(An Autonomous Institute Affiliated to Savitribai Phule Pune University)

A Project Report on

App Development for Specially- Abled(Deaf and Mute) Individuals

Submitted by,

Arya Khobragade	**(Exam Seat No. 202101070173)**
Dhawal Turkar	**(Exam Seat No. 202101040234)**
Manav Wadikar	**(Exam Seat No. 202101040211)**
Pranjali Shinde	**(Exam Seat No. 202101070197)**

Guided by,

Mrs. Aarti Deshpande

A Report submitted to MIT Academy of Engineering, Alandi(D), Pune, An Autonomous Institute Affiliated to Savitribai Phule Pune University in partial fulfillment of the requirements of

THIRD YEAR BACHELOR OF TECHNOLOGY in
Computer Engineering.

School of Computer Engineering
MIT Academy of Engineering
(An Autonomous Institute Affiliated to Savitribai Phule Pune University)
Alandi (D), Pune – 412105

(2023–2024)

MIT | Academy of Engineering

(An Autonomous Institute Affiliated to Savitribai Phule Pune University)

CERTIFICATE

It is hereby certified that the work which is being presented in the Third Year Project Design Report entitled **"App Development for Specially- Abled(Deaf and Mute) Individuals"**, in partial fulfillment of the requirements for the award of the Bachelor of Technology in Computer Engineering.and submitted to the **School of Computer Engineering** of **MIT Academy of Engineering, Alandi(D), Pune, Affiliated to Savitribai Phule Pune University (SPPU), Pune**, is an authentic record of work carried out during Academic Year **2023–2024** Semester **V**, under the supervision of **Mrs. Aarti Deshpande**, **School of Computer Engineering**

Arya Khobragade	**(Exam Seat No. 202101070173)**
Dhawal Turkar	**(Exam Seat No. 202101040234)**
Manav Wadikar	**(Exam Seat No. 202101040211)**
Pranjali Shinde	**(Exam Seat No. 202101070197)**

Mrs. Aarti Deshpande
Project Advisor

Mr. Amar More
Project Coordinator

Dr. Rajeshwari Goudar
Dean

External Examiner

DECLARATION

We the undersigned solemnly declare that the project report is based on our own work carried out during the course of our study under the supervision of **Mrs. Aarti Deshpande**.

We assert the statements made and conclusions drawn are an outcome of our project work. We further certify that

1. The work contained in the report is original and has been done by us under the general supervision of our supervisor.
2. The work has not been submitted to any other Institution for any other degree/diploma/certificate in this Institute/University or any other Institute/University of India or abroad.
3. We have followed the guidelines provided by the Institute in writing the report.
4. Whenever we have used materials (data, theoretical analysis, and text) from other sources, we have given due credit to them in the text of the report and giving their details in the references.

Arya Khobragade	**(Exam Seat No. 202101070173)**
Dhawal Turkar	**(Exam Seat No. 202101040234)**
Manav Wadikar	**(Exam Seat No. 202101040211)**
Pranjali Shinde	**(Exam Seat No. 202101070197)**

Abstract

Our project addresses the communication barrier between individuals who use sign language and those who are not proficient in it. We aim to provide a comprehensive solution for individuals without a voice by developing a cost-effective mobile app. The primary objective of our project is to establish a system capable of converting sign language into text and further converting that text to voice through the analysis of hand gestures. This innovative mobile application effectively eliminates the communication barrier between these distinct communities. Our overarching goal is to empower voiceless individuals, facilitating seamless communication with others.

Acknowledgment

We extend our heartfelt appreciation to our esteemed mentors for their pivotal role in guiding us throughout our journey. A special mention goes to Mrs. Aarti Deshpande, Mr. Sachin Shingade, Mr. Amar More, Dr. Dipti Sakhare (Dean of Electrical and Electronic Engineering), and Dr. Rajeshwari Goudar (Dean of Computer Engineering). Their unwavering support and mentorship have not only granted us the opportunity to partake in the major project course but have also been instrumental in elevating our software proficiency and refining our interpersonal skills. We are sincerely grateful for their invaluable contributions, which have played a key role in our holistic growth and development.

Arya Khobragade 202101070173

Dhawal Turkar 202101040234

Manav Wadikar 202101040211

Pranjali Shinde 202101070197

Contents

List of Figures

Chapter 1

Introduction

1.1 Background

In a world that thrives on communication, the ability to express oneself is a fundamental aspect of human interaction. However, not all individuals are granted the same channels for expression. In particular, those who are deaf and mute face unique challenges in conveying their thoughts and emotions verbally. Recognizing the pressing need to address this communication gap, our project, "App Development for Specially-Abled (Deaf and Mute) Individuals," takes root. This endeavour is fuelled by a commitment to inclusively and accessibility, seeking to empower individuals who are deaf and mute through innovative technological solutions. As we delve into the development of a specialized mobile application, our focus extends beyond conventional communication barriers. We aim to harness the power of technology to create a bridge between the worlds of those who can hear and speak and those who communicate through alternative means.

Our project is not just about developing an app; it's a testament to the transformative potential of technology in enhancing the lives of specially-abled individuals. By leveraging advancements in mobile application development, we aspire to provide a platform that transcends the limitations imposed by traditional communication methods. This application is envisioned to be a tool that not only facilitates communication but also fosters a sense of independence and empowerment among individuals who are deaf and mute. Through this introduction, we embark on a journey to explore the intersection of technology and empathy, aiming to create a solution that not only addresses the unique communication needs of the specially-abled but also contributes to a more inclusive and understanding society.

1.2 Motivation

The wealth of articles and documents available on Google and online platforms has played a pivotal role in deepening our comprehension of the challenges faced by individuals who are deaf or mute— a term we acknowledge has evolved to "deaf and mute" for more respectful representation. These resources have afforded us valuable insights into the daily hurdles they encounter when communicating with those who do not share their condition.

Communication for individuals who are deaf and mute diverges significantly from that of their hearing and speaking counterparts. Their reliance on alternative means, such as sign language, gestures, facial expressions, and written notes, necessitates mutual familiarity with these specific forms of communication. This creates a notable barrier between individuals who are deaf and mute and those who are unfamiliar with these alternative methods.

