**SIGN LANGUAGE TO AUDIO TRANSLATOR USING MACHINE LEARNING AND RASPBERRY PI**

**A PROJECT REPORT**

*Submitted by*

**ANJURALDEVIMUTHU M (962817104012)**

**DEVIKANITHIAN N H (962817104034)**

**GOPIKA A S (962817104040)**

*in partial fulfilment for the award of the degree*

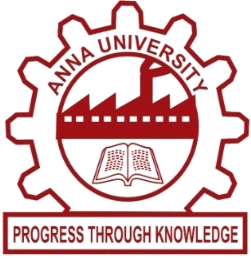
*Of*

**BACHELOR OF ENGINEERING**

**IN**

**COMPUTER SCIENCE AND ENGINEERING**

**UNIVERSITY COLLEGE OF ENGINEERING NAGERCOIL**



**ANNA UNIVERSITY: CHENNAI 600 025**

**APRIL 2021**

**BONAFIDE CERTIFICATE**

Certified that this report **“SIGN LANGUAGE TO AUDIO TRANSLATOR USING MACHINE LEARNING AND RASPBERRY PI”** is the bonafide work of **“ANJURALDEVIMUTHU M (962817104012),DEVIKANITHIAN N H (96281710403), and GOPIKA A S (962817104040)”** who carried out the project work under my supervision.

|  |  |
| --- | --- |
| **SIGNATURE** | **SIGNATURE** |
| **Dr. M. MUTHUSELVI, Ph.D.,**  **HEAD OF THEDEPARTMENT**  Computer Science and Engineering,  University College Of Engineering,  Nagercoil-629004 | **Dr. S. JEBA ANANDH, Ph.D.,**  **SUPERVISOR**  Teaching Fellow  Computer Science and Engineering,  University College Of Engineering,  Nagercoil-629004 |

Submitted to the project viva-voce examination held on\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**INTERNAL EXAMINER EXTERNAL EXAMINER**

**ACKNOWLEDGEMENT**

First and foremost, we thank **GOD** for his grace to enable us to complete this project work successfully.

We would like to extend our gratitude to our honourable Vice Chancellor, **Dr. M.K. SURAPPA, Ph.D.,** Anna University**,** Chennai for providing a better infrastructure for carrying out this project work.

Words cannot be expressed to show our gratitude to our respected Dean **Dr. T. SREE RENGA RAJA, Ph.D.,** for his support and freedom throughout our course.

We also express our deep gratitude to our respected Head of the Department **Dr. M. MUTHUSELVI,Ph.D.,** Assistant Professor, Department of Computer Science and Engineering for her valuable suggestions throughout the course of this project.

We thank our project coordinator **Dr. S. JEBA ANANDH, Ph.D.**, Department of Computer Science and Engineering.

We express our gratefulness and profound thanks to our Supervisor **Dr. S. JEBA ANANDH, Ph.D.,** Department of Computer Science and Engineering for his valuable guidance throughout this project.

And our sincere thanks to all **Department Teaching** and **Non-Teaching** **Staff Members** for their constant support throughout our course. Finally, we wish thank **Our Parents** for their blessings to complete project successfully.

**ABSTRACT**

Human beings interact with each other to convey their ideas, thoughts, and experiences to the people around them. But this is not the case for deaf-mute people. Sign language paves the way for deaf-mute people to communicate. Through sign language, communication is possible for a deaf-mute person without the means of acoustics. The aim behind this work is to develop a system for recognizing the sign language, which provides communication between people with speech impairment and normal people, thereby reducing the communication gap between them. Compared to other gestures, hand gesture plays an important role, as it expresses the user's views in less time. The main objective of this research is to develop an efficient sign language to voice conversion system using image processing and machine learning. The proposed system contains two parts: Training the model and integrating the trained model into the Raspberry Pi. It also contains four modules such as: pre-processing, Image processing, feature extraction, sign recognition and sign to text or voice. The machine learning algorithm is used for gesture recognition and recognized gesture is converted into audio format.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **CHAPTER NO** | **TITLE** | **PAGE NO** |
|  | **ABSTRACT** | **Iii** |
|  | **LIST OF FIGURES** | **Vii** |
| **1** | **INTRODUCTION** | **1** |
|  | 1.1 OBJECTIVE | 2 |
|  | 1.2 PROBLEM DEFINITION | 2 |
|  | 1.2.1 PURPOSE | 3 |
|  | 1.3 MOTIVATION | 3 |
|  | 1.4 LITERATURE REVIEW | 4 |
|  | 1.5 ORGANIZATION OF PROJECT REPORT | 14 |
| **2** | **SYSTEM DESCRIPTION** | **14** |
|  | 2.1 SYSTEM ANALYSIS | 14 |
|  | 2.1.1 EXISTING SYSTEM | 14 |
|  | 2.1.2 PROBLEM STATEMENT | 17 |
|  | 2.1.3 PROPOSED SYSTEM | 18 |
|  | 2.2 SYSTEM REQUIREMENTS | 19 |
|  | 2.2.1 HARDWARE REQUIREMENTS | 19 |
|  | 2.2.2 SOFTWARE REQUIREMENTS | 34 |
| **3** | **SYSTEM DESIGN** | **35** |
|  | 3.1 ARCHITECTURE DIAGRAM | 35 |
|  | 3.2 MODULE DESCRIPTION |  |
| **4** | 3.2.1 PRE-PROCESSING | **36** |
|  | 3.2.2 IMAGE PROCESSING | 36 |
|  | 3.2.3 FEATURE EXTRACTION | 36 |
|  | 3.2.4 SIGN RECOGNITION | 38 |
| **4** | **SYSTEM DEVELOPMENT** | 39 |
|  | Module Execution | 40 |
|  | System Testing | **42** |
| **5** | **RESULT AND DISUSSION** | **49** |
| **6** | **CONCLUSION AND FUTURE**  **ENHANCEMENTS** | **50** |
|  | 7.1 CONCLUSION FUTURE ENHANCEMENTS | **50** |
|  | **REFERENCES** | **51** |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **FIGURE NO** | **TITLE** | **PAGE NO** |
| 2.1 | Block Diagram Of Existing System |  |
| 2.2 | Raspberry Pi |  |
| 2.3 | Raspberry Pi Camera |  |
| 2.4 | Support vector machine Algorithm Example |  |
| 2.5 | Diagram of K-means Clustering |  |
| 3.1 | Architectural Diagram of Sign Language Translator |  |
| 4.1 | Execution of modules |  |
| 5.1 | Output Screenshot |  |

**CHAPTER 1**

**INTRODUCTION**

A sign is a form of non-verbal communication done with body parts, hand shapes, positions and movements of the hand, arms, facial expressions or movements of the lips and used instead of oral communication. Most people use both words and signs during communication. A sign language is a language that uses signs or action to communicate instead of sounds. A sign language has three major components. First important component is finger-spelling which means for each letter of the alphabet there is a corresponding sign. This type of communication is used mainly for spelling names sometimes for spelling the location names. Sometimes this can be used for expressing words for which no signs exist or for emphasizing or clarifying a particular word .Second vital component of any sign language is word level sign vocabulary which means for each word of the vocabulary there is a corresponding associated sign in the sign language. The most commonly used type of communication between people with hearing disabilities in combination with the facial expression is this type. Third essential component in sign communication is non-manual-features. This type of communication involves facial expressions, tongue, mouth, eyebrows and body position. Among all these components most used form of sign language by deaf community in reality is word level sign language. The most important part of Indian Sign Language (ISL) is it does not include grammar. Like the spoken languages and dialects, the sign language has designed differently depending on the region and culture. Sign Language used in India is different than American Sign Language. This system will act as an auxiliary tool for a deaf-mute to communicate with ordinary people through computer. Thus the major objective of this system is to make possible the communication between deaf people and the rest of the world in daily life will come into reality.

**1.1 OBJECTIVE**

To design an Independent communication system for person who is deaf and hard of hearing. Communication is an integral part of human life. But for people who are mute and hearing impaired, communication is a challenge. To understand them, one has to either learn their language i.e. sign language. The system proposed in this project aims at tackling this problem to some extent. In this project we train the system using Machine learning algorithms. The project uses a webcam to recognize the hand positions and outputs the Sign Language in PC onto the gesture made using Raspberry Pi. This will convert the gesture captured via webcam into audio output which will make normal people understand what exactly is being conveyed. Thus our project Sign Language to Audio Translator aims to convert the Sign Language into text and audio. The framework provides a helping-hand for speech-impaired to communicate with the rest of the world using sign language. This leads to the elimination of the middle person who generally acts as a medium of translation. This would contain a user-friendly environment for the user by providing speech/text output for a sign gesture input.

**1.2 PROBLEM DEFINITION**

Having to communicate between deaf people and normal public has become a difficult task now days and to implement a such as the society lacks a good translator for it and having an app for it in our mobile phones is like having a dream at day. But making an app for it is no simple task at it requires lot of efforts like memory utilization and a perfectly fined design to implement as such. What their application does is that they take a picture of a sign gesture and later converts is to a meaningful word. At first, they have compared the gesture using histogram that has been related to the sample test and moreover samples that are obliged to brief to basically reduce the weight on the CPU and its time. They have explained a process on which on their app, it’s very easy to add up a gesture and store it in their database for further and expand detection set. So lastly, they came strong with having an app as a translator instead of several applications that are being used lately by users.

