

MACHINE LANGUAGE BASED GESTURE RECOGNITION SYSTEM TO AID VISUALLY IMPAIRED PEOPLE

MENTOR: T. GOPHIKA

*Department of ECE,
Easwari Engineering College
Chennai, India
gophika.t@eec.srmrmp.edu.in*

UG STUDENT: A.S. Manoj

*Department of ECE,
Easwari Engineering College
Chennai, India
manomathesh077@gmail.com*

UG STUDENT: S.S. Meera Sahib Maraikar

*Department of ECE,
Easwari Engineering College Chennai, India
maraikarsahib15@gmail.com*

Abstract— The pace of technological progress is commonly known to be accelerating. However, despite this, the application of these technologies across various sectors is still limited. Communication is the primary means of expressing ideas or conveying a message yet individuals with disabilities may face challenges communicating with others. People with hearing or speech impairments may experience social isolation to some extent and using the same communication methods as others (such as sign language or text-to-speech) is a significant factor to consider. Acknowledging disability as an issue of equality is crucial to improving the lives of people with disabilities.

We are aware that blind people have a very difficult time navigating without human assistance. As a result, they are reliant on the other person to navigate with them. These physically disabled people communicate their feelings and thoughts to others by using sign language. Therefore, the empirical task of sign language recognition has emerged. The main objective is to create a deaf and dumb gesture recognition system using machine learning and image processing to enable communication between the deaf & dumb and the blind. The goal of this project is to make the communication between the Blind, Deaf, and Dumb people more simpler by introducing a new system prototype. The key actions performed are captured by the camera and The MATLAB is used to perform image processing. Both a voice command and an ultrasonic sensor are available for detecting objects in front of them.

Keywords— blind, gestures, micro-controller, MATLAB

INTRODUCTION

The treatment of students who are blind has changed as a result of discussions about visual impairment and education. According to a quote from Louis Braille, having access to communication in the broadest sense also means having access to knowledge, which is crucial for us if we want to stop being looked down upon or treated as inferior by sighted people. We don't require sympathy or a constant reminder of our frailty. The conversation is about how we can achieve treating each other as equals. India has 12.3 million people who have severe to total hearing loss. India is home to approximately 372 private deaf schools in addition to 478 government-funded institutions. Although science and technology have made people addicted to comfort, there is still a group of underprivileged people who are fighting to develop a novel method of communication that will benefit them.

Human-computer interfaces that are adaptive and customized are now valued highly in comparison to those that are created for "average" users in a wide range of applications. Currently, many HCI systems use machine learning

algorithms to automatically analyze body language and facial expressions. The disabled are a part of our society. Despite the rapid advancement of technology, little is being done to benefit blind people. Worldwide, there are about a billion people who are blind and deaf. It has always been difficult to communicate with a visually impaired person when you're deaf-mute.

People who are blind or deaf can communicate using sign language. A popular technology for assisting the dumb and particularly the blind is gesture recognition. The two important fields of computer vision and machine learning are related in this paper. The field of computer vision includes techniques for gathering, processing, and comprehending images. Additionally, it is employed in several different fields, including image reconstruction, human-computer interaction, and physics. Contrarily, the study of pattern recognition and computational learning in Artificial Intelligence gave rise to the field of machine learning, a branch of computer science. To accomplish its goal of gesture recognition, this paper employs computer vision and machine learning techniques.

LITERATURE SURVEY

A NOVEL APPROACH AS AN AID FOR BLIND, DEAF AND DUMB PEOPLE

Abstract: Communication is an essential aspect of expressing ideas and thoughts, but individuals with disabilities, such as deafness or muteness, face difficulties in interacting with others. These individuals must utilize the same modes of communication, such as sign language or text-to-speech, as non-disabled individuals to bridge this communication gap. Unfortunately, social isolation is a common issue experienced by individuals with disabilities. Recognizing disability as an equality issue is crucial to improving the lives of disabled individuals. To simplify communication between individuals with hearing, visual, and speech impairments, this project proposes a new prototype system that utilizes portable technology and Arduino circuit boards to provide a means of communication.

