Talking Hands & D and D (Dump & Deaf) Assistant

Ambily M[1], Shejina N M[2]

[1]PG Student, Dept. of Computer Science and Engineering., IES College of Engineering and Technology, Kerala Technological University, Chittilappilly, Thrissur
[2]Assistant Professor, Dept. of Computer Science and Engineering., IES College of Engineering and Technology, Kerala Technological University, Chittilappilly, Thrissur

ABSTRACT:
The largest issue facing the deaf and dumb community these days is communication with both the general public and fellow deaf and dumb people. Ordinary individuals can learn new information from spoken language, everyday conversations, and surrounding sounds. Individuals who are hard of hearing or deaf do not enjoy this luxury. These people will gain from this approach since it provides a means of communication. The term "assistive technology" describes gadgets that are rehabilitative, adaptive, and made for people with disabilities. This project aims to develop a system that can recognize sign language, bridging the communication gap by enabling those with speech impairments to interact with others who do not. Hand gestures are important because, in contrast to other techniques, they enable people to express their ideas rapidly. The current work aims to recognize signs using an image processing-based gesture recognition module and convert the corresponding speech to text using a text-to-speech synthesizer.

KEYWORDS: Gloves; Assistive Technology; Sign Language; Image Processing; ASL

I. Introduction

The "Talking Hands" concept uses gadgets like flex sensors and microcontrollers. With the Hand Glove's connection to flex sensors and a microprocessor, the device can convert human movements into audio or text. The converted sign language is displayed as text on an LCD and played back through speakers. D & D Assistant - The purpose of this project is to supply DEAF with assistive technology, which will employ state-of-the-art technologies to enhance the everyday quality of life for people who have hearing impairments. The speech is converted into text and shown on an LCD. The initial phase of an automatic translation system that can translate visual speech—which deaf people use—to text is shown in this work. Speech-to-text conversion is done on a laptop using Python programming. On a laptop, the picture processing and predicted sign language are presented as text and audio.