**1.2.1 PURPOSE**

The purpose is to develop a system that act as an interpreter for communication for deaf-mute people with normal people.

**1.3 MOTIVATION**

Inability to speak is considered to be true disability. People with this disability use different modes to communicate with others, there are n number of methods available for their communication one such common method of communication is sign language. Sign language allows people to communicate with human body language; each word has a set of human actions representing a particular expression. The motive of the paper is to convert the human sign language to Voice with human gesture understanding and motion capture. This is achieved with the help of Microsoft Kinect a motion capture device from Microsoft. There are a few systems available for sign language to speech conversion but none of them provide natural user interface. For consideration if a person who has a disability to speak can stand perform the system and the system converts the human gestures as speech and plays it loud so that the person could actually communicate to a mass crowd gathering. Also the system is planned in bringing high efficiency for the users for improved communication.

**1.4 LITERATURE SURVEY**

**Siming Heon, “Research of a Sign Language Translation System based on Deep Learning”**

Sign language is a visual language that uses hand gesture, change of hand shape and track information to express meaning, and is the main communication tool for people with hearing and language impairment. Sign language recognition can improve the problem that the number of people who need to use sign language is large but the popularity of sign language is poor, and provide a more convenient way of study, work and life for people with hearing and language impairment. Hand locating and sign language recognition methods can generally be divided into traditional methods and deep learning methods. In recent years, with the brilliant achievements of deep learning in the field of computer vision, it has been proved that the method based on deep learning has many advantages, such as rich feature extraction, strong modelling ability and intuitive training. Therefore, this paper studies hand locating and sign language recognition of common sign language based on neural network**.** Sign language recognition methods are easily affected by human movement, change of gesture scale, small gesture area, complex background, illumination and so on. And some sign language recognition methods must use gesture areas to input information. Therefore, robust hand locating is an important pre-treatment step in sign language recognition. Compared with basic gestures, gestures in sign language are characterized by complex hand shape, blurred movement, low resolution of small target area, mutual occlusion of hands and faces, and overlapping of left and right hands.

**Inferences:**

* Rich feature extraction.
* Strong modelling ability.
* Intuitive training.

**Issues:**

* The speed of hand locating and recognition of sign language words is slow.
* Still much to improve in sign language recognition.
* **Kanchan Dabre, Surekha Dholay, “Machine Learning Model for Sign Language Interpretation using Webcam images”.**

Human beings interact with each other either using a natural language channel such as words, writing, or by body language (gestures) e.g. hand gestures, head gestures, facial expression, and lip motion and so on. As understanding natural language is important, understanding sign language is also very important. The sign language is the basic communication method within hearing disable people. People with hearing disabilities face problems in communicating with other hearing people without a translator. For this reason, the implementation of a system that recognize the sign language would have a significant benefit impact on deaf people social live**.** Sign language recognition interfaces can also be used as a natural communication channel in between humans and machines and give rise to applications such as hardware-free remote controls, human-computer interaction in virtual reality, gaming and other human welfare applications.

**Inferences:**

* This method proves to be much faster than other methods.
* Convergence rate is also faster it improves the speed of SL interpretation

**Issues:**

* This system is not versatile.
* The Speech synthesis phase of sign recognition process sometimes give delayed response.

**Ankita Harkude, Sarika Namade, Anita Morey, “Audio to Sign Language Translation for Deaf People”.**

Deaf people always miss out the fun that a normal person does, may it be communication, playing computer games, attending seminars or video conferences, etc. Communication is the most important difficulty they face with normal people and also every normal person does not know the sign language. The aim of our project is to develop a communication system for the deaf people. It converts the audio message into the sign language. This system takes audio as input, converts this audio recording message into text and displays the relevant Indian Sign Language images or GIFs which are predefined. By using this system, the communication between normal and deaf people gets easier. It is said that Sign language is the mother language of deaf people. This includes the combination of hand movements, arms or body and facial expressions. There are 135 types of sign languages all over the world. Some of them are American Sign Language (ASL), Indian Sign Language (ISL), British Sign Language (BSL), Australian Sign Language and many more. We are using Indian Sign Language in this project.

**Inferences:**

* Anywhere anyone can use this system for understanding the sign language.
* This system allows the deaf community to enjoy all sort of things that normal people do from daily interaction to accessing the information.

**Issues:**

* In this project, facial expression is not included.
* It takes time to process the output**.**

**Eleni Efthimiou, Stavroula, Christian Voglier, “Sign Language recognition, generation and modelling; A research effort with applications in deaf communication”.**

Sign language and Web 2.0 applications are currently incompatible,

because of the lack of anonymisation and easy editing of online sign language contributions. This paper describes Dicta-Sign, a project aimed at developing

the technologies required for making sign language-based Web contributions

possible, by providing an integrated framework for sign language recognition,

animation, and language modelling. It targets four different European sign languages: Greek, British, German, and French. Expected outcomes are three Show case applications for a search-by-example sign language dictionary, a sign language-to-sign language translator, and a sign language-based Wiki. The development of Web 2.0 technologies has made the WWW a place where people constantly interact with another, by posting information (e.g. blogs, discussion forums), modifying and enhancing other people's contributions (e.g. Wikipedia), and sharing information (e.g., Facebook, social news sites). The choice of human computer interface plays a critical role in these activities.

**Inferences:**

* It is the first sign Language project to make use of shallow translation technologies.
* It will also demonstrate a web 2.0 application for the Deaf.

**Issues:**

* Lack of robustness
* Difficulties with incorporating results from linguistic research into recognition system.

**Taner Arsan and Oguz Ulgel, “Sign Language Converter”.**

The aim of this paper is to design a convenient system that is helpful for

the people who have hearing difficulties and in general who use very simple and effective method; sign language. This system can be used for converting sign language to voice and also voice to sign language. A motion capture system is used for sign language conversion and a voice recognition system for voice conversion. It captures the signs and dictates on the screen as writing. It also captures the voice and displays the sign language meaning on the screen as motioned image or video*.* The aim of this paper is to improve the communication with the people who has hearingdifficulties and using any sign language to express themselves. At the first sight, as an idea, howdifficult 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. To show how tough to working with any sign language, the general information about sign languages is given briefly. After have an idea about sign language linguistics, Microsoft Kinect Sensor XBOX 360 is decided to use for capturing abilities and technical features to the motion capture of sign to voice conversion. Google Voice Recognition is used for the voice to sign conversion. Google Voice Recognition is available only on android based programs. Eventually, the voice recognition program CMU Sphinx is chosen. This allows us to combine both components in Java. Conversion program is also designed and written in Java. Finally, Java based program is produced which can make voice recognition, motion capture and convert both of them to each other. So a deaf person easily speaks to in sign language in front of motion sensor, the person behind the screen can understand easily without ability to speak sign language and vice versa.

**Inferences:**

* With the flexibility of Sphinx it is possible to add new words to the system.

**Issues:**

* This device is not very much reliable.
* Microsoft Kinect sensor is very expensive.

**Gunashekaran K, Manikandan R R, “Sign Language to Speech Translation System using PIC Microcontroller”.**

The advancement in embedded system, provides a space to design and develop a sign language translator system to assist the dumb people. This paper mainly addresses to facilitate dumb person's lifestyle. Dumb people throughout the world use sign language to communicate with others, this is possible for those who has undergone special trainings. Common people also face difficult to understand the gesture language. To overcome these real time issues, this system is developed. Whenever the proposed system senses any sign language, it plays corresponding recorded voice. This reduces the communication gap between dumb and ordinary people. This proposed model consist of four modules, they are sensing unit, processing unit, voice storage unit and wireless communication unit. It is achieved by integrating flux sensor and APR9600 with PIC16F877A. The flux sensors are placed in gloves, which respond to gesture. By using suitable circuit response of the sensor is given to the microcontroller based on the response microcontroller plays the recorded voice using APR9600. A snapshot of the entire system, advantage over existing methods and simulation output of the process is discussed in this work. This system offers high reliability and fast response. This method is more precise on hand movement and different languages can be installed without altering the code in PIC microcontroller. Nowadays embedded system emerging as an important trend in all applications. More recently developed embedded applications are changing our lifestyle in a smart way. Sign language is an expressive and natural way for communication between normal and dumb people (information majorly conveys through the hand gesture). The intension of the sign language translation system is to translate the normal sign language into speech and to make easy contact with the dumb people. In order to improve the life style of the dumb people the proposed system is developed. Sign language uses both physical and non-physical communication. The physical gesture communication consist of hand gestures that convey respective meaning, the non-physical is head movement, facial appearance, body orientation and position. Sign language not a universal language and it is different from country to country. America developed American Sign Language (ASL), British developed British sign language system (BSL) and Thailand developed Thai sign language system (TSL). Most of spoken English countries follow same sign language but same sign represents the different meaning and depends upon to their own language.

**Inferences:**

* High reliability
* Fast response.

