International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Sign Recognition and Voice Conversion Device for Dumb

Vinu MS.[1], Santhosh A.[2], Gowtham K.[3], Kaviyarasu P.[4]

[1]Assistant Professor, [2,3,4] UG Students,
[1,2,3,4]Department of Artificial Intelligence and Data Science, Nehru Institute of Engineering and Technology, Anna University, Coimbatore, Tamil Nadu, India.
[1]nietmsaravanakumar@nehrucolleges.com ,[2]santhosan61087@gmail.com ,[3]kgowthamkandasamy@gmail.com ,[4]kaviyarasu.vij@gmail.com

ABSTRACT

This project proposes a comprehensive hardware setup for home automation and environmental monitoring. The system includes components such as a Bluetooth module for remote control, a voice module (APR9600) for audio feedback, an LCD for display purposes, relay modules for controlling electrical devices, bulbs with holders for lighting, a CPU fan for temperature regulation, speakers for audio output, and a 5V power supply for operation. The core functionality of the system revolves around the integration of flux sensors for environmental monitoring. When flux sensors detect specific conditions, predefined audio signals are played through the voice module. These conditions correspond to six distinct signs, facilitating real-time alerts and notifications. Moreover, the system offers remote control capabilities through Bluetooth connectivity, enabling users to manage lighting and fan operations conveniently. Additionally, temperature monitoring is implemented using a DHT11 sensor, triggering the activation of the CPU fan when temperature thresholds are exceeded. Furthermore, an LCD display enhances user interaction by providing visual feedback and status updates. The project utilizes Arduino UNO as the central processing unit and incorporates 12V 1A adapters for powering specific components. Connectors and power cables are utilized for seamless integration and electrical connectivity. Overall, this project presents an integrated solution for home automation and environmental monitoring, offering flexibility, convenience, and enhanced user experience.

Keywords: LCD, DHT11, Arduino Uno, APR9600, Sound Alert.

1. INTRODUCTION

The Communication is a fundamental human right, yet millions of speech-impaired individuals face challenges in expressing their thoughts and emotions in everyday interactions. Traditional methods, such as sign language, are effective but rely on mutual understanding between communicators. However, many people are unfamiliar with sign language, creating barriers to seamless communication.

This paper presents a Sign Recognition and Voice Conversion Device designed to bridge this gap by translating sign language gestures into audible speech in real time. Utilizing advanced technologies such as machine learning, computer vision, and natural language processing, the system captures hand gestures and facial expressions using gesture recognition sensors or camera-based tracking. The recognized signs are then converted into corresponding voice outputs, enabling effortless communication with non-sign language users.

Our approach integrates deep learning models trained on diverse sign language datasets, ensuring high accuracy in gesture interpretation. Additionally, the device supports multiple languages and dialects, making it adaptable for different regions and communities. The system also features customizable voice synthesis, allowing users to select their preferred voice output for a more personalized experience.

By enhancing accessibility and inclusivity, this innovation has the potential to empower speech-impaired individuals, improve social interactions, and foster a more inclusive society. This paper discusses the device's architecture, implementation, performance evaluation, and future scope, highlighting its impact on assistive communication technologies.

1.1 OBJECTIVES

- Develop a real-time system for accurate sign language recognition using computer vision and deep learning.
- Convert recognized gestures into clear and natural-sounding speech using text-to-speech (TTS) technology.
- Support multiple sign and spoken languages for broader accessibility and inclusivity.
- Design a user-friendly, portable, and affordable device for seamless daily communication.
- Enhance accuracy by incorporating facial expressions, complex gestures, and contextual understanding.

