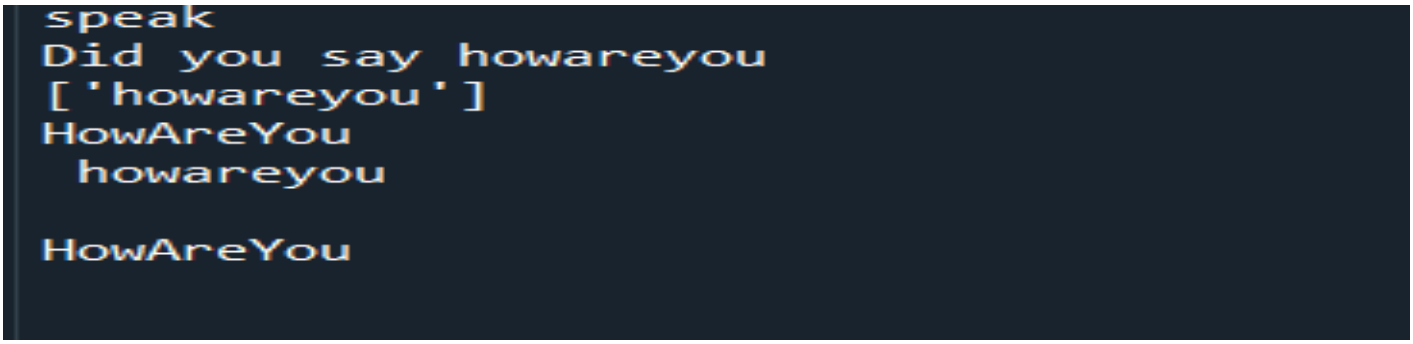


Figure 10.8: Convert speech to text

Play Video According to corresponding text:

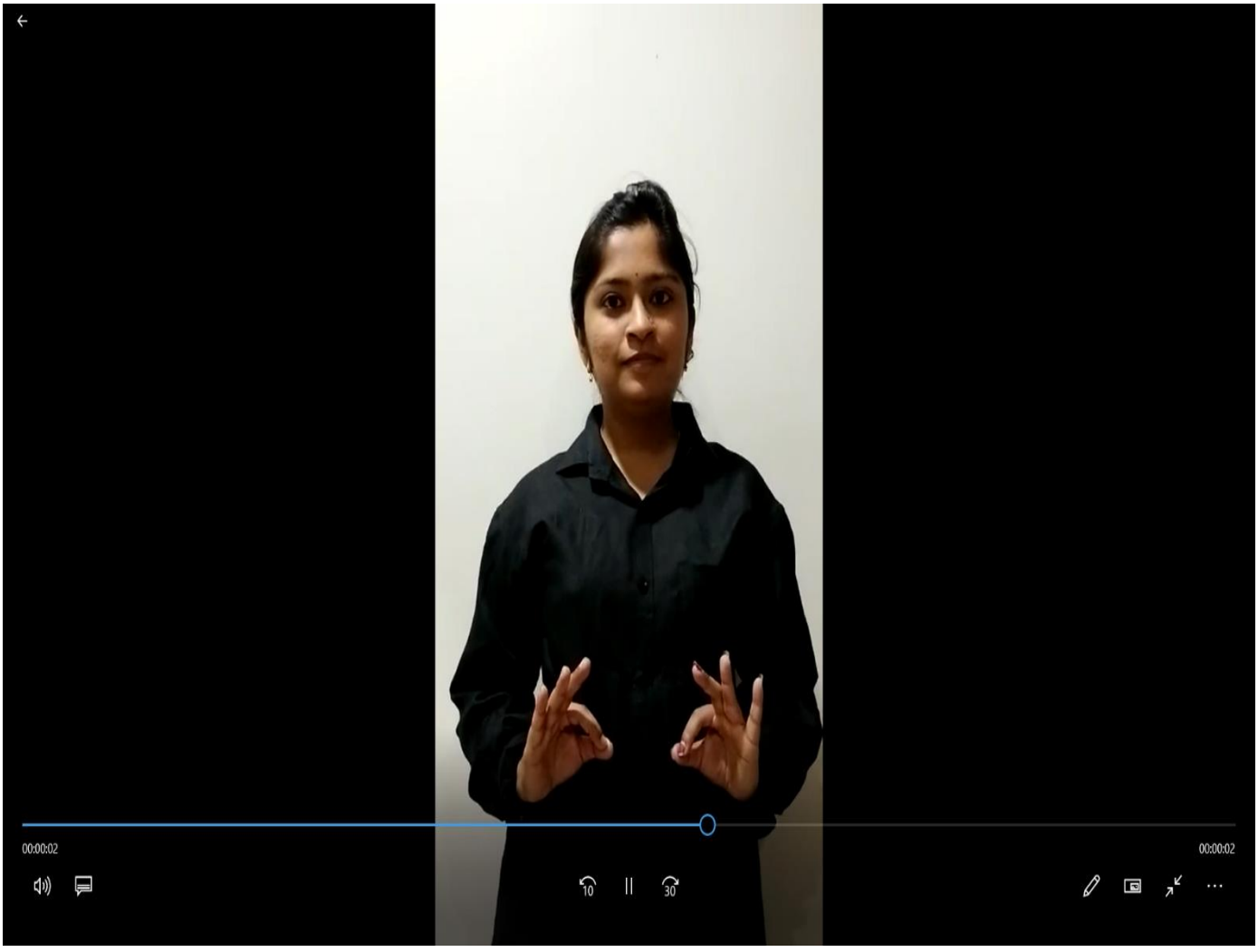


Figure 10.9: video from text

Convert speech to text:

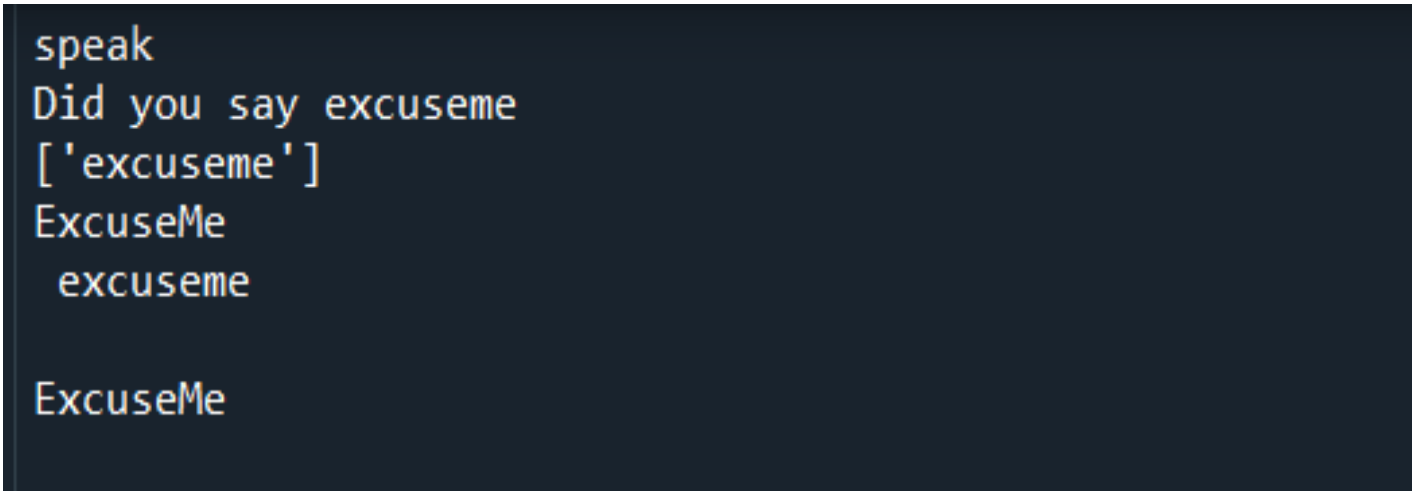


Figure 10.10: Speech to text

Play Video According to corresponding text:

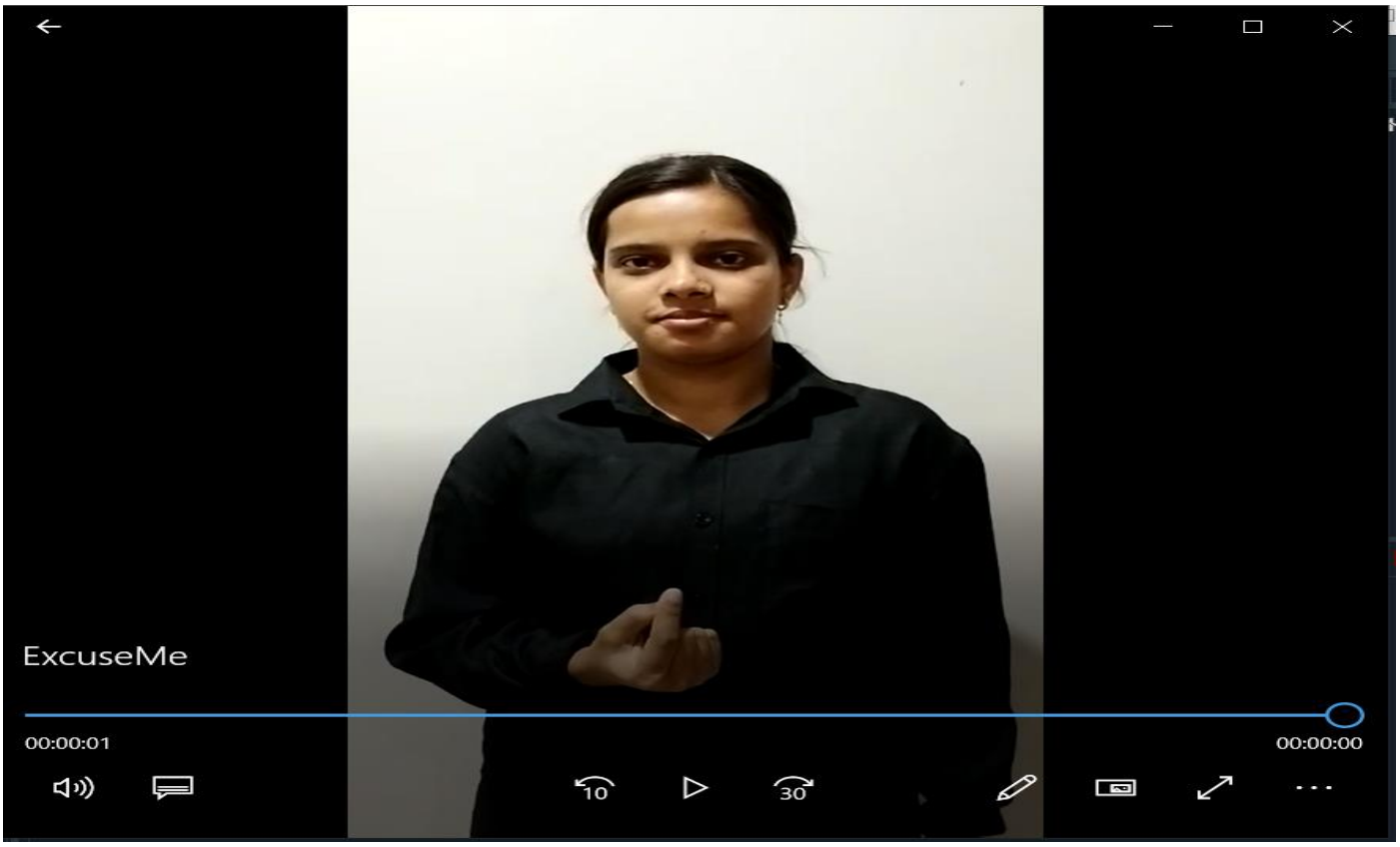
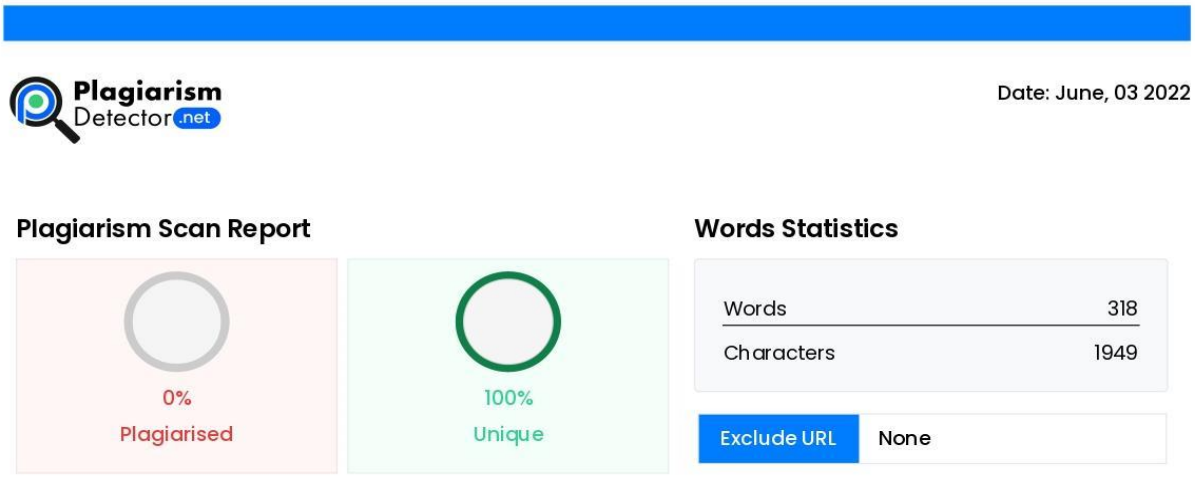


