Jyothi Hills, Panjal Road,
Vettikattiri PO, Cheruthuruthy , Thrissur,
Kerala 679531

Jyothi Engineering College
NAAC Accredited College with NBA Accredited Programmes*
Approved by AICTE & affiliated to APJ Abdul Kalam Technological University
A CENTRE OF EXCELLENCE IN SCIENCE & TECHNOLOGY BY THE CATHOLIC ARCHDIOCESE OF TRICHUR
JYOTHI HILLS, VETTIKATTIRI P.O, CHERUTHURUTHY, THRISSUR. PIN-679531 PH : +91- 4884-259000, 274423 FAX : 04884-274777
NBA accredited B.Tech Programmes in Computer Science & Engineering, Electronics & Communication Engineering, Electrical & Electronics Engineering and Mechanical Engineering valid for the academic years 2016-2022. NBA accredited B.Tech Programme in Civil Engineering valid for the academic years 2019-2022.

Talk2Mute

MAIN PROJECT REPORT

NEETHUU N (JEC17CS074)

SIDHARTH U (JEC17CS095)

SREEHARI (JEC17CS097)

VINCY ANTO (JEC17CS104)

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY (B.Tech)

in

COMPUTER SCIENCE & ENGINEERING

of

A P J ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Under the guidance of

Ms. SAJITHA I

CREATING TECHNOLOGY LEADERS OF TOMORROW

JANUARY 2021

Department of Computer Science & Engineering

Jyothi Hills, Panjal Road,
Vettikattiri PO, Cheruthuruthy , Thrissur,
Kerala 679531

Jyothi Engineering College
NAAC Accredited College with NBA Accredited Programmes*
Approved by AICTE & affiliated to APJ Abdul Kalam Technological University
A CENTRE OF EXCELLENCE IN SCIENCE & TECHNOLOGY BY THE CATHOLIC ARCHDIOCESE OF TRICHUR
JYOTHI HILLS, VETTIKATTIRI P.O, CHERUTHURUTHY, THRISSUR. PIN-679531 PH : +91- 4884-259000, 274423 FAX : 04884-274777
NBA accredited B.Tech Programmes in Computer Science & Engineering, Electronics & Communication Engineering, Electrical & Electronics Engineering and Mechanical Engineering valid for the academic years 2016-2022. NBA accredited B.Tech Programme in Civil Engineering valid for the academic years 2019-2022.

Talk2Mute

MAIN PROJECT REPORT

NEETHUU N (JEC17CS074)

SIDHARTH U (JEC17CS095)

SREEHARI (JEC17CS097)

VINCY ANTO (JEC17CS104)

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY (B.Tech)

in

COMPUTER SCIENCE & ENGINEERING

of

A P J ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Under the guidance of

Ms. SAJITHA I

CREATING TECHNOLOGY LEADERS OF TOMORROW

JANUARY 2021

Department of Computer Science & Engineering

Department of Computer Science and Engineering

JYOTHI ENGINEERING COLLEGE, CHERUTHURUTHY

THRISSUR 679 531

JANUARY 2021

BONAFIDE CERTIFICATE

This is to certify that the main project report entitled **Talk2Mute** submitted by **Neethuu N (JEC17CS074), Sidharth U (JEC17CS095), Vincy Anto (JEC17CS104), Sreehari (JEC17CS097)** and in partial fulfillment of the requirements for the award of **Bachelor of Technology** degree in **Computer Science and Engineering** of **A P J Abdul Kalam Technological University** is the bonafide work carried out by them under our supervision and guidance.

Ms. Sajitha I	**Mr. Shaiju Paul**	**Dr. Vinith R**
Project Guide	Project Coordinator	Head of The Dept
Assistant Professor	Assistant Professor	Associate Professor
Dept. of CSE	Dept. of CSE	Dept. of CSE

DEPARTMENT OF

COMPUTER SCIENCE & ENGINEERING

COLLEGE VISION

Creating eminent and ethical leaders through quality professional education with emphasis on holistic excellence.

COLLEGE MISSION

- To emerge as an institution par excellence of global standards by imparting quality engineering and other professional programmes with state-of- the-art facilities.
- To equip the students with appropriate skills for a meaningful career in the global scenario.
- To inculcate ethical values among students and ignite their passion for holistic excellence through social initiatives.
- To participate in the development of society through technology incubation, entrepreneurship and industry interaction.

DEPARTMENT OF

COMPUTER SCIENCE & ENGINEERING

DEPARTMENT VISION

Creating eminent and ethical leaders in the domain of computational sciences through quality professional education with a focus on holistic learning and excellence.

DEPARTMENT MISSION

- To create technically competent and ethically conscious graduates in the field of Computer Science & Engineering by encouraging holistic learning and excellence.
- To prepare students for careers in Industry, Academia and the Government.
- To instill Entrepreneurial Orientation and research motivation among the students of the department.
- To emerge as a leader in education in the region by encouraging teaching, learning, industry and societal connect

PROGRAMME OUTCOMES (POs)

1. **Engineering Knowledge**: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. **Problem Analysis**: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3. **Design/Development of Solutions**: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4. **Conduct Investigations of Complex Problems**: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
5. **Modern Tool Usage**: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
6. **The Engineer and Society**: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
7. **Environment and Sustainability**: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8. **Ethics**: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. **Individual and Team Work**: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
10. **Communication**: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. **Project Management and Finance**: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12. **Life-Long Learning**: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAMME EDUCATIONAL OBJECTIVES (PEOs)

1. The graduates shall have sound knowledge of Mathematics, Science, Engineering and Management to be able to offer practical software and hardware solutions for the problems of industry and society at large.
2. The graduates shall be able to establish themselves as practising professionals, researchers or Entrepreneurs in computer science or allied areas and shall also be able to pursue higher education in reputed institutes.
3. The graduates shall be able to communicate effectively and work in multidisciplinary teams with team spirit demonstrating value driven and ethical leadership.

PROGRAMME SPECIFIC OUTCOMES (PSOs)

1. An ability to apply knowledge of data structures and algorithms appropriate to computational problems.
2. An ability to apply knowledge of operating systems, programming languages, data management, or networking principles to computational assignments.
3. An ability to apply design, development, maintenance or evaluation of software engineering principles in the construction of computer and software systems of varying complexity and quality.
4. An ability to understand concepts involved in modeling and design of computer science applications in a way that demonstrates comprehension of the fundamentals and trade-offs involved in design choices.

COURSE OUTCOMES (COs)

C410.1 The students will be able to analyse a current topic of professional interest and present it before an audience.

C410.2 Students will be able to identify an engineering problem, analyse it and propose a work plan to solve it.

C410.3 Students will have gained thorough knowledge in design, implementations and execution of Computer science related projects.

C410.4 Students will have attained the practical knowledge of what they learned in theory subjects.

C410.5 Students will become familiar with usage of modern tools.

C410.6 Students will have ability to plan and work in a team.

ACKNOWLEDGEMENT

We take this opportunity to express our heartfelt gratitude to all respected personalities who had guided, inspired and helped us in the successful completion of this interim project. First and foremost, we express my thanks to **The Lord Almighty** for guiding us in this endeavour and making it a success.

We take immense pleasure in thanking the **Management** of Jyothi Engineering College and **Dr. Sunny Joseph Kalayathankal**, Principal, Jyothi Engineering College for permitting us to carry out this project. Our sincere thanks to **Prof.Dr.Vinith R**, Head of the Department of Computer Science and Engineering for permitting us to make use of the facilities available in the department to carry out the interim project successfully.

We express our sincere gratitude to **Mr. Shaiju Paul & Dr. Swapna B Sasi**, Project Coordinators for their invaluable supervision and timely suggestions. We are very happy to express our deepest gratitude to our mentor **Ms. Sajitha I**, Assistant Professor,Department of Computer Science and Engineering, Jyothi Engineering College for her able guidance and continuous encouragement.

Last but not least, we extend our gratefulness to all teaching and non-teaching staff who directly or indirectly involved in the successful completion of this interim project work and to all friends who have patiently extended all sorts of help for accomplishing this undertaking.

ABSTRACT

Speech impairment is a disability which affects one's ability to communicate with others. Such individuals use sign language to communicate with other people. Although it is an effective form of communication, there remains a challenge for people who do not understand sign language to communicate with speech impaired people. Sign Language Recognition is one of the most growing fields of research area. Many new techniques have been developed recently in this area. The Sign Language is mainly used for communication of deaf-dumb people.The people who know sign language can communicate with each other efficiently. However, when it comes to communicating with people who don't understand sign language it causes a lot of problems. Therefore, This system aims to translate sign language to English and Regional languages in the form of text and audio, thus aiding communication with sign language. The Linear Discriminant Analysis (LDA) algorithm was used for gesture recognition and recognized gesture is converted into text and voice format. The main goal of this project is to validate the usage of this methodology to identify the sign language gestures shown by a signer and translate it into its corresponding text and audio output.

Keywords: Sign Language Recognition, Linear Discriminant Analysis, Feature Extraction, Classification, Euclidean Distance, Eigen Values, Eigen vectors.

CONTENTS

List of Figures

List of Abbreviations

LDA	: *Linear Discriminant Analysis*
ASL	: *American Sign Language*
KNN	: *K Nearest Neighbour*
ISL	: *Indian Sign Language*
SDLC	: *Software Development Life Cycle*
SLR	: *Sign Language Recognition*
API	: *Application Programming Interface*
ROI	: *Region of Interest*
TTS	: *Text to Speech*
PCA	: *Principal Component Analysis*
HOG	: *Histogram of Oriented Gradients*
SVM	: *Support Vector Machine*

CHAPTER 1
INTRODUCTION

1.1 Overview

Sign language is an important part of life for deaf and mute people. They rely on it for everyday communication with their peers. A sign language consists of a well-structured code of signs, and gestures, each of which has a particular meaning assigned to it. They have their own grammar and lexicon. It includes a mixture of hand positioning, shapes and movements of the hand. The people who know sign language can communicate with each other efficiently. However, when it comes to communicating with people who don't understand sign language it causes a lot of problems. Communication is a very important part of our lives. We interact with our mates at offices, schools, hospitals and other public places. Deaf and mute people may find it difficult to express themselves in such situations because not everyone understands sign language. There are many highly talented people suffering from speech impediment. We feel that their disability should become a hindrance to achieve their goals. Adding them into the workforce will only improve the socio-economic development of the country. Deaf and mute people usually depend on sign language interpreters for communication. However, finding a good interpreter is difficult and often expensive. Thus, a computerized interpreter could be a reliable and cheaper alternative. A system that can translate sign language into plain text or audio can help in real-time communication. It can also be used to provide interactive learning of sign language.

