VISVESVARAYA TECHNOLOGICAL UNIVERSITY

JNANA SANGAMA, BELAGAVI – 590 018, KARNATAKA

A Project Report
on

SmartVoice

Submitted in partial fulfillment of the requirements for the VIII Semester of degree of ***Bachelor of Engineering in Information Science and Engineering*** *of Visvesvaraya Technological University, Belagavi*

by

Abhaya Simha S P	Adithya R Pai	Chinmay Hegde	Rakshita S
1RN19IS002	1RN19IS010	1RN19IS053	1RN19IS116

Under the Guidance of

Dr. Mamatha G

Professor

Department of ISE

Department of Information Science and Engineering

RNS INSTITUTE OF TECHNOLOGY

Dr. Vishnuvaradhan Road, Rajarajeshwari Nagar Post,
Channasandra, Bengaluru – 560 098

2022-2023

RNS INSTITUTE OF TECHNOLOGY

Dr. Vishnuvaradhan Road, Rajarajeshwari Nagar Post
Channasandra, Bengaluru – 560 098

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

CERTIFICATE

Certified that the project work entitled ***SmartVoice*** has been successfully completed by **Abhaya Simha S P (1RN19IS002), Adithya R Pai (1RN19IS010), Chinmay Hegde (1RN19IS053) and Rakshita S (1RN19IS116),** bonafide students of **RNS Institute of Technology, Bengaluru** in partial fulfillment of the requirements for the award of degree **Bachelor of Engineering** in **Information Science and Engineering** of **Visvesvaraya Technological University, Belagavi** during academic year **2022-2023**. The project report has been approved as it satisfies the academic requirements in respect of project work for the said degree.

Dr. Mamatha G Project Guide	**Dr. Prakasha S** Project Coordinator
Dr. Suresh L Professor and HOD	**Dr. Ramesh Babu H S** Principal

External Viva

Name of the Examiners	Signature with Date
1. _____	1. _____
2. _____	2. _____

DECLARATION

We, **ABHAYA SIMHA S P [USN: 1RN19IS002], ADITHYA R PAI [USN:1RN19IS010], CHINMAY HEGDE [USN: 1RN19IS053], RAKSHITA S [USN:1RN19IS116]** students of VIII Semester B.E. in Information Science and Engineering, RNS Institute of Technology hereby declare that the Project entitled ***SmartVoice*** has been carried out by us and submitted in partial fulfillment of the requirements for the ***VIII*** *Semester of degree of* ***Bachelor of Engineering in Information Science and Engineering*** *of Visvesvaraya Technological University, Belagavi* during academic year 2022-2023.

Place: Bengaluru
Date:

ABHAYA SIMHA S P **(1RN19IS002)**

ADITHYA R PAI **(1RN19IS010)**

CHINMAY HEGDE **(1RN19IS053)**

RAKSHITA S **(1RN19IS116)**

ABSTRACT

The project aims to address the communication challenges faced by individuals who are hearing-impaired, specifically those who are unable to speak and hear. These individuals often find it difficult to understand and convey messages to others, making it challenging for them to participate fully in society. The project proposes a solution that utilizes hand gestures, a crucial component of sign language, to facilitate communication between hearing-impaired individuals and the general population.

The system incorporates a webcam connected to a computer, which captures images of the user's hand movements. These images are then processed using contour feature extraction techniques to identify and recognize the specific hand gestures being performed. By analyzing the contours and patterns of the hand movements, the system can interpret the intended message behind each gesture.

The key innovation of this project lies in the automatic translation of these recognized hand gestures into text outputs. This translation feature enables individuals with disabilities to communicate effectively with individuals who do not understand sign language. By converting the hand gestures into text, the system bridges the communication gap between the hearing-impaired and the general population, allowing for equal participation and inclusion.

The proposed system holds great potential in enhancing the lives of individuals with special needs by providing them with a reliable and efficient means of communication. It empowers them to express their thoughts, needs, and emotions, enabling them to interact more confidently and independently with others.

Furthermore, this project aligns with the broader goal of fostering inclusivity and equality within society. By creating a technology that facilitates communication for individuals with disabilities, barriers are broken down, and opportunities for engagement and collaboration are increased. This not only benefits the hearing-impaired individuals themselves but also promotes a more inclusive and understanding society that values diversity.

In conclusion, the project's use of hand gesture recognition, coupled with automatic translation to text, offers a promising solution for improving communication between hearing-impaired individuals and the wider community. By providing an accessible and efficient means of communication, this project strives to empower individuals with special needs, enabling them to actively participate and contribute to society on an equal footing.

ACKNOWLEDGMENT

We would like to place our gratefulness to all those people who played an important role in guiding and making the project successful.

The fulfillment and rapture that go with the fruitful finishing of any assignment would be inadequate without specifying the people who made it conceivable, whose steady direction and support delegated the endeavors with success.

First of all, we would like to thank the Management of **RNS Institute of Technology** for providing such a healthy environment for the successful completion of project work.

I would like to express my thanks to our Director **Dr. M K Venkatesha** for his support and inspiring me towards the attainment of knowledge.

I would like to express my thanks to our Principal **Dr. Ramesh Babu H S** for his support and inspiring me towards the attainment of knowledge.

We are extremely grateful to our own and beloved Professor and Head of the Department of Information science and Engineering, **Dr. Suresh L**, for having accepted to patronize us in the right direction with all his wisdom.

We place our heartfelt thanks to **Dr. Mamatha G**, Professor, Department of Information Science and Engineering for having guided for project and all the staff members of our department for always helping us out.

We would like to thank Project coordinators **Dr. Prakasha S**, Associate Professor and **Dr. Bhagyashree Ambore**, Assistant Professor, Department of Information Science and Engineering for supporting and guiding us all through.

We thank our beloved friends and our parents for supporting and encouraging us throughout. We made an honest effort in this assignment.

ABHAYA SIMHA S P
ADITHYA R PAI
CHINMAY HEGDE
RAKSHITA S

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

API	American Sign Language
ASL	Application Programming Interface
HMM	Hidden Markov Models
ISL	Indian Sign Language
KNN	k-nearest neighbors
PCA	Principal Component Analysis
RDBMS	Relational Database Management System
SIFT	Scale-Invariant Feature Transform
SRS	Software Requirements Specification
SVM	Support Vector Machines

Chapter 1

INTRODUCTION

1.1 Background

Indian Sign Language (ISL) is a visual language that is used by the deaf community in India for communication. It is a complex language with its own grammar and syntax, and it is different from other sign languages used in other countries. Communication between the hearing and the deaf community in India is often limited due to the lack of understanding and familiarity with ISL.

The development of a machine learning model for ISL gesture detection involves several steps, including data collection, data preprocessing, feature extraction, and model training. One of the key challenges in this process is the lack of standardized ISL gestures, as different regions and communities may use different variations of the same gesture. This variability can make it difficult to develop a model that can accurately recognize all ISL gestures. As the technology and techniques continue to advance, the accuracy and robustness of ISL gesture detection models will continue to improve.

1.2 Existing Systems and their Limitations

Existing systems for Indian sign language detection using machine learning have made progress but face several limitations. These include the availability of limited and diverse datasets, the challenge of capturing the variability of gestures accurately, and the need for real-time recognition. Lighting and background conditions can also affect system accuracy, and scalability and adaptability to new gestures and individual differences pose challenges.

Some systems require specialized hardware, which limits their accessibility. Overcoming these limitations requires further research in data augmentation, transfer learning, and improved feature extraction methods. Advancements in hardware technology can contribute to real-time and scalable solutions. Collaboration between researchers, linguists, and the deaf community is crucial for inclusive and accurate systems.

1.3 Proposed System

The system's hardware is accessible, portable, and lightweight, making it convenient for users to carry it around and use it whenever needed. The software is user-friendly, interactive, and able to provide real-time feedback to the user to facilitate effective communication.

The proposed system is a sign language recognition system that uses MediaPipe, an open-source framework, to detect and recognize sign language gestures in real-time. and Random Forest Technique.

The Random Forest technique has been used over other neural classification algorithms because it is less computationally expensive and does not require a GPU to finish training. A random forest can give you a different interpretation of a decision tree but with better performance. Neural Networks will require much more data than an everyday person might have on hand to actually be effective.

1.4 Advantages of The Proposed System

- It provides an efficient and accurate means of interpreting sign language gestures, which can improve communication and inclusivity for individuals who are deaf or hard-of-hearing.

- The system is cost-effective and can be implemented using a standard webcam and a computer, making it accessible to a wider range of individuals.