2. LITERATURE REVIEW

Assistive communication technologies have significantly evolved, leveraging advancements in computer vision, deep learning, and natural language processing to aid speech-impaired individuals. Various studies have explored sign language recognition, primarily using computer vision-based and sensor-based approaches. Early methods relied on glove-based sensors, such as data gloves with accelerometers, to track hand movements. However, recent advancements favor camera-based recognition utilizing convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for real-time performance. Researchers have also employed depth sensors like Microsoft Kinect and hand pose estimation models, such as MediaPipe Hands, to enhance accuracy.
Machine learning and deep learning have played a crucial role in improving gesture recognition, with hybrid models combining CNNs and LSTMs demonstrating better performance in sequential gesture recognition. Attention mechanisms have further refined accuracy by addressing variations in signing speed and style. To enable voice output from recognized gestures, various text-to-speech (TTS) technologies have been employed. Deep learning-based TTS models like Tacotron and WaveNet generate natural-sounding speech, offering users customizable voice options. Some studies have even explored direct sign-to-speech conversion without text intermediaries, using deep generative models to

synthesize speech from gesture embeddings.

Recent research emphasizes multimodal approaches incorporating facial expressions, lip movements, and body posture to enhance recognition accuracy. Context-aware systems integrating sentiment analysis and natural language processing (NLP) have further improved expressiveness and comprehension. However, challenges persist in real-time processing, dataset diversity, and language adaptability. Many models are trained on limited datasets, affecting their generalizability across different sign languages. Addressing these challenges requires real-time optimization, low-power edge computing, and more inclusive datasets.

While significant progress has been made in sign language recognition and speech conversion, gaps remain in adaptability, multilingual support, and user personalization. This research aims to integrate deep learning, NLP, and multimodal processing to develop a practical and efficient assistive device that enhances communication for speech-impaired individuals.

The disadvantages of the existing methods:

1. **Regional and personal variations in sign language** – Different regions have different sign languages (e.g., ASL, BSL, ISL), and individuals may have unique ways of signing.
2. **Complex or rapid gestures** – Some devices struggle to recognize fast or nuanced gestures accurately.
3. **Environmental limitations** – Poor lighting, background noise, or crowded spaces can reduce the effectiveness of gesture recognition.

To overcome :

1. **AI-Powered Adaptive Learning** – Implement machine learning algorithms that continuously learn and adapt to the user’s unique signing style and different sign languages.
2. **Multi-Sensor Integration** – Use a combination of cameras, motion sensors, and gloves with haptic feedback to improve gesture detection accuracy.
3. **Cloud-Based Updates** – A system that updates regularly with new sign language data can enhance adaptability and coverage.
4. **Context-Based Prediction** – Incorporate natural language processing (NLP) to predict words based on context, reducing errors in interpretation.

3. PROPOSED WORK

The aim of this project is to develop an advanced sign recognition and voice conversion device that enhances communication for individuals with speech disabilities by accurately translating sign language into spoken words. This system will integrate AI-powered adaptive learning to recognize different sign languages, regional variations, and personal signing styles, improving accuracy over time. It will utilize a combination of computer vision, motion sensors, and smart gloves with haptic feedback to ensure precise gesture detection, even in challenging environments. A cloud-based machine learning model will continuously update with new sign data, making the device more adaptive and inclusive. Additionally, natural language processing (NLP) algorithms will be incorporated to predict and auto-correct words based on context, reducing misinterpretations. To enhance real-world usability, the device will feature offline functionality, allowing seamless operation without internet dependency, and real-time speech synthesis, providing instant voice conversion. The system will also support multi-language output, enabling users to communicate effectively with people from different linguistic backgrounds. The final goal is to create a lightweight, user-friendly, and cost-effective wearable or handheld device that ensures smoother, more inclusive communication for the speech-impaired community while promoting accessibility and independence.