This project aims to get the deaf and dumb people more involved to communicate and the idea of a camera-based sign language recognition system that would be in use for converting sign language gestures to text (English) and then to regional.The proposed system contains modules such as: pre-processing and hand segmentation, feature extraction, sign recognition and sign to text. By using image processing the segmentation can be done. Some of the features are extracted such as Eigen values and Eigen vectors which are used in recognition. The Linear Discriminant Analysis (LDA) algorithm was used for gesture recognition and recognized gesture is converted into text and voice format. The proposed system helps to dimensionality reduction.

1.2 Objectives

The main objective of this project is to identify the sign language gestures which are shown to a web camera using the LDA Algorithm in machine learning and given as text and speech output. Thereby all people can understand the intended meaning of each sign language gesture regardless of whether the person is trained or not.

1.3 Data Description

The different hand gesture for ISL (Indian Sign Language) are collected from internet. Then image segmentation and feature extraction are done followed by LDA algorithm and output is received as text and speech with regional language conversion. The data-set as a whole is generally divided into three categories: training, testing and validation. As in the case with a usual learning problem, training the model would be using pre-trained dataset and evaluating the performance with the test dataset.

1.4 Organization of the project

The report is organised as follow:

- **Chapter 1:Introduction** Gives an introduction to "Talk2Mute", our project which identifies the sign language gestures and converts it into text and speech.
- **Chapter 2:Literature Survey** Summarizes the various existiin techniques that helps in achieving the desired result.
- **Chapter 3: Problem Statement** Discuses about the need for the proposed system
- **Chapter 4:Project Management** Contains the effective project management model to be used for the project.
- **Chapter 5:Proposed System** Describes the various steps involved to produce this project.
- **Chapter 6:Conclusion** Concludes with the future scope of implementation.
- **References** Includes the references for the project.

CHAPTER 2
LITERATURE SURVEY

2.1 Sign Language Converter [2]

The aim of this paper is to improve the communication with the people who has hearing difficulties and using any sign language to express themselves. At the first sight, as an idea, how difficult could make a sign languages converter. After detailed research about sign language linguistics, it is figured out about 240 sign languages have exist for spoken languages in the world. Infrastructure of a sign language system consists of three main branches as Sign Language, Speech Recognition and Implementation with MS Kinect XBOX 360TM. Infrastructure of this sign language system consists of three main branches as Sign Language, Speech Recognition and Implementation with MS Kinect XBOX 360TM.

2.1.1 Implementation

Sign Language

It is easy to find a wide number of sign languages all over the world and almost every spoken language has its respective sign language, so there are about more than 200 languages available. American Sign Language (ASL) is well-known and the best studied sign language in the world. The different basic signs from the ASL dictionary have been taken and the parameters assigned to these characteristics as respect to the position, motion and plane.

Microsoft Kinect XBOX 360TM

Microsoft Kinect Sensor XBOX 360TM used for capturing abilities and technical features to the motion capture of sign to voice conversion [4]. Part A is a depth sensor or called as 3D sensor too. It is a combine of infrared laser projector with a CMOS sensor to let the Kinect sensor to process 3D scenes in any environmental light conditions. Part B is RGB camera which has 32 bits high color resolution. It can use 2 dimensional color video of the scene. Part C is motorized tilt which is concerned with the field of view. Part D contains an array of 4 microphones which is located along the horizontal bar. It is useful for speech recognition, ambient noise suppression and echo cancellation.

Open NI Middleware

Open NI is a multi-language that defines APIs for writing applications as a cross platform framework. Main purpose of Open NI is to form a standard API that enables communication with the sensors in the system such as vision and audio sensors. Open NI standard API let the natural interaction application developers to 3D scenes by utilizing data types like array of the pixels in a depth map.Open NI has 3 layers. Bottom layer contains devices that collect visual and audio data from the real world. Second layer contains Middleware components to analyze these data which is collected from the real world. The top layer contains software which implements natural application such as Sign Language Translator.

2.1.2 Methodology

There are three parts of methodology:

1. Database
2. Voice Recognition Procedure
3. Motion Capture Procedure

Database

Words for Speech Recognition, .gif images and Motions together create the database.

Voice Recognition Procedure

Speech processing is the field which works on the speech signals and the processing of them. Thesignals are usually processed in a digital representation, although the signals are analog. Once the user press the button to record the speech, computer's microphone starts to listen, and after catching the voice with the help of CMU Sphinx, it finds the meaning as the text. Then in Java it is matched with the proper .gif image, so that the other user will understand.

Motion Capture Procedure

In this procedure, image processing is really important. Image processing is used commonly in our life recently, and it seems that future will bring much more than that. the motion capturing is the part where Kinect Sensor is used. Once the user press the button to record the motion, Kinect sensor starts to capture motions, but it starts to record the motion with a specific "starting

motion". After the “Starting Motion”, Kinect captures the motions and it converts them to the text. On the computer, this text is converted to the voice and then the other user can hear the meaning of the sign.

2.1.3 Results

This paper is about a system that can support the communication between deaf and ordinary people. The aim of the study is to provide a complete dialog without knowing sign language. After this system, it is an opportunity to use this type of system in any places such as schools, doctor offices, colleges, universities, airports, social services agencies, community service agencies and courts, briefly almost everywhere.

2.2 Conversion of Sign Language to Text and Speech Using Machine Learning techniques [1]

Communication has been defined as an act of conveying intended meanings from one entity or group to another through the use of mutually understood signs and semiotic rules. It plays a vital role in the existence and continuity of human. For an individual to progress in life and coexist with other individuals there is the need for effective communication. Effective communication is an essential skill that enables us to understand and connect with people around us. It allows us to build respect and trust, resolve differences and maintain sustainable development in our environment where problem solving, caring and creative ideas can thrive. Poor communication skills are the largest contributor to conflict in relationships. subsectionPCL System

Research has shown that over nine billion people at intervals, all over the world are physically challenged in terms of communication; blind, deaf or mute. Academic and industrial researchers have recently been focusing on analyzing images of people and there has been a surge interest in recognizing human gestures. A research on scene segmentation of images was carried out using deep learning techniques; the classification yielded 53.8 percentage accuracy. In relation to conversion of sign language, there is the need to explore other image classification techniques to enhance accurate classification.

2.2.1 Methods and Materials

The aim of the study as earlier stated was to provide an unsupervised learning feature of signed hand gestures while the system returns corresponding output as text and speech. The following were the measurable methods employed in actualising the aim:

1. Segmentation of captured signed gestures of ASL as inputs
2. Feature extraction of the segmented images
3. UFL and classification of several images
4. Text and Speech synthesis of classified images

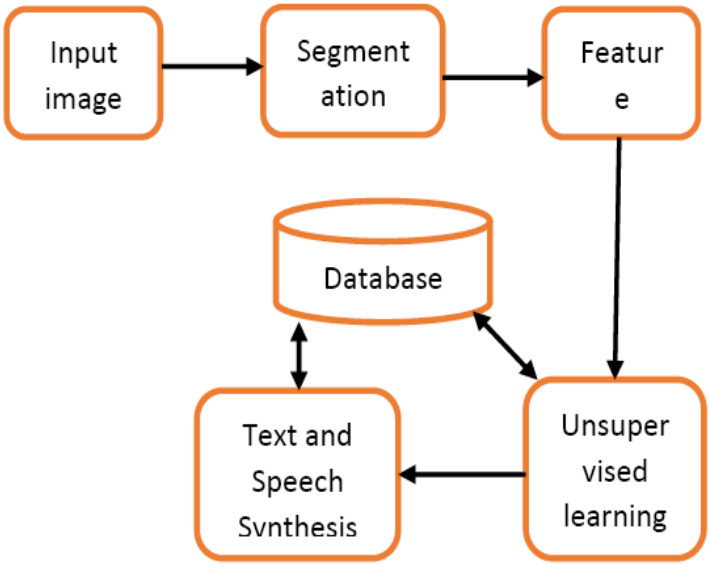

Figure 2.1: Overview of the system

- **Segmentation of captured signed gestures of ASL as inputs**:

  The aim of segmentation was to convert images into more meaningful and easy to analyse portions. Segmentation does the job of partitioning an image into multiple segments which help to locate the objects and boundaries (curves, arcs, lines, etc.) in an image in binary form. The set of images captured from the Kinect sensor using the Image Acquisition Tool in MATLAB would be selected and fed into the Image Segmenter in MATLAB which is then converted to grayscale image. The threshold of the images is then obtained by converting grayscale images into binary images to determine the high level contrast of the images. Such images can then be cropped or resized.[6]