Gratefully, these articles not only illuminate the challenges but also showcase various solutions explored to surmount these communication barriers. The overarching objective of these solutions is to offer inclusive and effective communication tools for individuals who are deaf and mute, facilitating improved interaction with the broader population.

Inspired by the ideas and information gleaned from these articles, our project introduces a novel approach that sets itself apart from existing initiatives. We have incorporated distinctive features not present in previous solutions, emphasizing cost-effectiveness and user-friendliness to ensure accessibility for a diverse user base.

By harnessing the knowledge and insights acquired from these articles, our endeavor seeks to create a solution tailored to the specific communication needs of individuals who are deaf and mute. Our ultimate aim is to dismantle the communication gap between these individuals and the wider society, empowering them to partake in meaningful interactions with confidence and ease.

1.3 Project Idea

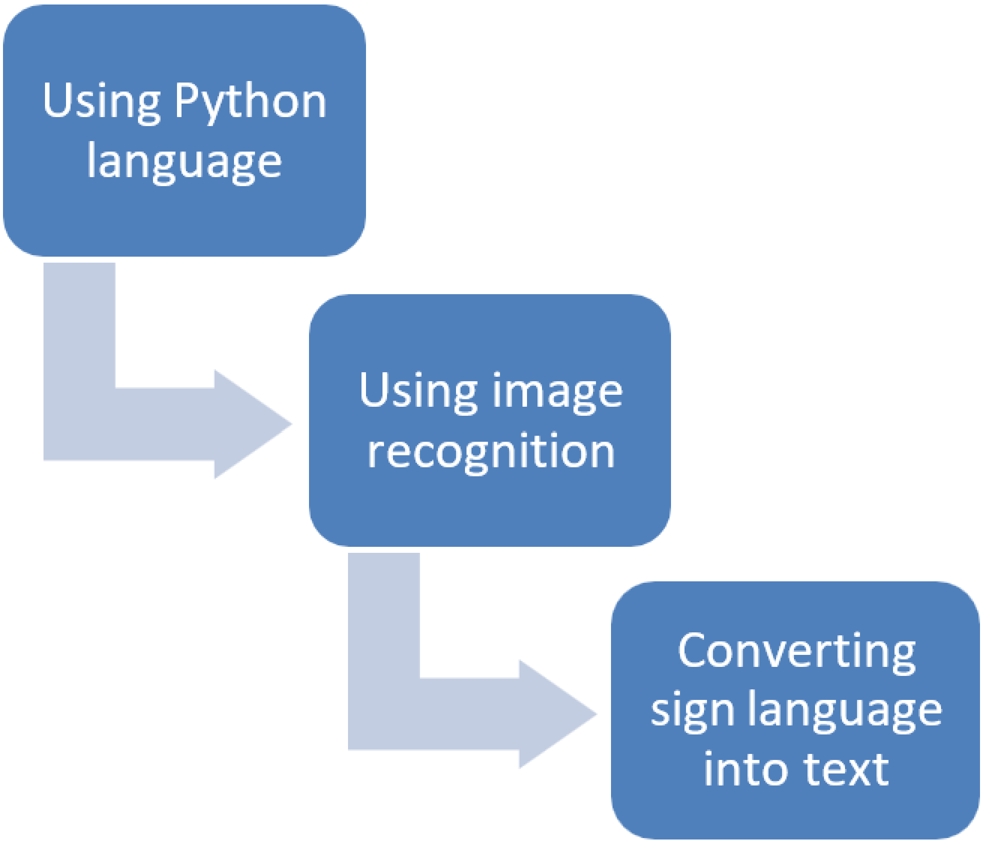

Figure 1.1: Flow chart of Project Idea

1.4 Proposed Solution

We envision the development of a revolutionary mobile application aimed at transcending the communication barriers faced by mute individuals in their interactions with the wider community. Our application acts as a real-time translator, converting the nuanced gestures and actions of mute individuals into easily understandable communication for others.

Leveraging the advanced capabilities of modern smartphones, our application utilizes the device's camera to capture and analyze the intricate sign language gestures of mute individuals. Through the integration of cutting-edge image processing and machine learning algorithms, the application ensures accurate recognition and interpretation of these gestures.

Upon successful identification, the application seamlessly generates corresponding text or visual representations to convey the intended message. This dynamic functionality allows for on-screen text display or voice output, offering flexibility based on user preferences and the communication context. The user-friendly interface has been meticulously designed to facilitate interaction for both mute individuals and those who can hear and speak.

Beyond breaking down the communication barrier, our mobile application aspires to cultivate inclusively, understanding, and enriched social interaction. It serves as a versatile tool, not only converting sign language to text but also incorporating a text-to-voice option. This dual functionality ensures that the application caters to a diverse range of users and promotes seamless communication between mute individuals and the broader community. Our overarching vision is to contribute to a world where communication is universally accessible, fostering a more inclusive and connected society.

Chapter 2

Literature Review

2.1 Related work And State of the Art (Latest work)

Currently, there exists a system that focuses on American Sign Language, providing real-time conversion to text. This innovation enables individuals who are deaf and mute to independently conduct their daily transactions without relying on an interpreter (Arora & Roy, 2018). Several projects have showcased the efficacy of Convolutional Neural Networks (CNN) in solving computer vision problems with remarkable accuracy (Ojha, Pandey, Maurya, Thakur, & Dayananda, 2020). Additionally, some projects have employed multiclass Support Vector Machines (SVM) to handle both static and dynamic gestures (Vedak, Zavre, Todkar, & Patil, 2019).

A noteworthy example is the development of a web-based application, such as Fingerspelling, which educates users on correct hand positions in real-time. Leveraging webcams, this app tracks hand movements while an algorithm assesses their accuracy [4]. Furthermore, certain projects specialize in translating sign language videos, utilizing smart algorithms to convert gestures into speech almost instantaneously [5].

These initiatives collectively highlight the diverse approaches taken to enhance communication and accessibility for individuals using sign language, showcasing advancements ranging from accurate gesture recognition to real-time translation and educational tools.