Published in: [2017 Third International Conference on Sensing, Signal Processing and Security \(ICSSS\)](#)

All Authors :: [B. Rajapandian](#); [V. Harini](#); [D. Raksha](#); [V. Sangeetha](#)

BRILLE ASSISTANCE SYSTEM FOR VISUALLY IMPAIRED, DEAF-MUTE PEOPLE IN INDOOR & OUTDOOR APPLICATION

Abstract: For people who are blind, visually impaired, or deaf-mute, getting around both inside and outside presents a challenge. Getting around inside is becoming more difficult for these individuals. Visual impairment is worse for those who are non-visually impaired according to observations. For decision-making support, people with visual impairments or blindness frequently rely on trained dogs, other people, or specialized equipment. As a result, it has become essential for blind people to have an assistive device that will allow them to freely navigate. With the help of the ARM LPC-2148 and the Braille keypad, which is an application that is user-friendly for these people, various sensors and actuators are here interfaced and Utilizing technology as much as possible helps to minimize the issues faced by blind people. Multiple sensors are built into the walking stick that blind people use, allowing for the addition of more features and technological advancements. The primary functions include obstacle detection for collision avoidance, along with a few other sensors for pit hole, fire, and water detection. A panic switch is an emergency button that activates the GSM module, which then sends an SMS to the carer with the GPS coordinates of the person who is blind, visually impaired, or deaf and mute. Giving the blind the safest route is the task at hand.

Published in: [2019 4th International Conference on Recent Trends on Electronics](#)

All Authors::[Sunil Kumar KN](#); [R Sathish](#); [S Vinayak](#); [TarkeshworParasadPandit](#)

EMBEDDED SIGN LANGUAGE INTERPRETER SYSTEM FOR DEAF AND DUMB PEOPLE

Abstract: The foundation of dumb and deaf conversation is sign language that represents words with hands and fingers. The important goal of the project is to create a tool that enables dumb and deaf people to communicate with hearing people. There are two main components to the system. Speech recognition comes after hand sign

recognition in the second step. By doing this, two-way communication is made possible without an interpreter's help. The hand gesture identification unit includes a webcam for capturing images, a microcontroller, a display device linked to the controller, and a Bluetooth module. The digitized signal is processed and then subjected to statistical modeling and matching to recognize the gesture in the speech recognition stage. The speech recognition part is available in a user-friendly Android app. To make the system applicable on both mobile and embedded platforms, a high-performance and energy-efficient ARM Cortex A8 Processor board is used. The Haar classifier is utilized for image classification, and 1-D HMM is used for speech conversion. Sign language identification has proven to be significant in various fields, including Human-Computer Interaction (HCI), robotic controls, and home automation.

Published in: [2017 International Conference on Innovations in Information, Embedded and Communication Systems \(ICIECS\)](#)

HAND GESTURE RECOGNITION AND VOICE CONVERSION SYSTEM FOR DUMB PEOPLE

Abstract: Communication between dumb people and sighted people is greatly facilitated by sign language. Mute people find it very challenging to communicate with non-mute people because most people are not trained in hand sign language. It is very difficult for them to communicate during times of emergency. The answer to this issue is to translate sign language into a voice that people can hear. The two main methods for detecting hand motion or a gesture are vision and non-vision techniques, and the raspberry pi can translate the detected data into voice. For gesture detection, the vision-based technique uses a camera, while the non-vision-based technique uses sensors. This project will employ a method that is not vision-based. Deafness is common among the dumb. Thus, the voices of regular people can be translated into sign language. The message will automatically be sent to their family or friends in an emergency.