II. Related work

The purpose of this work is to present an operation armature for a hand glove that uses Bluetooth, GSM-CDMA, and Internet modules to record the gestures performed by individuals who are speech- and hail-bloodied, restate those movements into meaningful text, and shoot the text to distant locales. It features a graphic stoner interface that shows all of the data transferred and entered between two druggies, trans-receiver modules for data transmission and event, an AVR 2560 microcontroller for gesture processing, and five flex, three contact, and one three-axis accelerometer detectors that act as input channels [1]. This exploration proposes a wireless hand gesture discovery glove for real-time Taiwanese sign language restatement. Flex and inertial detectors are integrated into the glove to enable demarcation between colorful hand gestures. These three parameters — cutlet posture, win exposure, and hand stir — are pivotal in the Taiwanese Language and must be honored. The suggested system uses the stirring line attained by the gyroscope, the win exposure attained by the G-detector, and the cutlet flexion postures attained by the flex detectors as input signals. The input signals will be recorded and reviewed regularly to determine whether or not they're licit sign language gestures. The tried signal is considered licit if it exceeds a destined number of timepiece cycles. It's also transferred to a mobile device via Bluetooth for voice restatement and gesture demarcation. The delicacy of gesture recognition with the suggested armature and algorithm is relatively good. trials show that it's possible to reach a delicacy rate of over 94 on perceptivity for gesture recognition, which supports the superiority of the suggested design [2]. In this paper, according to how the flex detectors identify a certain bending angle, the exploration composition" stoner- acquainted Cutlet- Gesture Glove Controller with Hand Movement Visualization Using Flex Detectors and Digital Accelerometer" established a control medium employing recorded countries. The countries have a defined range because they're grounded on a destined threshold. With the help of five flex detectors and the Y-axis of a three-axis accelerometer, 64 possible movements were reduced to 256 countries by the current configuration, which doesn't include the fresh contact detector countries. This backed in the creation of textual coequals for all 26 English rudiments [3] This study presents the conception of a smart glove that can restate sign language affairs into voice. Flex detectors and an Inertial Measurement Unit (IMU) are integrated inside the glove to describe the gesture. A brand-new state estimation fashion has been created to track a hand's movement in three confines. The feasibility of rephrasing the Indian subscribed language-to-speech affair using the prototype was examined. Despite being designed to restate sign language to speech, the glove has multiple uses in the gaming, robotics, and medical fields [5]. This exploration aims to produce and dissect a technology that can fete sign language. Flex detectors, accelerometers, and gyroscope signals make up all of the input data used by the Multilayer Perceptron, which does the recognition. A neural network instinctively is estimated with three different parameters changed a) the quantum of neurons in the middle subcaste solely; b) the literacy rate between the middle and input layers; and c) the literacy rate between the middle and affair layers. Following testing, confirmation, and training, the network achieves a megahit rate of roughly 96.1. It's suggested as volition to the accessible results presently available for deaf individuals, offering good delicacy at a lower cost than those formerly on the request [6]. In this work, They created a detector-grounded contrivance that can fete hand gestures used to represent the English ABC. We suggest that a mute person may wear this wearable device — a hand glove — and it would nearly rightly identify the 26 letters. The standard hand movements used by silent people as sign language are enforced by this suggested system. They also pay close attention to the need that the system to take the form of a wearable device with its data collection module that can collude data to the applicable sign and also restate that sign into the ABC each in one go. This task is completed without the need for a large CPU running Matlab software or a fresh processing unit. A mute existent can hold a single, specially designed device that houses our entire system [6]. In this paper, the thing is to produce a system that's affordable, reliable, and movable. Because of this, the system is affordable and can be bought by anyone because it makes use of readily available, locally-produced electronics. Detector gloves, the foundation of the technology, restate sign language into speech. The gloves included an accelerometer and flex detectors to describe the movement and exposure of both hands. Bluetooth is used to dissect and multiplex all of the detector data wirelessly. When the gestures' detector values match, the audio module plays the corresponding sound train. The system can go into its resting mode and save energy when no sign is entered. The system armature avoids complicated calculations by using a straightforward system that directly maps movements via a lookup procedure [7]. This paper will enable silent persons to fete colourful hand gestures while wearing gloves, and these gestures will be restated into applicable addresses for non-silent people. Flexible detectors are pivotal in this situation. Along with the accelerometer value of the hand's slant position relative to the ground, the resistance value change that results from the detectors' degree of curve is also covered. Through Bluetooth communication, this collected data is further handled by a microcontroller module and transferred to any stoner of a smartphone. An operation that has been developed can also be used to convert the data into text grounded on the hand shape that's honoured and will induce a voice signa l[8]. This article presents the framework that was created to support them and make it easier for them to communicate their message. Such a device requires expertise in both computer engineering and electronics. Flex sensors, an Arduino, a Bluetooth module, an accelerometer, a glove, and an Android app to show the outcomes are all included. The framework aims to lower the communication gap that exists between a normal person and a deaf or dumb person. The objective is also to identify the choices and gaps in this field to enhance the framework even more [9]. The study focuses on the indications of the Dzongkha alphabet, which is Bhutan's original tongue. This study builds a system that senses bend, orientation, and contact using an assemblage of flex sensors, a gyroscope, an accelerometer, and contact sensors. Each sign has a purpose, and the design conveys that purpose to the audience using audio that clarifies the meaning of that specific sign [10]. The development of a glove translator capable of translating American Sign Language motions into text and speech is the aim of this research. Using flex sensors and an accelerometer, the prototype can translate the alphabet, numerals one through ten, and fifty common words and phrases in three seconds. The translation's result can be viewed using an Android application. For wireless operation, the glove translator will be connected via Bluetooth. The accuracy rating of the prototype is 95%. The purpose of this research is to facilitate communication between the deaf-mute community and non-sign language users [11]. This study presents a revolutionary way to use hand gestures to connect with computers and other digital devices. The technique relies on using a glove that has been specially designed and fitted with flex sensors, which can detect finger movements and wirelessly transfer them to the device that has to be controlled. The flex sensors on the glove must be coated with polydimethylsiloxane (PDMS) to enable precise and dependable hand gesture detection. PDMS helps to dampen and cancel out any vibration caused by the user's movements. Several tests were conducted using a variety of hand gestures from users to assess the method's efficacy. The results demonstrate that cross-correlation approaches may distinguish between diverse hand motions even when some of them are intentionally designed to look similar to one another. Additionally, the method's resistance to noise and other forms of interference made it suitable for usage in real-world scenarios [12].