- The use of machine learning models and real-time processing allows for the system to adapt and improve over time, increasing its accuracy and usability.

Chapter 2

LITERATURE SURVEY

A Literature survey describes the idea of how the concept has emerged, how it has been implemented and what is the current status.

[Paper-01] Deaf Mute Communication Interpreter - A Review

Authors: Sunitha K. A, Anitha Saraswathi.P, Aarthi.M, Jayapriya. K, Lingam Sunny

This paper [1] provides a comprehensive review of systems designed to facilitate communication for individuals with hearing and speech impairments.

The authors emphasize the importance of effective communication for deaf and mute individuals in order to enable their full participation in society. They highlight the challenges faced by these individuals and the need for technological solutions to bridge the communication gap. The review covers various aspects of Deaf Mute Communication Interpreter systems, including hand gesture recognition and translation techniques. The authors discuss different approaches used in these systems, such as contour feature extraction, wavelet transform, and principal component analysis (PCA).

The paper explores the preprocessing steps involved in capturing hand gestures, such as hand segmentation, which separates the hand region from the background. Feature extraction techniques are then employed to extract relevant information from the hand images. These features serve as input for gesture classification algorithms, which interpret the gestures and generate corresponding text or speech outputs.

Furthermore, the paper emphasizes the need for continuous research and development in this field to improve the recognition accuracy and expand the vocabulary of recognized gestures. The authors also stress the importance of usability and user-friendliness in designing communication aids for deaf and mute individuals.

[Paper 02] An Efficient Framework for Indian Sign Language Recognition Using Wavelet Transform

Authors: Mathavan Suresh Anand, Nagarajan Mohan Kumar, Angappan Kumaresan

In this paper [2], The authors start by highlighting the significance of ISL recognition systems in enabling communication between hearing-impaired individuals and the general population. They emphasize the need for efficient and accurate recognition techniques to facilitate effective communication.

The proposed framework utilizes the wavelet transform as a feature extraction technique for analyzing hand gestures. The authors describe the preprocessing steps,

which involve segmenting the hand region from the background and extracting relevant features using wavelet transform coefficients. The extracted features are then used as input for the classification stage. The authors employ a support vector machine (SVM) classifier to recognize different ISL gestures. They discuss the training and testing processes of the classifier and the evaluation metrics used to assess the performance of the system.

The paper presents experimental results to demonstrate the effectiveness of the proposed framework. The authors compare the recognition accuracy of their approach with other existing techniques and show improved results. They also discuss the computational efficiency of the system, highlighting its real-time capability.

Furthermore, the authors discuss the challenges and limitations of the proposed framework. They address issues such as variations in lighting conditions, hand orientations, and occlusions that can affect the recognition performance. They suggest potential solutions and avenues for future research to overcome these challenges.

[Paper-03] Hand Gesture Recognition Using PCA

Authors: Mandeep Kaur Ahuja, Amardeep Singh

In this paper [3] , the authors present a method for hand gesture recognition based on Principal Component Analysis (PCA). The approach involves preprocessing the hand gesture images, applying PCA for feature extraction, and using a k-nearest neighbors (KNN) algorithm for classification. Experimental results demonstrate that the proposed method achieves competitive recognition accuracy. The system shows potential for real-time applications in sign language interpretation and human-computer interaction. The use of PCA allows for dimensionality reduction, simplifying the representation of hand gestures and improving efficiency. However, the paper acknowledges challenges such as occlusions, lighting variations, and hand shape differences. Suggestions for future work include exploring other feature extraction techniques and classification algorithms to further enhance recognition performance. Overall, the paper contributes to the field of hand gesture recognition and provides insights into an efficient and accurate method for interpreting hand gestures using PCA.

[Paper-04] Hand gesture recognition system for dumb people

Authors: Sagar P.More, Prof. Abdul Sattar

In this paper [4] , the authors focuses on the development of a hand gesture recognition system specifically designed for individuals who are unable to speak, referred to as "dumb" people in the paper. The authors address the communication challenges faced by these individuals and propose a system that can interpret hand gestures as a means of communication. The system utilizes image processing techniques to capture and analyze hand gestures captured by a camera. The paper

describes the various stages of the system, including image acquisition, hand segmentation, feature extraction, and gesture classification. Feature extraction techniques are employed to extract relevant information from the hand images, and a classification algorithm is used to interpret and recognize different gestures.

The authors present experimental results to demonstrate the effectiveness of their system in recognizing hand gestures accurately. The system shows potential in facilitating communication for individuals who are unable to speak, enabling them to convey their messages effectively.

In conclusion, the paper presents a hand gesture recognition system designed specifically for individuals who are unable to speak. The proposed system utilizes image processing techniques and classification algorithms to interpret and recognize hand gestures accurately. This research has the potential to provide a valuable means of communication for individuals who face challenges in expressing themselves verbally.

[Paper -05] An Automated System for Indian Sign Language Recognition

Authors: Chandandeep Kaur, Nivit Gill

In this paper[5], the authors focus on developing a system that can understand and interpret ISL gestures, enabling effective communication between hearing-impaired individuals and the general population. The system employs image processing techniques to capture and analyze hand gestures using a camera. The paper describes the various stages involved, including hand region detection, feature extraction, and gesture classification. The authors present experimental results to demonstrate the accuracy and effectiveness of the system in recognizing ISL gestures. This research contributes to the development of technology that promotes inclusivity and accessibility by facilitating communication for individuals who rely on sign language.

Paper-06] Hand Gesture Recognition for Sign Language Recognition: A Review

Authors: Pratibha Pandey, Vinay Jain

This paper [6], provides a comprehensive review of hand gesture recognition techniques specifically focused on sign language recognition.

The authors recognize the significance of sign language in facilitating communication for individuals with hearing impairments and emphasize the importance of accurate hand gesture recognition for effective sign language interpretation. The paper reviews various approaches and algorithms proposed by different researchers for hand gesture recognition, covering aspects such as hand segmentation, feature extraction, and classification. The review highlights the strengths and limitations of different techniques, including template matching, neural networks, Hidden Markov Models (HMMs), and Support Vector Machines (SVMs). The authors analyze the performance metrics used to evaluate the recognition systems and discuss challenges faced in real-time implementation.

shapes and sizes. The paper also emphasizes the need for robust and efficient recognition systems that can handle complex sign language gestures and accommodate dynamic variations. It concludes by suggesting future research directions, such as the integration of multiple modalities, such as vision and sensors, and the exploration of deep learning techniques for improved accuracy and robustness in sign language recognition systems.

Overall, the paper provides a comprehensive review of hand gesture recognition techniques for sign language recognition, offering insights into the existing approaches and highlighting areas for further research and development in this important field.

[Paper-07] Design Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People

Authors: Nakul Nagpal,Dr. Arun Mitra., Dr. Pankaj Agrawal

In this paper [7], the authors address the design challenges and presents a proposed implementation of a communication aid for individuals who are deaf and dumb.

The authors recognize the communication barriers faced by this specific group and aim to develop a system that assists them in expressing their thoughts and understanding others. The paper discusses the design considerations involved in creating such a system, including the need for simplicity, ease of use, and adaptability to individual user needs.

The proposed implementation involves the use of a graphical user interface (GUI) and a combination of image and text-based communication methods. The system allows users to input text or select predefined symbols and phrases, which are then converted into visual or textual outputs, enabling communication with non-sign language users.

The paper highlights the importance of user testing and feedback to refine the system's design and functionality. It also discusses potential future enhancements, such as incorporating natural language processing and expanding the symbol database.

In conclusion, the paper addresses the design challenges of developing a communication aid for individuals who are deaf and dumb. The proposed implementation offers a user-friendly interface and utilizes a combination of images and text to facilitate effective communication. The research contributes to improving accessibility and inclusivity for this specific user group by providing a tool that enhances their ability to express themselves and interact with others.

[Paper-8] Real Time Detection And Recognition Of Indian And American Sign Language Using Sift

Authors: Neelam K. Gilorkar, Manisha M. Ingle

The paper [8] outlines the different stages of the system, including hand detection, hand segmentation, and SIFT feature extraction. The authors describe the extraction of key points and descriptors from the hand gesture images, which are then used for matching and recognition.

Experimental results presented in the paper demonstrate the effectiveness of the proposed system in accurately recognizing hand gestures from both ISL and ASL. The real-time performance of the system highlights its practical applicability in various domains, such as sign language interpretation and assistive technologies.