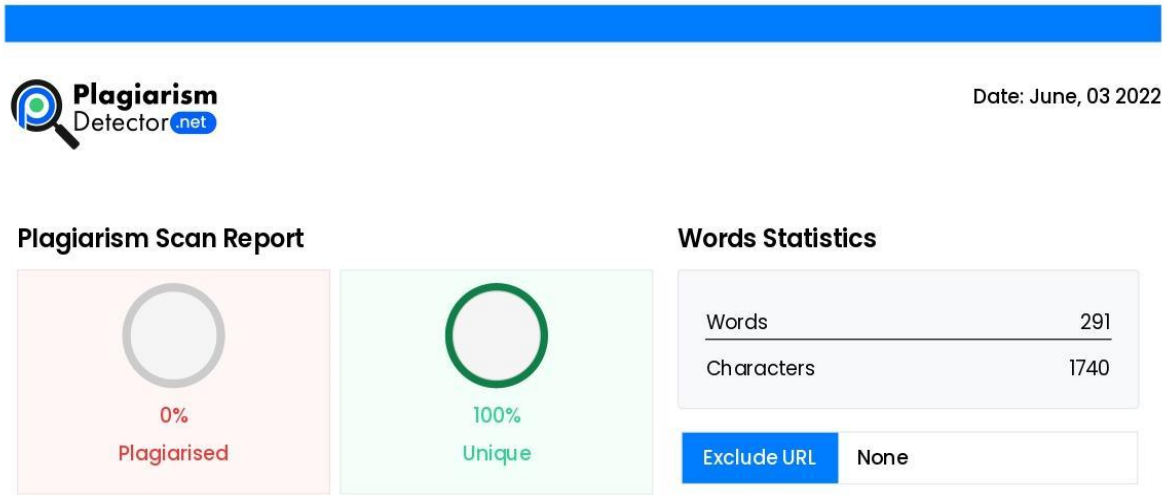
Figure 10.11: Video from text

11. Plagiarism Report



Content Checked For Plagiarism

c. Detailed Description of Methods 1. browse_button (): This method is used to select a directory to browse the video. It will print the name of the selected video file from the directory. 2.play_video (): This method is used to select the video which will be played.With the help of os module it will play the browsed video. 3.extract button ():This method extracts frames. Initialize variable frame rate to 0.25. cvtColor() method used to convert image frames into grayscale. Resize the frames using resize() method. Initialize variables as y = 300, h = 350, x = 50, w = 300. Frames are resized to 350*650 pixels.Declare variable count to 0 to count the number of frames extracted. Declare variable success to True which tells whether frames are extracted or not. 4.most_frequent (List): Input to this method is a list of probabilities of features. This method returns a column index which has maximum probability comparsed to others. 5.predict_button (): This method takes output from most_frequent method to print maximum predicted column.Also prints action corresponding to the maximum predicted column. 6.videoToSpeech_button (): This method converts predicted video to audio format. Declare variable mytext which is the text that you want to convert to audio i.e. value of action variable. Pass the text and language to the gTTS() method to convert into audio. The converted audio is saved in a mp3 file named welcome. With the help of the os module play the welcome.mp3 file. 7.speechToVideo_button (): In this, the SpeakText() method takes the input as speech. Microphone () is give as a source for input. The user's input is listened and using recognize_google() audio is recognized. Recognized audio is converted into text and that will be displayed. The text is compared with the index file using the casefold() method. If both matches then with the help of os module search video corresponding to the text in the index file.



Content Checked For Plagiarism

d. Implementation Details 1. Module -1 (Sign video to audio): i. Browse Video: 1. Select a directory to browse the video using method askopenfilename () method. ii. Extraction of Frames: 1. Initialize variable frame rate to 0.25. 2. Convert frames to grayscale using cvtColor() method. 3. Resize the frames into 300 * 350 pixel using resize() method. 4. Declare variable count to 0 to count the number of frames extracted. 5. Declare variable success to True which tells whether frames are extracted or not. 6. Save frames as JPEG files. 7. Display number of frames extracted. iii. Prediction of frames: 1. Read model from json file. 2. Declare variable num which takes value of the most predicted column. 3. Print maximum predicted column. 4. Print action corresponding to the maximum predicted column. iv. Conversion of video to audio: 1. Declare variable mytext which is value of action variable. 2. Pass the text and language to the gTTS() method to convert into audio. 