III. Proposed System

TALKING HANDS

1. Language Detection System

This device records the user's hand movements using a glove. The gloves' thumbs and fingers include flex sensors running the length of them. Flex sensors generate a stream of data that varies with the bend's degree. The analog outputs from the sensors are then sent to the microcontroller. It processes the signals after converting them from analog to digital. The motion is recognized, as is the textual information that goes along with it. The user needs to spend two seconds focusing on each sign and be familiar with the signs for the various alphabets. The newly released gesture must be supported by the system. These sensors are affixed to the thumb and fingers. The degree of flexion in the fingers and thumb determines the output. When converted to analog form, it produces the required voice based on the voltage variation. Mute people can interact with the general public in the required language by wearing gloves and sensors.

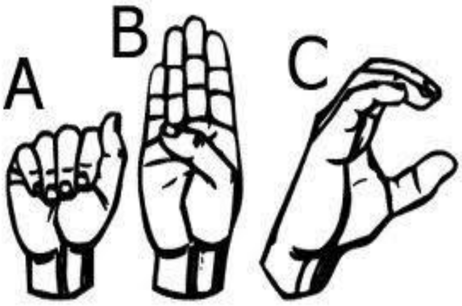

Fig 1: Basic Sign Language Gestures

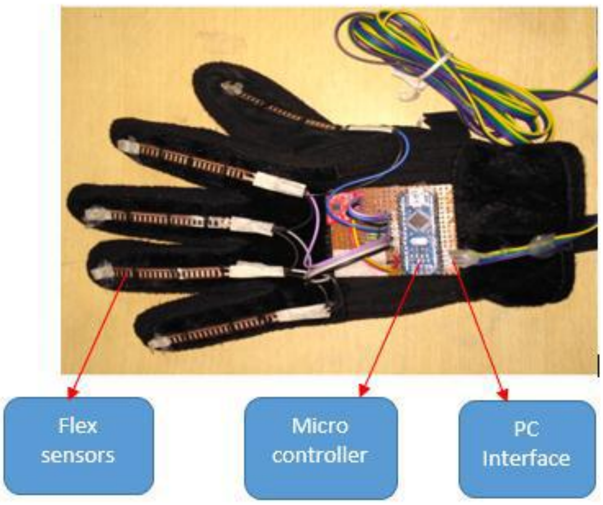

Fig 2: Experiment Glove Model

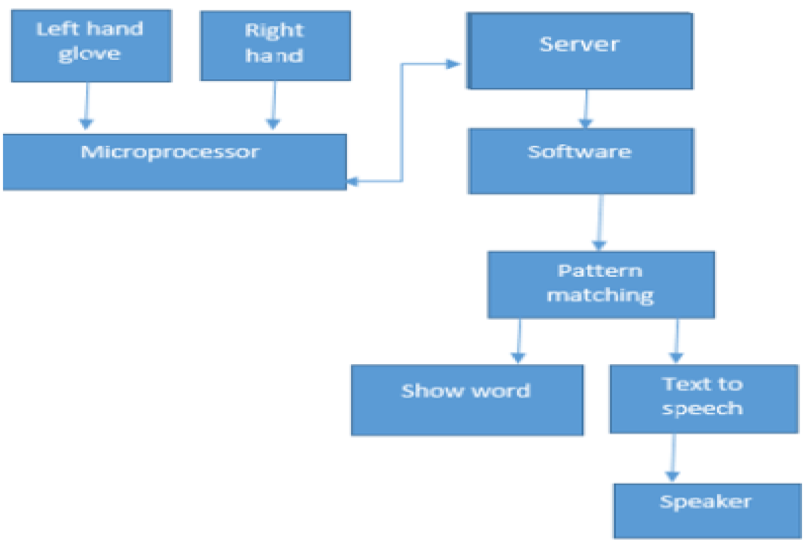

Fig 3: Block Diagram for Talking Gloves

2. Arduino Microcontroller

An Arduino microcontroller board, a simple board that enables open-source development, is used in this system. You can use it to build computers that run useful and artistic programs. This microcontroller combines all of the flex sensor values before sending the input data to the server.