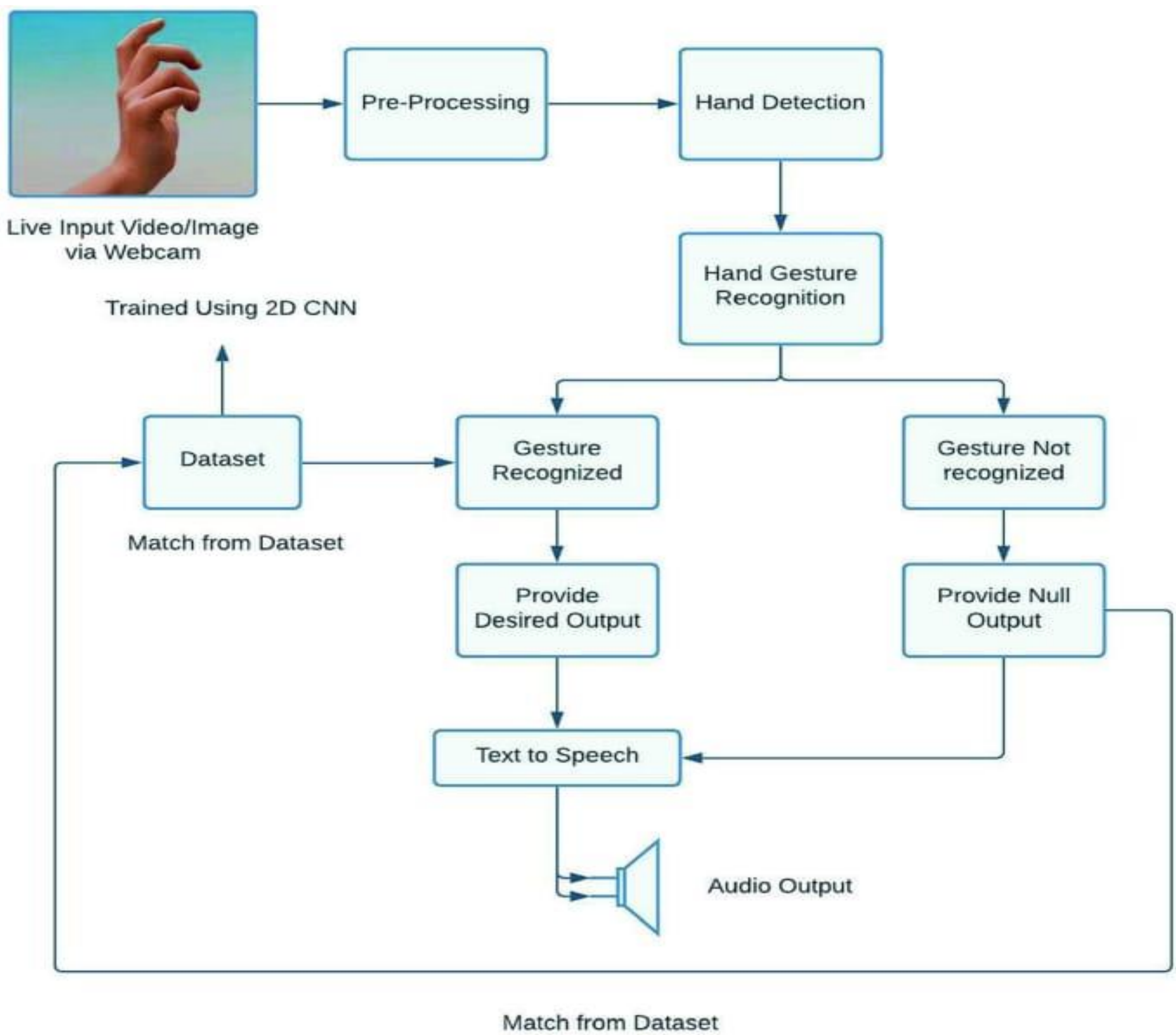


Fig1. Block Diagram

NAME OF THE COMPONENTS	SPECIFICATIONS	MODEL
ARDUINO UNO	Data from several drivers and motors can be processed by Arduino UNO	ATmega328P microcontroller
SENSORS	9-axis IMU (Accelerometer, Gyroscope, Magnetometer) for hand movement tracking	MPU-9250, MPU-6050
RELAY	It is a switch controlled by electricity,working with different voltage,current and contact types to turn devices on or off	SRD-05VDC-SL-C – 5V, 10A, SPDT
LCD	It is an 5v screen that displays 16 characters on 2 line with LED backlights	16x2 LCD
VOICE MODULE	It is a small audio player with 3.3v-5v power, serial communication ,support for WAV files and amplifier for playback	Apr9600
BATTERY	Rechargeable Li-ion, 3.7V, 2000mAh+	Adafruit 3.7V 2500mAh LiPo, Panasonic NCR18650B
WI-FI/BLUETOOTH MODULE	Wireless connectivity for cloud-based updates and pairing	ESP8266, HC-05 Bluetooth Module
NATURAL LANGUAGE PROCESSING (NLP) API	AI-powered text prediction and context analysis	Google Dialogflow, OpenAI GPT API
CASING AND WEARABLE FRAME	Lightweight, durable, and ergonomic design	Custom 3D-printed case using PLA/ABS