**Issues:**

* The length of the program will be big of using microcontroller
* The system has a complex structure

**Abey Abraham and Rohini.V, “Real time conversion of sign language to speech and prediction of gestures using artificial neural network”.**

Sign language is generally used by the people who are unable to speak, for communication. Most people will not be able to understand the Universal Sign Language (unless they have learnt it) and due to this lack of knowledge about the language, it is very difficult for them to communicate with mute people. A device that helps to bridge a gap between mute persons and other people forms the crux of this paper. This device makes use of an Arduino Uno board, a few flex sensors and an Android application to enable effective communication amongst the users. Using the flex sensors, gestures made by the wearer is detected and then according to various pre-defined conditions for the numerous values generated by the flex sensors, corresponding messages are sent using a Global System for Mobile (GSM) module to the wearer’s android device, which houses the application that has been designed to convert text messages into speech. The GSM module is also used to send the sensor inputs to a cloud server and these values are taken as input parameters into the neural network for a time series based prediction of gestures. The system is designed to be a continually learning device and improve reliability by monitoring every individual’s behavior at all times. There are a lot of mute people around us who use sign languages for communication. Sign language is simply a collection of various gestures shown using hands and each of these gestures have certain meaning. Since they use sign languages and most of the other people who can speak does not know sign language which makes it difficult for communication. In this work, a product “SMART TALK” which is a glove with flex sensors is being developed. The implementation of American Sign Language gestures using these flex sensors is also done. Each gesture is being acquired accurately and those values have certain messages according to sign languages. These messages are send to the person’s android device and through the android application that we have developed, the text is converted to speech which will be out through speaker. Four flex sensors are used and the data from flex sensors is taken as input for a Back-Propagation neural network to predict the gestures of future. These gestures have certain meaning and by predicting the sensor values (gesture) we are predicting the mute person’s needs.

**Inferences:**

* The Neural network model will adjust their weights and make prediction accurate.

**Issues:**

* There was no provision by storing the sensor values for further study of data.

**Neha Poddou, Shrushti Rao, Shruti Sawant, “Study of Sign Language Translation using Gesture recognition”.**

Communication is an integral part of human life. But for people who are mute & hearing impaired, communication is a challenge. To understand them, one has to either learn their language i.e. sign language or finger language. The system proposed in this project aims at tackling this problem to some extent. In this paper, the motivation was to create an object tracking application to interact with the computer, and develop a virtual human computer interaction device. The motivation behind this system is two-fold. It has two modes of operation: Teach and Learn. The project uses a webcam to recognize the hand positions and sign made using contour recognition [3] and outputs the Sign Language in PC onto the gesture made. This will convert the gesture captured via webcam into audio output which will make normal people understand what exactly is being conveyed. Thus our project Sign Language to Speech Converter aims to convert the Sign Language into text and audio. The project aims to develop mobile interactive application software for translation of American Sign Language to speech and text and vice versa, to assist communication between hearing and speech impaired people with normal people. The language translator should be able to translate 26 alphabet gestures from the American Sign Language into speech and text and vice versa. We will basically acquire images using the inbuilt camera of the device and perform vision analysis, functions in the operating system and provide speech output through the inbuilt audio device.

**Inferences:**

* High Scalability and robustness.
* High Computational Efficiency

**Issues:**

* Low accuracy.
* This Project is more compatible with mobile phones.

**Nithyakalyani K, S Ramkumar, K Manikandan, “Design and implementation of Sign Language Translator using micro touch sensor”.**

Background: Use of sign language alone can impair the communication between the aurally impaired and the rest of the world. Lack of understanding by the unimpaired and environmental conditions such as darkness also affect communication when sign language is used. Objective: To develop a sign language translator for the aurally impaired, that will translate the signs to speech and display the output for better communication. Methodology: A glove incorporated with micro touch switch is used for the gesture recognition of Indian Sign Language (ISL). The output of the micro touch switch is given as an input to the Arduino based unit for converting it into speech and text. Results: The developed system was able to successfully detect signs for gestures such as I, Hello, Thank you, Know, Sorry, Yes, and No. Conclusion: This system was efficient in detecting the gestures with minimal time delay. This device was efficient in translation, and performed satisfactorily during the bench testing. However, a feasibility study with the intended patient group is required, where the efficacy of this device can be evaluated. Sign language is a visual language initially used by people who are aurally impaired by making gestures with hand and facial expressions. There are about 300 different gestures in Practice around the world which have resemblance and difference with each other. British and American sign languages are considered as different spoken languages. Some countries have more than one sign language. This diversity leads to confusion when people travel from one country to other. Even in their own country, normal people may not be familiar with the sign language which may limit the communication with the aurally impaired people, especially in dark it becomes even hard to interpret the gestures. Several projects to translate the sign language have been developed in these years but are poorly deployed.

**Inferences:**

* The system is capable of reducing the time limit for translation**.**

**Issues:**

* The system is expensive.
* Low reliability and Flexibility.

**Oi Mean Foong, Tang Jung Low, Wai Wan La, ”Voice to Sign Language Translation System for Malaysian deaf people”.**

The process of learning and understand the sign language may be

cumbersome to some, and therefore, this paper proposes a solution to this problem by providing a voice (English Language) to sign language translation system using Speech and Image processing technique. Speech processing which

includes Speech Recognition is the study of recognizing the words being spoken, regardless of whom the speaker is. This project uses template-based recognition as the main approach in which the V2S system first needs to be trained with speech pattern based on some generic spectral parameter set. These spectral parameter set will then be stored as template in a database. The system will perform the recognition process through matching the parameter set of the input speech with the stored templates to finally display the sign language in video format. Empirical results show that the system has 80.3% recognition rate. There are at least 70 million people around the globe who suffer from speech and hearing disabilities, either at birth or by accident. It is somehow difficult for us to interact with them because of the unfamiliar communication means used. Sign Language (SL) is a common form of communication which is widely used by the speech and hearing impaired. Thus, probably the only way of easier communication/ interaction with them is by learning their language - the sign language. We may have friends or family members who have hearing or speech disabilities. Such disabilities may be from birth, or by accident. Surely it is difficult for us to communicate with them if we do not know their language – the Sign Language. It is also difficult for them as well to communicate with us since they have such disability. One may be interested to learn up this language; however it may be costly to attend tuition classes to learn this language. Furthermore, tuition classes exhibits time constraint, where one does not have the flexibility in time on whether or not to attend the tutorial. He/she may prefer to have a self-tutorial, however there is no such inexpensive software that can self-taught them. These may contribute to the negligence of the public to learn the Sign Language to better communicate with those with hearing or speech impairment.

**Inferences:**

* Low cost
* The system has high recognition rate.

**Issues:**

* Training the system is more complex**.**

**1.5 ORGANIZATION OF PROJECT REPORT**

The project report is organized as follows:

Chapter 1 gives an introduction to the project, the objective, literature review, purpose and motivation.

Chapter 2 discuss about system requirements which includes software requirements and also the system analysis which includes requirement analysis as well as analysis of existing and proposed system.

Chapter 3 discusses about the system design which includes Architecture diagram.

Chapter 4 comprises of module wise description of the project.

Chapter 5 includes experimental results and discussion/test cases.

Chapter 6 concludes the project and discusses about future enhancement

**CHAPTER 2**

**SYSTEM DESCRIPTION**

**2.1 SYSTEM ANALYSIS**

System analysis is a method of studying a system by examining its component parts and their interactions. It is analysing the flow of information with data flow diagrams. System design is the process of defining architecture, components, and data of a system to satisfy specified requirements. System Analysis is an exciting, active field in which analysis continually learn new techniques and approaches to develop systems more effectively and efficiently. However, there is a core set of skills that all analysis needs to know no matter what approach or methodology is used. All information systems projects move through the four phases of planning, analysis, design and implementation.

**2.1.1 Existing System**

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. “Application research on face detection technology uses Open CV technology in mobile augmented reality” introduces the typical technology. Open source computer vision library, Open CV for short is a cross-platform library computer vision based on open source distribution. The Open CV, with C language provides a very rich visual processing algorithm to write it part and combined with the characteristics of its open source. Data gloves and Vision based method are commonly used to interpret gestures for human computer interaction. The sensors attached to a glove that finger flexion into electrical signals for determining the hand posture in the data gloves method. The camera is used to capture the image gestures in the vision based method. The vision based method reduces the difficulties as in the glove based method. “Hand talk-a sign language recognition based on accelerometer and semi data” this paper introduces American Sign Language conventions. It is part of the “deaf culture” and includes its own system of puns, inside jokes, etc. It is very difficult to understand understanding someone speaking Japanese by English speaker. The sign language of Sweden is very difficult to understand by the speaker of ASL. ASL consists of approximately 6000 gestures of common words with spelling using finger used to communicate obscure words or proper nouns. “Hand gesture recognition and voice conversion system for dumb people” proposed lower the communication gap between the mute community and additionally the standard world. The projected methodology interprets language into speech. The system overcomes the necessary time difficulties of dumb people and improves their manner. Compared with existing system the projected arrangement is simple as well as compact and is possible to carry to any places. This system converts the language in associate text into voice that's well explicable by blind and ancient people. The language interprets into some text kind displayed on the digital display screen, to facilitate the deaf people likewise. In world applications, this system is helpful for deaf and dumb of us those cannot communicate with ancient person. Conversion of RGB to grey scale and grey scale to binary conversion introduced in the intelligent sign language recognition using image processing. Basically any colour image is a combination of red, green, blue colour. A computer vision system is implemented to select whether to differentiate objects using colour or black and white and, if colour, to decide what colour space to use (red, green, blue or hue, saturation, luminosity).