The authors discuss the advantages of using SIFT for sign language recognition, including its invariance to scale, rotation, and affine transformations. They also acknowledge the challenges associated with variations in hand shapes, lighting conditions, and occlusions.

In conclusion, the paper presents a real-time sign language recognition system that utilizes SIFT for detecting and recognizing hand gestures in ISL and ASL. The research contributes to the advancement of technology in improving communication accessibility for individuals with hearing impairments.

[Paper-9] Sign Pro-An Application Suite for Deaf and Dumb

Authors: Ashish Sethi, Hemanth ,Kuldeep Kumar, Bhaskara Rao ,Krishnan R

In this paper[9], the authors acknowledge the communication challenges faced by this specific group and propose Sign Pro as a solution to facilitate effective communication. The application suite includes various modules that address different aspects of communication, such as text-to-speech conversion, sign language recognition, and gesture-based input.

The paper outlines the functionalities of each module in the Sign Pro application suite. The text-to-speech module converts written text into spoken words, enabling users to understand written content. The sign language recognition module allows users to input sign language gestures, which are then recognized and translated into textual or auditory outputs. The gesture-based input module enables users to interact with the system through hand gestures, providing an alternative means of input.

In conclusion, the paper introduces Sign Pro as an application suite aimed at assisting individuals who are deaf and dumb in their communication. The suite incorporates modules for text-to-speech conversion, sign language recognition, and gesture-based input, offering a comprehensive solution to overcome communication barriers. The research contributes to the development of technology that enhances the accessibility and inclusivity of individuals with hearing and speech impairments.

[Paper-10] A Review on Feature Extraction for Indian and American Sign Language
Authors: Neelam K. Gilorkar, Manisha M. Ingle

In this paper[10], the authors recognize the importance of accurate feature extraction in sign language recognition systems, as it plays a crucial role in capturing relevant information from hand gesture images. The paper reviews various feature extraction methods proposed by different researchers for ISL and ASL, highlighting their strengths and limitations.

The review covers techniques such as Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), Shape Context, and Discrete Cosine Transform (DCT). The authors discuss the advantages and disadvantages of each method, considering factors like computational complexity, robustness to variations, and discriminatory power.

Additionally, the paper addresses the challenges associated with hand shape and appearance variations, dynamic hand movements, occlusions, and lighting conditions that impact the effectiveness of feature extraction methods in sign language recognition.

The authors conclude by suggesting the need for further research to explore hybrid feature extraction techniques, incorporating multiple modalities such as hand shape, motion, and texture. They emphasize the significance of feature selection and dimensionality reduction to enhance the efficiency and accuracy of sign language recognition systems.

Overall, the paper provides a comprehensive review of feature extraction techniques for ISL and ASL, highlighting their potential applications and limitations. This research serves as a valuable resource for researchers and practitioners working in the field of sign language recognition, facilitating the development of more accurate and robust systems for effective communication with the deaf and hard-of-hearing communities.

Chapter 3

ANALYSIS

3.1 Problem Statement

Communication between deaf-dumb and normal person have been always a challenging task. The communication gap between deaf and dumb individuals and the wider community due to the lack of understanding needs to be addressed.

Communication is a fundamental human right and is crucial for social integration, education, and employment opportunities. The lack of communication leads to social exclusion and is a significant barrier to achieving equality and inclusion for deaf and dumb individuals.

Interpretation services and note-taking services are available but are not always accessible, affordable, or reliable. These services also require a third party, which can be inconvenient for spontaneous interactions.

3.2 Objectives

To overcome this challenge, a system that can facilitate communication between deaf and dumb individuals and people is needed. It must provide solutions for bridging the communication gap between deaf and dumb individuals and the wider community are limited.

The system must be able to translate spoken or written language into ISL and vice versa, using video or other visual aids to ensure effective communication. Additionally, the system should be easy to use and accessible to a wide range of users, including those with limited technology skills.

3.2.1 Aims of Project

- Develop a machine learning model that can accurately recognize and interpret ISL gestures from video.
- Improve communication and accessibility for the deaf community in India by providing a more effective means of communication.
- Develop a user-friendly application that allows for real-time ISL gesture detection and interpretation, which can be used by both the deaf and hearing communities for improved communication.
- Investigate the potential of using the ISL gesture detection technology for language modeling and generation, by creating a dataset of ISL gestures and using it to train a generative machine learning model..

3.3 Methodology

The Indian Sign Language Detection System involves several steps. Firstly, the system captures video input using a webcam, which is then processed using MediaPipe's hand tracking model to identify the user's hands' positions and movements accurately.

Once the hand tracking has been completed, MediaPipe's landmark detection model is applied to identify specific points on the user's hand, such as fingertips, knuckles, and wrist joints. This step helps to identify the exact hand shape and movement, which is crucial in recognizing specific sign language gestures accurately.

After identifying the landmarks, the system uses a custom-trained machine learning model to recognize specific sign language gestures. This machine learning model is trained using Random Forest on a dataset of sign language gestures, allowing it to identify and recognize the correct gesture being performed by the user accurately.

The Google API helps translate the given voice input into the text and based on that text the proposed system will prompt the appropriate ISL images

3.4 Software Requirements Specification

A Software Requirements Specification (SRS) is a description of a software system to be developed. It lays out functional and non-functional requirements and may include a set of use cases that describe user interactions that the software must provide.

3.4.1 Software Specifications

- Operating System: Windows 7 or above
- Software: Visual Studio Code
- Programming Language: Python

3.4.2 Hardware Specifications

- Processor: Intel i5 (any processor)
- Processor Speed: 2.4 GHz
- RAM: 4 GB,
- Webcam: 720p
- Storage: 32 Gb

3.5 Functional Requirements

The following are some functional requirements for better performance of the model:

1. Data Collection: Collect a dataset of Indian Sign Language signs. The dataset should include images of signs along with their corresponding labels.

2. Data Preprocessing: Preprocess the dataset by extracting relevant features from the images. The features should be selected based on their ability to differentiate between different signs.

3. Training and Testing: Split the dataset into training and testing sets. Train the random forest classifier on the training set and evaluate its performance on the testing set.

4. Model Selection: Compare the performance of the random forest classifier with other machine learning algorithms for ISL recognition. Select the best-performing model for deployment.

5. Deployment: Deploy the trained model to a user-friendly interface such as a web application or mobile application. The interface should allow users to input videos or images of ISL signs and receive the predicted label for each sign.

6. Continuous Improvement: Collect user feedback and improve the model's performance over time. This could involve adding more signs to the dataset, refining the feature extraction process, or fine-tuning the hyperparameters of the random forest classifier.

3.6 Non-Functional Requirements

Here are some non-functional requirements for Indian Sign Language System:

1. Performance: The system should be designed to provide high performance, low latency, and high throughput to ensure that signs in live video can be processed quickly and efficiently.

2. Scalability: The system should be scalable to accommodate increasing numbers of users and ability to handle large volumes of data for reliable performance.

3. Reliability: The system should be designed to be highly reliable and available, with minimal downtime or service interruptions. This can include redundancy, failover mechanisms, and proactive monitoring.

4. Usability: The system should be user-friendly and intuitive, with clear instructions and documentation for users to manage and use the system efficiently.

Chapter 4

SYSTEM DESIGN

Systems design is the process of defining the architecture, modules, interfaces, and data for a system to satisfy specified requirements. Systems design could be seen as the application of systems theory to product development. Systems design implies a systematic approach to the design of a system. It may take a bottom-up or top-down approach, but either way, the process is systematic wherein it takes into account all related variables of the system that needs to be created—from the architecture, to the required hardware and software, right down to the data and how it travels and transforms throughout its travel through the system. Systems design then overlaps with systems analysis, systems engineering, and systems architecture.

4.1 System Architecture

System architecture refers to the design and organization of a computer system, including its components, modules, interfaces, and the way they interact with each other to fulfill specific functionality or provide desired services. The architecture defines the structure, behavior, and communication patterns within the system.

In the context of , the system architecture typically involves multiple layers and components that work together to support the capturing of images, storage, analysis, and management of images.

4.1.1 OpenCV

OpenCV (Open Source Computer Vision) is an open-source computer vision library that provides a wide range of tools and functions for developing computer vision applications. It was first developed by Intel in 1999 and is now maintained by the OpenCV community. OpenCV supports various programming languages, including C++, Python, Java, and MATLAB, making it a popular choice for computer vision researchers and developers.

OpenCV provides a wide range of functionalities for image and video processing, including image and video capture, image filtering and enhancement, object detection and recognition, and camera calibration. It also includes machine learning algorithms for tasks such as classification and clustering. OpenCV is designed to be highly efficient and scalable, making it suitable for both desktop and mobile platforms.