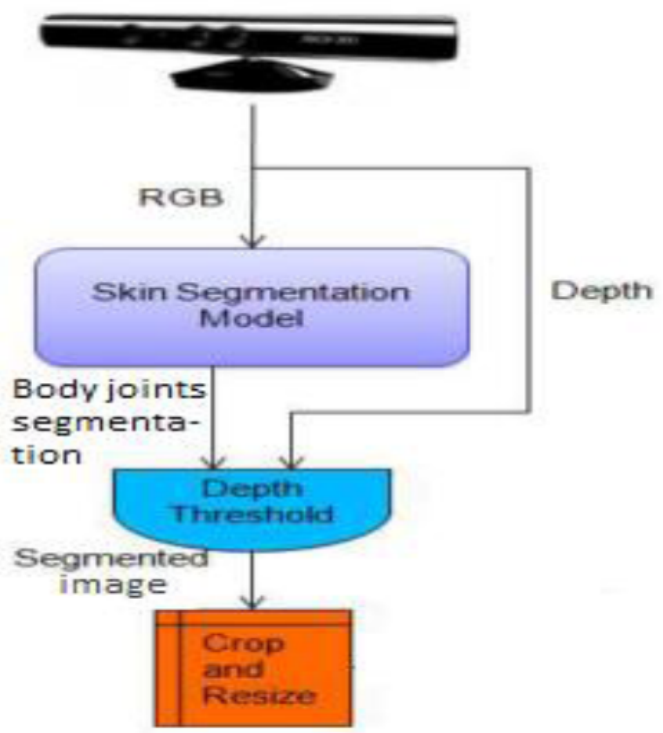

Figure 2.2: Image Segmentation

- **Feature extraction of the segmented images**:

  Sign Language usually involves movement in the upper part of the body; head, shoulder, hands and elbows coordinates are retained while other parts are discarded. To satisfy this need, key points corresponding to high-contrast locations such as object edges and corners were used. These features are intended to be non-redundant, informative and relevant for the intended use. Extracting ROI from images has been very much challenging as it is the base for further image analysis, interpretation and classification. A rectangular ROI whose outline consists of four segments joining the four corner points is used to make computational statistics feasible.Batch segmentation for all training samples was carried out to convert the coloured images into binary form with MATLAB Image Segmenter and Batch Processor toolbox. Basically only the set of binarised images are useful in feature detection and extraction.

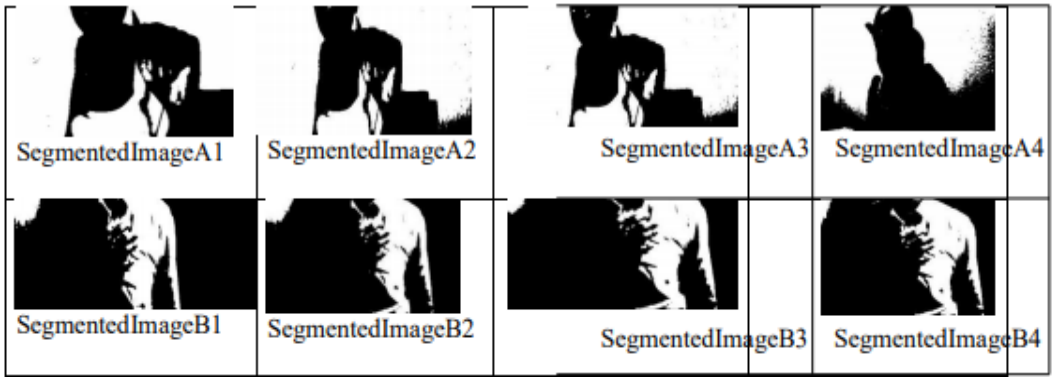

Figure 2.3: Segmented ImageSet of Coloured Images

The vertices of an ROI outline may be positioned anywhere with respect to the array of image pixels, so the same Rectangular ROI superimposed on the pixel array may appear.

- **UFL and classification of several images**:

The identification of interest points present within the space of an image is important in the determination of the image's ROI, therefore the method being proposed in this paper maximizes the number of interest points detected within a sample image through the use of the combination of FAST corner detector and SURF detector

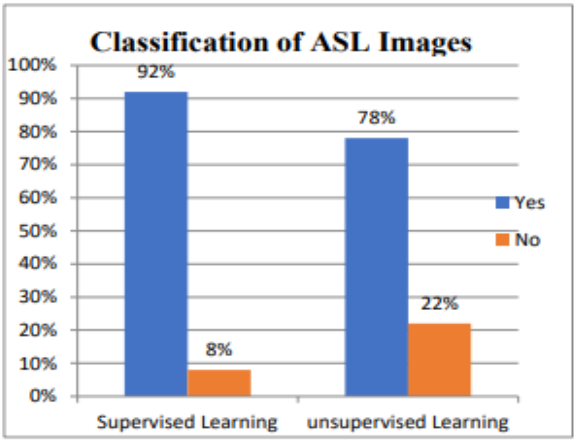

Figure 2.4: Classification of ASL Images

- **Text and Speech synthesis of classified images**:

Text-to-Speech (TTS) refers to the ability of computers to read text aloud. A TTS Engine converts written text to a phonemic representation, and then converts the phonemic representation to waveforms that can be output as sound. Speech synthesis is the

artificial production of human speech. A computer system used for this purpose is called a speech synthesizer, and can be implemented in software or hardware. After the successful classification of these features, the important task is to generate appropriate text and speech output for every input image using MATLAB Speech Synthesizer.[3]

2.2.2 Results

Sample images of different ASL signs were collected using the Kinect sensor using the image acquisition toolbox on MATLAB. About five hundred (500) data samples (with each sign count five and ten (5-10)) were collected as the training data. The reason for this is to make the algorithm very robust for images of the same database in order to reduce the rate of misclassification.

The study findings show that American Sign Language (ASL) is commonly used in Nigeria by the hearing impaired hence; five hundred (500) ASL images were collected as training set. From the collection of images, a database of forty-nine (49) different signs was used. Having subjected the set of images to batch segmentation, features of each signs were detected and extracted from specific bounding-box of Region of Interest (ROI) to aid supervised learning. The combination FAST and SURF with a KNN of 10 also showed that unsupervised learning classification could determine the best matched feature from the existing database. In turn, the best match was converted to text as well as speech. The introduced system achieved a 92% accuracy of supervised feature learning and 78% of unsupervised feature learning.

2.3 Conversion of Sign Language into Text [8]

Sign Language is the most natural and expressive way for the hearing impaired people. People, who are not deaf, never try to learn the sign language for interacting with the deaf people. This leads to isolation of the deaf people. But if the computer can be programmed in such a way that it can translate sign language to text format, the difference between the normal people and the deaf community can be minimized. Indian sign language (ISL) uses both hands to represent each alphabet and gesture. ISL alphabets are derived from British Sign Language (BSL) and French Sign Language (FSL). Most of the researchers in this area concentrate on the recognition of American Sign Language (ASL) since most of the signs in ASL are single handed and thus, complexity is less. Another attractive feature is that ASL already has a standard database that is available for use. When compared with ASL, Indian Sign Language relies on both hands and thus, an ISL recognition system is more complex.

Deaf and Dumb people rely on sign language interpreters for communications. A

real time Sign Language Recognition system was designed and implemented to recognize 26 gestures from the Indian Sign Language by hand gesture recognition system for text generation. The signs are captured by using web cam. This signs are processed for feature extraction using some colour model. The extracted features are compared by using pattern matching algorithm. In order to calculate the sign recognition, the features are compared with testing database. Finally, recognized gesture is converted into text. This system provides an opportunity for a deaf-dumb people to communicate with non-signing people without the need of an interpreter.Nowadays, people are not interested to speak in ASL when having a deaf relative or friend, or even classmate/acquaintance. Hence, deaf people are often trapped and isolated. ASL requires the use of a person's hands so if something happens where a wrist was sprained and it disables that person from talking. For example, there was a mother who strained her wrist from signing all of her life for her deaf daughter. The doctor also made her stop signing. This caused the communication with her deaf daughter to decrease, since she had to read lips from then on.

ASL vocabulary dictionary contains thousands of sign justlike words. It is very easy to get two completely different signs mixed up which leads to bad miscommunication. For example, the sign for "chocolate" and "cleve land" are similar, and they definitely don't mean the same thing, or even close. It is very hard to follow when a conversation has and something gets mixed up.

Related works:
A various hand gestures were recognized with different methods by different researchers in which were implemented in different fields. The recognition of various hand gestures were done by vision based approaches, data glove based approaches, soft computing approaches like Artificial Neural Network, Fuzzy logic, Genetic Algorithm and others like PCA, Canonical Analysis, etc. The recognition techniques are divided into three broad categories such as Hand segmentation approaches, Feature extraction approaches and Gesture recognition approaches.

2.3.1 LDA Approach

The Generalization of the Fisher's linear discriminant(FLD) is known as Linear discriminant analysis (LDA). LDA mainly used in statistics, pattern recognition and machine learning. It is used to find a linear combination of features that characterizes or separates two or more classes of objects or events. The LDA and FLD are used linear classifier. Its combination also used for dimensionality reduction before later classification. [10]

LDA is also closely resembles to principal component analysis (PCA) and factor analysis. Both PCA and factor analysis is linear combinations of variables and they describe the data

in a better manner. LDA explains to model the difference between the classes of data. PCA cannot consider the difference in class but factor analysis builds the feature combinations based on differences rather than similarities. There is a difference between Discriminant analysis and factor analysis in that it is not an interdependence technique: a distinction between independent variables and dependent variables (also called criterion variables) must be made. In LDA, the measurements made on independent variables for each observation are continuous quantities. Discriminant correspondence analysis is used to deal with categorical independent variables in LDA.

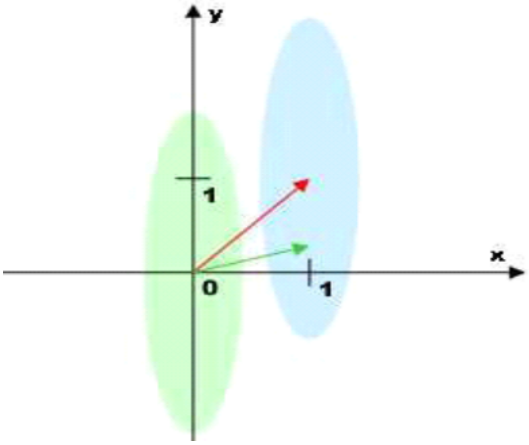

Figure 2.5: LDA Approach

The above figure 2.9 explains about the LDA approach.