Published in: [2019 5th International Conference on Advanced Computing And Communication Systems \(ICACCS\)](#)

All Authors:: [S. Vigneshwaran](#); [M. ShifaFathima](#); [V. Vijay Sagar](#); [R. SreeArshika](#)

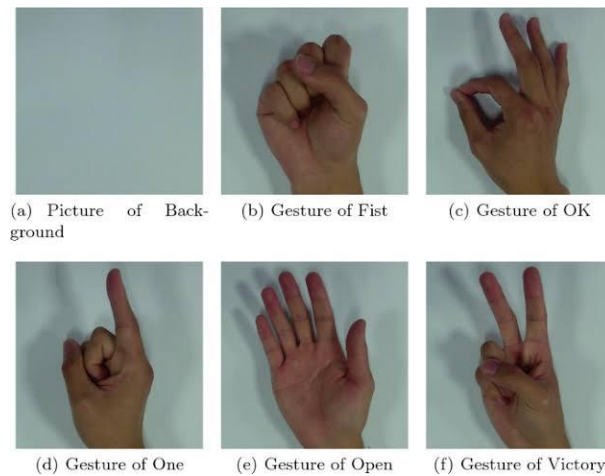
PROPOSED SYSTEM

We propose this system for blind and deaf people to address the shortcomings of the system previously described. Using a Camera, the system takes pictures. The performed action is recognized by the MATLAB section. The proposed system allows only a limited set of actions. Recognizing the action that was performed, the micro controller is given the information. Even though they can see the action, the micro controller provides a voice command so that people can understand it. The micro controller is linked to ultrasonic sensors. It will recognize the object. If the object is in close proximity, the micro controller informs the user via voice commands. When blind people trip, the system will send an alert message to the carer via IOT along with the location because IoT is a rapidly evolving technology, we have included it in our proposed design to improve it significantly.

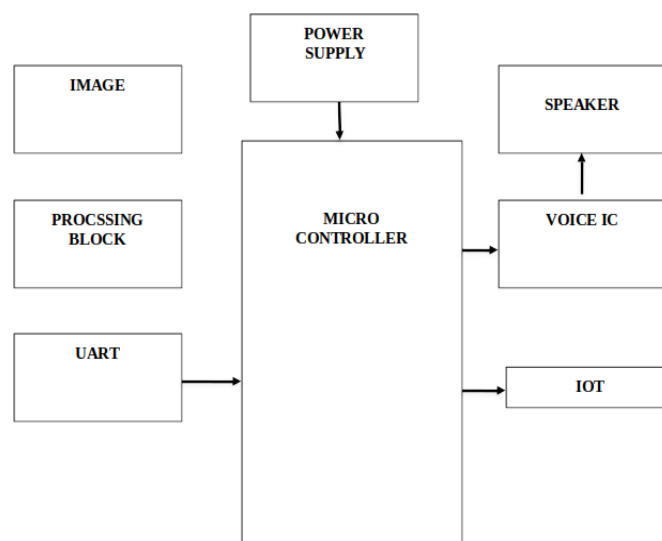
METHODOLOGY

This system first captures the image or video and goes through some preliminary processing, tracking the hand in the video's subsequent frames to develop a reliable system that can recognize and follow hands with various skin tones in a range of lighting, the process of removing key features that represent key aspects of the gesture throughout the images or video then storing these features. Find features that represent shape, motion, size, reflectivity, and other significant properties. To recognize multiple gestures with a small set of features, features must be generative rather than discriminative.

This step entails identifying and categorizing the gesture that was performed. It consists of two phases: the training phase, in which the system is trained using datasets; and the classification phase, in which performed gestures are classified using CNN and converted to voice. The speakers are then given the corresponding voice so that the blind can hear. This classification achieves high accuracy in a short amount of time.

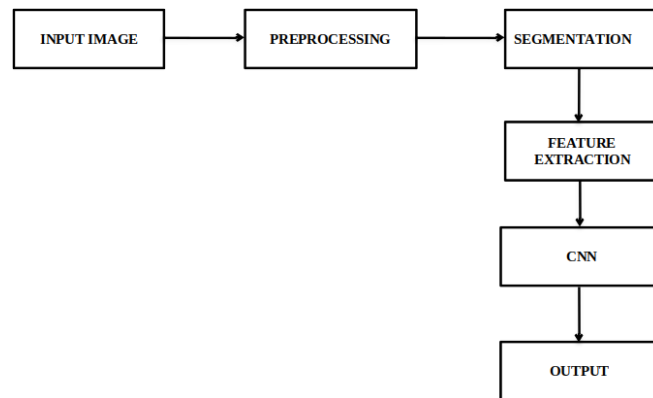


SYSTEM ARCHITECTURE



The PC will have a MATLAB program to identify hand gestures by pre-processing images obtained from webcams and sending the ID to the microcontroller using a universal asynchronous receiver and transmitter cable. Following processing, the alphanumeric liquid crystal display and FN-M16P MP3 are sent the output of the gestures the micro-controller received via the UART cable. The output will be displayed on the Active Matrix Liquid Crystal Display (ALCD) based on the ID of the gesture that the microcontroller has received.

