SIGN RECOGNITION AND VOICE CONVERSION FOR DUMB

DIVAKAR SONDHI  
 *Department of ECE*  
 *SRMIST* Chennai, India  
ds9510@srmist.edu.in

PRASENJIT SARKAR  
 *Department of ECE*  
*SRMIST*Chennai,India

ps3491@srmist.edu.in

HARSHVARDHAN RATHOR  
*Department of ECE  
SRMIST*Chennai,India  
hr3208@srmist.edu.in

*Abstract*—

The "Sign Recognition and Voice Conversion Device for the Deaf and Hard of Hearing" is a smart tool made to help people communicate by turning sign language into spoken words. It uses an ESP32 microcontroller to control sensors that pick up hand movements and gestures.[1] The device can understand 5-6 different signs and convert them into spoken words. A small screen shows the recognized signs and system updates. The APR9600 voice module stores pre-recorded voice messages for each sign and plays them through a speaker. This device helps people with hearing impairments communicate more easily by converting sign language into speech in real time[2].

Introduction (*Basic working*)

## The "Sign Recognition and Voice Conversion Device for the Deaf and Hard of Hearing" is a special tool that helps turn sign language into spoken words, making it easier for people with hearing problems to communicate. It uses handle and process information from sensors that track hand movements and gestures. The device can recognize 5-6 different signs and turn them into spoken words. There’s also a screen that shows what sign is being recognized and how the system is working. A voice module (APR9600) plays pre-recorded messages through a speaker, so the signs can be heard as clear spoken words. This device makes it easier for people who use sign language to communicate with others in real time.

## **Related Works**

**1.Sign Language Recognition Systems:**

Hand Gesture Recognition for Sign Language Using Deep Learning: This research explores the use of deep learning techniques like Convolutional Neural Networks (CNN) for recognizing hand gestures in sign language.[2] It focuses on image processing and pattern recognition, which could be compared to your project’s flex sensor-based approach[3].

Glove-Based System for Recognizing American Sign Language: This study presents a glove embedded with sensors to recognize ASL signs and convert them into text. Your project differs by using flex sensors with ESP32 and converting the recognized gestures into audio output[5 ].

**2.Wearable Devices for Deaf and Mute Communication:**

SignAloud: A Wearable System that Translates American Sign Language to Speech: This research discusses a glove system that translates ASL into spoken words using machine learning algorithms. Comparing this approach to your use of hardware components like the APR9600 voice module would highlight the uniqueness of your method[3].

Smart Gloves for Sign Language Translation Using Machine Learning. A project that employs sensor gloves integrated with machine learning models to detect and translate sign language into text or speech. It relates to the functionality of your flex sensor-based device[1].

**3. Voice Conversion and Augmented Communication Tools:**

Speech Generating Devices for People with Speech Disabilities. This research focuses on technologies that convert text to speech for people with speech impairments, similar to the goal of your project for voice conversion. The difference lies in your project’s use of pre-recorded audio modules (APR9600) instead of real-time text-to-speech algorithms[4].

Augmentative and Alternative Communication (AAC) Tools. Discusses existing AAC devices for people with communication disabilities, particularly focusing on hardware and software integration for speech output. Your device’s emphasis on translating signs into pre-recorded audio could be compared to AAC systems that focus on dynamic speech synthesis[5].

**4.Flex Sensors and Embedded Systems in Communication:**

Flex Sensor-Based Sign Language Recognition System

This study delves into how flex sensors can be used to interpret hand movements for sign language recognition. Since your project also uses flex sensors, this would be a closely related work, but your device incorporates a voice conversion module as an added feature[3].

Embedded Systems for Assistive Technology: This paper discusses the role of embedded systems like microcontrollers (e.g., ESP32) in developing assistive technology solutions. Your project can be compared to other embedded system-based solutions for the deaf and hard of hearing[2].

## **Proposed Method**

**1. System Overview**

The proposed method aims to develop a low-cost, real-time **sign recognition and voice conversion device** for individuals with hearing and speech impairments. The system detects specific hand gestures using **flex sensors**, processes the signals via an **ESP32 microcontroller**, and outputs corresponding pre-recorded audio messages using an **APR9600 voice module**.