- Compute the d-dimensional mean vectors for the different classes from the dataset
- Compute the scatter matrices (between class and within-class scatter matrix)
- Compute the eigenvectors (e1, e2, ..., ed) and corresponding eigenvalues (1, 2, ..., d) for the scatter matrices
- Sort the eigenvectors by decreasing Eigen values and choose k eigenvectors with the largest Eigen values to form a d×k-dimensional matrix W (where every column represents an eigenvector)
- Use this d×k eigenvector matrix to transform the samples onto the new subspace. This can be summarized by the equation Y = X × W (where X is an n×d-dimensional matrix; the ith row represents the ith sample, and Y is the transformed n×k-dimensional matrix with the n samples projected into the new subspace)

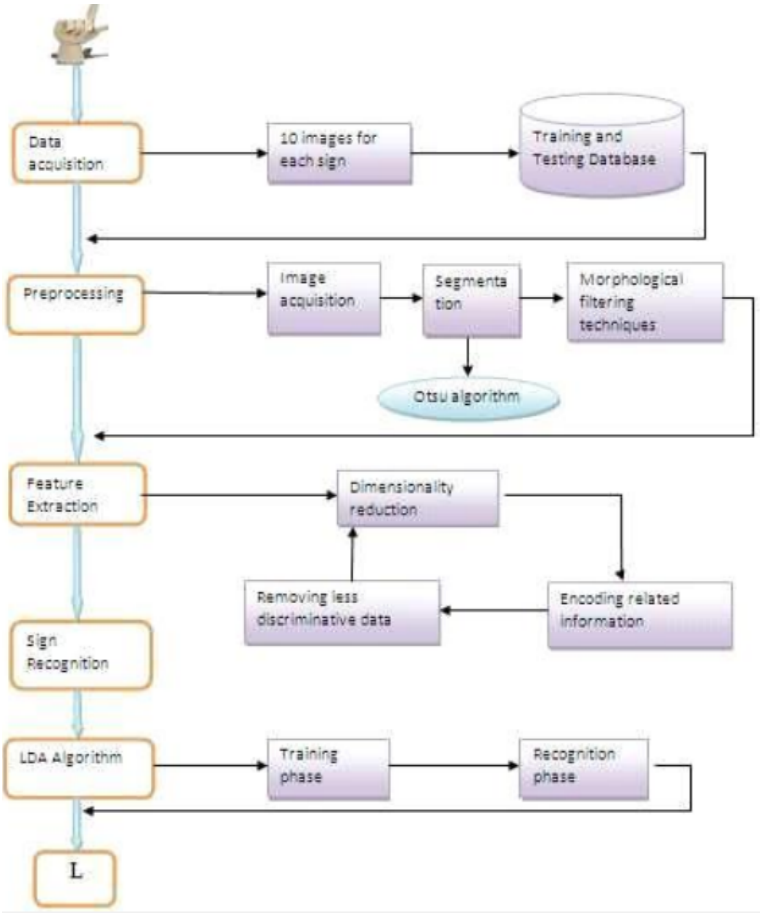

Figure 2.6: System Overview of Indian Sign Language

2.3.2 Data Acquisition

To achieve a high accuracy for sign recognition in sign language recognition system, 10 images will be taken for each 26 signs. These images are included in training and testing database. The captured image at a distance is adjusted by the signer to get the required image clarity

2.3.3 Pre-Processing

Pre-processing consist image acquisition, segmentation and morphological filtering methods.

The above figure 2.9 explains about the LDA approach.

- **Image acquisition**

  This is the first step of pre-processing. This is the process of sensing of an image. So in an

image acquisition, image is sensed by “illumination”. It will also involve pre-processing such as scaling. In image acquisition the image will be taken from database.

- **Segmentation**
  Segmentation is the process in which image is converted into small segments so that the more accurate image attribute can be extracted. If the segments are properly autonomous (two segments of an image should not have any identical information) then representation and description of image will be accurate and while taking rugged segmentation, the result will not be accurate. Here the Segmentation of hands is carried out to separate object and the background. Otsu algorithm is used for segmentation purpose. The segmented hand image is represented certain features. The following figure 4 shows the segmented of hand image.

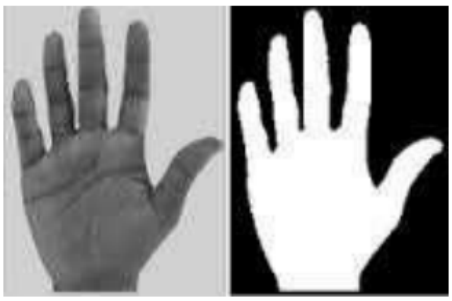

Figure 2.7: Segmented Image

- **Morphological Filtering**
  The image components are extracted by Morphological Filtering tools which are useful for representation and description of shape. Definitely the output of this process is image attribute. The following figure 5 shows the filtered form of segmented image.

  The features extracted from the segmentation operation used for gesture recognition. The smooth contour is obtained by removing the noise from the images with Morphological filtering techniques. The pre-processing operation is done on the stored database.

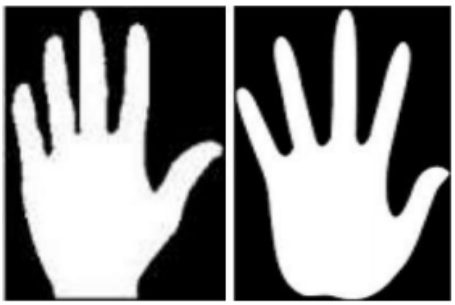

Figure 2.8: Morphological Filtered Image

2.3.4 Feature Extraction

The reduction of data dimensionality by encoding related information in a compressed representation and removing less discriminative data is called as Feature extraction Technique. Feature extraction is vital to gesture recognition performance. Therefore, the selection of which features to deal with and the extraction method are probably the most significant design decisions in hand motion and gesture recognition development. Here principal component is used as main features

Figure 2.9: Feature extraction method

2.3.5 Sign Recognition

Sign recognition using LDA is a dimensionality reduction technique based on extracting the desired number of principal components of the multi-dimensional data. The gesture recognition using LDA algorithm that involves two phases.

- **Training Phase** Each gesture is represented as a column vector in the training phase. These gesture vectors are then normalized with respect to average gesture. Next, the algorithm finds the eigenvectors of the covariance matrix of normalized gestures by using a speed up technique that reduces the number of multiplications to be performed. The corresponding gesture space projections were obtained by the eigenvector matrix then multiplied by each of the gesture vectors.

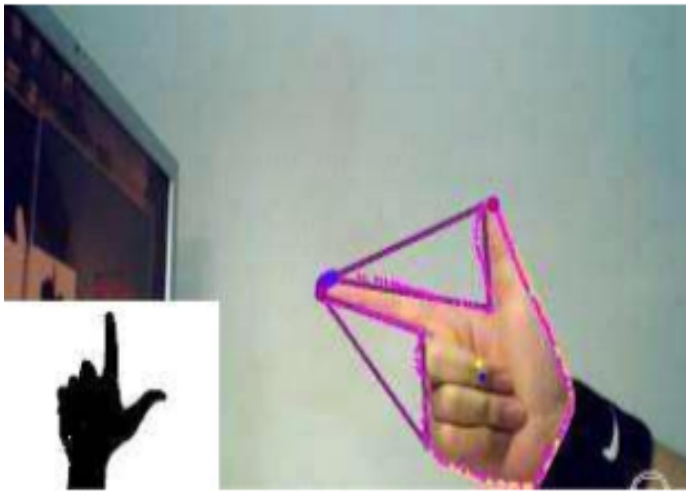

Figure 2.10: Dimensionality reduction

- **Recognition Phase** In the recognition phase, a subject gesture is normalized with respect to the average gesture and then projected onto gesture space using the eigenvector matrix. Lastly, Euclidean distance is computed between this projection and all known projections. The minimum value of these comparisons is selected for recognition during the training phase. Finally, recognized sign is converted into appropriate text and voice which is displayed on GUI.

Currently, research works have focused mainly on the recognition of static signs of ISL from images or video sequences that have been recorded under controlled conditions. By using LDA algorithm for sign recognition operation the dimensionality will be reduced. Due to dimensionality reduction the noise will be reduced and with high accuracy. In future this project will be enhanced by determining the numbers which will be shown in words. Using various concepts of image processing and fundamental properties of image we tried to developed this system. By using LDA algorithms recognition of gesture has done successfully. Every God creature has an importance in the society, remembering this fact, let us try to include hearing impaired people in our day to day life and live together.

2.4 Sign Language Interpreter using Image Processing and Machine Learning [9]

Speech impairment is a disability which affects one's ability to speak and hear. Such individuals use sign language to communicate with other people. Although it is an effective form of communication, there remains a challenge for people who do not understand sign

language to communicate with speech impaired people.Sign language is an important part of life for deaf and mute people. They rely on it for everyday communication with their peers. A sign language consists of a well-structured code of signs, and gestures, each of which has a particular meaning assigned to it. They have their own grammar and lexicon. It includes a mixture of hand positioning, shapes and movements of the hand. The people who know sign language can communicate with each other efficiently.hearing impaired people. People, who are not deaf, never try to learn the sign language for interacting with the deaf people. This leads to isolation of the deaf people. But if the computer can be programmed in such a way that it can translate sign language to text format, the difference between the normal people and the deaf community can be minimized. Indian sign language (ISL) uses both hands to represent each alphabet and gesture. ISL alphabets are derived from British Sign Language (BSL) and French Sign Language (FSL). Most of the researchers in this area concentrate on the recognition of American Sign Language (ASL) since most of the signs in ASL are single handed and thus, complexity is less. Another attractive feature is that ASL already has a standard database that is available for use. When compared with ASL, Indian Sign Language relies on both hands and thus, an ISL recognition system is more complex. A few research works carried out by the researchers in the recognition of ISL. Currently, more researchers have started doing research in ISL. Here this proposed system is able to recognize the various alphabets of Indian Sign Language; this will reduce the noise and give accurate result.