Table 1. Components & Specification

Arduino UNO

It is the main controller for processing inputs and controlling output devices. It interfaces with sensors like flex sensors or accelerometers for sign language recognition and the voice module for audio output. The board communicates with an LCD for visual feedback and can trigger relays to control external devices. It processes data from sensors, runs algorithms, and sends commands for voice conversion or gesture interpretation. The 5V Arduino Uno ensures smooth operation of the entire system, powered either through USB or a DC jack.

Fig 2. Arduino UNO

LCD Display

An LCD screen can display useful information to the user such as the current command, system status or errormessages. This feedback helps users understand the system's operation and troubleshoot any issues. Some systems incorporate audible alerts or LEDs to indicate status, providing feedback. Fig 3. Shows the LCD Display 2 x 16.

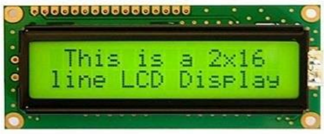


Fig 3. LCD Display 2 x 16

Sensors

The flux sensor detects the bending of a surface, commonly used in wearable devices like gloves for sign language recognition. It changes its resistance as it bends, and the Arduino reads this data to interpret hand movements. The DHT11 sensor measures temperature and humidity, providing digital readings for environmental monitoring. It operates on a 5V supply and communicates over a single-wire interface. Both sensors are low-cost, widely used in DIY projects for environmental sensing and motion detection.

Fig 4. Flux and DHT11 Sensors

Relay

A relay is an electrically operated switch used to control high-power devices with a low-power signal. It has a coil that, when energized, creates a magnetic field to move the contacts and open or close a circuit. Common relay types include SPDT (Single Pole Double Throw) and DPDT (Double Pole Double Throw), with typical voltage ratings of 5V, 12V, or 24V for the coil. The contacts can handle up to 10A for switching high-power devices like motors or lights. Relays are widely used in automation, home control systems, and safety circuits.

Fig 5. Relay

Battery

The DC convertor manages the system power needs, converting the energy from a 12V battery to a suitable

level for the robotic arm's components. This stable power source ensures consistent performances and reducesthe risk of power related malfunctions.

Bluetooth Module

The initial point of interaction is the microphone through which it captures a voice commands. The microphone should be in a good condition. It should capture a command properly so that a voice-controlledrobot can move correctly according to it. The Bluetooth module has connected the robotic arm with the external device, providing a wireless communication. It operates in a range of 10-15 meters.

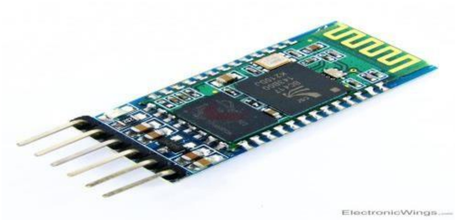


Fig 6. Bluetooth module HC-05 **Fig 7.** Bluetooth Module HC-05 Pin Diagram

Fig 6. shows Bluetooth Module HC-05 & Fig 7. shows Bluetooth Module HC-05 Pin Diagram. The Bluetoothserial module HC-05 facilitates communication between serial-enabled devices via Bluetooth. It features six pins:

1. **Key/EN**: This pin toggles the HC-05 module between command mode and data mode. When set high, it activates command mode (AT commands), while low sets it to data mode. By default, it operates

in data mode at 9600bps; in command mode, it operates at 38400bps.