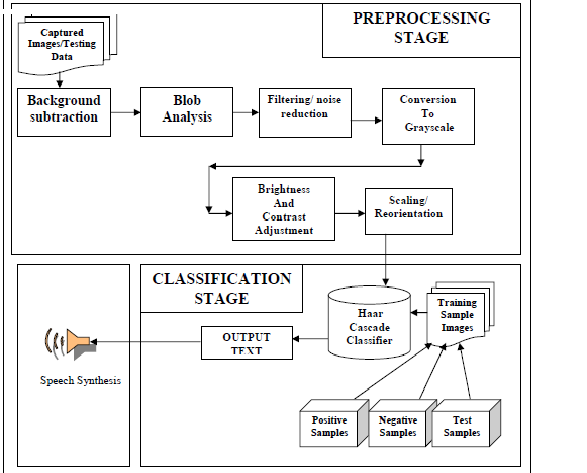


Fig 2.1 Block Diagram of Existing System

**2.1.2 PROBLEM STATEMENT**

There are 1.1 million deaf-mute people in India. 98% of these people are illiterate, hence sign language is the only way available to them for communication. When communication happen between deaf-mute people and normal people, there will be less transparency due to lack of understanding. So it is necessary to develop a system that has to be act as an interpreter between them. Each country has their own sign language. Thus the major objective of this system is to make possible the communication between deaf people and the rest of the world in daily life will come into reality. Until now there is no efficient system available for translating Indian sign language. So we proposed a system which translates Indian Sign language to audio. This system will act as an auxiliary tool for a deaf-mute to communicate with ordinary people through computer. We have proposed this system using Machine learning under Python platform as well as Raspberry pi. The output audio will be played in Raspberry pi and the text will be shown in the system.

**2.1.3 Proposed System**

Our objective is to help people suffering from the problem of hearing. There have been many projects done on the sign languages that convert sign language as input to text or audio as output. But audio to sign language conversion systems have been rarely developed. It is useful to both normal and deaf people. In the proposed system we are using Raspberry Pi which is portable, less cost when compared to other boards. The models are trained using machine learning algorithms. In this system we are using Indian Sign Language and it is converted to audio. This project work proposes a camera-based assistive sign detection and processing some functions like audio request or text using Raspberry Pi. Normally in deaf and dump schools they are using sign language for their communication .Our proposed system is mainly used for deaf and dump peoples, if they show some signs, each sign does some functions like audio output or text output. To provide information access and services to deaf people in Indian sign language. To develop a scalable project which can be extended to capture whole vocabulary of ISL through manual and non-manual signs.

* 1. **SYSTEM REQUIREMENTS**

**2.2.1 Hardware Requirements**

* Raspberry Pi 3
* USB
* HDM1

**Raspberry Pi 3:**

The Raspberry Pi is a series of small single-board computers developed in the United Kingdom by the Raspberry Pi Foundation to promote the teaching of basic computer science in schools and in developing countries. The original model became far more popular than anticipated, selling outside its target market for uses such as robotics. It does not include peripherals (such as keyboards, mice and cases). However, some accessories have been included in several official and unofficial bundles. According to the Raspberry Pi Foundation, over 5 million Raspberry Pi’s were sold by February 2015, making it the best-selling British computer. By November 2016 they had sold 11 million units, and 12.5m by March 2017, making it the third best-selling "general purpose computer". In July 2017, sales reached nearly 15 million. In March 2018, sales reached 19 million. Several generations of Raspberry is have been released. All models feature a Broadcom system on a chip (SoC) with an integrated ARM compatible central processing unit (CPU) and on-chip graphics processing unit (GPU). Processor speed ranges from 700 MHz to 1.4 GHz for the Pi 3 Model B+; on-board memory ranges from 256 MB to 1 GB RAM. Secure Digital (SD) cards are used to store the operating system and program memory in either SDHC or Micro SD HC sizes. The boards have one to four USB ports. For video output, HDMI and composite video are supported, with a standard 3.5 mm phono jack for audio output. Lower-level output is provided by a number of GPIO pins which support common protocols like I²C. The B-models have an 8P8C Ethernet port and the Pi 3 and Pi Zero W have on-board Wi-Fi 802.11n and Bluetooth. Prices range from US$5 to $35.

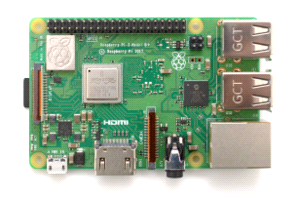


Fig 2.2 Raspberry Pi

The organization behind the Raspberry Pi consists of two arms. The first two models were developed by the Raspberry Pi Foundation. After the Pi Model B was released, the Foundation set up Raspberry Pi Trading, with Eben Upton as CEO, to develop the third model, the B+. Raspberry Pi Trading is responsible for developing the technology while the Foundation is an educational charity to promote the teaching of basic computer science in schools and in developing countries.The Foundation provides Raspbian, a Debian-based Linux distribution for download, as well as third-party Ubuntu, Windows 10 IoT Core, RISC OS, and specialized media centre distributions. It promotes Python and Scratch as the main programming language, with support for many other languages. The default firmware is closed source, while an unofficial open source is available. The Raspberry Pi hardware has evolved through several versions that feature variations in memory capacity and peripheral-device support.Raspberry Pi 3 V1.1 included a quad-core Cortex-A7 CPU running at 900 MHz and 1 GB RAM. It was described as 4–6 times more powerful than its predecessor. The GPU was identical to the original. In parallelised benchmarks, the Raspberry Pi 3 V1.1 could be up to 14 times faster than a Raspberry Pi 1 Model B+.

While operating at 700 MHz by default, the first generation Raspberry Pi provided a real-world performance roughly equivalent to 0.041 GFLOPS. On the CPU level the performance is similar to a 300 MHz Pentium II of 1997–99. The GPU provides 1 G pixel/s or 1.5 G Texel/s of graphics processing or 24 GFLOPS of general purpose computing performance. The graphical capabilities of the Raspberry Pi are roughly equivalent to the performance of the Xbox of 2001.The LINPACK single node compute benchmark results in a mean single precision performance of 0.065 GFLOPS and a mean double precision performance of 0.041 GFLOPS for one Raspberry Pi Model-B board. A cluster of 64 Raspberry Pi Model B computers, labelled "Iridis-pi", achieved a LINPACK HPL suite result of 1.14 GFLOPS (n=10240) at 216 watts for c. US$4000. The Raspberry Pi Foundation recommends the use of Raspbian, a Debian-based Linux operating system. Other third-party operating systems available via the official website include Ubuntu MATE, Windows 10 IOT Core, RISC OS and specialized distributions for the Kodi media centre and classroom management. Many other operating systems can also run on the Raspberry Pi.

**USB:**

Universal Serial Bus (USB) is an [industry standard](https://en.m.wikipedia.org/wiki/Technical_standard) that establishes specifications for cables and connectors and [protocols](https://en.m.wikipedia.org/wiki/Communication_protocol) for connection, communication and power supply between computers, [peripherals](https://en.m.wikipedia.org/wiki/Peripheral) and other computers. A broad variety of [USB hardware](https://en.m.wikipedia.org/wiki/USB_hardware) exists, including eleven different [connectors](https://en.m.wikipedia.org/wiki/USB_hardware#Connector_types), of which [USB-C](https://en.m.wikipedia.org/wiki/USB-C) is the most recent. Released in 1996, the USB standard is maintained by the [USB Implementers Forum](https://en.m.wikipedia.org/wiki/USB_Implementers_Forum) (USB-IF). There have been four generations of USB specifications: [USB 1.*x*](https://en.m.wikipedia.org/wiki/USB#USB_1.x), [USB 2.0](https://en.m.wikipedia.org/wiki/USB#USB_2.0), [USB 3.*x*](https://en.m.wikipedia.org/wiki/USB_3.0), and [USB4](https://en.m.wikipedia.org/wiki/USB4). USB was designed to standardize the connection of [peripherals](https://en.m.wikipedia.org/wiki/Peripheral) to personal computers, both to communicate with and to supply electric power. It has largely replaced interfaces such as [serial ports](https://en.m.wikipedia.org/wiki/Serial_port) and [parallel ports](https://en.m.wikipedia.org/wiki/Parallel_port), and has become commonplace on a wide range of devices. Examples of peripherals that are connected via USB include computer keyboards and mice, video cameras, printers, portable media players, disk drives, and network adapters. USB connectors have been increasingly replacing other types as charging cables of portable devices. The Universal Serial Bus was developed to simplify and improve the interface between personal computers and peripheral devices, when compared with previously existing standard or ad hoc proprietary interfaces. From the computer user's perspective, the USB interface improves ease of use in several ways:

* The USB interface is self-configuring, eliminating the need for the user to adjust the device's settings for speed or data format, or configure [interrupts](https://en.m.wikipedia.org/wiki/Interrupt), input/output addresses, or direct memory access channels.
* USB connectors are standardized at the host, so any peripheral can use most available receptacles.
* USB takes full advantage of the additional processing power that can be economically put into peripheral devices so that they can manage themselves. As such, USB devices often do not have user-adjustable interface settings.
* The USB interface is [hot-swappable](https://en.m.wikipedia.org/wiki/Hot_swapping) (devices can be exchanged without rebooting the host computer).
* Small devices can be powered directly from the USB interface, eliminating the need for additional power supply cables.
* Because use of the USB logo is only permitted after [compliance testing](https://en.m.wikipedia.org/wiki/Conformance_testing), the user can have confidence that a USB device will work as expected without extensive interaction with settings and configuration.
* The USB interface defines protocols for recovery from common errors, improving reliability over previous interfaces.
* Installing a device that relies on the USB standard requires minimal operator action. When a user plugs a device into a port on a running computer, it either entirely automatically configures using existing [device drivers](https://en.m.wikipedia.org/wiki/Device_driver), or the system prompts the user to locate a driver, which it then installs and configures automatically.