However, when it comes to communicating with people who don't understand sign language it causes a lot of problems. Communication is a very important part of our lives. We interact with our mates at offices, schools, hospitals and other public places. Deaf and mute people may find it difficult to express themselves in such situations because not everyone understands sign language. There are many highly talented people suffering from speech impediment. We feel that their disability should become a hindrance to achieve their goals. Adding them into the workforce will only improve the socio-economic development of the country. The aim is to develop an application which will translate sign language to English in the form of text and audio, thus aiding communication with sign language. The application acquires image data using the webcam of the computer, then it is pre-processed using a combinational algorithm and recognition is done using template matching. The translation in the form of text is then converted to audio. The database used for this system includes 6000 images of English alphabets. We used 4800 images for training and 1200 images for testing. The system produces 88% accuracy.

Deaf and mute people usually depend on sign language interpreters for communication. However, finding a good interpreter is difficult and often expensive. Thus, a computerized interpreter could be a reliable and cheaper alternative. A system that can translate sign language

into plain text or audio can help in real-time communication. It can also be used to provide interactive learning of sign language. There is no universal sign language for deaf people. Different countries use their own sign language, although there some striking similarities among them. It is yet unclear how many sign languages exist in the world. Some languages have got legal recognition and some have not. India's National Association of Deaf estimates that there are 18 million people in India with hearing impairment. Here discusses the implementation of a system which translates Indian Sign Language gestures to its English language interpretation.

2.4.1 Proposed Method

The proposed system consists of two main stages: (1) Segmentation of hand (2) Recognition of hand sign. The block diagram shows the working of the proposed system. The features of a hand are an important criterion for the classifier to differentiate between the hand gestures. These characteristics must be able to adapt to different hand and gestures by different people. In this system, we have used histograms of oriented gradients (HOG) as a feature descriptor. It is better than other descriptors because it can adapt to changing illuminations and rotation of objectives. It does not consider an image as a whole, but divides it into smaller cells and then for the pixels within the cells edge or gradient direction histogram is calculated. This approach creates a bin, and clubs the histograms of different samples based on magnitude and angle. In the proposed system, we are first segmenting the hand using YCbCr color space and then processing the image through HOG and then provide it to the model. We trained the SVM classifier using 5000 images and developed a model.

2.4.2 Dataset

We have created a dataset of 26 English language alphabets in Indian Sign Language. Each sign gesture is performed by two different individuals with different hand structure in varied lighting conditions. The videos were recorded on a camera and then each video was broken down frame by frame to images and adjusted to 100 frames and then augmented to get about 250 images for each sign. The data was then divided into 4800 images for training and 1200 images for testing

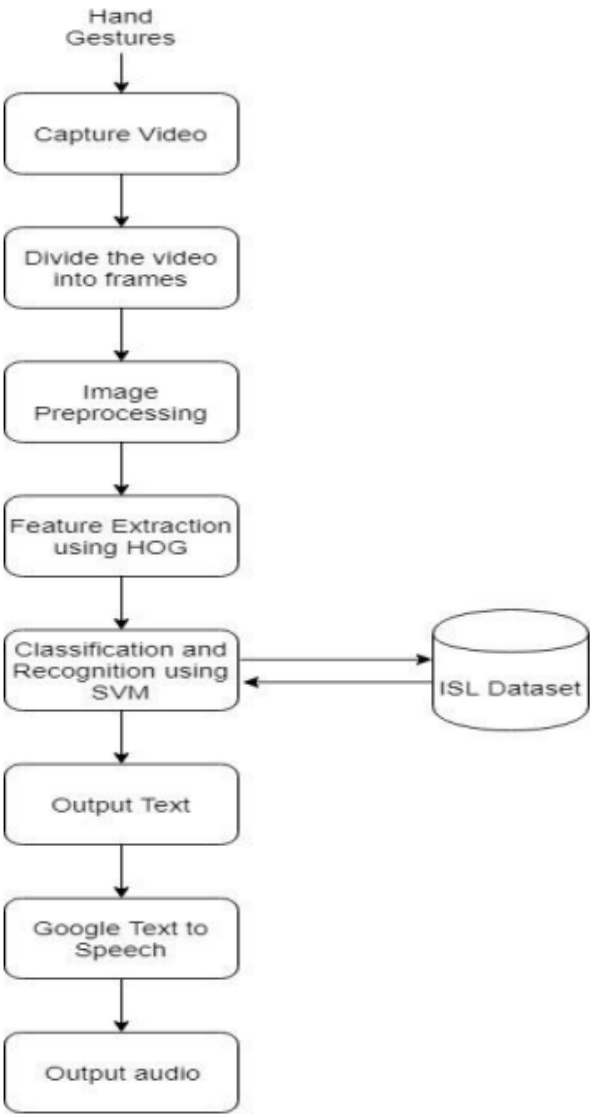

Figure 2.11: Sign language interpreter flowchart

- **Image preprocessing and edge detection**
  Before performing feature extraction, the images must be processed in such a way that only the useful information is considered and the redundant, distracting noise and superficial data are neglected. The images are first converted to 100 * 100 pixel size for faster computations. The image is then converted to grayscale and finally transform into a binary image. Simultaneously, skin color is detected using YCbCr model. Finally, edge detection is performed using Canny edge detector. The process is illustrated in figure below.

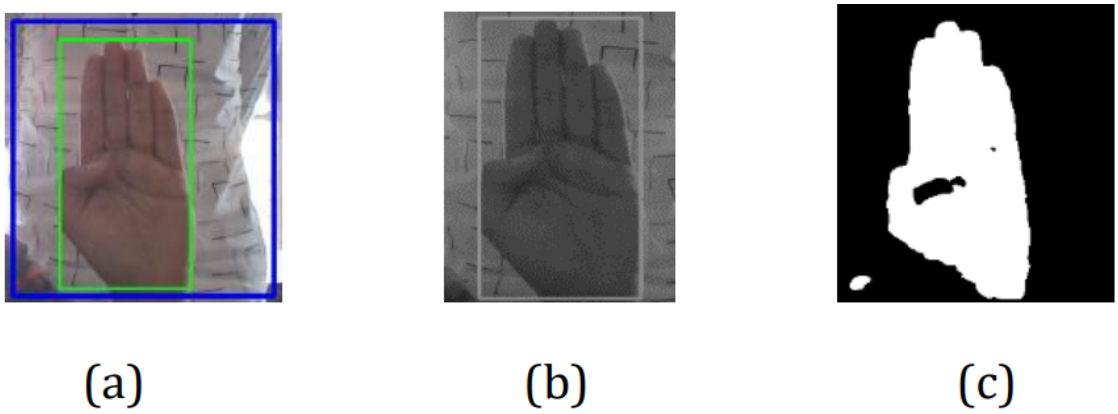

Figure 2.12: a) Original cropped image, b) Grey scale converted image, c) Skin colour detection output

- **Feature extraction**

  Extracting features from the data is a critical part of any object detection system. It can be performed using various methods such as Fourier Descriptor, Scale Invariant Feature Transform (SIFT) or Principle Component Analysis (PCA). Histogram of Oriented Gradients (HOG) is another successful method of feature extraction. In this paper, we have used the HOG method for feature extraction. The basic idea behind HOG is that an object or shape within an image can be represented by intensity distribution gradients or edge directions. In this method, the image is divided into small cells and a histogram of gradient is calculated for each cell. It creates a bin and combines the histogram of different samples based on magnitude and angle. For improving the accuracy, all the cells in the image are normalized so that they are not affected by variations in lighting. Finally, HOG feature vector is calculated for the entire image.

- **Template matching and sign recognition**

  The feature vector produced in the above step is fed into an image classifier. In this paper, we have used Support Vector Machine (SVM) for classification. By using SVM classifier we can maximize accuracy and avoid overfitting of data. In SVM data items are plotted in n-dimensional space where n is the number of features. Each feature is associated with a coordinate value. Then it finds a hyperplane that differentiates the classes. The model is saved for real-time sign language recognition.[5]

- **Text to Speech**

  We have used Google's Text to Speech API for transforming the sign language into audio. It is one of the best text to speech API available. Unlike other TTS APIs, this API generates human-like voice. The sign language is interpreted using the above steps and then the result is fed to text to speech function which converts it to audio. In this system,

we can see and hear the sign language translation at the same time which makes it very convenient to use.

2.4.3 Results

In this system, still hand image frame is captured using a webcam. These frames are processed to get enhanced features. Then feature extraction and classification algorithms are used to translate the sign language into English text. This translation is converted to speech using text to speech API. The system has implemented using the above algorithms to get the final output. The proposed model is evaluated by a dataset containing 26 signs from two different people. The results show the overall accuracy of the system to be 88%. Our future research will work towards implementing this model on a mobile application.

2.5 Sign Language Translator using Machine Learning [7]

Machine learning provides a versatile and robust environment to work on. The machine learning subject also eliminates the need for the coder to write updates whenever a new sign is read, this will be done by the machine itself. This system aims to get the deaf and dumb people more involved to communicate and the idea of a camera-based sign language recognition system that would be in use for converting sign language gestures to text (English) and then to regional. A product that is versatile and robust has to be found. Here, the objective is to design a solution that is intuitive and simple which simplifies the communication for the majority of people with deaf and dumb. There are many methods to convert the sign language which often use Kinect as the basic system to get the inputs and work on them for conversion. Kinect methods are complicated in so many aspects.The discussing approach will be simple. Here, used simpler ways to capture the inputs and process them. We have used common and easily available libraries in our system

The main objective of this project is to recognizing the gestures and displaying the correspondent word. The first phase involves capturing the gesture using a webcam along with pose estimation library.The webcam captures the image and image is processed with pose estimation algorithm in tensor-flow utility. The skeleton mapped on the image is the result of the pose estimation library. The skeleton obtained provides the values for creating the data set; the data set is a collection of the values of the coordinates of the end points of the skeleton. These values are labelled accordingly and are appended to the machine for predicting when the input is taken.