The ID is transmitted to the FN-MI6P unit, which compares the gestures ID with the commands stored on the SD card and outputs the voice of the corresponding command through the speaker. The equipment devices, such as the ZigBee module, are used by both the beneficiary and the transmitter. In order for dumb, deaf, and blind people to interact with normal people and communicate with each other, accelerometer sensors are used to detect gesture-based structures. The system centre operation manager (SCOM) buffer receives the received ASCII X value and converts it to an integer value using the ATOI command in the programming and then it first stores the information in a temporary buffer, compares it to the threshold value, and then outputs a voice command.

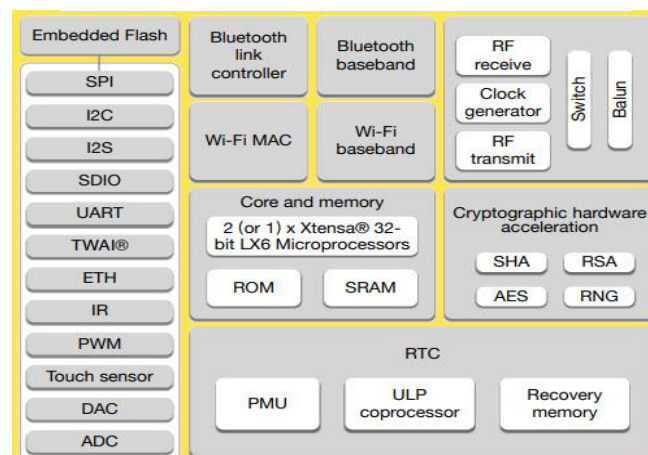


HARDWARE DESCRIPTION

ESP32 MICROCONTROLLER:

The ESP32 System on Chip (SoC) Microprocessor is a cost-effective alternative to the well-known ESP8266 SoC, developed by Espressif System. This chip is equipped with Tensilica's 32-bit Xtensa LX6 Microprocessor, which incorporates Bluetooth and Wi-Fi capabilities, unlike the ESP8266 SoC. The ESP32 is available in both single-core and dual-core versions. Similar to the ESP8266, the ESP32 includes RF components such as a low-noise receiver, power amplifier, 23 amps, filters, an RF balun, and an antenna switch. The ESP32 requires minimal additional components for building hardware, making it easy to use. Additionally, TSMC's ultra-low-power 40 nm technology is used to produce the ESP32, which is ideal for developing battery-powered applications such as smartwatches, wearables, baby monitors, audio and video equipment, and more.

The ESP32 series of microcontrollers is an affordable and low-power system on a chip option that is highly recommended due to its built-in Wi-Fi and dual-mode Bluetooth capabilities. In addition to an integrated antenna, the ESP32 includes 24 switches, an RF balun, a power amplifier, a low-noise receive amplifier, filters, and power control modules. Espressif Systems, a Chinese company based in Shanghai, developed and manufactured the ESP32 using TSMC's 40 nm manufacturing process.



ULTRASONIC SENSOR:

The ultrasonic sensor sends out ultrasonic pulses, and by timing how long it takes for those pulses to return to the transducer after leaving the item, it can determine its location. A reflection from a surface causes the transducer's sonic vibrations to return to it. The ultrasonic sensor will transition to receive mode once it has finished emitting sound waves. In direct proportion to the object's distance from the sensor, the delay between emitting and receiving increases.