2.2 Limitation of State of the Art techniques

The existing systems developed so far are predominantly tailored for American Sign Language [4]. A limitation observed in certain systems lies in their reliance on histograms for image recognition, restricting their applicability to images with pre-saved histograms (Arora & Roy, 2018). Additionally, several systems exhibit inaccuracies, particularly when the user's gesture posture deviates from the expected form, resulting in incorrect predictions (Ojha et al., 2020).

In some models, the necessity of wearing gloves for sign detection poses a constraint on user convenience (Ojha et al., 2020). Affordability emerges as a prevalent issue, with many existing systems being deemed financially inaccessible [5]. Moreover, a notable limitation is the incapacity of certain systems to accurately recognize signs that involve dynamic hand movements (Narvekar, Mungekar, Pandey, & Mahadik, n.d.).

These identified constraints underscore the need for advancements in sign language recognition systems, calling for solutions that overcome limitations related to language variations, posture accuracy, user comfort, affordability, and dynamic sign gestures. Addressing these challenges is essential for the development of inclusive and effective communication tools for individuals using sign language.

2.3 Discussion and future direction

In our project, we will meticulously curate a diverse dataset encompassing various categories such as alphabets, numbers, common phrases, idioms, and more. To train our model, we've chosen to leverage the user-friendly Google Teachable Machine platform, facilitating the efficient training of machine learning models. Our approach involves collecting an extensive array of images for each specific sign, aiming to enhance the accuracy and robustness of our model.

A significant advantage of our system is its independence from users wearing gloves for successful image recognition. This eliminates the need for specialized equipment or accessories, significantly improving convenience and ease of use for individuals who are mute.

Notably, our project addresses the limitations prevalent in existing systems, emphasizing improvements in both accuracy and usability. Our goal is to offer a more comprehensive and accessible solution for communication between mute individuals and others. Through our innovative approach and meticulous attention to accuracy, we aspire to overcome existing barriers and develop a system that provides reliable and precise image recognition results, all without the necessity for additional accessories or compromising the user experience.

2.4 Concluding Remarks

Our project focuses on bridging the communication gap for individuals who are deaf or hearing impaired by introducing a cost-effective mobile application. This app aims to automate the process of capturing, recognizing, and translating sign language into speech, benefiting the deaf community. Conversely, it also analyzes hand gestures and converts them into text displayed on the screen for the benefit of the hearing impaired.

The mobile app will leverage the capabilities of modern smartphones, utilizing their camera to capture and analyze sign language gestures in real-time. Through advanced image processing and machine learning algorithms, the app will accurately recognize and interpret these gestures. The recognized gestures will then be translated into spoken language, allowing deaf individuals to understand and participate in conversations more easily.

Conversely, for individuals who are hearing impaired, the app will analyze spoken language input and convert it into text displayed on the screen. This facilitates effective communication with others, even in situations where verbal communication may be challenging or impossible.

Our goal is to provide an accessible and affordable solution that empowers individuals who are deaf or hearing impaired to engage in seamless communication with the broader community. By leveraging mobile technology and advanced algorithms, we aim to enhance inclusivity, break down communication barriers, and foster understanding and connection between individuals with different hearing abilities.

Chapter 3

Problem Definition and Scope

3.1 Problem statement

The communication challenges faced by specially-abled individuals, specifically those who are deaf and mute, represent a significant barrier to their full participation in daily interactions. Existing solutions often fall short in providing inclusive and user-friendly methods of communication, with limitations ranging from dependency on specialized equipment to the exclusion of certain sign languages. The identified problem is the insufficient availability of accessible and accurate communication tools tailored to the unique needs of deaf and mute individuals. This inadequacy hinders their ability to engage effectively with the broader community and impedes their participation in various aspects of daily life. The objective of our project is to address this critical problem by developing a mobile application that offers a comprehensive and universally applicable solution, overcoming the limitations of current communication tools and fostering greater inclusivity for specially-abled individuals.

3.2 Goals and Objectives

1. To enhance communication accessibility for individuals with disabilities by bridging the gap between the disabled and able-bodied individuals.

2. To develop advanced language interpretation technology that can accurately

convert spoken language into written text, enabling effective communication for people with hearing impairments or those who prefer written communication.

3. To create intuitive and user-friendly tools or applications that facilitate learning and understanding of sign language for beginners. This can include interactive tutorials, visual aids, and feedback mechanisms to make the learning process more engaging and efficient.

4. To prioritize affordability and reliability in the development of the application, ensuring that it is accessible to a wide range of users, regardless of their financial capabilities. This can be achieved through partnerships with organizations, offering subsidized pricing, or exploring sustainable funding models to keep the app affordable while maintaining its reliability and quality.

3.3 Scope and Major Constraints

Main Users: Common people

3.4 Hardware and Software Requirements

1. Operating System: Windows

2. Programming Language: Python

3. Computer Vision Libraries: OpenCV is required for image processing, feature extraction, and object detection.

4. Machine Learning Framework: TensorFlow, Keras are used for training and inference tasks.

5. Neural Network Architectures: Convolutional neural networks (CNNs) is used for image processing.

6. Development Environment: Jupyter Notebook.

3.5 Expected Outcomes

1. Accurate Sign Language Recognition: The system should be capable of accurately detecting and recognizing sign language gestures and movements performed by users. It should have a high degree of accuracy in identifying and understanding different signs and their meanings.

2. Real-Time Translation: The system should provide real-time translation of sign language into spoken or written language. It should be able to process sign language gestures swiftly and produce corresponding translations in a timely manner, enabling seamless communication between sign language users and non-sign language users.

3. Continuous Improvement: The project should have mechanisms in place for continuous improvement and refinement of the sign language translator system. This can involve collecting user feedback, updating the system with new sign language patterns or gestures, and incorporating machine learning techniques to enhance accuracy over time.