2. **VCC**: Connect either 5V or 3.3V to power the module.
3. **GND**: Ground connection for the module.
4. **TXD**: Transmit serial data wirelessly received by the Bluetooth module.
5. **RXD**: Receive serial data transmitted wirelessly by the Bluetooth module.
6. **State**: Indicates the connection status of the module.

Voice Module

A voice module (like the DFPlayer Mini) is a small audio playback device used in electronics projects. It can play MP3/WAV files stored on a microSD card via a serial interface (UART). The module supports 3.3V-5V power, has a built-in amplifier, and can output sound through speakers. It can be controlled by microcontrollers like Arduino to trigger sound playback based on inputs. The DFPlayer Mini also supports volume control and looping tracks for audio applications in interactive projects.

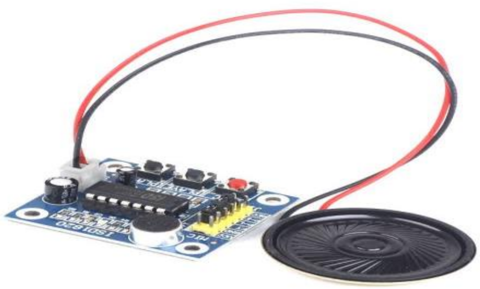


Fig 8. Voice Module

NLP(Natural Language Processing)

Natural Language Processing (NLP) is a branch of artificial intelligence (AI) focused on enabling computers to understand, interpret, and generate human language. It involves tasks like speech recognition, text analysis, translation, and sentiment analysis. NLP uses various techniques such as tokenization, part-of-speech tagging, named entity recognition (NER), and machine learning algorithms to process language data. It plays a key role in technologies like chatbots, virtual assistants, automatic summarization, and language translation services. NLP bridges the gap between human communication and machine understanding.

Casing and Wearable frame

Casing and wearable frames are protective and structural components designed to house electronic device

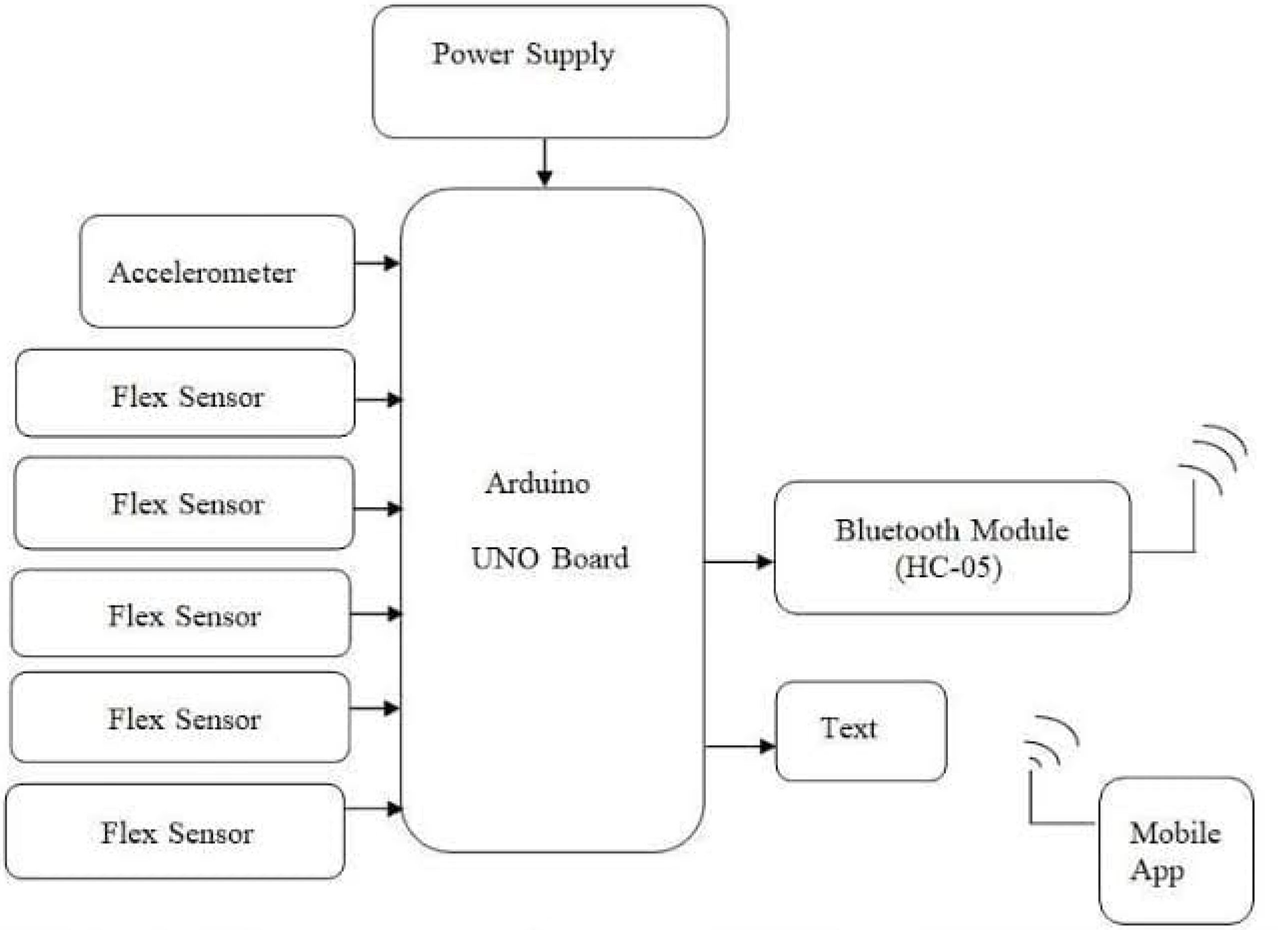
While ensuring comfort and functionality in wearable projects.