Overall, OpenCV is a powerful and versatile library for computer vision applications that provides a wide range of functionalities and supports multiple programming languages. Its open-source nature and active community make it an ideal choice for researchers, developers, and hobbyists alike who are interested in exploring the field of computer vision.

4.1.2 Mediapipe

MediaPipe is an open-source cross-platform framework for building machine learning-based applications for real-time perception, analysis, and processing of multimedia data such as video and audio. It was developed by Google and released in 2019. MediaPipe offers a wide range of pre-built, customizable building blocks that can be used to construct pipelines for various computer vision and machine learning tasks, such as object detection, face and hand tracking, pose estimation, and augmented reality.

MediaPipe is designed to be flexible and efficient, with a focus on real-time performance and low-latency processing. It supports various platforms, including desktop, mobile, and embedded devices. MediaPipe offers a variety of interfaces for different programming languages, including C++, Python, and JavaScript, making it easy to integrate into existing applications.

Overall, MediaPipe is a powerful and versatile framework for building machine learning-based applications for real-time multimedia processing. Its pre-built building blocks and support for various programming languages and platforms make it easy to get started with, while its flexibility and efficiency make it suitable for a wide range of applications. As a result, MediaPipe is increasingly being used in various fields, including augmented reality, robotics, and video analytics.

4.1.3 SQLite:

SQLite is a lightweight, open-source relational database management system (RDBMS) that is widely used in various applications, particularly in mobile devices, embedded systems, and small-scale web applications. SQLite is known for its simplicity, ease of use, and low overhead, making it a popular choice for developers who need a simple and reliable database solution.

One of the main advantages of SQLite is its portability. Since SQLite is a self-contained database engine, it can be easily integrated into applications without the need for separate installations or configuration. SQLite also requires minimal system resources, making it suitable for resource-constrained environments such as mobile devices or embedded systems.

SQLite is widely used in various applications, including web browsers, mobile operating systems, and embedded systems. It is also used in various programming languages, including C/C++, Java, Python, and PHP. SQLite has gained popularity due to its ease of use, low overhead, and portability, making it an ideal choice for developers who need a simple and reliable database solution that can be easily integrated into their applications

4.1.4 Random Forest Classifier

Random Forest Classifier is a popular machine learning algorithm used for classification tasks. It is an ensemble learning method that combines multiple decision trees to make a final prediction. Each decision tree in the random forest is trained on a subset of the data, and the final prediction is made by aggregating the predictions of all the trees. This process helps to improve the accuracy of the predictions and reduces the risk of overfitting.

One of the main advantages of Random Forest Classifier is its ability to handle missing data and noisy data. Random Forest Classifier can automatically detect and handle missing values and noisy data, making it a robust algorithm for real-world data. Random Forest Classifier can also provide feature importance scores, which can help identify the most important features in the dataset.

Random Forest Classifier is a powerful and versatile algorithm for classification tasks. Its ability to handle missing data and noisy data, as well as its ability to identify important features, make it a popular choice for machine learning practitioners. Random Forest Classifier has been successfully used in various applications, including image classification, text classification, and fraud detection.

4.1.5 Google Cloud Speech-to-Text API

The Google Cloud Speech-to-Text API is a machine learning-based service that allows developers to convert audio to text using powerful neural network models. This service can recognize speech from various sources, including microphones, audio files, and streaming audio. The API supports multiple languages and dialects, making it suitable for a wide range of applications.

The Google Cloud Speech-to-Text API for Python provides a simple and easy-to-use interface for developers to integrate speech recognition capabilities into their applications. The API allows developers to send audio files or streaming audio to the Speech-to-Text API and receive a text transcription in real-time. Developers can also specify various parameters, such as language, audio encoding, and model selection, to improve the accuracy of the transcription.

The Google Cloud Speech-to-Text API for Python is based on the REST API, which means that developers can use any programming language that supports HTTP requests to interact with the API. This makes it easy to integrate the API into existing applications and workflows. The API also provides client libraries for various programming languages, including Python, Java, and Go, which simplify the integration process even further.

The Google Cloud Speech-to-Text API for Python is a powerful and flexible tool for adding speech recognition capabilities to applications. Its ability to recognize speech from various sources and support multiple languages and dialects make it suitable for a wide range of use cases. The API's simple and easy-to-use interface, along with the availability of client libraries for various programming languages, makes it easy for developers to integrate the API into their applications.

4.1.6 Tkinter:

Tkinter is a Python library that provides a simple and powerful way to create graphical user interfaces (GUIs) for desktop applications. It is based on the popular Tk GUI toolkit, which was developed by John Ousterhout at the University of California, Berkeley in the late 1980s. Tkinter provides a wide range of widgets, such as buttons, labels, text boxes, and menus, that can be easily placed on a GUI window and customized to meet the needs of the application.