Chapter 4

System Requirement Specification

4.1 Overall Description

4.1.1 Block diagram/ Proposed System setup

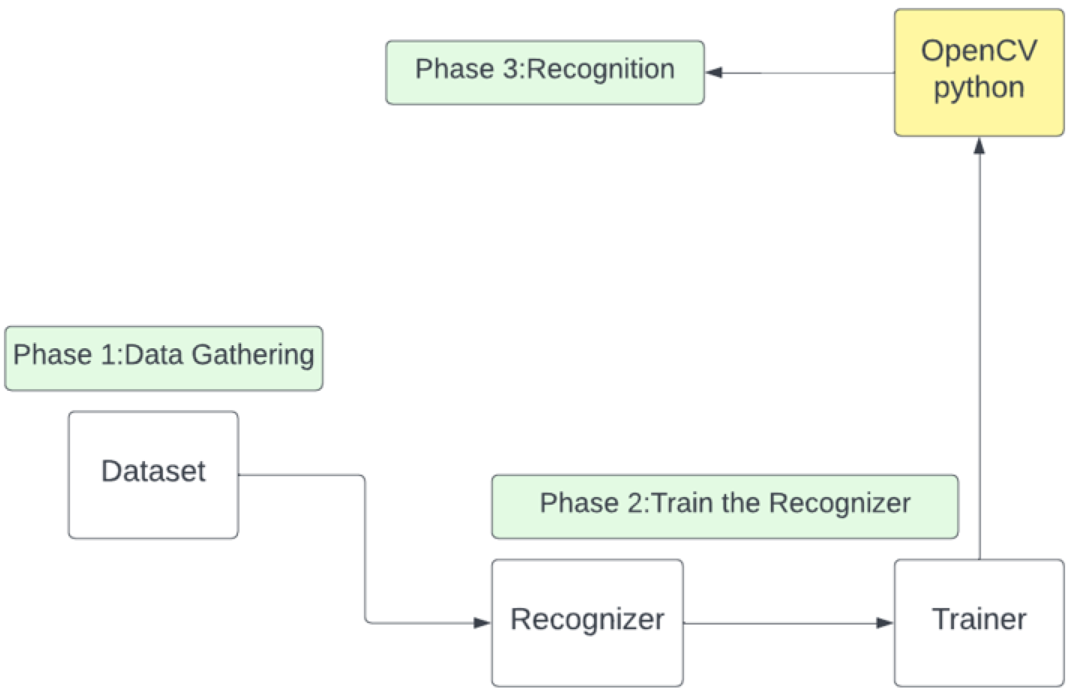

Figure 4.1: Block Diagram

4.1.2 Use Case Diagram

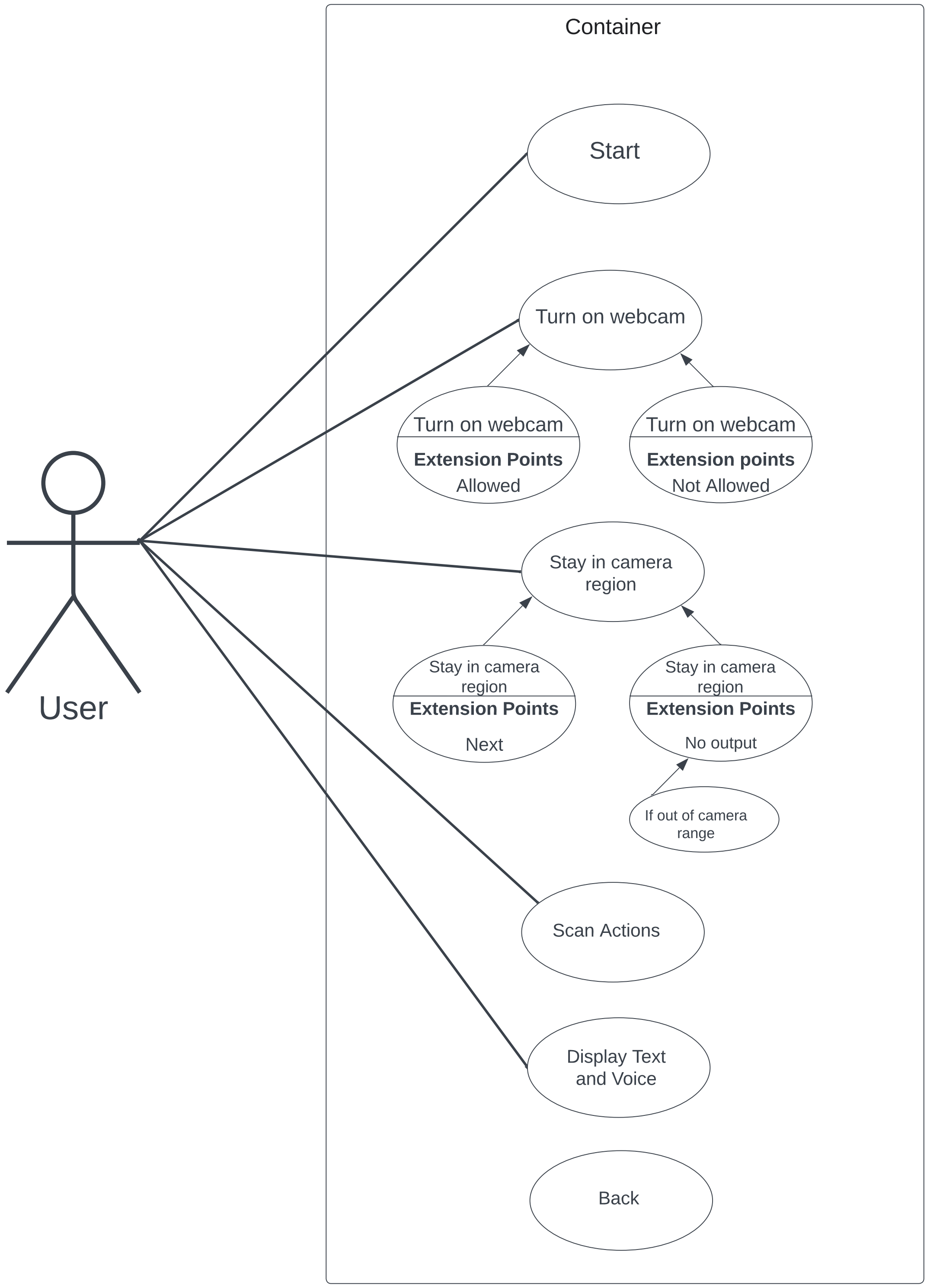

Figure 4.2: Use Case Diagram

4.1.3 Sequence Diagram

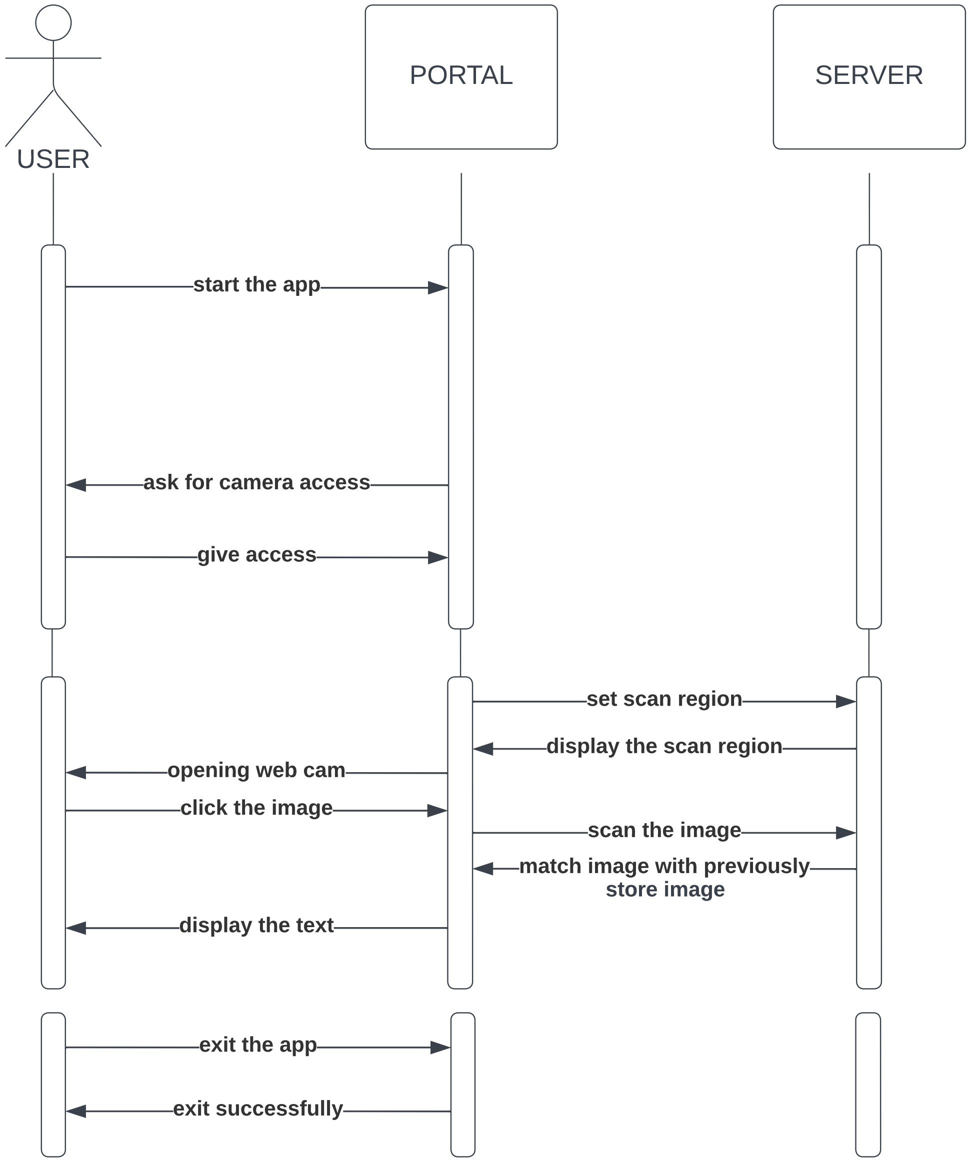

Figure 4.3: Sequence diagram

4.1.4 Activity Diagram

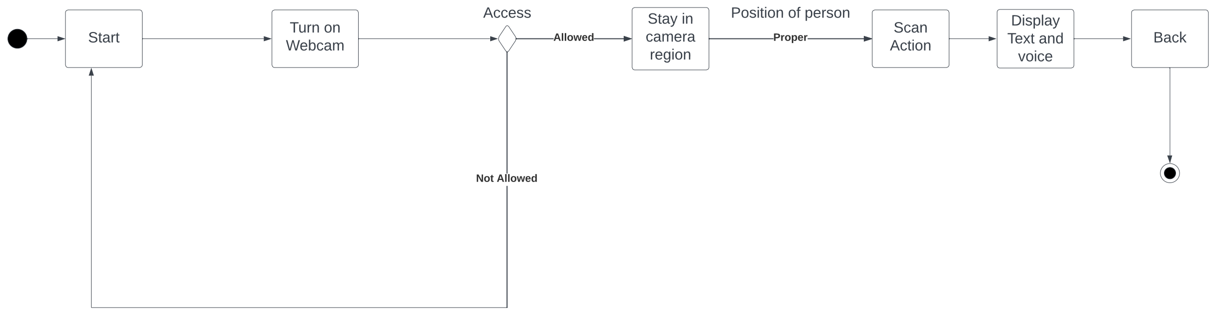

Figure 4.4: Activity Diagram

4.1.5 Hardware and Software Requirements

Hardware Requirement

1. Mobile
2. Webcam

Software Requirement

1. Python libraries – open-cv, numpy, tensorflow, mediapipe.
2. Python compiler – jupyter notebook.
3. Android Studio.

Chapter 5

Proposed Methodology

5.1 System Architecture

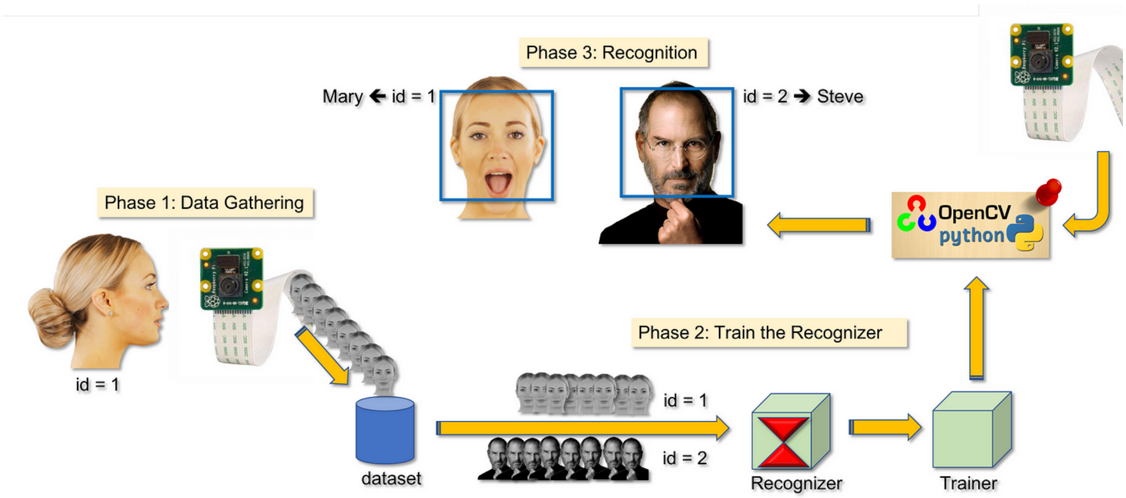

Figure 5.1: System Architecture

5.2 Mathematical Modeling

Convolution Neural Networks (CNN):

Computer vision is a field of Artificial Intelligence that focuses on problems related to images and videos. CNN combined with Computer vision is capable of performing complex problems.