- **Casing**: The casing is a protective enclosure made from materials like plastic, silicone, or metal, designed to keep electronics safe from physical damage, dust, and moisture. It often includes mounting points for attaching components and may have ventilation holes or openings for sensors.
- **Wearable Frame**: A wearable frame is designed to comfortably hold sensors, displays, and other electronics on the body. It is typically made from lightweight materials like flexible plastic, fabric, or foam to ensure comfort. It may be in the form of a glove, bracelet, or band for easy integration with the body, especially in projects like gesture recognition or health monitoring.

Fig 8. Casing and wearable frame

4. WORKING PRINCIPLE

The In this project, wearable sensors (like flex sensors and accelerometers) detect hand movements and gestures, which are then processed by the Arduino Uno. The data from these sensors is interpreted to recognize sign language or specific gestures, with the help of programmed algorithms. The Arduino also controls a voice module that converts the recognized gestures into speech, allowing real-time communication for individuals with speech impairments. An LCD display provides visual feedback of the gestures, and relays can be used to trigger external devices if needed. This combination of sensors, processing, and output mechanisms enables seamless sign language translation and voice conversion in a wearable, interactive system.



Programming in Arduino UNO

Programming our Arduino UNO is very much important because it makes our project to move and do thenecessary job that is given. Fig 11. shows the Arduino Code.