4.2 Sequence Diagram

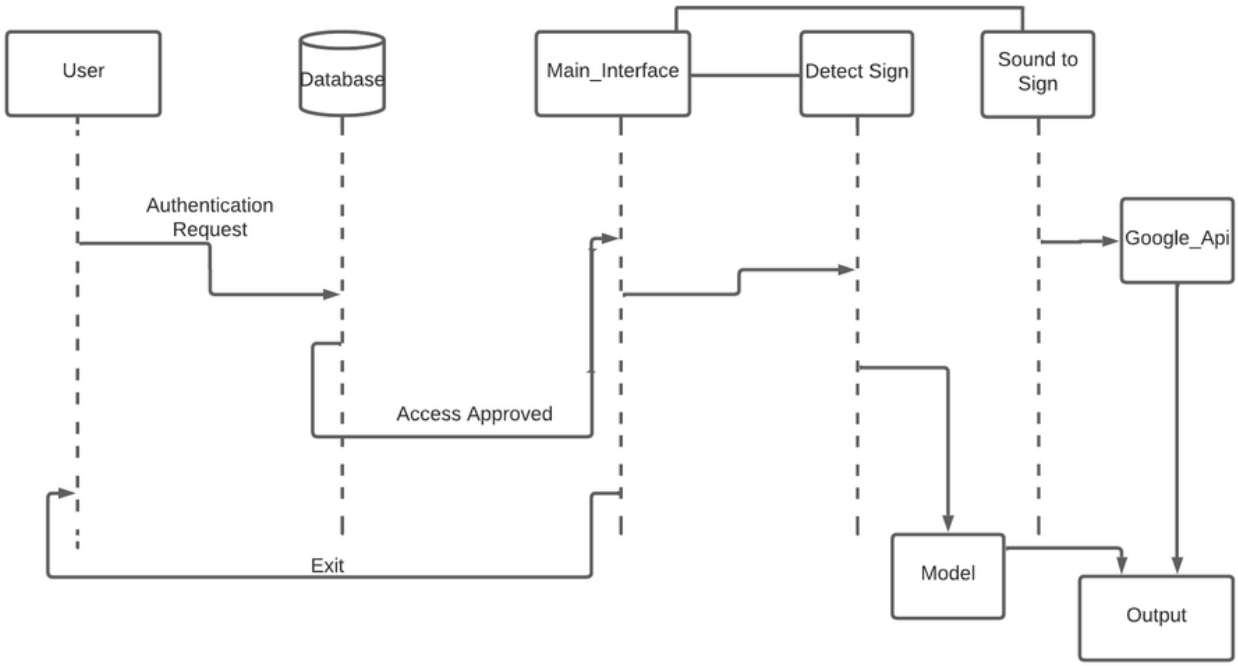

Fig 4.1 Sequence Diagram

4.3 Flowchart

1. The input image is scanned for any gestures done by the user through the webcamera.

2. The landmark points on the hands are detected for gestures made.

3. The gestures shown are compared with the trained dataset to find the corresponding sign that is being made.

4. The result is displayed in the form of text representing the letter being gestured.

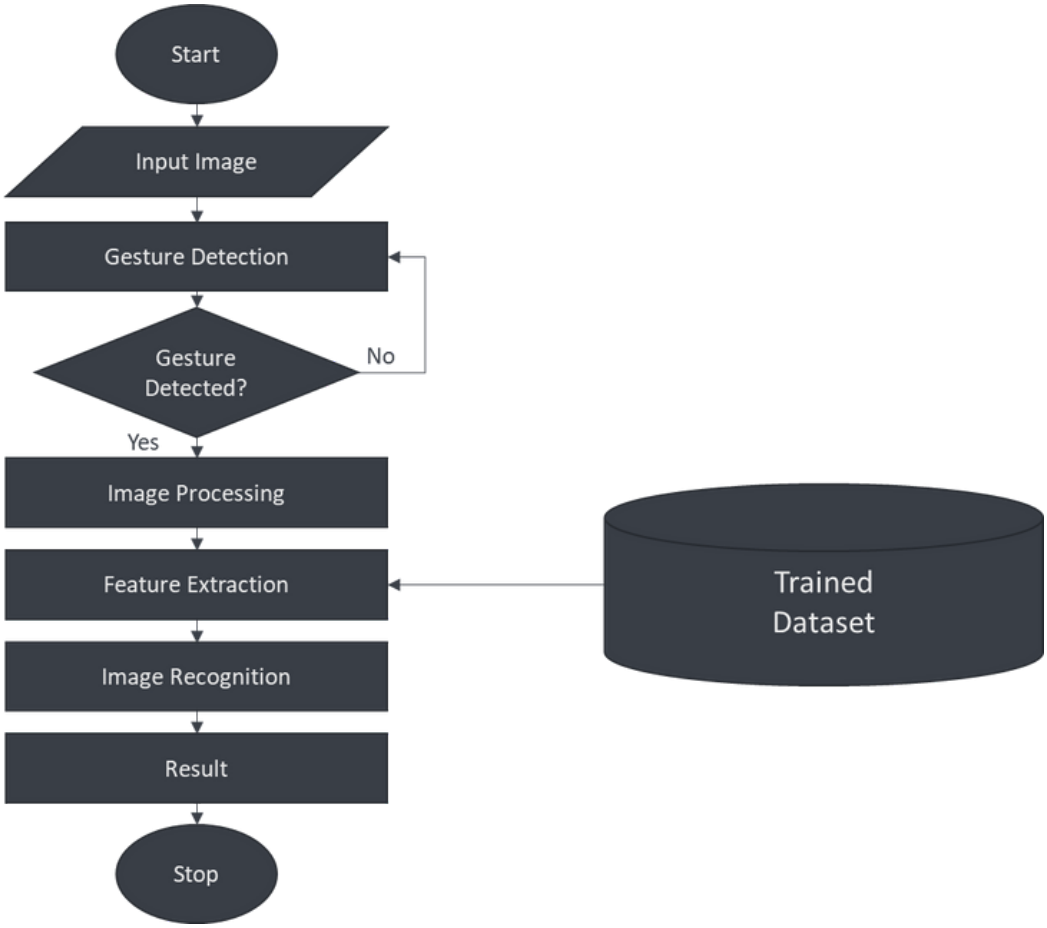

Fig 4.2 Flowchart

Chapter 5

IMPLEMENTATION

5.1 Pseudocode

```
cap = cv2.VideoCapture(0)
for j in range(number_of_classes):
    if not os.path.exists(os.path.join(DATA_DIR, str(j))):
        os.makedirs(os.path.join(DATA_DIR, str(j)))

    print('Collecting data for class {}'.format(j))

    done = False
    while True:
        ret, frame = cap.read()
        cv2.putText(frame, 'Ready? Press "Q" ! :)', (100, 50), cv2.FONT_HERSHEY_SIMPLEX, 1.3, (0, 255,
0), 3,
                    cv2.LINE_AA)
        cv2.imshow('frame', frame)
        if cv2.waitKey(25) == ord('q'):
            break

    counter = 0
    while counter < dataset_size:
        ret, frame = cap.read()
        cv2.imshow('frame', frame)
        cv2.waitKey(25)
        cv2.imwrite(os.path.join(DATA_DIR, str(j), '{}.jpg'.format(counter)), frame)

        counter += 1

cap.release()
```

Fig 5.1 Collecting Images

```
def createdb():
    conn = sqlite3.connect('files/users_info.db')
    c = conn.cursor()
    c.execute(
        "CREATE TABLE IF NOT EXISTS users (name TEXT , passs TEXT,sqltime TIMESTAMP DEFAULT CURRENT_TIMESTAMP
NOT NULL)")
    conn.commit()
    conn.close()


createdb()
```

Fig 5.2 Creating Database

```
def saveadmin():
    name_err = name_entry.get()
    pass_err = pass_entry.get()
    if name_err == "":
        messagebox.showinfo("Invalid input", "Username can't be Empty")
    elif pass_err == "":
        messagebox.showinfo("Invalid input", "Password can't be Empty")
    else:
        conn = sqlite3.connect("files/users_info.db")
        c = conn.cursor()
        c.execute("INSERT INTO users(name,passs) VALUES(?,?) ", (name_entry.get(), pass_entry.get()))
        conn.commit()
        messagebox.showinfo("Information", "New User has been Added")
```

Fig 5.3 Adding New User To Database

```
def loggin():
    while True:
        a = name2_entry.get()
        b = pass2_entry.get()
        with sqlite3.connect("files/users_info.db") as db:
            cursor = db.cursor()
        find_user = ("SELECT * FROM users WHERE name = ? AND passs = ?")
        cursor.execute(find_user, [(a), (b)])
        results = cursor.fetchall()
```

Fig 5.4 Fetching Data From Database

```
window2 = Tk()
                f1 = Frame(window2)
                f2 = Frame(window2)
                f3 = Frame(window2)
                f4 = Frame(window2)
                f5 = Frame(window2)
                def swap(frame):
                    frame.tkraise()
                for frame in (f1, f2, f3, f4,f5):
                    frame.place(x=0, y=0, width=500, height=500)
                window2.geometry("500x500+520+170")
                window2.resizable(False, False)
                label3 = Label(f1, text="Welcome To SmartVoice", font=("arial", 20, "bold"), bg="grey16",
fg="white",
                               relief=SUNKEN)
                label3.pack(side=TOP, fill=X)
                label4 = Label(f2, text="                                              Indian Sign
Language Recognition System", font=("arial", 10, "bold"), bg="grey16",
                               fg="white")
                label4.pack(side=BOTTOM, fill=X)
                statusbar = Label(f1, text="                                              Indian Sign
Language Recognition System", font=("arial", 8, "bold"),
                                  bg="grey16", fg="white", relief=SUNKEN, anchor=W)
                statusbar.pack(side=BOTTOM, fill=X)
```

Fig 5.5 Purchasing NFT from marketplace

Fig 5.5 Tkinter Ui Code Snippet

```
with sr.Microphone() as source:
        messagebox.showinfo('Info','Speak Now')
        audio=voice.listen(source)
        try:
            messagebox.showinfo('Info', 'Recognizing...')
            text=voice.recognize_google(audio)
            messagebox.showinfo('Info', 'You said: '+str(text))
        except:
            messagebox.showerror("error","Your voice is not clear")
    text.lower()
    try:
        for l in text:
            if l is not ' ':
                img=imageio.imread(path+str(l)+'.jpg')
                display(img,l)
    except:
        messagebox.showerror("error","There was an error reading the input")
```

Fig 5.6 Voice To ISL

```
        results = hands.process(img_rgb)
        if not (results.multi_hand_landmarks is None):
                n = len(results.multi_hand_landmarks)
                if n == 2:
                    try:
                        for hand_landmarks in
results.multi_hand_landmarks:for i in range(len(hand_landmarks.landmark)):
                                    x= hand_landmarks.landmark[i].x
                                    y= hand_landmarks.landmark[i].y
                                    x_.append(x)
                                    y_.append(y)
                                for i in range(len(hand_landmarks.landmark)):
                                     x = hand_landmarks.landmark[i].x
                                     y = hand_landmarks.landmark[i].y
                                     data_aux.append(x - min(x_))
                                     data_aux.append(y - min(y_))
                        data.append(data_aux)
                        labels.append(dir_)
                    except:
                            data_aux(np.zeros([1,189], dtype=int)[0])
```

Fig 5.7 Creating Data Set For Two Hands

```
       results = hands.process(img_rgb)
        if not (results.multi_hand_landmarks is None):
                n = len(results.multi_hand_landmarks)
                if n == 1:
                    try:
                        for hand_landmarks in
results.multi_hand_landmarks:for i in range(len(hand_landmarks.landmark)):
                                    x= hand_landmarks.landmark[i].x
                                    y= hand_landmarks.landmark[i].y
                                    x_.append(x)
                                    y_.append(y)
                                for i in range(len(hand_landmarks.landmark)):
                                     x = hand_landmarks.landmark[i].x
                                     y = hand_landmarks.landmark[i].y
                                     data_aux.append(x - min(x_))
                                     data_aux.append(y - min(y_))
                        data.append(data_aux)
                        labels.append(dir_)
                    except:
                            data_aux(np.zeros([1,63], dtype=int)[0])
```

Fig 5.8 Creating Dataset For One Hand

```
data_dict = pickle.load(open('./onehand.pickle', 'rb'))
data_dict2 = pickle.load(open('./twohand.pickle', 'rb'))
data_one = np.asarray(data_dict['data'])
data_two = np.asarray(data_dict2['data'])
labels1 = np.asarray(data_dict['labels'])
labels2 = np.asarray(data_dict2['labels'])


x_train, x_test, y_train, y_test, = train_test_split(data_one, labels1, test_size=0.2, shuffle=True,
stratify=labels1)
x_train2, x_test2, y_train2, y_test2, = train_test_split(data_two, labels2, test_size=0.2, shuffle=True,
stratify=labels2)

model1 = RandomForestClassifier()
model2 = RandomForestClassifier()

model1.fit(x_train, y_train)
model2.fit(x_train2, y_train2)

y_predict = model1.predict(x_test)
y_predict2 = model2.predict(x_test2)
```

Fig 5.9 Training Models

Chapter 6

TESTING

6.1 Unit Testing

Unit Testing is defined as a type of software testing where individual components of a software are tested.