Fig 4: Arduino NANO

3. Flex Sensor:

Within this system, the Flex Sensor is an essential component. Flex sensors' amount of resistance varies with their degree of bend. Resistance levels will result in voltage changes, which will give the required voice when converted to analog form. In this case, the voltage is split between two resistors linked in series in direct proportion to their resistance, according to the voltage divider principle.

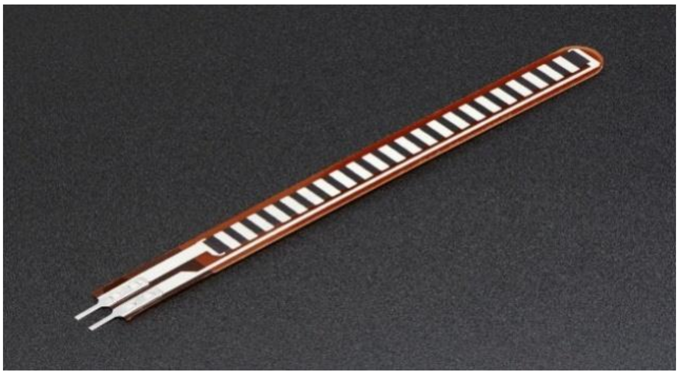

Fig 5: Flex Sensor

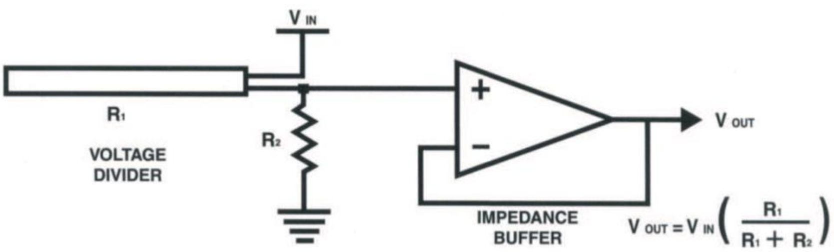

Fig 6: Circuit For Voltage Divider Principle

4. LCD Display

An LCD is a compact, electrically controlled flat screen. It is made up of one or more monochrome or colored pixels that are placed in front of a backlight or reflector and filled with liquid crystals. Its exceptionally low electric power consumption makes it a common component of battery-operated electronic devices. Each pixel in an LCD is made up of a layer of molecules organized between two transparent electrodes and two polarizing filters, each of whose axes of transmission are (mostly) perpendicular to the other. If there was no actual liquid crystal between the polarizing filters, the second (crossing) polarizer would prevent light from passing through the first filter.