When an ultrasonic wave was released by the ultrasonic transmitter, timing began. When it came across obstructions on its path, ultrasonic spread in the air and instantly turned around. Eventually, when it picks up the reflected wave, the ultrasonic receiver will cease timing. It is calculated how far away the sensor is from the intended item. With high precision and stable readings, it provides superb non-contact range detection in a convenient package. Sunlight or dark objects have no impact on how it operates. The sensor's power source voltage is 5 VDC. Two pins on the sensor, designated trig, and echo, are connected to the controller to provide digital data.



UART COMMUNICATION:

A UART is a physical circuit present in microcontrollers or standalone ICs that enables asynchronous transmission and reception of serial data. To transmit data, the UART receives it from the data bus, which is also utilized by other devices like processors, memory, or microcontrollers.

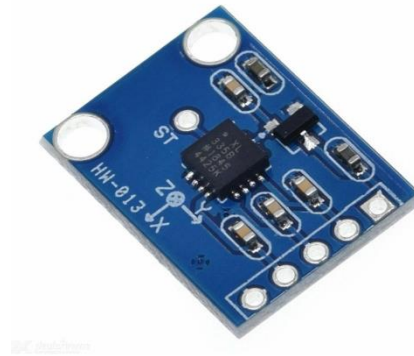
The data is sent in parallel from the data bus to the transmitting UART during data transfer. The transmitting UART then converts this parallel data into a serial data packet by appending a start bit, a parity bit, and a stop bit. The serial data packet is then transmitted bit by bit through the Transmitting port. On the receiving end, the UART acquires the serial data packet bit by bit through the Receiving port, which is then transformed back to parallel form by the receiving UART, and the stop bits are removed after processing the start bit and parity bit.

UART messages contain a start bit, five to nine data bits, an optional parity bit, one or two stop bits, and up to nine data bits. The UART line's voltage level is typically set to high when it is not transmitting data. To begin data transmission, the transmitting UART pulls the transmission line from high to low for one clock cycle. After detecting this high-to-low voltage transition, the receiving UART begins to read the bits in the data frame at the baud rate frequency. Finally, to indicate the end of the data packet, the sending UART changes the voltage on the data transmission line from low to high for at least two-bit durations.

MEMS ACCELEROMETER:

The accelerometer has several features such as temperature compensation, self-test, 0g-Detect, and g-Select. It is a small capacitive micromachined accelerometer with low energy consumption and a low profile. A 1-pole low pass filter and signal conditioning are also integrated. The accelerometer comes pre-configured with zero-g offset and sensitivity, so no additional equipment is needed. Additionally, it includes a Sleep Mode, which makes it ideal for portable battery-powered devices.

How fast speed changes is gauged by acceleration. A speedometer and an accelerometer are both meters that detect speed and acceleration, respectively. Accelerometers can be used for purposes that require orientation or to detect vibrations in systems. One, two, or three axes can all be used by accelerometers to detect acceleration. As the expense of developing 3-axis units falls, their use is expanding. The capacity of an accelerometer to detect acceleration can be used to measure a wide range of things that are very helpful for electronic and robotic projects.



SOFTWARE DESCRIPTION

MATLAB SOFTWARE:

Matrix Laboratory is the abbreviation for the word MATLAB. The initial purpose of MATLAB was to make it simple to access the matrix computer programs created by the EISPACK and LINPACK projects. MATLAB utilizes the LAPACK and ARPACK projects' computer programs, which are considered the most advanced software for matrix computation. In MATLAB, there is a category of application-specific solutions called toolboxes. Most MATLAB users rely on toolboxes as they provide access to specialized technologies. Toolboxes consist of a collection of MATLAB functions (M-files) that are tailored for specific problem domains and are used to extend the capabilities of the MATLAB environment. These toolboxes are available for a wide range of fields, including but not limited to signal processing, control systems, fuzzy logic, wavelets, neural networks, and simulation.

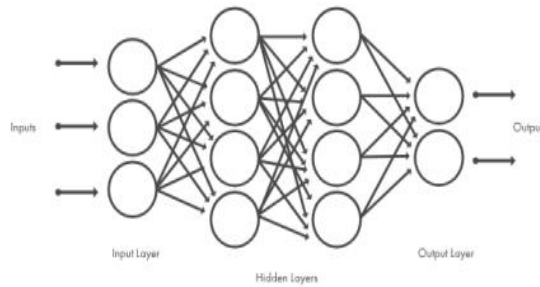
A high-performance language for complex processing is called MATLAB. In a simple-to-use environment, it supports system-level design, visualization, and computing while expressing issues and answers using well-known mathematical terms.