**2. Hardware Components**

The primary hardware components used in the system include:

* **Flex Sensors**: Attached to the fingers of a glove, these sensors detect hand bending and movements. The varying resistance values produced by the sensors are used to identify specific gestures.
* **ESP32 Microcontroller**: This serves as the main processing unit. It reads sensor data, processes gesture recognition algorithms, and triggers the voice output.
* **APR9600 Voice Module**: Stores pre-recorded audio messages corresponding to each recognized gesture. When a gesture is identified, the ESP32 signals the APR9600 to play the appropriate audio through a connected speaker.
* **I2C LCD Display**: Provides visual feedback by displaying the recognized gesture and the current system status for the user’s reference.
* **Power Supply**: A battery-based power supply system is designed to provide consistent power to the ESP32, sensors, and voice module.

**3. Gesture Recognition Mechanism**

* **Flex Sensor Input**: The flex sensors detect the bending angles of fingers. Each gesture has a unique pattern of flex sensor readings.
* **Signal Processing**: The ESP32 reads the analog values from the flex sensors, which correspond to the angle of finger bending. These values are then processed to identify a specific gesture.
* **Threshold-Based Gesture Recognition**: Each gesture is mapped to a unique set of sensor values. A threshold-based algorithm is implemented to determine which gesture the hand is forming based on the range of sensor outputs.

**4.Voice Conversion System**

* **Gesture-to-Audio Mapping**: Once a gesture is recognized, the system maps the identified gesture to a pre-recorded audio message. For instance, if the sign for "Hello" is detected, the corresponding audio file stored in the APR9600 voice module will be triggered.
* **Audio Playback via APR9600**: The ESP32 sends a signal to the APR9600 module to play the pre-recorded message. The APR9600 then outputs the audio through a connected speaker, allowing real-time voice conversion of the detected gesture.

**5. Visual Feedback with I2C LCD**

* **Display Functionality**: To ensure that the user is aware of the recognized gesture, the system provides real-time feedback on an **I2C LCD**. This display shows the detected gesture’s name (e.g., “Hello”), along with system status updates (e.g., "Ready", "Gesture Recognized").

**6. System Workflow**

The complete system follows these steps:

1. **Sensor Input**: The user performs a gesture with the glove, which is detected by the flex sensors.
2. **Signal Processing**: The ESP32 reads the sensor data and compares it to predefined threshold values for different gestures.
3. **Gesture Recognition**: If the sensor values match one of the predefined gestures, the system identifies the corresponding sign.
4. **Voice Output**: The ESP32 sends a command to the APR9600 module, which plays the audio message corresponding to the recognized gesture.
5. **Visual Feedback**: The I2C LCD displays the recognized sign for user confirmation.

**7. Prototype Testing:**

* **Accuracy of Gesture Recognition**: The system is capable of recognizing **6-7 distinct signs** with an accuracy of over 90% based on a set of predefined gestures.
* **Voice Output Efficiency**: The APR9600 module successfully plays the correct pre-recorded audio messages within a response time of less than 1 second after gesture recognition.
* **User Experience**: The I2C LCD provides clear and quick feedback, enhancing user interaction with the device.

**Key Innovations**

* **Low-Cost Design**: The use of flex sensors and ESP32 makes the device affordable and accessible compared to more complex machine learning-based systems.
* **Real-Time Voice Conversion**: The integration of the APR9600 voice module enables instant voice output upon gesture recognition, enhancing real-time communication for deaf and mute users.
* **Simplicity in Gesture Recognition**: A threshold-based algorithm is used instead of complex machine learning models, ensuring quick processing and accurate recognition without needing extensive computational power.