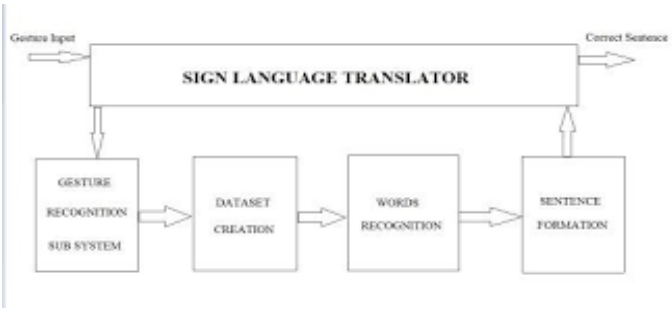

Figure 2.13: Block diagram of the SLR model

2.5.1 Extraction of gestures:

Capturing signs from real world and translating them is the core objective of this work. The real-world signs are read using a webcam which captures both static and moving images of the objects in front of it. The deaf and dumb person who is signing is made to stand in front of the webcam and the image captured from this is processed with the tf-pose-estimation library to map out the skeleton of the person signing. Figure 2.2 is an example of how the skeleton is mapped on the system. Tf-pose-estimation basically sketches out a stick figure of the body. When the webcam is running the pose estimation algorithm identifies the key points on the subject's body such as elbow joints, wrist, knee joints etc and connects them as one skeleton. The key points namely the end points of the skeleton are labelled with x, y and z co-ordinates for every frame captured. As such 17 key points are identified from the pose-estimation algorithm. The value of these coordinates change for different gestures and the relative distance between the key points is different for different people (as size changes from person to person). These coordinates are the main component to form the data set for training.

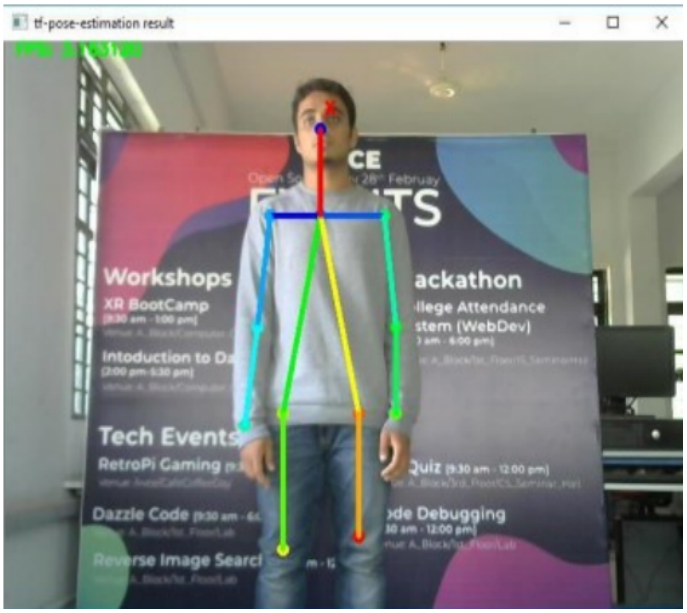

Figure 2.14: Tf-pose-estimation result

Creating Data set:

Each gesture captured has its coordinates values stored in a csv (comma separated file) file and the corresponding labels are written in another csv file. Each frame as 17 key points and each of these points has 3 coordinates; therefore, there exists 17x 3 values for each frame. This entire set of 51 values is labelled together as one gesture. We are assigning numbers for these gestures in the corresponding csv file as it is easy to handle integers while training the machine. Here '0' is the uses to label certain gestures for example in Figure 2.2 a boy standing 'Idle' is recorded figures representing the coordinates of his skeleton, and '0' is the label for this posture. Later '0' is substituted for "Idle" while displaying the result. Training the machine requires several sets of values, therefore a single person has to record many frames for a single gesture and the same gesture has to be signed by different people. Different people are required to sign for the same posture as the size of the skeleton varies from person to person.Several frames have to be recorded for a single gesture by a single person and several people have to record the same gesture to provide a better data set for training[

2.5.2 Training

The data set created is taken up in the training platform and Decision tree algorithm is used to train the machine. The set of values stored for a specific gesture is referred by the machine in its training and makes it possible to predict the gesture when the input is taken.The input values will be run through the tree and the final answer will be displayed along with its value and the

corresponding label. The corresponding values for labels are then substituted with words and are displayed in the result. Every new gesture has several frames recorded for it and trained using the decision tree algorithm. More the number of frames recorded better the efficiency of the system in predicting the gesture.

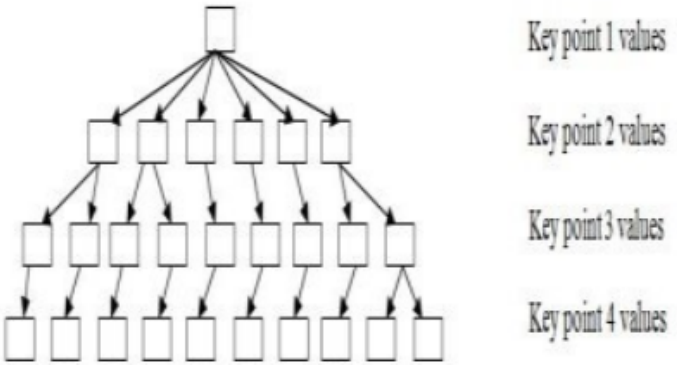

Figure 2.15: Decision Tree Values

The system, when provided with the proper gestures, gives out the corresponding words. The system can provide proper results even when there are some slight variations in gestures. There will be different kinds of variations from different kinds of persons performing the gestures. The system recognizes multiple gestures one after the after and the gives out the respective words.The requirement of machine-based sign language translator is very important in the present scenario. Even though we have found initial success in this regard, lot of work needs to be done.The main area where this can be used is in public places like ticket issuing counters, hospitals etc. This can be even used to teach the sign language to normal people. Further this can be used to take words and display the gesture for the same. Recognizing fingers will widen the training set for the machine.

CHAPTER 3
PROBLEM STATEMENT

The project "Talk2Mute" aims at helping the people who are not able to speak or hear, to communicate to others in a more comfortable and easier way. Deaf and mute people usually depend on sign language interpreters for communication. However, finding a good interpreter is difficult and often expensive. Speech impairment is a disability which affects one's ability to communicate with others. Such individuals use sign language to communicate with other people. Although it is an effective form of communication, there remains a challenge for people who do not understand sign language to communicate with speech impaired people. The people who know sign language can communicate with each other efficiently. However, when it comes to communicating with people who don't understand sign language it causes a lot of problems. Therefore, This system aims to translate sign language to English and Regional languages in the form of text and audio, using machine learning techniques, thus aiding communication with sign language.

CHAPTER 4
PROJECT MANAGEMENT

4.1 Introduction

Project management is the discipline of planning, organizing, securing, managing, leading, and controlling resources to achieve specific goals. A project is a temporary endeavor with a defined beginning and end (usually time-constrained, and often constrained by funding or deliverables), undertaken to meet unique goals and objectives, typically to bring about beneficial change or added value. The temporary nature of projects stands in contrast with business as usual (or operations), which are repetitive, permanent, or semi-permanent functional activities to produce products or services. In practice, the management of these two systems is often quite different, and as such requires the development of distinct technical skills and management strategies.

In our project we are following the typical development phases of an engineering project

1. Initiation
2. Planning and Design
3. Execution and Construction
4. Monitoring and Controlling Systems
5. Completion

4.1.1 Initiation

The initiating processes determine the nature and scope of the project. The initiating stage should include a plan that encompasses the following areas :

1. Analysing the business needs/requirements in measurable goals
2. Reviewing of the current operations
3. Financial analysis of the costs and benefits including a budget
4. Stakeholder analysis, including users, and support personal for the project

5. Project charter including costs, tasks, deliverables, and schedule

4.1.2 Planing and design

After the initiation stage, the project is planned to an appropriate level of detail. The main purpose is to plan time, cost and resources adequately to estimate the work needed and to effectively manage risk during project execution. As with the initiation process, a failure to adequately plan greatly reduces the project's chances of successfully accomplishing its goals.

- Determining how to plan
- Developing the scope statement
- Selecting the planning team
- Identifying deliverables and creating the work breakdown structure
- Identifying the activities needed to complete those deliverables
- Developing the schedule
- Risk planning

4.1.3 Execution

Executing consists of the processes used to complete the work defined in the project plan to accomplish the project's requirements. The execution process involves coordinating people and resources, as well as integrating and performing the activities of the project in accordance with the project management plan. The deliverables are produced as outputs from the processes performed as defined in the project management plan and other frameworks that might be applicable to the type of project at hand.

4.1.4 Monitoring & controlling

Monitoring and controlling consists of those processes performed to observe project execution so that potential problems can be identified in a timely manner and corrective action can be taken, when necessary, to control the execution of the project. The key benefit is that project performance is observed and measured regularly to identify variances from the project management plan.

4.2 System Development Life Cycle

The Systems development life cycle (SDLC), or Software development process in systems engineering, information systems, and software engineering, is a process of creating or altering information systems, and the models and methodologies that people use to develop these systems. In software engineering, the SDLC concept underpins many kinds of software development methodologies. These methodologies form the framework for planning and controlling the creation of an information system.

The SDLC phases serve as a programmatic guide to project activity and provide a flexible but consistent way to conduct projects to a depth matching the scope of the project. Each of the SDLC phase objectives is described in this section with key deliverables, a description of recommended tasks, and a summary of related control objectives for effective management. The project manager must establish and monitor control objectives during each SDLC phase while executing projects. Control objectives help to provide a clear statement of the desired result or purpose and should be used throughout the entire SDLC process.

4.2.1 Spiral Model

We have used the Spiral model in our project. The Spiral model incorporates the best characteristics of both- waterfall and prototyping model. In addition, the Spiral model also contains a new component called Risk Analysis, which is not there in the waterfall and prototype model. In the Spiral model, the basic structure of the software product is developed first. After the basic structure is developed, new features such as user interface and data administration are added to the existing software product. This functionality of the Spiral model is similar to a spiral where the circles of the spiral increase in diameter. Each circle represents a more complete version of the software product. The spiral is a risk-reduction oriented model that breaks a software project up into main projects, each addressing one or major risks. After major risks have been addressed the spiral model terminates as a waterfall model. Spiral iteration involves six steps:

1. Determine objectives, alternatives and constraints.
2. Identify and resolve risks.
3. Evaluate alternatives.
4. Develop the deliverables for the iteration and verify that they are correct.
5. Plan the next iteration.