Test Case no.	Description	Test Steps	Expected Result	Actual Result	Pass/Fail
1	Signing up of users	Press the 'Signup' button and enter the details	Stores the given details in the database.	Successfully stored data in database.	**Pass**
2	Logging in of users	Press the 'Login' button and enter the details	Checks for the given details in the database and logs in if matched	Successfully matched user in database and logged in.	**Pass**
				User not found in database hence login failed	**Pass**
3	Sign prediction	Press the 'Predict Sign' button and show the signs	Predicts the signs made by the user and must output the appropriate letter.	Successfully predicts the sign and outputs the appropriate letter.	**Pass**
4	Landmark of the hands	Press the 'Show Landmarks' button and display hands.	Shows the Landmark points on the hands in real time.	Successfully shows Landmark points.	**Pass**
5	Voice recognition of users	Press the 'Sound Recognition' button and start talking after the prompt.	Voice of the user must be recognized and converted into appropriate text and display the corresponding signs.	Successfully displays the accurate signs for the recognized speech.	**Pass**

Fig 6.1 Unit Testing

Chapter 7

RESULTS AND DISCUSSIONS

The application provides an interface for users to detect Indian Sign Language and provide ISL signs for the given word detected through speech recognition. The user can register into the database and login to provide a custom experience.

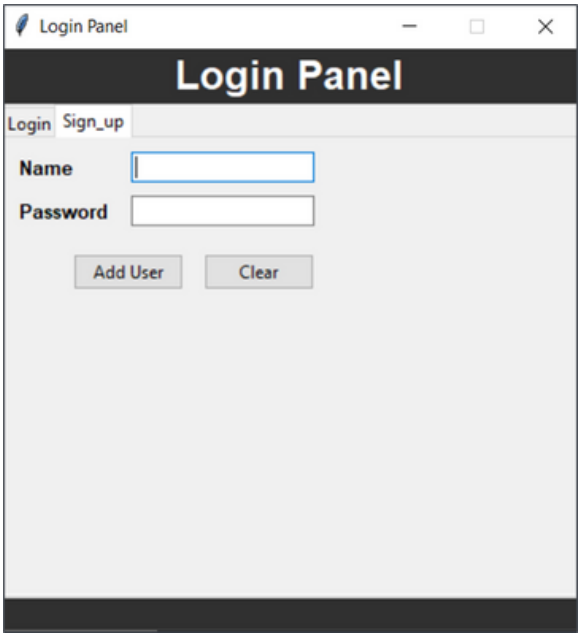

Fig 7.1 Sign-up Panel

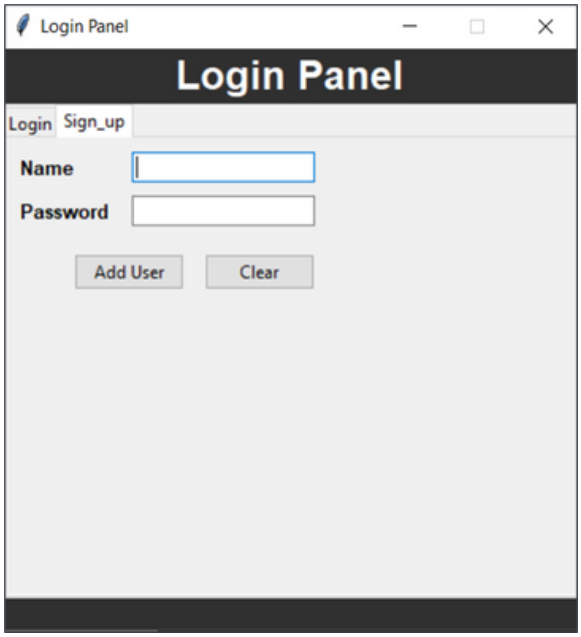

Fig 7.2 Log-in Panel

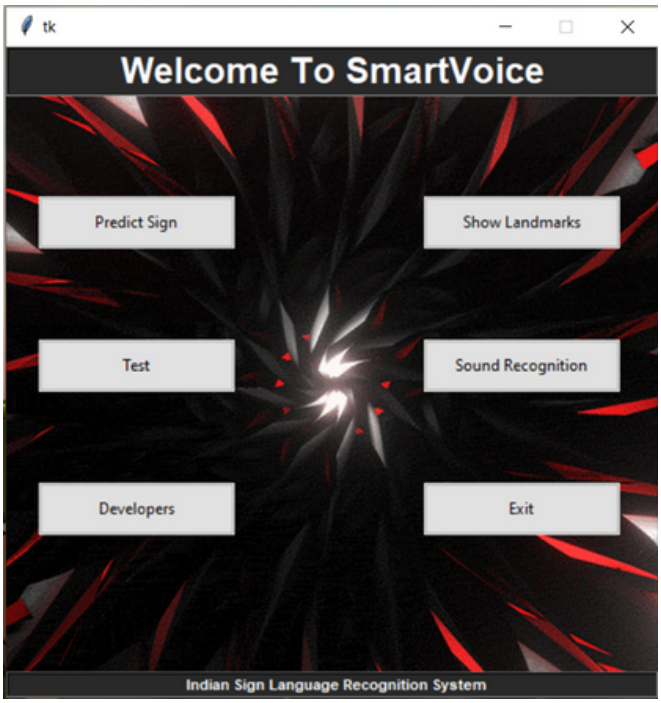

Fig 7.3 Interface

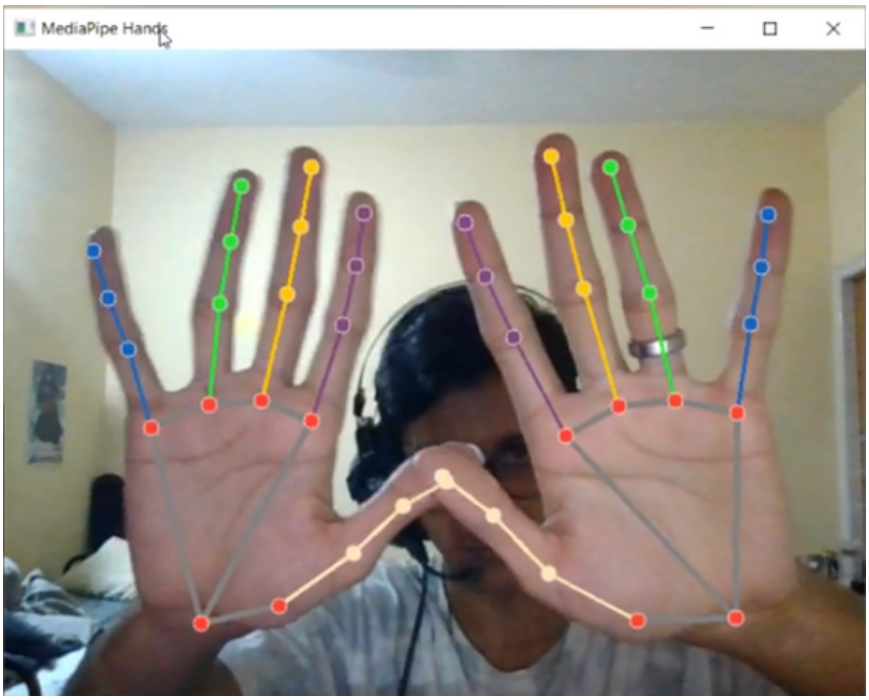

Fig 7.4 Landmark mapping

Show Landmarks option displays the landmark points used to map the joints of the fingers. These points when trained with the right data set allows for the detection of signs for letters made by the user and display the respective alphabet.