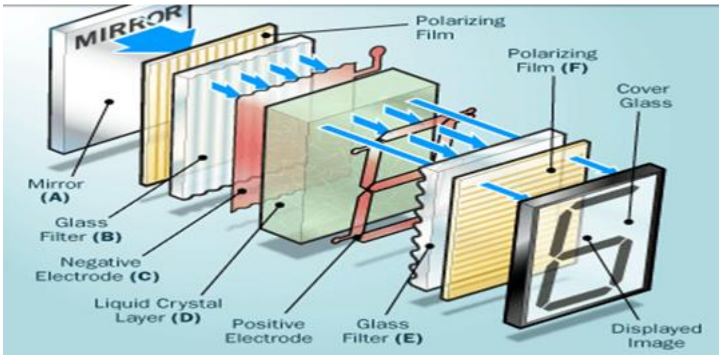

Fig 7: Working of LCD Display

5. Jhd162a:16 X 2 Alphanumeric Lcd Module

Furthermore, the JHD162A dot-matrix liquid crystal display controller and driver LSI display symbols in addition to alphanumeric and Japanese kana characters. It can be configured to run a dot-matrix liquid crystal display using a 4- or 8-bit microprocessor. Since it has all the internal parts required to drive a dot-matrix liquid crystal display, including display RAM, and a character, a minimal system can connect with this controller/driver.

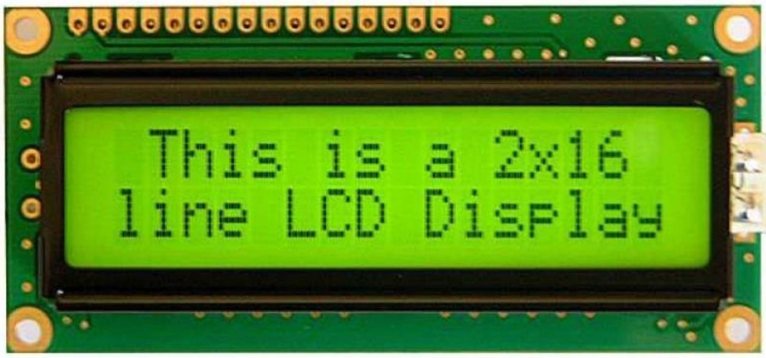

Fig 8: Jhd162a:16 X 2 Alphanumeric Lcd Module

6. Voice Recorder And Playback Module 8 Channel (Apr33a3)

The APR33A3 is an 8-channel speech recorder and audio playback board that is integrated with the APR33A series IC, a powerful audio processor, and high-performance digital-to-analog and analog-to-digital converters (ADCs and DACs). The IC is a fully integrated solution with outstanding performance, thanks to its superb combination of analog input, digital processing, and analog output features. The basic key trigger was the primary consideration when developing the APR33A series. The user can record and play one, two, four, or eight voice messages on average by flicking a switch. By altering the resistor values, one can also adjust the sampling rate. It functions effectively in systems like answering machines, toys, and leave-message systems that have simple interfaces or message length limitations.

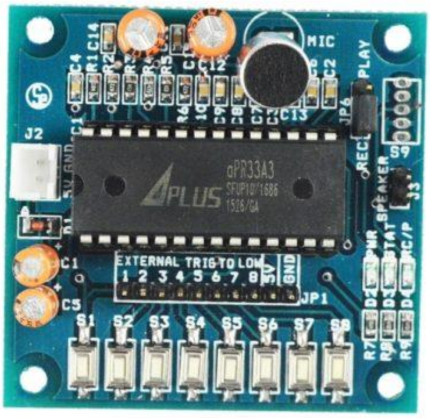

Fig 9: APR33A3

D AND D ASSISTANT (DUMP AND DEAF)

1. Speech To Text

The purpose of this project is to supply DEAF with assistive technology, which will make use of cutting-edge technologies to enhance the everyday lives of people who have hearing impairments. The speech is converted into text and shown on an LCD. The initial phase of an automatic translation system that can translate visual speech—which deaf people use—to text is shown in this work. Text-to-speech conversion is achieved with Raspberry Pi running Python programming.

2. Sign Language To Speech

A speech impairment affects both speaking and hearing. These folks use sign language to interact with one another.

It can be challenging for non-sign language speakers to communicate with people who have speech problems, even though sign language is a helpful tool for communication. The purpose of this project is to offer DUMB with assistive technology that will translate text and audio from sign language to English, allowing users of sign language to communicate more easily. As a result, this technology makes it easier to communicate by generating speech equivalents for the many signs that the DUMP uses. The application takes pictures using the computer's webcam, pre-processes them with a combinational technique, and then identifies them.