The USB standard also provides multiple benefits for hardware manufacturers and software developers, specifically in the relative ease of implementation:

* The USB standard eliminates the requirement to develop proprietary interfaces to new peripherals.
* The wide range of transfer speeds available from a USB interface suits devices ranging from keyboards and mice up to streaming video interfaces.
* A USB interface can be designed to provide the best available [latency](https://en.m.wikipedia.org/wiki/Latency_(engineering)) for time-critical functions or can be set up to do background transfers of bulk data with little impact on system resources.
* The USB interface is generalized with no signal lines dedicated to only one function of one device.

**HDMI:**

HDMI (High-Definition Multimedia Interface)is a [proprietary](https://en.m.wikipedia.org/wiki/Proprietary_hardware) audio/video [interface](https://en.m.wikipedia.org/wiki/Interface_(computing)) for transmitting [uncompressed video](https://en.m.wikipedia.org/wiki/Uncompressed_video) data and compressed or uncompressed [digital audio](https://en.m.wikipedia.org/wiki/Digital_audio) data from an HDMI-compliant source device, such as a [display controller](https://en.m.wikipedia.org/wiki/Display_controller), to a compatible [computer monitor](https://en.m.wikipedia.org/wiki/Visual_display_unit), [video projector](https://en.m.wikipedia.org/wiki/Video_projector), [digital television](https://en.m.wikipedia.org/wiki/Digital_television), or [digital audio](https://en.m.wikipedia.org/wiki/Digital_audio) device. HDMI is a digital replacement for [analog video](https://en.m.wikipedia.org/wiki/Analog_video)  standards. HDMI implements the [EIA/CEA-861](https://en.m.wikipedia.org/wiki/CEA-861) standards, which define video formats and waveforms, transport of compressed and uncompressed [LPCM](https://en.m.wikipedia.org/wiki/Linear_pulse-code_modulation) audio, auxiliary data, and implementations of the [VESA EDID](https://en.m.wikipedia.org/wiki/Extended_display_identification_data). CEA-861 signals carried by HDMI are electrically compatible with the CEA-861 signals used by the [Digital Visual Interface](https://en.m.wikipedia.org/wiki/Digital_Visual_Interface) (DVI). No signal conversion is necessary, nor is there a loss of video quality when a DVI-to-HDMI adapter is used. The [CEC](https://en.m.wikipedia.org/wiki/Consumer_Electronics_Control) (Consumer Electronics Control) capability allows HDMI devices to control each other when necessary and allows the user to operate multiple devices with one handheld [remote control](https://en.m.wikipedia.org/wiki/Remote_control) device. Several versions of HDMI have been developed and deployed since the initial release of the technology, but all use the same cable and connector. Other than improved audio and video capacity, performance, resolution and colour spaces, newer versions have optional advanced features such as [3D](https://en.m.wikipedia.org/wiki/3D_television), [Ethernet](https://en.m.wikipedia.org/wiki/Ethernet) data connection, and CEC (Consumer Electronics Control) extensions. Production of consumer HDMI products started in late 2003. In Europe, either DVI-[HDCP](https://en.m.wikipedia.org/wiki/High-bandwidth_Digital_Content_Protection) or HDMI is included in the [HD ready](https://en.m.wikipedia.org/wiki/HD_ready) in-store labelling specification for TV sets for HDTV, formulated by [EICTA](https://en.m.wikipedia.org/wiki/European_Information,_Communications_and_Consumer_Electronics_Technology_Industry_Associations) with [SES Astra](https://en.m.wikipedia.org/wiki/SES_Astra) in 2005. HDMI began to appear on [consumer](https://en.m.wikipedia.org/wiki/Consumer) [HDTVs](https://en.m.wikipedia.org/wiki/HDTV) in 2004 and [camcorders](https://en.m.wikipedia.org/wiki/Camcorder) and [digital still cameras](https://en.m.wikipedia.org/wiki/Digital_still_camera) in 2006. As of January 6, 2015 (twelve years after the release of the first HDMI specification), over 4 billion HDMI devices have been sold.

**PI CAMERA**

The **Pi camera module** is a portable light weight camera that supports Raspberry Pi. It communicates with Pi using the MIPI camera serial interface protocol. It is normally used in image processing, machine learning or in surveillance projects. It is commonly used in surveillance drones since the payload of camera is very less. Apart from these modules Pi can also use normal USB webcams that are used along with computer. All current models of Raspberry Pi have a port for connecting the Camera Module. The Raspberry Pi Camera Module V1 is based on Omni Vision OV5647 camera sensor, which has 5 Megapixels resolution. The features of Raspberry Pi are:

* + - 5MP colour camera module without microphone for [Raspberry Pi](https://components101.com/microcontrollers/raspberry-pi-3-pinout-features-datasheet)
    - Supports both Raspberry Pi Model A and Model B
    - MIPI Camera serial interface
    - Omni vision 5647 Camera Module
    - Resolution: 2592 \* 1944
    - Supports: 1080p, 720p and 480p
    - Light weight and portable (3g only)

The Raspberry Pi Camera Board is a custom designed add-on module for Raspberry Pi hardware. It attaches to Raspberry Pi hardware through a custom CSI interface. The sensor has 5 megapixel native resolution in still capture mode. In video mode it supports capture resolutions up to 1080p at 30 frames per second.

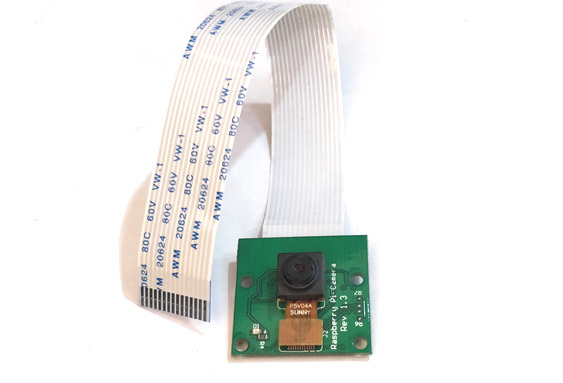


Fig 2.3. Raspberry Pi Camera

**2.2.2 Software Requirements**

* Raspbian OS
* Python

**Raspbian OS:**

Raspbian is a Debian-based computer operating system for Raspberry Pi. There are several versions of Raspbian including Raspbian Stretch and Raspbian Jessie. Since 2015 it has been officially provided by the Raspberry Pi Foundation as the primary operating system for the family of Raspberry Pi single-board computers. Raspbian was created by Mike Thompson and Peter Green as an independent project. The initial build was completed in June 2012. The operating system is still under active development. Raspbian is highly optimized for the Raspberry Pi line's low-performance ARM CPUs. Raspbian uses PIXEL, Pi Improved Window’s Environment, Lightweight as its main desktop environment as of the latest update. It is composed of a modified LXDE desktop environment and the Open box stacking window manager with a new theme and few other changes. The distribution is shipped with a copy of computer algebra program Mathematics and a version of Minecraft called Minecraft Pi as well as a lightweight version of Chromium as of the latest version.

**Raspbian Stretch with Desktop**

* This is a version of Debian Linux, complete with graphical desktop (GUI)
* Easier for beginners to navigate
* Uses more Raspberry Pi’s resources
* More software applications
* Python and Thone IDE preinstalled

**Raspbian Stretch Lite**

* This is a version of Debian Linux, with no graphical desktop (GUI) and is command line driven
* Lite is a minimal version of the Raspbian image for the Raspberry Pi
* X-server, and its components, are not installed
* Less software installed and fewer modules will load with the kernel
* Operating system using a lot less of the Raspberry Pi’s resources
* Uses less power (current consumption)
* Perform a little faster for very specific tasks
* Ideal where you need a Pi and never need to use a monitor

**Python:**

IDLE is Python’s Integrated Development and Learning Environment. IDLE has been criticized for various usability issues, including losing focus, lack of copying to clipboard feature, lack of line numbering options, and general user interface design; it has been called a "disposable" IDE, because users frequently move on to a more advanced IDE as they gain experience.Author Guido van Rossum says IDLE stands for "Integrated Development Environment", and since van Rossum named the language Python partly to honour British comedy group Monty Python, the name IDLE was probably also chosen partly to honour Eric Idle, one of Monty Python's founding members

IDLE has the following features:

* Coded in 100% pure Python, using the tkinter GUI toolkit.
* Cross-platform: works mostly the same on Windows, UNIX, and Mac OS X.
* Python shell window (interactive interpreter) with colorizing of code input, output, and error messages.
* Multi-window text editor with multiple undo, Python colorizing, smart indent, call tips, auto completion, and other features.
* Search within any window, replace within editor windows, and search through multiple files (grep).
* Debugger with persistent breakpoints, stepping, and viewing of global and local namespaces.
* Configuration, browsers, and other dialogs.