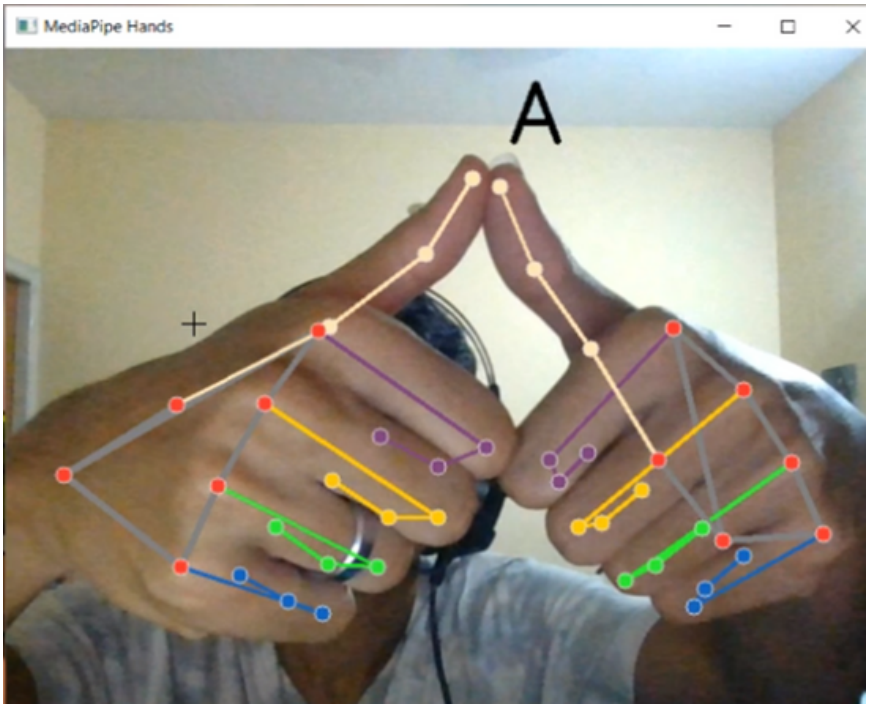

Fig 7.5 Recognition of letter 'A'

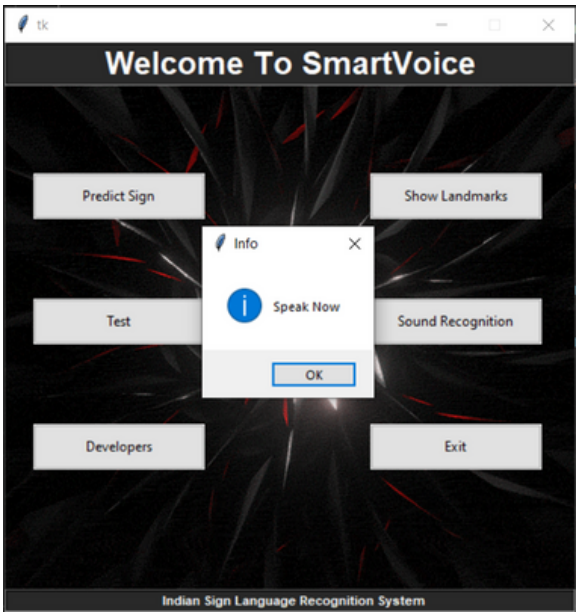

Fig 7.6 Sound Recognition

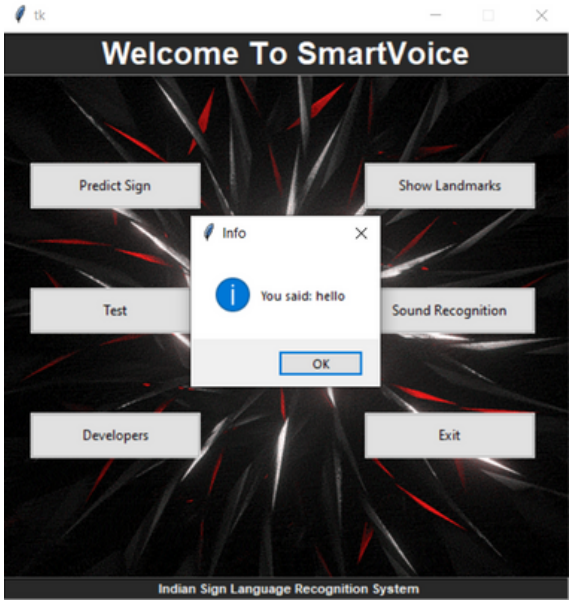

Fig 7.7 Sound Recognition Successful

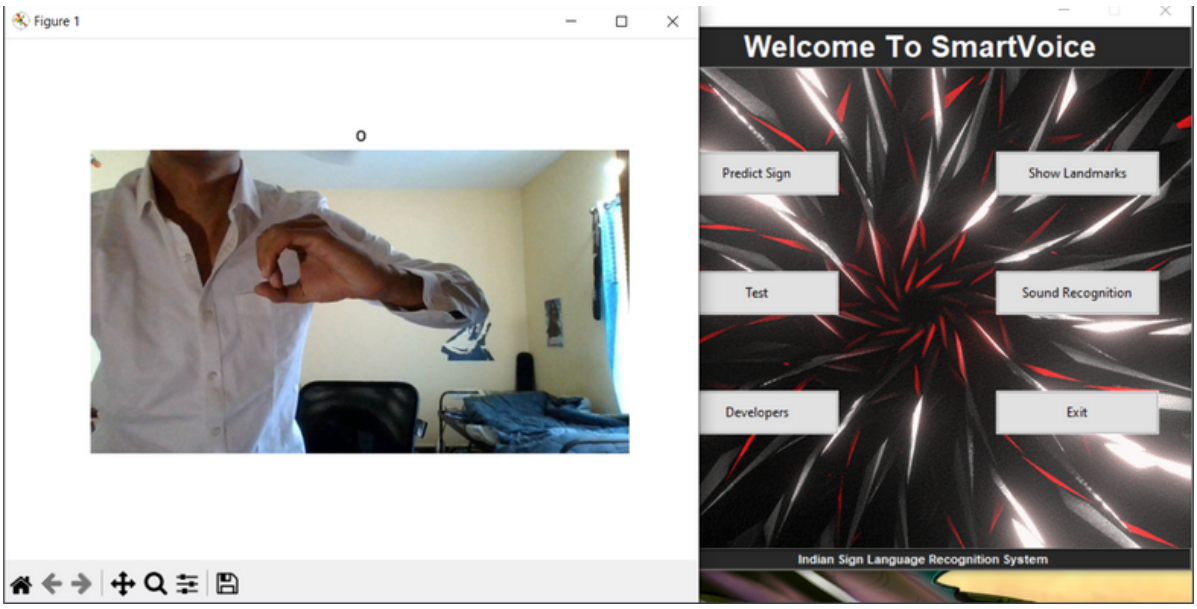

Fig 7.8 ISL representation of detected speech

Sound recognition allows for detection of words from speech and display the appropriate sign for each letter in the detected word.

Chapter 8

CONCLUSION

The Indian Sign Language Detection System has emerged as an innovative solution to overcome the communication barrier for the hearing-impaired community. It recognizes and interprets sign language gestures and converts them into textual or audible outputs, enabling the users to communicate effectively with the wider world.

The system's accuracy and reliability are critical factors, and the use of advanced machine learning algorithms has greatly improved its performance. It can now identify and interpret various sign language gestures with high accuracy and minimal false positives.

The system's scalability is equally important, and it is be capable of handling large volumes of data for reliable performance. Modular architecture, efficient resource utilization, and support for multiple platforms are key to achieving this.

The system's impact on society has been immense, particularly in promoting inclusivity for the hearing-impaired. It has opened up new avenues for communication, education, and employment opportunities, making it easier for the hearing-impaired community to engage with the wider world.

In conclusion, the Indian Sign Language Detection System is a remarkable technological innovation that has revolutionized the way the hearing-impaired community communicates with the wider world. Its accuracy, scalability, and impact on society make it a vital tool for promoting inclusivity and enhancing the lives of the hearing-impaired.