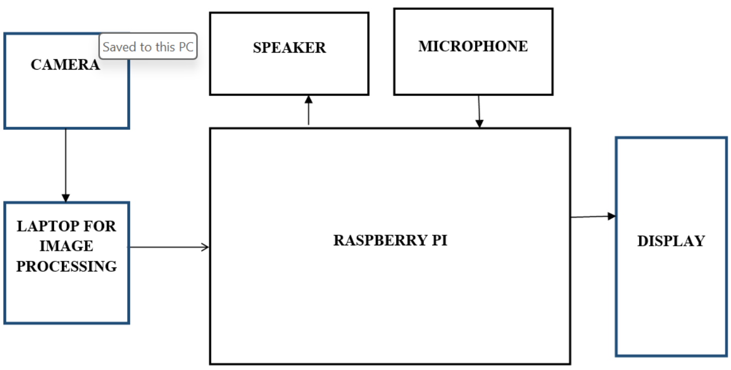

Fig 10: Block Diagram for D AND D ASSISTANT Model

3. Image Processing using CNN

The machine is trained using CNN, and images are captured using OpenCV, resulting in text output. Convolutional neural networks (CNNs) are artificial neural networks designed primarily to process input that resembles a grid, including images and videos. It employs specialized layers, including as convolutional, pooling, and fully connected layers, to automatically and hierarchically extract patterns and properties from the input data. Because CNNs are so adept at seeing spatial patterns in the data, they excel at tasks like object detection, image categorization, and image recognition. They use convolution methods to scan and extract local patterns, which enables powerful and efficient feature extraction from complicated visual data. Although several other projects have offered methods for partially recognizing signs, the goal of this project is to fully adopt American Sign Language, which consists of 26 letters. Video is captured using a camera and sent through a CNN model that has been trained to recognize hand gestures. The majority of ASL letters are static, however a select handful are dynamic. Thus, this project's goal was to use feature extraction from finger and hand motions to distinguish between static and dynamic gestures. Potential uses for this project include communication tools, assistive technology, and instructional resources for the deaf and hard-of-hearing population.