3. Save the converted audio in mp3 file. 4. Play mp3 file. 2. Module -2 (Speech to Sign video): 1. Initialize the recognizer to listen to recognize the voice. 2. Use the microphone as a source for input. 3. Wait for the second to let recognizer adjust the threshold based on the surrounding noise level. 4. Declare a variable which takes listen() method to listen to the user's input. 5. Declare variable myText to Recognize audio using google speech API and audio file is converted into text using recognize_google() method. 6. Convert text to lowercase using lower() method and remove the whitespaces. 7. Compare text with line in index file having class name. 8. Search video with respect to text from the database. 9. Play the respective video.

12. Ethics

Creative Common License
Attribute-Share Alike 4.0 International (CC BY SA 4.0)
This is human-readable summary of (and not a substitute for) the license.

You are free to:
Share- copy and redistribute the material in any medium or format.
Adapt- remix, transform and build upon the material for any purpose, even commercially. The licensor cannot revoke these freedoms as long as you follow the license terms.

Using the following terms:

Attributes you must give appropriate credit, provide link to the license, and indicate if changes were made. You may do so in any responsible manner, but not in any way that suggests the licensor endorses you or your use.

Share Alike- you remix, transforms and build upon the material for any purpose, you must distribute your contributions under the same license as the original.

No additional restrictions- You must apply legal terms or technological measures that legally restricts others from doing anything the license permits.

Notices:
You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an application exception or limitation. No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

Ethical Practices for CSE Students:

As Computer Sc. & Engineering student, I believe it is unethical to:

- Take credit for someone else's work.
- Hire someone to write an assignment.
- Purchase or submit a research or team paper from the internet to a class as one's own work.
- Cheat on a graded assignment.
- Cheat on an exam.
- Plagiarize other people's work without citing or referencing the work.
- Add the name of a non-contributing person as an author in a project/research study.
- Copy and paste material found on the internet for an assignment without acknowledging the authors of the material.
- Deliberately provide inaccurate reference for a project or research study.
- Knowingly permit student work done by done student to be submitted by another student.
- Surf the internet for personal interest and non-class related purposes during classes.
- Make a copy of software for personal or commercial use.
- Make a copy of software for a friend.
- Loan CDs of software to friends.
- Download pirated software from the internet.
- Distribute pirated software from the internet.
- Buy software with a single user license and then install it on multiple computers.
- Share a pirated copy of software.
- Install a pirated copy of software.