The following are examples of typical applications:

- Modeling, simulation, and prototype creation, as well as data analysis, exploration, and visualization.
- Application creation, which includes creating graphical user interfaces, Scientific and engineering graphics.

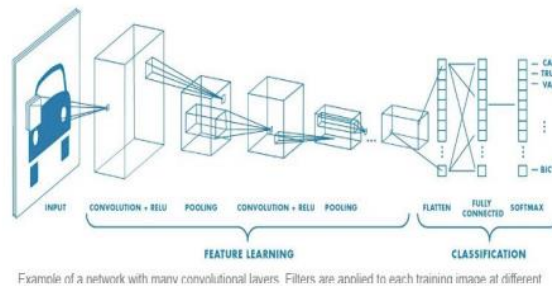
CONVOLUTIONAL NEURAL NETWORK:

Among the most widely used deep learning methods, convolutional neural networks (CNN or ConvNet), enable models to learn to perform classification tasks directly from images, videos, texts, or sounds. Deep learning is a type of supervised learning. Using Convolutional Neural Networks (CNNs) to detect patterns in images can be very useful in identifying objects, faces, and scenes. In the UART data transmission process, the received data is first converted back to parallel form by the receiving UART, and then the stop bits are removed. The data packet is then transmitted in parallel with the data bus from the sending UART to the receiving UART. The UART data is divided into segments, with each message containing a start bit, 5 to 9 data bits, a potential parity bit, 1 or 2 stop bits, and up to 9 data bits. They derive their knowledge straight from the image data, classifying images with the aid of patterns rather than manually extracting features. CNNs play a key role in several applications that require computer vision and object recognition, including self-driving cars and facial recognition software. You can create a CNN from inception or use an existing model that has already been trained using your dataset, depending on your application.



There may be hundreds and hundreds or even large numbers of layers in a convolutional neural network, and each one gains the ability to recognize various features in a picture. At various resolutions, filters are attached to each training picture, and the output of each convolved picture serves as the input for the following layer. The filters may begin with very basic characteristics, like brightness and edges, and then become more complicated to include characteristics that specifically identify the object. Using CNNs, features in images, text, music, and video are identified and categorized. An input layer, an output layer, and numerous buried layers are present in a CNN, just like in other neural networks.

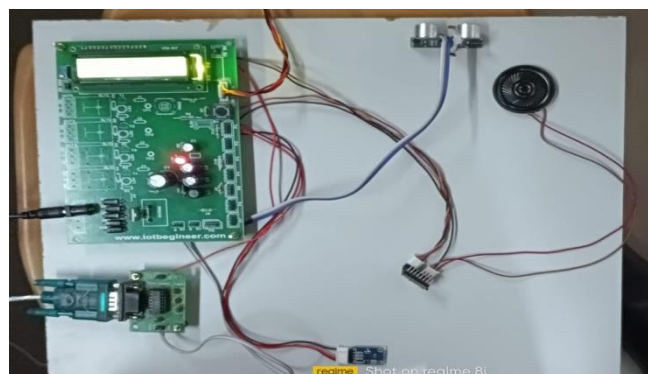
These modules carry out functions to acquire characteristics unique to the data. Convolution, activation or ReLU, and sharing are three of the most popular layers. Convolution runs a series of convolutional filters through the raw images, activating different aspects of the images with each filter. Rectified linear unit (ReLU), which maintains positive values while transferring negative values to zero, enables quicker and more efficient training. Since only the activated characteristics are carried over to the following layer, this is sometimes referred to as activation. By conducting nonlinear downsampling on the output, pooling reduces the number of parameters the network needs to learn. Over tens or hundreds of layers, these processes are repeated, and each layer learns to identify different features .