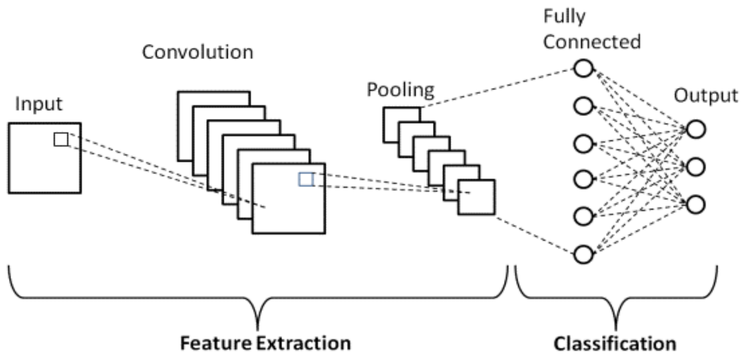

Fig 11: CNN Architecture

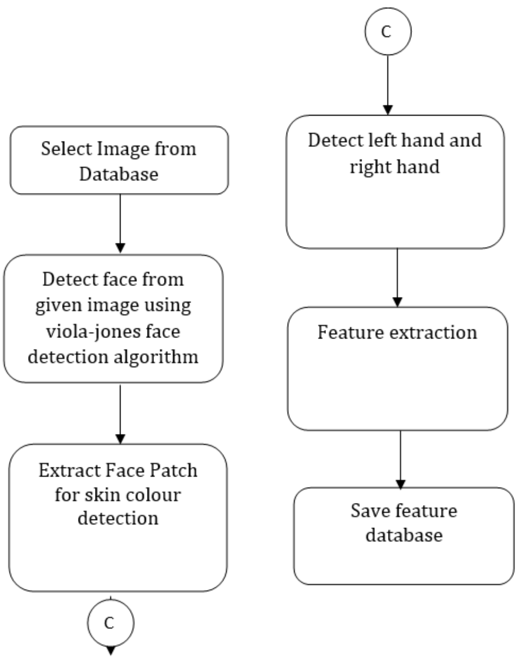

Fig 12: Block Diagram for Image Processin

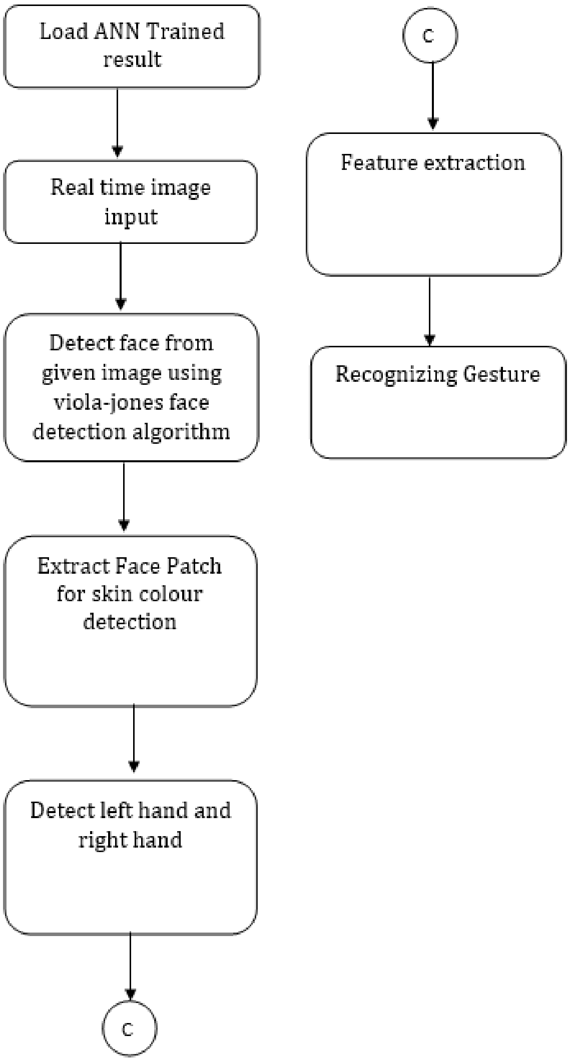

Fig 13: ANN Testing – Gesture Detection

4. Viola-Jones Face Detection Algorithm

Given a picture (this algorithm operates on grayscale images), it examines numerous smaller subregions and searches each one for particular traits in an attempt to locate a face. An image may have multiple faces of varying sizes, therefore it must verify a wide range of positions and scales. In this technique, Viola and Jones used Haar-like properties to detect faces.

The Viola-Jones algorithm has four main steps

- Selecting Haar-like features
- Creating an integral image
- Running AdaBoost training
- Creating classifier cascades

Add up and compare the pixel values of the two sections to find out which is lighter or darker. There will be fewer values overall for all the pixels in the darker area than there are for all the pixels in the brighter area. If one side is lighter than the other, it might symbolize an eyebrow's edge. On rare occasions, the centre section may have more gloss than the surrounding boxes, indicating the outline of a nose. This can be accomplished by using Haar-like features, which we can then utilize to understand the different facial qualities. The integral image helps us quickly finish these intricate calculations and decide if a feature or combination of features satisfies the requirements. "Integral image" describes both the data structure and the method (also known as a summed-area table) that generates it. This approach is a quick and efficient way to determine the total pixel values in an image or the rectangular portion of an image. We use the machine learning algorithm AdaBoost. The 24 × 24 detector frame contains about 160,000 features, although only a small portion of these features are essential for face identification. We use the AdaBoost algorithm to identify the most important features from the 160,000 features. As a result, when we train AdaBoost to discover key features, we feed it training data and allow it to learn from the knowledge to forecast. Ultimately, the algorithm determines a lower bound to determine if a given input meets the criteria for being considered a relevant feature. The cascade's job is to quickly remove non-faces to reduce computation time and waste. As a result, the speed needed to detect faces in real time is reached. We set up a cascaded system and divided the face identification procedure into multiple parts. In other words, the subregion goes through the best features in the first stage, like the feature that recognizes the nasal bridge or our best features comprise the classifier in the first stage.