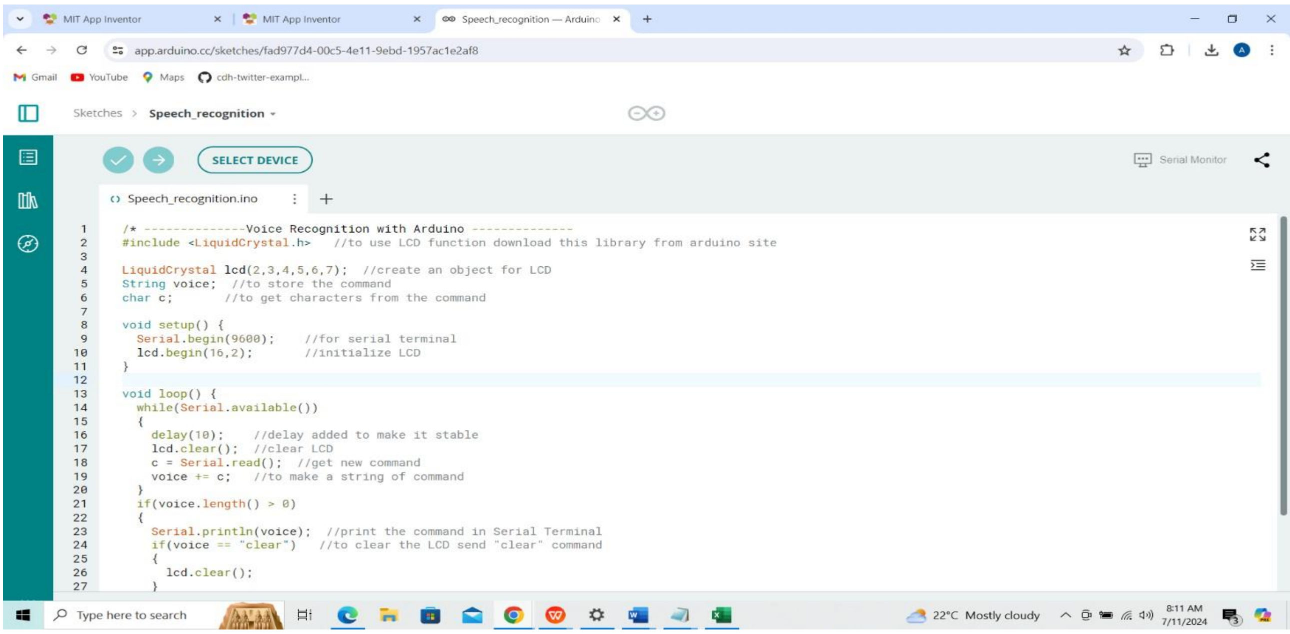


Fig 9. Arduino Code

5. RESULTS

The proposed model of our project is displayed below. Fig 14. Shows the proposed model

Fig 10. Proposed model

The LCD will display recognized sign language gestures or converted speech text, providing visual feedback for users. It will show messages like "Gesture Recognized: [Sign]" or "Voice: [Text]" based on the system's output.

Fig 11. Command displaying in LCD

The next step is to integrate the sensors (flex and accelerometer) with the Arduino Uno, ensuring accurate gesture detection. Then, program the Arduino to interpret the gestures and trigger the voice module or display output accordingly.

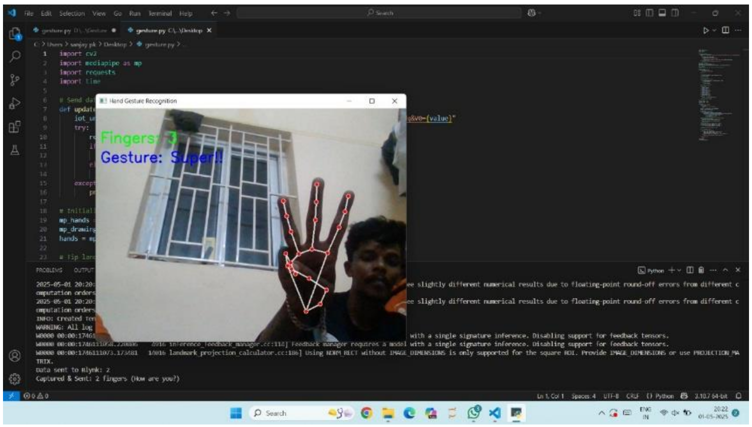


Fig 12. Sensor Integration and Gesture Interpretation

The next step is to connect the voice module to the Arduino and program it to convert recognized gestures into speech. Then, test and calibrate the system for accurate gesture-to-speech translation.
The next step is to integrate the LCD display to show real-time feedback for gesture recognition and voice conversion. Then, test the entire system to ensure smooth communication between the sensors, Arduino, voice module, and display.

Fig 12. Voice Output and Display Integration

The final step is to optimize the system's performance, ensuring reliable gesture recognition, accurate voice conversion, and clear display output. Then, enclose the components in a suitable wearable frame or casing for comfort and durability.