13. References

[1] P. V. V. Kishore and P. Rajesh Kumar Video Based Indian Sign Language Recognition System (INSLR) Using Wavelet Transform and Fuzzy Logic , IACSIT International Journal of Engineering and Technology, Vol. 4, No. 5, October 2012 A

[2] ELGHOUL, M.J.O. An avatar-based approach for automatic interpretation of text to Sign language. Chall. Assist. Technol. AAATE 2007, 20, 266 Shravani K Sree Lakshmi A, Sri Geethika M, Dr.Sapna B Kulkarni Indian Sign Language Character Recognition.

[3]B . Gupta, P. Shukla, and A. Mittal. K-nearest correlated neighbor classification for Indian sign language gesture recognition using feature fusion. In 2016 International Conference on Computer communication and Informatics (ICCCI), pages 1–5, 2016. Francois Chollet Deep Learning with Python.

[4] Canny. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI8(6):679–698, 1986.

[5]. K. Dixit and A. S. Jalal. Automatic Indian sign language recognition system. In 2013 3rd IEEE International Advance Computing Conference (IACC), pages 883–887, 2013.

[6] Yu-Gang Jiang, Jun Yang, Chong-Wah Ngo, and Alexander Hauptmann. Representations of keypoint-based semantic concept detection: A comprehensive study. Multimedia, IEEE Transactions on, 12:42 – 53, 02 2010.

[7]. David G. Lowe. Distinctive image features from scale-invariant key points. Int. J. Comput. Vision, 60(2):91–110, November 2004.

[8]. Dutta, Kusumika& K, Satheesh& S, Anil & Sunny, Breeze. (2015). Double handed Indian Sign Language to speech and text. 374- 377. 10.1109/ICIIP.2015.7414799.

[9] Kishore, P.; Kumar, P.R. A video based Indian sign language recognition system (INSLR) using wavelet transform and fuzzy logic. Int. J. Eng. Technol. 2012, 4, 537.

[10]. Goyal, L.; Goyal, V. Automatic translation of English text to Indian sign language synthetic animations. In Proceedings of the 13th International Conference on Natural Language Processing, Varanasi, India, 17–20 December 2018; pp. 144–153.

[11]. San-Segundo, R.; Barra, R.; Córdoba, R.; D'Haro, L.; Fernández, F.; Ferreiros, J.; Lucas, J.M.; Macías-Guarasa, J.; Montero, J.M.; Pardo, J.M. Speech to sign language translation system for Spanish. Speech Commun. 2008, 50, 1009–1020.

[12]. López-Ludeña, V.; San-Segundo, R.; Morcillo, C.G.; López, J.C.; Muñoz, J.M.P. Increasing adaptability of a speech into sign language translation system. Exp. Syst. Appl. 2013, 40, 1312–1322.

[13]. Halawani, S.M.; Zaitun, A. An avatar-based translation system from Arabic speech to Arabic sign language for deaf people. Int. J. Inf. Sci. Educ. 2020, 2, 13–20.

[14] Papadogiorgaki, M.; Grammalidis, N.; Tzovaras, D.; Strintzis, M.G. Text-to-sign language synthesis tool. In Proceedings of the IEEE 13th European Signal Processing Conference, Antalya, Turkey, 4–8 September 2019; pp. 1–4.

[15]. Elliott, R.; Glauert, J.R.; Kennaway, J.; Marshall, I.; Safar, E. Linguistic modelling and language-processing technologies for Avatar-based sign language presentation. Universe. Access Inf. Soc. 2018, 6, 375–391.