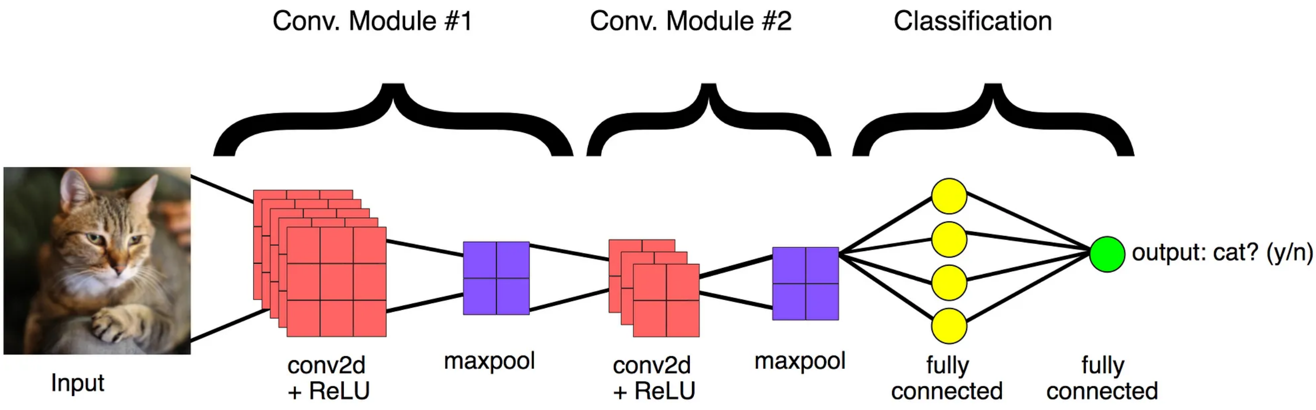

Figure 5.2: Working of CNN

The Convolution Neural Networks has two main phases namely feature extraction and classification. A series of convolution and pooling operations are performed to extract the features of the image. The size of the output matrix decreases as we keep on applying the filters. Size of new matrix = (Size of old matrix — filter size) +1
A fully connected layer in the convolution neural networks will serve as a classifier. In the last layer, the probability of the class will be predicted. The main steps involved in convolution neural networks are:

1.Convolution

2. Pooling

3. Flatten

4. Full connection

1. **Convolution**:

   Convolution is nothing but a filter applied to an image to extract the features from it. We will use different filters to extract features like edges, highlighted patterns in an image. The filters will be randomly generated. What this convolution does is, creates a filter of