FUTURE WORK

- For future development, it will require further research in advanced mathematical methods in image processing
- This includes new detection algorithms from a mathematical point of view and methods to calculate lighting offsets and adjustments based on the user operating environment
- Additional research and investments into hardware components such as infra-red cameras or two side-by-side webcams will also help to solve the problem
- Implementing text-to-speech in regional languages

REFERENCES

[1] Sunitha K. A, Anitha Saraswathi.P, Aarthi.M, Jayapriya. K, Lingam Sunny, "Deaf Mute Communication Interpreter- A Review", International Journal of Applied Engineering Research, Volume 11, pp 290-296, 2016.

[2] Mathavan Suresh Anand, Nagarajan Mohan Kumar, Angappan Kumaresan, " An Efficient Framework for Indian Sign Language Recognition Using Wavelet Transform" Circuits and Systems, Volume 7, pp 1874-1883, 2016.

[3] Mandeep Kaur Ahuja, Amardeep Singh, "Hand Gesture Recognition Using PCA", International Journal of Computer Science Engineering and Technology (IJCSET), Volume 5, Issue 7, pp. 267-27, July 2015.

[4] Sagar P.More, Prof. Abdul Sattar, "Hand gesture recognition system for dumb people". International Journal of Science and Research (IJSR).

[5] Chandandeep Kaur, Nivit Gill, "An Automated System for Indian Sign Language Recognition", International Journal of Advanced Research in Computer Science and Software Engineering.

[6] Pratibha Pandey, Vinay Jain, "Hand Gesture Recognition for Sign Language Recognition: A Review", International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 3, March 2015.

[7] Nakul Nagpal,Dr. Arun Mitra., Dr. Pankaj Agrawal, "Design Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People", International Journal on Recent and Innovation Trends in Computing and Communication ,Volume: 3 Issue: 5,pp- 147 – 149.

[8] Neelam K. Gilorkar, Manisha M. Ingle, "Real Time Detection And Recognition Of Indian And American Sign Language Using Sift", International Journal of Electronics and Communication Engineering & Technology (IJECET), Volume 5, Issue 5, pp. 11-18 , May 2014.

[9] Neelam K. Gilorkar, Manisha M. Ingle, "A Review on Feature Extraction for Indian and American Sign Language", International Journal of Computer Science and Information Technologies, Volume 5 (1) , pp314- 318, 2014.

REFERENCES

[10] Ashish Sethi, Hemanth ,Kuldeep Kumar, Bhaskara Rao ,Krishnan R, "Sign Pro-An Application Suite for Deaf and Dumb", IJCSET , Volume 2, Issue 5, pp-1203-1206, May 2012.

[11] Priyanka Sharma, "Offline Signature Verification Using Surf Feature Extraction and Neural Networks Approach", International Journal of Computer Science and Information Technologies, Volume 5 (3) , pp 3539-3541, 2014.

RNS INSTITUTE OF TECHNOLOGY

(AICTE APPROVED, VTU AFFILIATED AND NAAC 'A+' GRADE ACCREDITED)

(UG Programs - CSE, ISE, ECE, EIE and EEE have been Accredited by NBA up to 30.06.2025)

CHANNASANDRA, Dr. VISHNUVARDHAN ROAD, BENGALURU - 560 098, KARNATAKA, INDIA

TEL : +91 80 28611880/ 1, FAX: +91 80 28611882, WEBSITE: WWW.RNSIT.AC.IN

ESTD:2001

Dr. R N Shetty
Founder

Vision of Institution : Building RNSIT into a world- class Institution

NCCT 2023

Certificate of Participation

Presented to

Paper ID PI227

Mr./Ms. Abhaya Simha S P of ISE, RNSIT for presenting paper entitled INDIAN SIGN LANGUAGE DETECTION USING MACHINE LEARNING in the National Conference on Computing Technology (NCCT - 2023), jointly organized by departments of CSE, ISE, AI&ML and MCA held on May 08, 2023.

Dr. Kiran P
HoD, CSE

Dr. Suresh L
HoD, ISE

Dr. Harsha S
HoD, AI & ML

Dr. N P Kavya
HoD, MCA

Dr. Ramesh Babu H S
Principal

Dr. M K Venkatesha
Director

ISBN

In Association with

INSTITUTION'S INNOVATION COUNCIL

ARIIA TOI Outlook DATAQUEST

RNS INSTITUTE OF TECHNOLOGY

(AICTE APPROVED, VTU AFFILIATED AND NAAC 'A+' GRADE ACCREDITED)

(UG Programs - CSE, ISE, ECE, EIE and EEE have been Accredited by NBA up to 30.06.2025)

CHANNASANDRA, Dr. VISHNUVARDHAN ROAD, BENGALURU - 560 098, KARNATAKA, INDIA

TEL : +91 80 28611880/ 1, FAX: +91 80 28611882, WEBSITE: WWW.RNSIT.AC.IN

ESTD:2001

Dr. R N Shetty
Founder

Vision of Institution : Building RNSIT into a world- class Institution

NCCT 2023

Certificate of Participation

Presented to

Paper ID PI227

Mr./Ms Chinmay Hegde of ISE, RNSIT for presenting paper entitled INDIAN SIGN LANGUAGE DETECTION USING MACHINE LEARNING in the National Conference on Computing Technology (NCCT - 2023), jointly organized by departments of CSE, ISE, AI&ML and MCA held on May 08, 2023.

Dr. Kiran P
HoD, CSE

Dr. Suresh L
HoD, ISE

Dr. Harsha S
HoD, AI & ML

Dr. N P Kavya
HoD, MCA

Dr. Ramesh Babu H S
Principal

Dr. M K Venkatesha
Director

ISBN

In Association with

INSTITUTION'S INNOVATION COUNCIL

ARIIA TOI Outlook DATAQUEST

RNS INSTITUTE OF TECHNOLOGY

(AICTE APPROVED, VTU AFFILIATED AND NAAC 'A+' GRADE ACCREDITED)

(UG Programs - CSE, ISE, ECE, EIE and EEE have been Accredited by NBA up to 30.06.2025)

CHANNASANDRA, Dr. VISHNUVARDHAN ROAD, BENGALURU - 560 098, KARNATAKA, INDIA

TEL : +91 80 28611880/ 1, FAX: +91 80 28611882, WEBSITE: WWW.RNSIT.AC.IN

ESTD:2001

An Institute with a Difference

Dr. R N Shetty
Founder

Vision of Institution : Building RNSIT into a world- class Institution

NCCT 2023

Certificate of Participation

Presented to

Paper ID
PI227

Mr./Ms. Rakshita S of ISE, RNSIT for presenting paper entitled INDIAN SIGN LANGUAGE DETECTION USING MACHINE LEARNING in the National Conference on Computing Technology (NCCT - 2023), jointly organized by departments of CSE, ISE, AI&ML and MCA held on May 08, 2023.

Dr. Kiran P
HoD, CSE

Dr. Suresh L
HoD, ISE

Dr. Harsha S
HoD, AI & ML

Dr. N P Kavya
HoD, MCA

Dr. Ramesh Babu H S
Principal

Dr. M K Venkatesha
Director

ISBN

In Association with

INSTITUTION'S INNOVATION COUNCIL

ARIIA TOI Outlook DATAQUEST

RNS INSTITUTE OF TECHNOLOGY

(AICTE APPROVED, VTU AFFILIATED AND NAAC 'A+' GRADE ACCREDITED)

(UG Programs - CSE, ISE, ECE, EIE and EEE have been Accredited by NBA up to 30.06.2025)

CHANNASANDRA, Dr. VISHNUVARDHAN ROAD, BENGALURU - 560 098, KARNATAKA, INDIA

TEL : +91 80 28611880/ 1, FAX: +91 80 28611882, WEBSITE: WWW.RNSIT.AC.IN

Dr. R N Shetty
Founder

Vision of Institution : Building RNSIT into a world- class Institution

NCCT 2023

Certificate of Participation

Presented to

Paper ID
PI227

Mr./Ms. Adithya R Pai of ISE, RNSIT for presenting paper entitled INDIAN SIGN LANGUAGE DETECTION USING MACHINE LEARNING in the National Conference on Computing Technology (NCCT - 2023), jointly organized by departments of CSE, ISE, AI&ML and MCA held on May 08, 2023.

Dr. Kiran P
HoD, CSE

Dr. Suresh L
HoD, ISE

Dr. Harsha S
HoD, AI & ML

Dr. N P Kavya
HoD, MCA

Dr. Ramesh Babu H S
Principal

Dr. M K Venkatesha
Director

ISBN

In Association with

INSTITUTION'S INNOVATION COUNCIL

ARIIA TOI Outlook DATAQUEST