CONCLUSION AND FUTURE WORK

Given that sign language is a way for people who are Deaf or Dumb to express themselves, this technology will bolster that medium's durability and usefulness. Here, the system will convert the sign language into both text and speech using the gloves. We've provided the option to expand the gesture database to improve and make gesture recognition easier. This system can have more sensors added to it to detect sign language more accurately and completely. Furthermore, the system can be designed to translate words between different languages and to be language-neutral. Any portable device can have this added by connecting it to a mobile app. It may be possible to enhance the D AND D ASSISTANT system approach to translate text and voice from any sign language to voice and vice versa.

REFERENCES

1. Dhananjai Bajpai, Uddaish Porov Gaurav Srivastav ,Nitin Sachan,2015 Fifth International Conference on Communication Systems and Network Technologies,2015 Conference Paper,IEEE, DOI: 10.1109/CSNT.2015.121
2. Lih-Jen Kau, Wan-Lin Su ,Pei-Ju Yu ,Sin-Jhan Wei, 2015 IEEE 58th International Midwest Symposium on Circuits and Systems (MWSCAS),Conference Paper IEEE,DOI: 10.1109/MWSCAS.2015.7282137
3. Carlos Emmanuel A.Quiapo, Katrina Nicole M.Ramos, 2016 IEEE Region 10 Conference (TENCON), Conference Paper IEEE, DOI: 10.1109/TENCON.2016.7848321
4. K Abhijith Bhaskaran, Anoop G. Nair,K Deepak Ram ,Krishnan Ananthanarayanan, H.R. Nandi Vardhan ,2016 International Conference on Robotics and Automation for Humanitarian Applications (RAHA), Conference Paper ,IEEE , DOI: 10.1109/RAHA.2016.7931887
5. Brunna Silva, Wesley Calixto , Geovanne Furriel, 2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), Conference Paper, IEEE, DOI: 10.1109/CHILECON.2017.8229734
6. Abhishek B. Jani, Nishith A. Kotak , Anil K. Roy 2018 IEEE SENSORS ,Conference Paper,IEEE, DOI: 10.1109/ICSENS.2018.8589574

7. Hina Shaheen, Tariq Mehmood, 2018 IEEE Region Ten Symposium (Tensymp), Conference Paper, IEEE, DOI: 10.1109/TENCONSpring.2018.8692069.
8. Anirbit Sengupta, Tausif Mallick , Abhijit Das, 2019 Devices for Integrated Circuit (DevIC), Conference Paper, IEEE, DOI: 10.1109/DEVIC.2019.8783574
9. Ajay Suri, Sanjay Kumar Singh, Rashi Sharma ,Pragati Sharma ,Naman Garg ,Riya Upadhyaya, 2020 International Conference on Smart Electronics and Communication (ICOSEC), Conference Paper IEEE, DOI: 10.1109/ICOSEC49089.2020.9215392
10. Basudev Subedi, Kelzang Ugyen Dorji ,Pema Wangdi ,Tshering Dorji ,Kazuhiro Muramatsu, 2021 Innovations in Power and Advanced Computing Technologies (i-PACT), Conference Paper, IEEE, DOI: 10.1109/i-PACT52855.2021.9696641
11. Ana Antoniette C. Illahi, Maria Fides M. Betito, Christine Caryl F.Chen, Christian Vittorio A. Navarro ,Irene Victoria L.Or, 2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Conference Paper, IEEE, DOI: 10.1109/HNICEM54116.2021.9731835
12. Amir R.Ali, Jessica Magdy, Aya ElNahas, Marina Samir, N. M. Samantha Kamel, Mariam A. Fathy, 2023 International Telecommunications Conference (ITC-Egypt),Conference Paper, IEEE, DOI: 10.1109/ITC-Egypt58155.2023.10206218