some size says 3x3 which is the default size. After creating the filter, it starts performing the element-wise multiplication starting from the top left corner of the image to the bottom right of the image. The obtained results will be extracted feature.

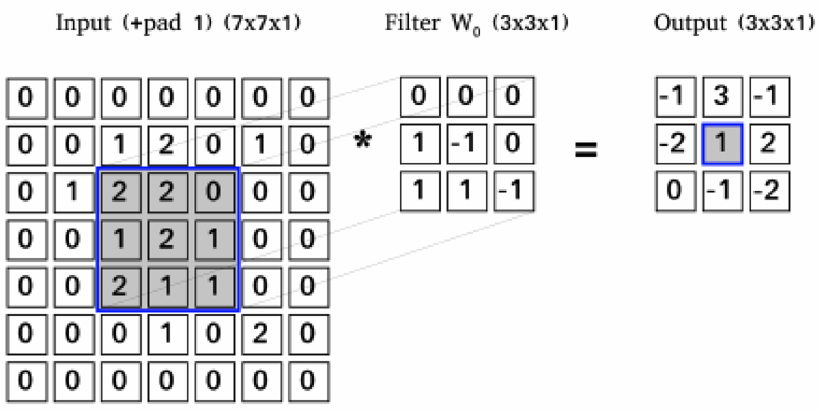

Figure 5.3: Convolution

2. **Pooling**:

   After the convolution operation, the pooling layer will be applied. The pooling layer is used to reduce the size of the image

   There are two types of pooling:

   (a). Max Pooling

   (b). Average Pooling

   (a) Max pooling:

   Max pooling is nothing but selecting the maximum pixel value from the matrix. This method is helpful to extract the features with high importance or which are highlighted in the image.

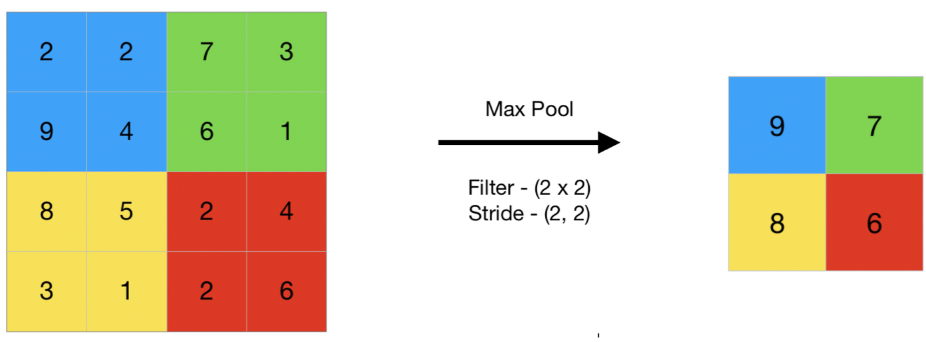

Figure 5.4: Max pooling

(b) Average pooling:

Unlike Max pooling, the average pooling will take average values of the pixel. In most cases, max pooling is used because its performance is much better than average pooling.