**Menus:**

IDLE has two main window types, the Shell window and the Editor window. It is possible to have multiple editor windows simultaneously. Output windows, such as used for Edit / Find in Files, are a subtype of edit window. They currently have the same top menu as Editor Windows but a different default title and context menu. IDLE’s menus dynamically change based on which window is currently selected. Each menu documented below indicates which window type it is associated with.

Python's developers strive to avoid premature optimization, and reject patches to non-critical parts of the Python reference implementation that would offer marginal increases in speed at the cost of clarity. When speed is important, a Python programmer can move time-critical functions to extension modules written in languages such as C, or use PyPy, a just-in-time compiler. Cython is also available, which translates a Python script into C and makes direct C-level API calls into the Python interpreter. An important goal of Python's developers is keeping it fun to use. This is reflected in the language's name—a tribute to the British comedy group Monty Python—and in occasionally playful approaches to tutorials and reference materials, such as examples that refer to spam and eggs (from a famous Monty Python sketch) instead of the standard foo and bar. A common neologism in the Python community is pythonic, which can have a wide range of meanings related to program style. To say that code is pythonic is to say that it uses Python idioms well, that it is natural or shows fluency in the language, and that it conforms to Python's minimalist philosophy and emphasis on readability. In contrast, code that is difficult to understand or reads like a rough transcription from another programming language is called unpythonic. Users and admirers of Python, especially those considered knowledgeable or experienced, are often referred to as Pythonistas.

**NumPy**:

NumPy is a Python library used for working with arrays. It also has functions for working in domain of linear algebra, Fourier transform, and matrices. NumPy was created in 2005 by Travis Oliphant. It is an open source project and you can use it freely. NumPy stands for Numerical Python. NumPy is a library for the Python programming language, adding support for large, multidimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

**SciPy**

SciPy is a free and open-source Python library used for scientific computing and technical computing. SciPy contains modules for optimization, linear algebra, integration, interpolation, special functions, signal and image processing, and other tasks common in science and engineering. It is built on the NumPy extension and allows the user to manipulate and visualize data with a wide range of high-level command.

**Matplotlib**

Matplotlib is a python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. Matplotlib tries to make easy things easy and hard things possible. You can generate plots, histograms, power spectra, bar charts, error charts, scatterplots, etc., with just a few lines of code.

**Scikit-Learn**

Scikit-learn (formerly scikits.learn and also known as sklearn) a free software machine learning library for the Python programming language which provide stable Python modules for machine learning and data mining. It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

**TensorFlow**

TensorFlow is an end-to-end open source platform for machine learning. TensorFlow is a rich system for managing all aspects of a machine learning system; however, this class focuses on using a particular TensorFlow API to develop and train machine learning models. TensorFlow APIs are arranged hierarchically, with the high-level APIs built on the low-level APIs. Machine learning researchers use the low-level APIs to create and explore new machine learning algorithms. In this class, you will use a high-level API named tf.keras to define and train machine learning models and to make predictions. tf.keras is the TensorFlow variant of the open-source [Keras](https://keras.io/) API.

**Keras**

Keras is one of the leading high-level neural networks APIs. It is written in Python and supports multiple back-end neural network computation engines. Keras was created to be user friendly, modular, easy to extend, and to work with Python. The API was “designed for human beings, not machines,” and “follows best practices for reducing cognitive load.” Neural layers, cost functions, optimizers, initialization schemes, activation functions, and regularization schemes are all standalone modules that you can combine to create new models. New modules are simple to add, as new classes and functions. Models are defined in Python code, not separate model configuration file.

**2.2.3 ALGORITHMS**

* Support Vector Machine
* K-means Clustering Algorithm

**Support Vector Machine Algorithm**

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane. SVM chooses the extreme points/vectors that help in creating the hyperplane. Hyperplanes are decision boundaries that help classify the data points. Data points falling on either side of the hyper plane can be attributed to different classes. Also, the dimension of the hyper plane depends upon the number of features. If the number of input features is 2, then the hyper plane is just a line. If the number of input features is 3, then the hyper plane becomes a two-dimensional plane. It becomes difficult to imagine when the number of features exceeds 3. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine.SVM can be of two types**:**

* **Linear SVM:** Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier is used called as Linear SVM classifier.
* **Non-linear SVM:** Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier used is called as Non-linear SVM classifier.



**K-means Clustering Algorithm**

K-Means Clustering is an [Unsupervised Learning algorithm](https://www.javatpoint.com/unsupervised-machine-learning), which groups the unlabelled dataset into different clusters. It is an iterative algorithm that divides the unlabelled dataset into k different clusters in such a way that each dataset belongs only one group that has similar properties. It allows us to cluster the data into different groups and a convenient way to discover the categories of groups in the unlabelled dataset on its own without the need for any training. It is a centroid-based algorithm, where each cluster is associated with a centroid. The main aim of this algorithm is to minimize the sum of distances between the data point and their corresponding clusters. The algorithm takes the unlabelled dataset as input, divides the dataset into k-number of clusters, and repeats the process until it does not find the best clusters. The value of k should be predetermined in this algorithm.

The k-means [clustering](https://www.javatpoint.com/clustering-in-machine-learning) algorithm mainly performs two tasks:

* Determines the best value for K centre points or centroids by an iterative process.
* Assigns each data point to its closest k-centre. Those data points which are near to the particular k-centre, create a cluster.

Hence each cluster has data points with some commonalities, and it is away from other clusters.



The steps to be followed for the implementation of K-means Clustering algorithm are given below:

* Data Pre-processing
* Finding the optimal number of clusters using the elbow method
* Training the K-means algorithm on the training dataset
* Visualizing the clusters

**CHAPTER 3**

**SYSTEM DESIGN**

Systems design is the process of defining elements of a system like modules, architecture, components and their interfaces and data for a system based on the specified requirements. It is the process of defining, developing and designing systems which satisfies the specific needs and requirements of a business or organization. The purpose of the System Design process is to provide sufficient detailed data and information about the system and its system elements to enable the implementation consistent with architectural entities as defined in models and views of the system architecture.

**3.1 ARCHITECTURE DIAGRAM**

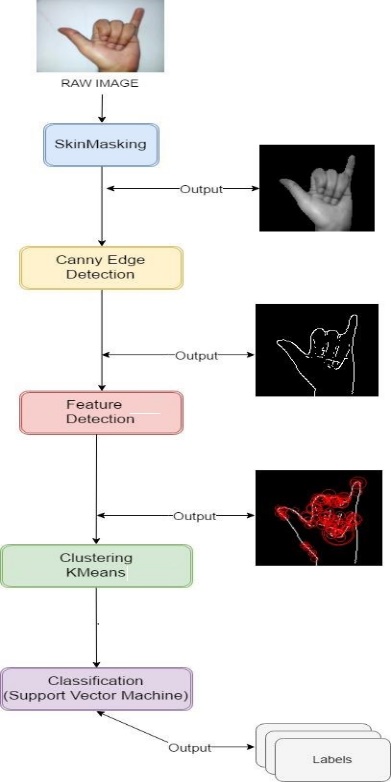
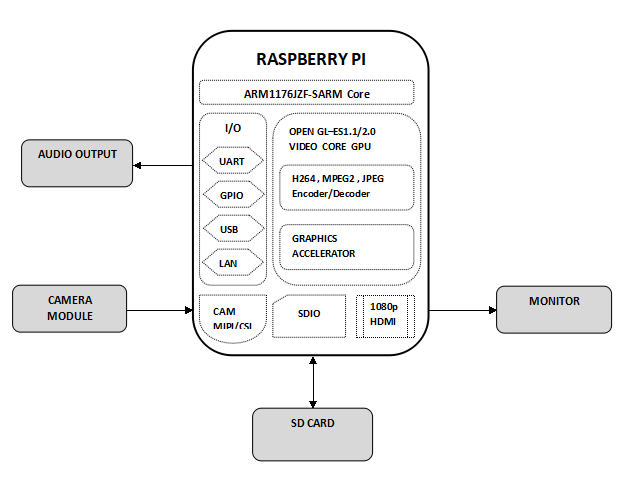


Fig 3.1: Architecture Diagram of Sign Language Translator

The architectural diagram represents the overall process included in the system. At first the raw image captured by the Pi Camera is processed for finding skin-colored image regions from images using Skin masking. The second step is canny edge detection which is a multi-step algorithm that can detect edges with noise supressed at the same time, and then the features will be detected and are extracted from the image. The next step is clustering of required datasets using k-means clustering algorithm and then classifies those data using support vector machine algorithm. This data models are then trained using machine learning techniques and frameworks so that it can be integrated into the Raspberry Pi for our future output.

**3.2 MODULES DESCRIPTION**

**3.2.1 PREPROCESSING**

Procedure of transforming raw data into a format that is more suitable for further analysis and interpretable for the user.  When it comes to creating a Machine Learning model, data pre-processing is the first step marking the initiation of the process. Typically, real-world data is incomplete, inconsistent, inaccurate (contains errors or outliers), and often lacks specific attribute values/trends. This is where data pre-processing enters the scenario – it helps to clean, format, and organize the raw data, thereby making it ready-to-go for Machine Learning models. A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data pre-processing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

It involves below steps:

* Getting the dataset
* Importing libraries
* Importing datasets
* Finding Missing Data
* Encoding Categorical Data
* Splitting dataset into training and test set
* Feature scaling

1. Acquire the dataset

To build and develop Machine Learning models, you must first acquire the relevant dataset. This dataset will be comprised of data gathered from multiple and disparate sources which are then combined in a proper format to form a dataset. Dataset formats differ according to use cases. For instance, a business dataset will be entirely different from a medical dataset. While a business dataset will contain relevant industry and business data, a medical dataset will include healthcare-related data.