6. Commit to an approach for the next iteration.

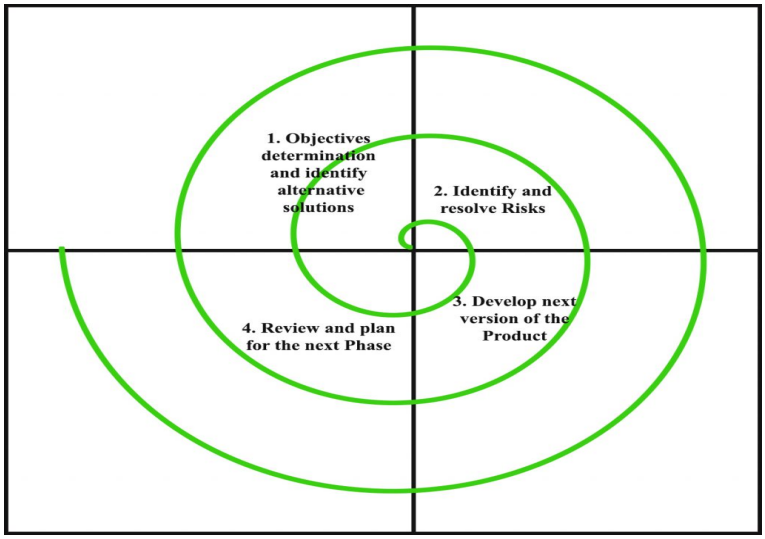

Figure 4.1: Spiral Model

CHAPTER 5
METHODOLOGY

5.1 System Requirements & Specifications

5.1.1 Tensorflow

TensorFlow is a free and open-source software library for machine learning. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. TensorFlow provides stable Python and C++ APIs, as well as non-guaranteed backward compatible API for other languages.

5.1.2 Windows 10

Windows 10 is a series of personal computer operating systems produced by Microsoft as part of its Windows NT family of operating systems. It is the successor to Windows 8.1 and was released to manufacturing on July 15, 2015, and to retail on July 29, 2015. Windows 10 receives new builds on an ongoing basis, which are available at no additional cost to users. Mainstream builds of Windows 10 are labeled version YYMM with YY representing the year and MM representing the month of release. For example, the latest mainstream build of Windows 10 is Version 1809. There are additional test builds of Windows 10 available to Windows Insiders. Devices in enterprise environments can receive these updates at a slower pace, or use long-term support milestones that only receive critical updates, such as security patches, over their ten-year lifespan of extended support.

5.1.3 Python 3.6.2

Python is a dynamic object-oriented programming language that can be used for many kinds of software development. It offers strong support for integration with other languages and tools, comes with extensive standard libraries, and can be learned in a few days. Many Python programmers report substantial productivity gains and feel the language encourages the development of higher quality, more maintainable code.

Python runs on Windows, Linux/Unix, Mac OS X, OS/2, Amiga, Palm Handhelds, and

Nokia mobile phones. Python has also been ported to the Java and .NET virtual machines. Python is distributed under an OSI-approved open source license that makes it free to use, even for commercial products.

5.1.4 Jupyter Environment

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more. JupyterLab is a web-based interactive development environment for Jupyter notebooks, code, and data. JupyterLab is flexible: configure and arrange the user interface to support a wide range of workflows in data science, scientific computing, and machine learning. JupyterLab is extensible and modular: write plugins that add new components and integrate with existing ones.

5.1.5 Visual Studio Code

Visual Visual Studio Code is a free source-code editor made by Microsoft for Windows, Linux and macOS.[8] Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git. Users can change the theme, keyboard shortcuts, preferences, and install extensions that add additional functionality.

5.1.6 API

An application programming interface (API) is a computing interface that defines interactions between multiple software intermediaries. It defines the kinds of calls or requests that can be made, how to make them, the data formats that should be used, the conventions to follow, etc. It can also provide extension mechanisms so that users can extend existing functionality in various ways and to varying degrees.[1] An API can be entirely custom, specific to a component, or designed based on an industry-standard to ensure interoperability. Through information hiding, APIs enable modular programming, allowing users to use the interface independently of the implementation.

5.2 Proposed System

Modules

5.2.1 Data Acquisition Module

To achieve a high accuracy for sign recognition in sign language recognition system, 6 images will be taken for each 26 signs and others. These images are included in training and testing database. The captured image at a distance is adjusted by the signer to get the required image clarity.. In my project data set we build a new dataset of 700 images of 70% training section and 30% testing section.

5.2.2 Data Preprocessing Module

Pre-processing consists of segmentation and morphological filtering methods.

1. **Segmentation:**

   Segmentation is the process in which image is converted into small segments so that the more accurate image attribute can be extracted. If the segments are properly autonomous (two segments of an image should not have any identical information) then representation and description of image will be accurate and while taking rugged segmentation, the result will not be accurate. Here the Segmentation of hands is carried out to separate object and the background.

2. **Morphological Filtering:**

   The image components are extracted by Morphological Filtering tools which are useful for representation and description of shape. Definitely the output of this process is image attribute. The features extracted from the segmentation operation used for gesture recognition. The smooth contour is obtained by removing the noise from the images with Morphological filtering techniques. The preprocessing operation is done on the stored database.

5.2.3 Feature Extraction

The reduction of data dimensionality by encoding related information in a compressed representation and removing less discriminative data is called as Feature extraction Technique. Feature extraction is vital to gesture recognition performance. Therefore, the selection of which

features to deal with and the extraction method are probably the most significant design decisions in hand motion and gesture recognition development. Here principal component is used as main features.

5.2.4 Sign Recognition

Sign recognition using LDA is a dimensionality reduction technique based on extracting the desired number of principal components of the multi-dimensional data. The gesture recognition using LDA algorithm that involves two phases

- **Training Phase** Each gesture is represented as a column vector in the training phase. These gesture vectors are then normalized with respect to average gesture. Next, the algorithm finds the eigenvectors of the covariance matrix of normalized gestures by using a speed up technique that reduces the number of multiplications to be performed. The corresponding gesture space projections were obtained by the eigenvector matrix then multiplied by each of the gesture vectors.
- **Recognition Phase** In the recognition phase, a subject gesture is normalized with respect to the average gesture and then projected onto gesture space using the eigenvector matrix. Lastly, Euclidean distance is computed between this projection and all known projections. The minimum value of these comparisons is selected for recognition during the training phase. Finally, recognized sign is converted into appropriate text and voice which is displayed on GUI.

5.2.5 Translation of output text into regional language and audio

Here a translator API is used to convert original text into Regional Language. A translation API can dynamically translate text between language pairs. State-of-the-art text translation APIs support thousands of language pairs. These APIs are based on statistical machine translation and machine learning. Text-to-Speech (TTS) refers to the ability of computers to read text aloud. A TTS Engine converts written text to a phonemic representation, and then converts the phonemic representation to waveforms that can be output as sound. Speech synthesis is the artificial production of human speech. A computer system used for this purpose is called a speech synthesizer, and can be implemented in software or hardware.

ALGORITHM IDENTIFIED

Algorithm identified for classification is Linear discriminant analysis (LDA)

ALGORITHM USED: LDA

Linear Discriminant Analysis is a dimensionality reduction technique which is commonly used for the supervised classification problems. It is used for modeling differences in groups. i.e. separating two or more classes. The main goal of dimensionality reduction techinques is to reduce the dimensions by removing the reduntant and dependent features by transforming the features from higher dimensional space to a space with lower dimensions.

Linear Discriminant Analysis projects the features in higher dimension space onto a lower dimensional space. This can be achieved in three steps :

The first step is to calculate the separability between different classes(i.e the distance between the mean of different classes) also called as between-class variance

$$S_b = \sum_{i=1}^{g} N_i(\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T$$

Second Step is to calculate the distance between the mean and sample of each class,which is called the within class variance

$$S_w = \sum_{i=1}^{g} (N_i - 1)S_i = \sum_{i=1}^{g} \sum_{j=1}^{N_i} (x_{i,j} - \bar{x}_i)(x_{i,j} - \bar{x}_i)^T$$

The third step is to construct the lower dimensional space which maximizes the between class variance and minimizes the within class variance.Let P be the lower dimensional space projection,which is called Fisher's criterion.

$$P_{lda} = \arg\max_{P} \frac{\left| P^T S_b P \right|}{\left| P^T S_w P \right|}$$

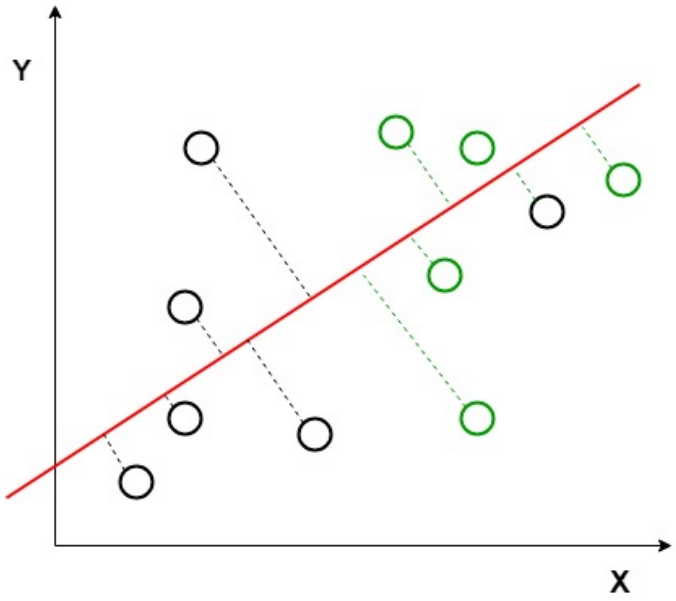

Figure 5.1: Linear Discriminant Analysis

Linear Discriminant Analysis is a simple and effective method for classification. Because it is simple and so well understood, there are many extensions and variations to the method. Some popular extensions include:

- **Quadratic Discriminant Analysis (QDA)**
  Each class uses its own estimate of variance (or covariance when there are multiple input variables).
- **Flexible Discriminant Analysis (FDA)**
  Where non-linear combinations of inputs is used such as splines.
- **Regularized Discriminant Analysis (RDA)**
  Introduces regularization into the estimate of the variance (actually covariance), moderating the influence of different variables on LDA.