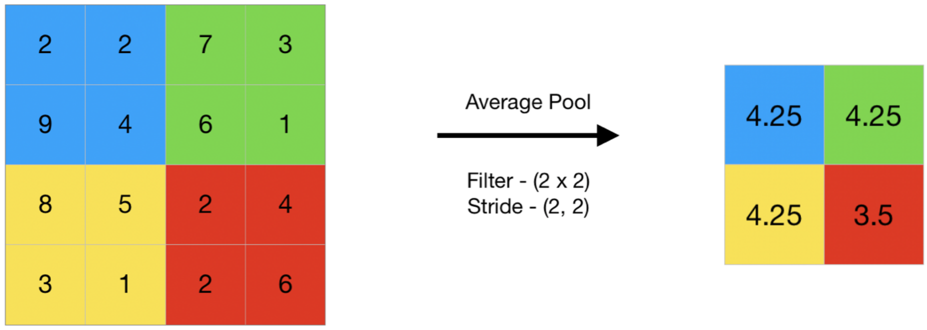

Figure 5.5: Average pooling

3. **Flatten**:

The obtained resultant matrix will be in multi-dimension. Flattening is converting the data into a 1-dimensional array for inputting the layer to the next layer. We flatten the convolution layers to create a single feature vector.

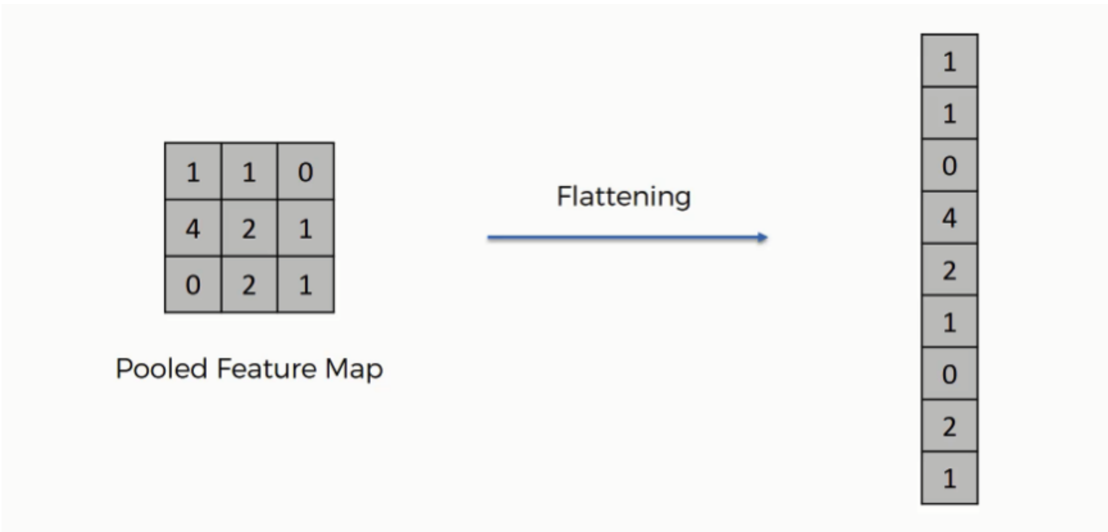

Figure 5.6: Flatten

4. **Fully connection**:

   A fully connected layer is simply a feed-forward neural network. All the operations will be performed and prediction is obtained. Based on the ground truth the loss will be calculated and weights are updated using gradient descent back propagation algorithm

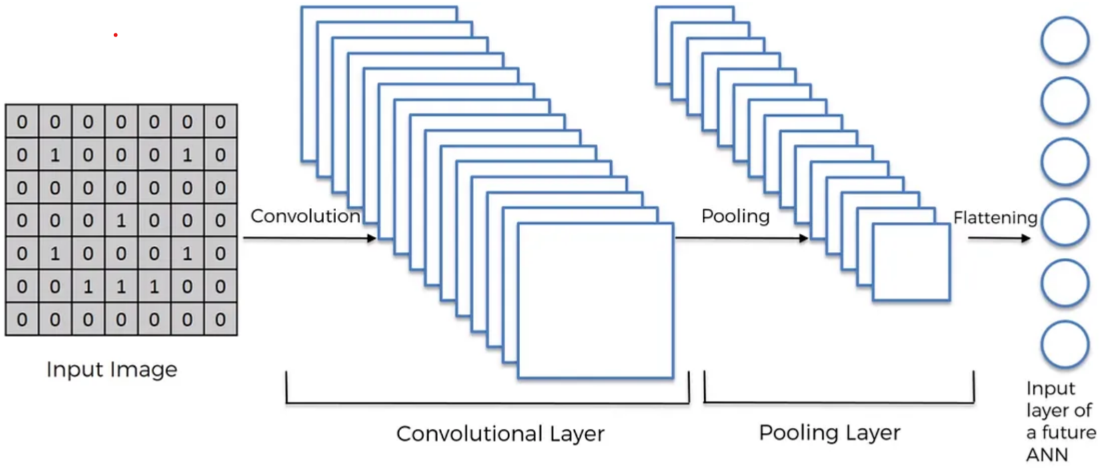

Figure 5.7: Fully connected

MediaPipe:

MediaPipe, developed by Google, is an open-source, cross-platform library for cutting-edge machine learning solutions. Its versatility spans iOS, Linux, Windows, and Android, including low-powered devices like Raspberry Pi.

With a focus on accessibility and efficiency, MediaPipe provides solutions for face detection, iris detection, human pose estimation, and instant motion tracking.

One of its standout features is cross-platform compatibility, ensuring seamless performance across various environments, from high-powered systems to resource-constrained devices like Raspberry Pi.