### Import all the crucial libraries

### 

### The predefined Python libraries can perform specific data pre-processing jobs. The three core Python libraries used for this data pre-processing in Machine Learning are: Numpy, Pandas, and Matplotlib.

### **Import the dataset**

### 

### In this step, you need to import the dataset/s that you have gathered for the ML project at hand. However, before you can import the dataset/s, you must set the current directory as the working directory.

### Identifying and handling the missing values

### In data pre-processing, it is pivotal to identify and correctly handle the missing values, failing to do this, you might draw inaccurate and faulty conclusions and inferences from the data. Needless to say, this will hamper your ML project. Basically, there are two ways to handle missing data: Deleting a particular row and calculating the particular mean.

### **Encoding the categorical data**

### Categorical data refers to the information that has specific categories within the dataset. In the dataset cited above, there are two categorical variables – country and purchased. Machine Learning models are primarily based on mathematical equations. Thus, you can intuitively understand that keeping the categorical data in the equation will cause certain issues since you would only need numbers in the equations.

### Splitting the dataset

### Every dataset for Machine Learning model must be split into two separate sets – training set and test set. Training set denotes the subset of a dataset that is used for training the machine learning model. Here, you are already aware of the output. A test set, on the other hand, is the subset of the dataset that is used for testing the machine learning model. The ML model uses the test set to predict outcomes.

### Feature scaling

### Feature scaling marks the end of the**data pre-processing in Machine Learning.** It is a method to standardize the independent variables of a dataset within a specific range. In other words, feature scaling limits the range of variables so that you can compare them on common grounds.

* + 1. **IMAGE PROCESSING**

 Image processing is a method to perform some operations on an image, order to get an enhanced image or to extract some useful information from it. Image processing and analysis often require fixed sequences of local operations to be performed at each pixel of an image. Such sequences of operations can be performed in parallel using a pipeline of processors, each operating on the output of the preceding one. The first processor performs the first operation on the image, pixel by pixel. As soon as the first pixel and its neighbours have been processed, the second processor begins to perform the second operation, and so on. Since each processor has available to it the output value of its operation at every pixel, it can also compute statistics of these values, if desired.

**3.2.3 FEATURE EXTRACTION**

Feature extraction is a part of the dimensionality reduction process, in which, an initial set of the raw data is divided and reduced to more manageable groups. So when you want to process it will be easier. The most important characteristic of these large data sets is that they have a large number of variables. These variables require a lot of computing resources to process them. So Feature extraction helps to get the best feature from those big data sets by select and combine variables into features, thus, effectively reducing the amount of data. These features are easy to process, but still able to describe the actual data set with the accuracy and originality. The technique of extracting the features is useful when you have a large data set and need to reduce the number of resources without losing any important or relevant information. Feature extraction helps to reduce the amount of redundant data from the data set. In the end, the reduction of the data helps to build the model with less machine’s efforts and also increase the speed of learning and generalization steps in the machine learning process. Feature extraction increases the accuracy of learned models by extracting features from the input data. This phase of the general framework reduces the dimensionality of data by removing the redundant data. Of course, it increases training and inference speed. The methods of feature extraction obtain new generated features by doing the combinations and transformations of the original feature set [30]. The type of features that can be extracted from the medical images is color, shape, texture or due to the pixel value. Some of the medical images, such as X-ray images, do not contain any color information and have few objects. However, the texture and shape feature extraction techniques are very suitable. The examples of the texture feature extraction techniques are gray level co-occurrence matrices and LBP. On the other hand, the examples of the shape feature extraction techniques are the canny edge and Laplacian operators.

**3.2.4 SIGN RECOGNITION**

Sign consists of three main parts: Manual features involving gestures made with the hands (employing hand shape and motion to convey meaning), Non-manual features such as facial expressions or body posture, which can both form part of a sign or modify the meaning of a manual sign, and Finger spelling, where words are spelt out gestural in the local verbal language. Naturally this is an oversimplification, Sign language is as complex as any spoken language, each sign language has many thousands of signs, each differing from the next by minor changes in hand shape, motion, position, non-manual features or context. Since signed languages evolved alongside spoken languages, they do not mimic their counterparts. E.g. British Sign Language (BSL) loosely follows the sequence of time-line, location, subject, object, verb and question. It is characterised by topic-comment structure where a topic or scene is set up and then commented on. It uses its own syntax which makes use of both manual and non-manual features, simultaneous and sequential patterning and spatial as well as linear arrangement. Signs can be described at the sub-unit level using phonemes. These encode different elements of a sign. Unlike speech they do not have to occur sequentially, but can be combined in parallel to describe a sign. Studies of American Sign Language (ASL) by Liddell and Johnson model sign language on the movement-hold system. Signs are broken into sections where an aspect is changing and sections where a state is held steady. This is in contrast to the work of Stoke where different components of the sign are described in different channels; the motion made by the hands, the place at which the sign is performed, the hand shapes, the relative arrangement of the hands and finally the orientation of both the hands and fingers to explain the plane in which the hands sit. Both of these models are valid in their own right and yet they encode different aspects of sign. Within SLR both the movement hold, sequential information from Liddell and Johnson and the parallel forms of Stoke are desirable annotations. Below are described a small subset of the constructs of sign language.

(a) Adverbs modifying verbs; signers would not use two signs for ‘run quickly’ they would modify the sign for run by speeding it up.

(b) Non-manual features (NMFs); facial expressions and body posture are key in determining the meaning of sentences, e.g. eyebrow position can determine the question type. Some signs are distinguishable only by lip shape, as they share a common manual sign.

(c) Placement; pronouns like ’he’, ’she’ or ’it’ do not have their own sign, instead the referent is described and allocated a position in the signing space. Future references point to the position, and relationships can be described by pointing at more than one referent.

(d) Classifiers; these are hand shapes which are used to represent classes of objects, they are used when previously described items interact. E.g. to distinguish between a person chasing a dog and vice versa.

(e) Directional verbs; these happen between the signer and referent(s), the direction of motion indicates the direction of the verb. Good examples of directional verbs are ‘give’ and ‘phone’. The direction of the verb implicitly conveys which nouns are the subject and object.

(f) Positional Signs; where a sign acts on the part of the body descriptively. e.g. ‘bruise’ or ‘tattoo’.

(g) Body Shift; represented by twisting the shoulders and gaze, often used to indicate role-shifting when relating a dialogue.

(h) Iconicity; when a sign imitates the thing it represents, it can be altered to give an appropriate representation. E.g. the sign for getting out of bed can be altered between leaping out of bed with energy to a recumbent who is reluctant to rise.

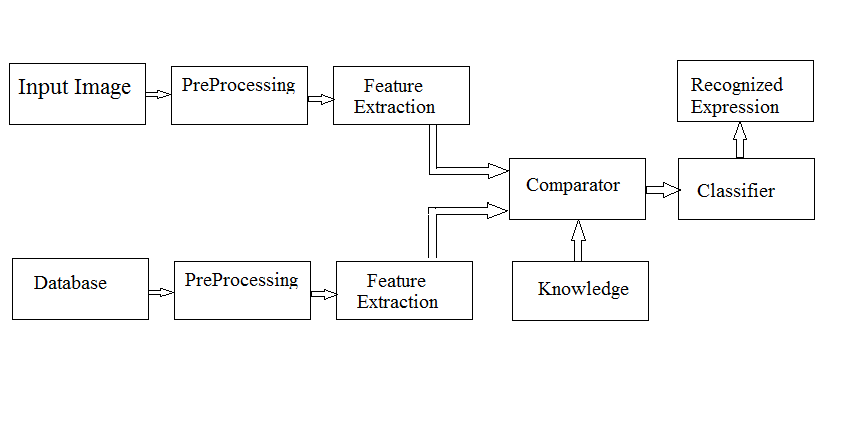
(i) Finger spelling; where a sign is not known, either by the signer or the recipient, the local spoken word for the sign can be spelt explicitly by finger

**CHAPTER 4**

**SYSTEM DEVELOPMENT**

**MODULE EXECUTION:**

# In the stage of pre-processing, input image is made compatible with the developed code. For instance, conversion of the input colored image into the gray scale values and low pass filtering to avoid the effect of blur and to reduce the noise if any from the input image to get the exact expected results. For Feature extraction certain features of the face to be compared with the existing database are extracted. The database consists of the training or the reference images. This database could be one of the standard databases or the self-prepared database. The functions of pre-processing and feature extraction are also operated on the database images to obtain the average values for the sake of comparison with the values obtained by the input images. Knowledge consists of the required information for the classifier to compare with the training and testing images like some mean values or some thresholds or some standards to be followed while making the comparison. The classified data will be recognized using sign detection and Sign recognition modules.