Example of a network with many convolutional layers. Filters are applied to each training image at different

The functionality of a CNN changes to segmentation after learning features through multiple levels. The next-to-last layer, which is a completely connected layer, generates a vector with K dimensions, where K is the total number of classes that the network can predict. The possibilities for each class of any picture that is being classified are contained in this vector. The categorization output is provided by a classification layer, such as softmax, in the CNN architecture's top layer.

HARDWARE SETUP:



LOCATION MONITORING:

Sensor ID	status	latitude	longitude	Updated on
140151	FELL DOWN OCCURS	13.0334873	80.2124365	2022-12-04 23:54:32
140150	FELL DOWN OCCURS	13.0334873	80.2124365	2022-12-04 23:54:18
140149	FELL DOWN OCCURS	13.0334873	80.2124365	2022-12-04 23:54:10
140142	FELL DOWN OCCURS	13.0334873	80.2124365	2022-12-04 23:52:40
140137	FELL DOWN OCCURS	13.0334873	80.2124365	2022-12-04 23:52:33

CONCLUSION

This project proposes a basic framework to assist individuals with various combinations of disabilities such as blindness, deafness, and dumbness to overcome communication barriers. The system allows people to communicate and convey messages based on their skills and preferences. Individuals who are dumb and deaf can use American Sign Language, and those who cannot understand Sign Language can utilize the device to receive output in audio, Braille, or regular text displayed on an LCD. This enables them to interact with the outside world despite their disabilities.

REFERENCES:

- [1] Rajamohan, Hemavathy R., Dhanalakshmi M. "Deaf-Mute Communication Interpreter", International Journal of Scientific Engineering and Technology Volume 2 Issue 5, pp: 336-341 (ISSN: 2277-1581)
- [2] Kuldeep Singh, Shashank Deshpande, Uma Mudenagudi, "INTERACTIVE ACCELEROMETRIC GLOVE FOR HEARING IMPAIRED".
- [3] Channaiah Chandana K1, Nikhita K2, Nikitha P3, Bhavani N K4, Sudeep J5, "Hand Gestures Recognition System for Deaf, Dumb and Blind People", International journal of innovative research in computer and communication engineering volume 5, issue 5, may 2017.
- [4] Nikolaos Bourbakis, Anna Esposito, D. Kabraki, "Multi-modal Interfaces for Interaction Communication between Hearing and Visually Impaired Individuals: Problems & Issues", 19th IEEE International Conference on Tools with Artificial Intelligence.
- [5] Tahir Khan, "Hand gesture recognition based on digital image processing using MATLAB" International Journal of Scientific & Engineering Research, Volume 6, Issue 9, September 2015.
- [6] Netchanok Tanyawiwat and Surapa Thiemjarus, Design of an Assistive Communication Glove using Combined Sensory Channels, 2012, Ninth International Conference on Wearable and Implantable Body Sensor Networks.
- [7] N. Bourbakis, An SPNG based method for image to NL text conversion, PR Journal.
- [8] [6] G. Grimes, Digital Data Entry Glove Interface Device, AT & T Bell Labs, 1983.
- [9] D. Sturman and D. Zelter, -A survey of glove-based input, II IEEE Computer Graphics and Applications, vol. 14, no. 1, pp. 30-39, 1994.
- [10] M. Mohandes and S. Buraiky, -Automation of the Arabic sign language recognition using the powerglove, II AIML Journal, vol. 7, no. 1, pp. 41-46, 2007.
- [11] S. Sidney and E. Geoffrey, -Glove talk-a neural network interface between a data-glove and a speech synthesizer, II IEEE Transactions on Neural Networks, vol. 4, no. 1, pp. 2-8.
- [12] M. V. Lamar, M. S. Bhuiyan, and A. Iwata, "Hand gesture recognition using T-CombNet: A new neural network model," IEICE Trans.
- [13] W. K. Chung, X. Wu, and Y. Xu, "A realtime hand gesture recognition based on haar wavelet representation," in Proc.
- [14] Hiroomi Hikawa, Member, IEEE, and Keishi Kaida "Novel FPGA Implementation of Hand Sign Recognition System With SOM-Hebb Classifier" IEEE Transactions on Circuits And Systems For Video Technology, Vol. 25, No. 1, January 2015.