Being open-source encourages collaboration, allowing developers to customize and contribute to the library's continual evolution. MediaPipe finds applications in computer vision, augmented reality, robotics, and healthcare, making it a dynamic tool in the realm of machine learning.

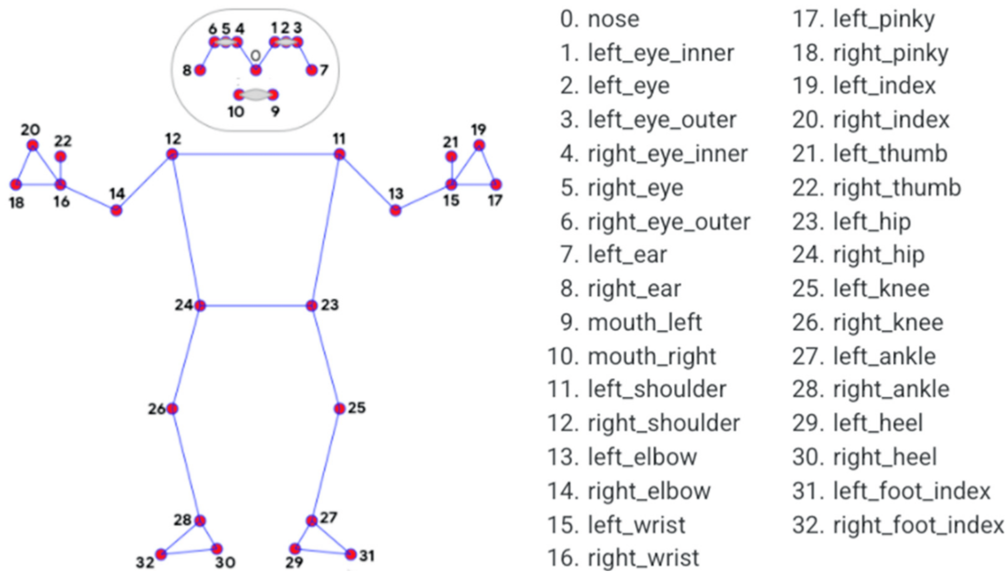

Figure 5.8: 33 Pose Landmarks

Chapter 6

Conclusion

6.1 Conclusion

In summary, our system for converting sign language into text represents a breakthrough in simplicity and cost-effectiveness, setting it apart from alternative methods that often rely on intricate hardware. By leveraging the straightforward input mechanism of capturing hand gestures through images in real-time, our system eliminates the need for complex setups, making it an accessible and user-friendly solution. Notably, the system's ability to yield accurate results even with partially captured inputs ensures its robustness and practicality in diverse real-world scenarios, minimizing processing overhead and maximizing reliability.

The significance of our innovation extends beyond technical efficiency to the profound impact it can have on the lives of the deaf and mute community. This system, with its simplicity and precision, holds the potential to empower individuals, fostering self-reliance and enabling increased societal participation. By providing an accessible means of communication, our solution contributes not only to the technological landscape but also to the broader goal of fostering inclusivity, independence, and overall societal development.

6.2 Future Scope

The future scope for this project is promising, with several avenues for enhancement and expansion. First and foremost, further refinement of the image recognition algorithms can be pursued to improve the system's accuracy and expand its vocabulary of recognized sign gestures. Implementing machine learning techniques, such as deep learning models, can contribute to continuous improvement by enabling the system to learn and adapt to variations in sign language expressions over time.

Additionally, incorporating real-time feedback mechanisms and user interactions can enhance the system's responsiveness and user experience. This could involve integrating natural language processing capabilities to interpret context and refine the accuracy of text output. Collaborating with the deaf and mute community for user feedback and involvement in the development process will be essential to ensure that the system addresses their specific needs and preferences.

Exploring mobile application development can extend the reach of the system, allowing users to access the sign language to text conversion on portable devices. Integration with other assistive technologies and communication platforms can further amplify the system's impact, fostering seamless communication for the deaf and mute individuals across various contexts. Lastly, ongoing research and development efforts should focus on maintaining compatibility with emerging technologies and standards, ensuring the system remains adaptive and relevant in the rapidly evolving landscape of assistive technologies.

Appendices

Appendix A

OpenCV

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today'ssystems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

GETTING STARTED (HOW TO READ IMAGES)

1. Open Jupyter Notebook
2. Import cv2
3. Paste a test image in the directory.
4. Create variable to store image using imread() function
5. Display the image using imshow() function.
6. Add a delay using a waitkey() function

   1.import cv2

   2.LOAD AN IMAGE USING 'IMREAD'

   3.img = cv2.imread("Resources/lena.png")

   4.DISPLAY

5.cv2.imshow("Lena Soderberg", img)

6.cv2.waitKey(0)

ACCESSING LIVE FEED FROM WEBCAM

1. Open Jupyter Notebook.
2. Import cv2.
3. Create variable to store video using VideoCapture() function.
4. Pass parameter 0 in VideoCapture(0) to access webcam.
5. Create an infinite while loop to display each frame of the webcam's video continuously.
6. Display the live feed using imshow() function.
7. Add a delay of infinity using waitKey(0).

   1.import cv2

   2.Width = 640

   3.Height = 480

   4.cap = cv2.VideoCapture(0)

   5.cap.set(3, frameWidth)

   6.cap.set(4, frameHeight)

   7.cap.set(10, 150)

   8.while True: success, img = cap.read()

   9 .cv2.imshow("Result", img)

References

(n.d.).

An ai app that converts sign language into text, speech. (n.d.). The Economics Time.

Arora, S., & Roy, A. (2018). Recognition of sign language using image processing. *International Journal of Business Intelligence and Data Mining*, *13*(1-3), 163–176.

Narvekar, C., Mungekar, S., Pandey, A., & Mahadik, M. (n.d.). Sign language to speech conversion using image processing and machine learning.

Ojha, A., Pandey, A., Maurya, S., Thakur, A., & Dayananda, P. (2020). Sign language to text and speech translation in real time using convolutional neural network. *Int. J. Eng. Res. Technol.(IJERT)*, *8*(15), 191–196.

Vedak, O., Zavre, P., Todkar, A., & Patil, M. (2019). Sign language interpreter using image processing and machine learning. *International Research Journal of Engineering and Technology (IRJET)*.