**SYSTEM TESTING**

Software testing is an investigation conducted to provide stakeholders with information about the quality of the product or service under test. Software testing can also provide an objective, independent view of the software to allow the business to appreciate and understand the risks of software implementation. Software testing can be stated as the process of validating and verifying that a computer program/application/product: b meets the requirements that guided its design and development, works as expected, can be implemented with the same characteristics, and satisfies the needs of stakeholders. Software testing is a critical element of quality assurance and represents the ultimate previews of specification, design and coding. Testing represents and interesting anomaly for the software. Doing the earlier definition and development phase it was attempted to build software from an abstract concept to a tangible implementation.

**5.1 THE BOX APPROACH**

Software testing methods are traditionally divided into white- and black-box testing. These two approaches are used to describe the point of view that a test engineer takes when designing test cases.

**5.1.1 WHITE BOX TESTING**

White-box testing (also known as clear box testing, glass box testing, and transparent box testing and structural testing) tests internal structures or workings of a program, as opposed to the functionality exposed to the end user. In white-box testing an internal perspective of the system, as well as programming skills, are used to design test cases.

**5.1.2 BLACK BOX TESTING**

Black-box testing treats the software as a "black box", examining functionality without any knowledge of internal implementation. The testers are only aware of what the software is supposed to do, not how it does it.

**5.2 FUNCTIONAL TESTING**

Functional testing refers to activities that verify a specific action or function of the code. These are usually found in the code requirements documentation, although some development methodologies work from use cases or user stories. Functional tests tend to answer the question of "can the user do this" or "does this particular feature work."

**5.3 NON-FUNCTIONAL TESTING**

Non-functional testing refers to aspects of the software that may not be related to a specific function or user action, such as scalability or other performance, behaviour under certain constraints, or security Non-functional requirements tend to be those that reflect the quality of the product, particularly in the context of the suitability perspective of its users.

**5.4 TOP-DOWN AND BOTTOM-UP**

Testing is an approach to integrated testing where the lowest level components (modules, procedures, and functions) are tested first, then integrated and used to facilitate the testing of higher level components. After the integration testing of lower level integrated modules, the next level of modules will be formed and can be used for integration testing. The process is repeated until the components at the top of the hierarchy are tested.

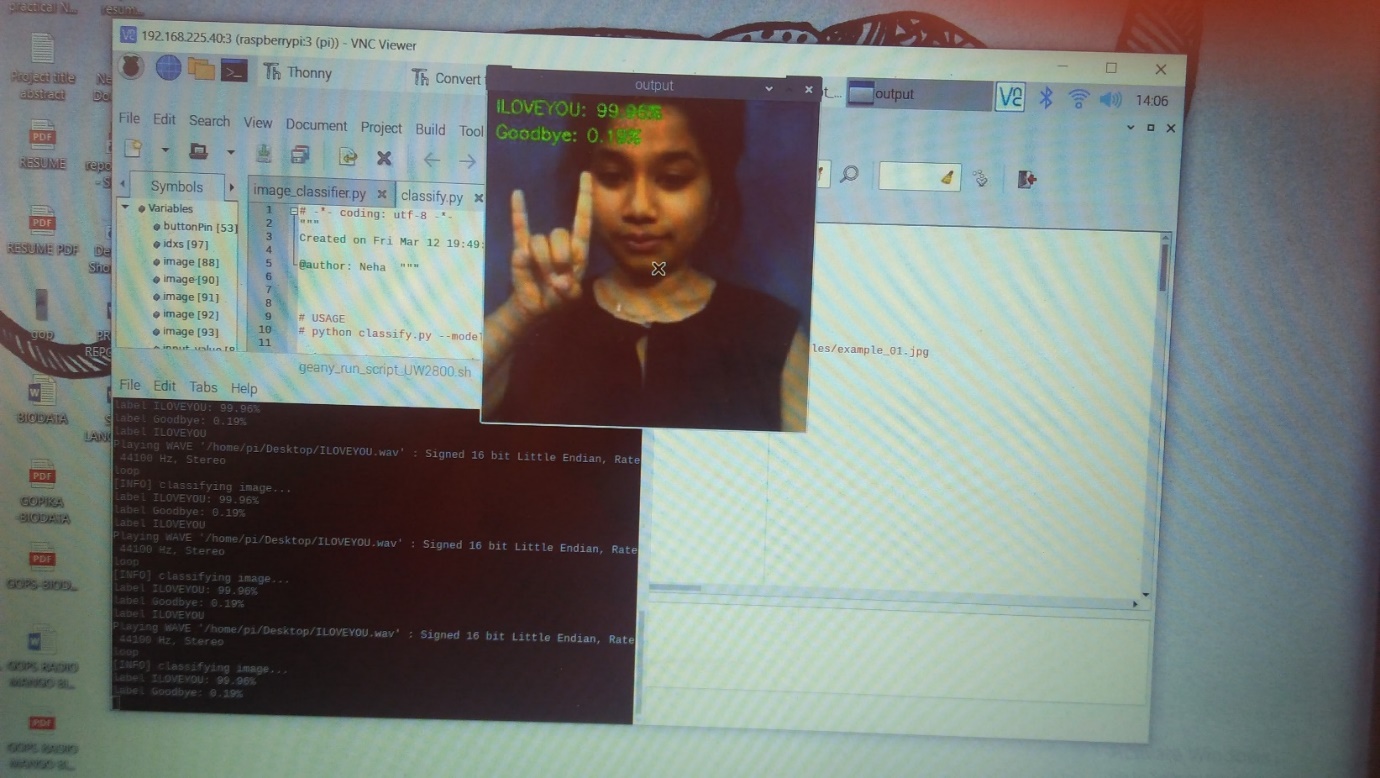
**5.13 OUTPUT TESTING**

After performing the validation testing. The next step is output testing of the proposed system since no system could be useful if it does not produce the required output generated or considered into two ways; one is no screen and another is printed format. The output format on the screen is found to be correct as the format was designed in the system design phase according to the user needs. For the hard copy also, the output comes out as the specified requirements by the user. Hence output testing does not result in any correction in the system.

**CHAPTER 5**

**RESULT AND DISCUSSION**

Thus we proposed a system called Sign language to audio translator for deaf mute people using Raspberry Pi. Thus the required datasets are trained in the system using Support Vector Machine Algorithm, K-means Clustering Algorithm. Support Vector Machine is a supervised machine learning algorithm capable of performing classification, regression and even outlier detection. K-Means Clustering is an Unsupervised Learning Algorithm, which groups the unlabeled dataset into different clusters. The input will be compared with the dataset which have been feed in the system and then translate into text and give the audio output via Raspberry pi.



**CHAPTER 6**

**CONCLUSION AND FUTURE ENHANCEMENT**

Sign language translator is very useful in various areas. In schools, colleges, hospitals, universities, airports, courts anywhere anyone can use this system for understanding the sign language to communicate. It makes communication between a normal hearing person and a hard to hearing person easier.

Since deaf people are usually deprived of normal communication with other people, they have to rely on an interpreter or some visual communication. The system can be extended to incorporate the knowledge of facial expressions and body language too so that there is a complete understanding of the context and tone of the input speech. A mobile and web based version of the application will increase the reach to more people. Integrating hand gesture recognition system using computer vision for establishing 2-way communication system.

**REFERENCES**

[1] Siming He, “Research of a Sign Language Translation System Based on Deep Learning”, 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), Ridley College, St. Catharines, Canada

[2] Kanchan Dabre and Surekha Dholay, “Machine Learning Model for Sign Language Interpretation using Webcam Images”, 2014 International Conference on Circuit Systems, Communication and Information Technology Applications.

[3] Ankita Harkude, Sarika Namade, Anita Morey, “Audio to Sign Language Translation for Deaf People”, International Journal of Engineering and Innovative Technology (IJEIT) Volume 9, Issue 10, April 2020.

[4] Eleni Efthimiou , Stavroula , Christian Voglier, “Sign Language Recognition, Generation , and Modelling: A Research Effort with Applications in Deaf Communication”, Universal Access in HCI, Part I, HCII 2009, LNCS 5614, pp. 21–30, 2009, University Hamburg.

[5] Taner Arsan and Oguz Ulgel, “Sign Language Converter”, International Journal of Computer Science & Engineering Survey (IJCSES) Vol.6, No.4, August 2015, Kadir Has University, Turkey.

[6] Gunashekaran K, Manikandan R R , “Sign Language to Speech Translation System using PIC Microcontroller”, International Journal of Engineering and Technology Volume 5 No 2 April –May 2013.

[7] Abey Abraham and Rohini.V, “Real time conversion of sign language to speech and prediction of gestures using Artificial neural network”, 8th International Conference on Advances in Computing and Communication (ICACC-2018).

[8] Neha Poddou, Shrushti Rao,Shruti Sawant, “Study of Sign Language Translation using Gesture Recognition”,International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2, February 2015.

[9] Nithyakalyani K, S Ramkumar, K Manikandan, “Design and Implementation of Sign Language Translator Using Micro touch Sensor”, International Journal of Scientific & Technology Research Volume 9, issue 01, January 2020.

[10] Oi Mean Foong, Tang Jung Low, Wai Wan La, “ Voice to Sign Language Translation System for Malaysian Deaf People”, IVIC 2009, LNCS 5857, pp. 868–876, 2009. © Springer-Verlag Berlin Heidelberg 2009.