In the field of Computer Vision, hand gesture recognition recognition is a very popular application in which each gesture is represented by a very large number of pixel values. Linear discriminant analysis (LDA) is used here to reduce the number of features to a more manageable number before the process of classification. Each of the new dimensions generated is a linear combination of pixel values, which form a template. The linear combinations obtained using Fisher's linear discriminant are called Fisher faces.

5.3 Data Flow Diagrams

5.3.1 Data Flow Diagram- Level 0

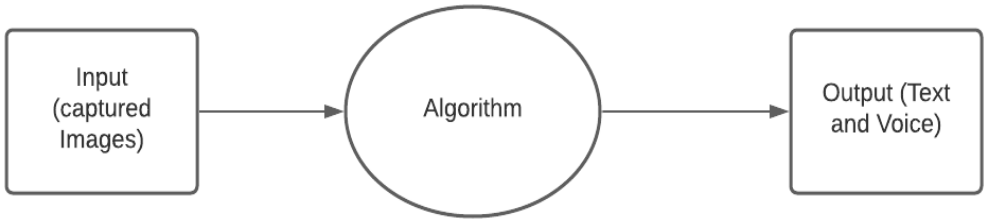

Figure 5.2: DFD- Level 0

5.3.2 Data Flow Diagram- Level 1

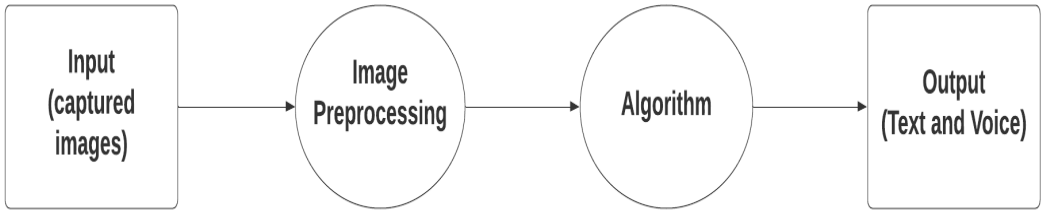

Figure 5.3: DFD- Level 1

5.3.3 Data Flow Diagram- Level 2

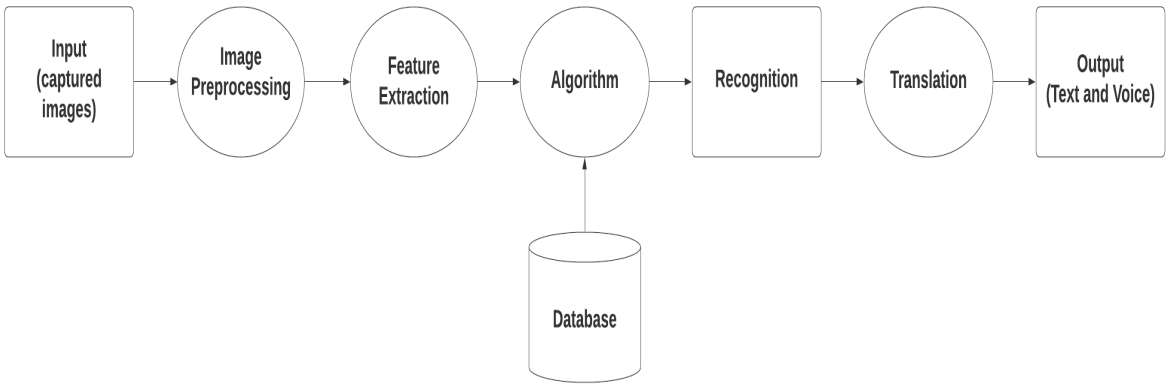

Figure 5.4: DFD- Level 2

5.4 Architecture

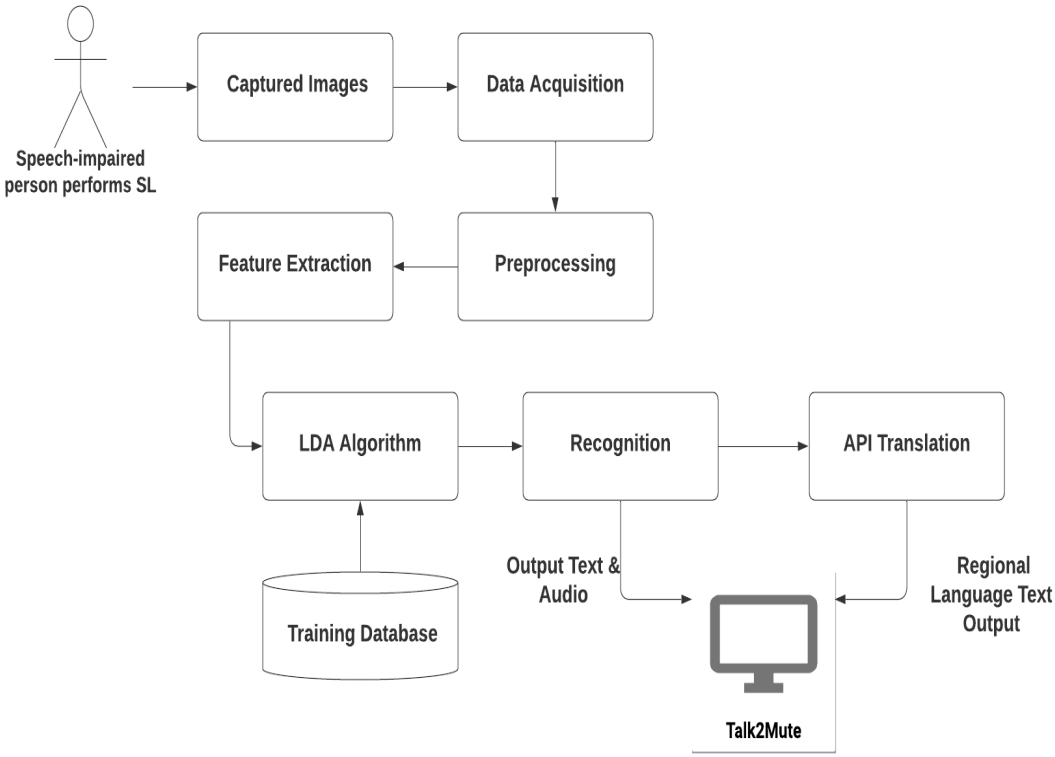

Figure 5.5: Architecture

CHAPTER 6

RESULTS

Sign Language Recognition is one of the most growing fields of research area. Many new techniques have been developed recently in this area. The Sign Language is mainly used for communication of deaf-dumb people.The people who know sign language can communicate with each other efficiently. However, when it comes to communicating with people who don't understand sign language it causes a lot of problems. Therefore, This system aims to translate sign language to English and Regional languages in the form of text and audio, thus aiding communication with sign language.By using image processing the segmentation can be done. Some of the features are extracted such as Eigen values and Eigen vectors which are used in recognition. The Linear Discriminant Analysis (LDA) algorithm was used for gesture recognition and recognized gesture is converted into text and voice format. The proposed system helps to dimensionality reduction.

CHAPTER 7
CONCLUSION AND FUTURE WORKS

Currently, research works have focused mainly on the recognition of static signs of ISL from images or video sequences that have been recorded under controlled conditions. By using LDA algorithm for sign recognition operation the dimensionality will be reduced. Due to dimensionality reduction the noise will be reduced and with high accuracy. In future this project will be enhanced by determining the numbers which will be shown in words.

Using various concepts of image processing and fundamental properties of image we are trying to develop this system. By using LDA algorithms recognition of gesture has done successfully. In future this project can be turned into a learning platform for understanding sign languages and also with some extra functionalities like video calling feature, this project can make lives of the speech deaf and mute more easier.

REFERENCES

[1] Victoria A Adewale and Adejoke O Olamiti. Conversion of sign language to text and speech using machine learning techniques. 2018.

[2] Taner Arsan and Oğuz Ülgen. Sign language converter. *International Journal of Computer Science & Engineering Survey (IJCSES)*, 6(4):39–51, 2015.

[3] Thierry Dutoit and Milos Cernak. Ttsbox: A matlab toolbox for teaching text-to-speech synthesis. In *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 5, pages v–537. IEEE, 2005.

[4] Jungong Han, Ling Shao, Dong Xu, and Jamie Shotton. Enhanced computer vision with microsoft kinect sensor: A review. *IEEE transactions on cybernetics*, 43(5):1318–1334, 2013.

[5] Huan Li, Fu-lai Chung, and Shitong Wang. A svm based classification method for homogeneous data. *Applied Soft Computing*, 36:228–235, 2015.

[6] Oge Marques. *Practical image and video processing using MATLAB*. John Wiley & Sons, 2011.

[7] Amit Moryossef, Ioannis Tsochantaridis, Roee Aharoni, Sarah Ebling, and Srini Narayanan. Real-time sign language detection using human pose estimation. In *European Conference on Computer Vision*, pages 237–248. Springer, 2020.

[8] Mahesh Kumar NB. Conversion of sign language into text. *International Journal of Applied Engineering Research*, 13(9):7154–7161, 2018.

[9] Omkar Vedak, Prasad Zavre, Abhijeet Todkar, and Manoj Patil. Sign language interpreter using image processing and machine learning. *International Research Journal of Engineering and Technology (IRJET)*, 2019.

[10] Jieping Ye and Qi Li. Lda/qr: an efficient and effective dimension reduction algorithm and its theoretical foundation. *Pattern recognition*, 37(4):851–854, 2004.