Fig 12. System Optimization and Enclosure

6.DISCUSSSION

This project presents an innovative solution to bridge communication for individuals with speech impairments through a wearable system. By integrating flex sensors and accelerometers, the system can recognize sign language gestures and convert them into audible speech using the voice module. The Arduino Uno plays a critical role as the main processing unit, receiving data from the sensors, interpreting the gestures, and activating the voice output. The system is designed to be user-friendly, offering both visual feedback on an LCD screen and voice conversion, making communication more accessible for users with limited speech capabilities. Additionally, the integration of relays for controlling external devices adds versatility to the project, enabling potential use cases in home automation or assistive technology.

Despite the promising features, challenges include ensuring accurate gesture recognition, especially with varying hand movements and different sign language systems. The system's performance may also depend on external factors like lighting or sensor placement. However, with further development and fine-tuning, such wearable technology can have a significant impact on individuals with speech impairments, offering them greater independence and an easier means of communication. Future improvements could focus on enhancing the system's accuracy, making the wearable frame more comfortable, and increasing the range of gestures recognized for broader accessibility.

7. CONCLUSION

In conclusion, this project explores the potential of wearable technology to assist individuals with speech impairments by converting sign language gestures into audible speech using a combination of flex sensors, accelerometers, an Arduino microcontroller, and a voice module. The system captures hand movements through the sensors, processes the data with the Arduino, and outputs speech via the voice module. Additionally, an LCD display provides visual feedback, enhancing the user experience by displaying recognized gestures or text translations.

Although the system shows promise, challenges remain in ensuring precise gesture recognition and optimizing the accuracy of voice conversion. Environmental factors, sensor placement, and variations in sign language could impact the system's effectiveness, but these hurdles can be addressed through further calibration and fine-tuning. This project has the potential to make a significant impact in the field of assistive technology, offering individuals with speech impairments a more accessible and practical communication tool. Future work could focus on expanding the system's gesture vocabulary, improving its portability and comfort in wearable form, and refining the voice output for clearer and more natural speech. By enhancing these features, the system could promote greater independence, improving the quality of life and enabling more inclusive interactions for users with speech disabilities.

REFERENCE

1. Shinde, Shweta S., Rajesh M. Autee, and Vitthal K. Bhosale presented a novel method titled "Realtime bidirectional communication approach for individuals with hearing impairment and speech impediments leveraging image processing" at the 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), organized by the Institute of Electrical and Electronics Engineers (IEEE).
2. Shangeetha, R. K., V. Valliammai, and S. Padmavathi presented a research paper titled "Computer vision-based approach for Indian Sign Language character recognition" at the 2012 International Conference on Machine Vision and Image Processing (MVIP), organized by IEEE.
3. Sood, Anchal, and Anju Mishra presented a communication system named AAWAAZ designed specifically for individuals who are deaf and mute. This innovative system was introduced at the 2016 5th International Conference on Reliability, Infocom Technologies, and Optimization (ICRITO), organized by IEEE.
4. Ahire, P.G., and colleagues authored a paper titled "Bidirectional Communication System for Deaf and Mute Individuals and Hearing Individuals," which was presented at the 2015 International Conference on Computing Communication Control and Automation (ICCUBEA), organized by IEEE.
5. Kumari, Sonal, and Suman K. Mitra conducted a study on human action recognition utilizing Discrete Fourier Transform (DFT). Their research was presented in a paper titled "Human action recognition using DFT" at the 2011 Third National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics (NCVPRIPG), organized by the Institute of Electrical and Electronics Engineers (IEEE).
6. "Designing and Developing a Hand Gesture Recognition System for Individuals with Speech Impairments" by R. Vinitha and A. Theerthana.
7. Pantic, M., Nijholt, A., Pentland, A., and Huang, T.S. (Year). Advancing Towards Human-Centered Intelligent Human-Computer Interaction (HCI2): Current Progress and Future Prospects.
8. International Journal of Autonomous and Adaptive Communications Systems (IJAACS), Volume 1, Issue 2, published in 2008, with pages numbered 168 to 187. The Digital Object Identifier (DOI) for this publication is 10.1504/IJAACS.2008.019799.