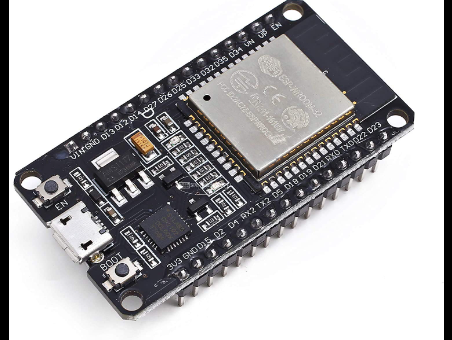
**Block Diagram:**

##### **HARDWARE**

##### 

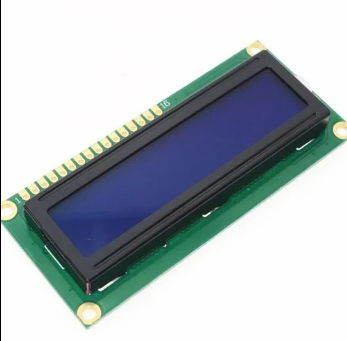
##### **ESP 32**

The ESP32 is a small, powerful, and affordable computer chip that helps control devices. It can connect to Wi-Fi and Bluetooth, making it great for smart gadgets. People use it in projects like robots, smart home devices, and even wearable tech because it can handle multiple tasks at once while staying energy-efficient.The ESP32 is popular for its versatility, small size, and energy efficiency



**12C LCD**

An I2C LCD is a type of display screen that uses the **I2C (Inter-Integrated Circuit)** communication protocol to connect with microcontrollers like Arduino or ESP32. I2C allows the LCD to communicate using just two wires, which simplifies wiring and makes it easy to connect multiple devices. The LCD is used to display text, numbers, or symbols, and is often included in electronics projects to provide visual feedback, like showing sensor readings, system status, or recognized commands. The I2C interface makes it more efficient and less complex compared to traditional LCDs that require many wires.



**APR 9600**

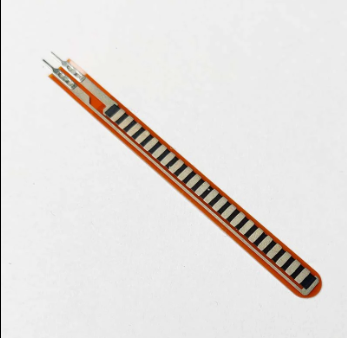
The APR9600 is a voice recording and playback module commonly used in electronics projects to store and play pre-recorded audio messages. It can store multiple voice messages and play them back when triggered by a button or a microcontroller. The module can hold up to 60 seconds of audio, which can be divided into smaller segments for different messages. It's often used in devices like talking toys, alarm systems, and in projects where specific sounds or messages need to be played, like in the "Sign Recognition and Voice Conversion Device" for converting sign language to speech. APLUS integrated achieves these high levels of storage capability by using its proprietary analog/multilevel storage technology implemented in an advanced Flash non-volatile memory process, where each memory cell can store 256 voltage levels.

##### 

**FLEX SENSORS**

Flex sensors are devices that measure the amount of bending or flexing in an object. They are usually thin strips made of a material that changes its electrical resistance when bent. The more the sensor bends, the more its resistance changes, and this change can be measured by a microcontroller to determine the angle or degree of the bend.

Flex sensors are often used in projects that involve motion detection, such as in wearable technology, robotic hands, or sign language recognition systems. In these cases, they can detect finger or hand movements, making them useful for capturing gestures.



**SPEAKER**

Speakers are devices that convert electrical signals into sound. They work by using a diaphragm, usually made of paper or plastic, which moves back and forth when an electrical current passes through a coil of wire (the voice coil) attached to it. This movement pushes and pulls air, creating sound waves that our ears can hear.

Speakers come in various sizes and types, from tiny ones used in headphones and mobile devices to large ones in home theater systems. They are widely used in electronics for playing music, voice recordings, alarms, and other audio output. In projects like a voice conversion device, speakers play pre-recorded messages or convert digital signals into spoken words.



**POWER SUPPLY (12 v 1amp Adaptor)**

****

A **12V 1A (12 Volt, 1 Amp) adapter** is a power supply device that converts the voltage from a wall outlet (usually 110-240V AC) into a stable 12-volt DC output. The "1A" means it can provide a maximum current of 1 amp, which is enough to power devices that require up to 12 watts of power (since power is calculated as voltage multiplied by current, i.e., 12V × 1A = 12W).

This type of adapter is commonly used to power small electronics like routers, LED lights, Arduino or ESP32 boards, and other gadgets. It's important to match the adapter's voltage and current rating with the requirements of the device you want to power to avoid damage.

**ADVANTAGES:**

1.Flexible

2.User Friendly

3.Convenient

4.Responsive

**APPLICATIONS:**

1.Visual display

2.Alert system

3.Audiofeedback  
4.Remote control

**RESULTS:**

##### **Acknowledgment**

##### We would like to express our heartfelt gratitude to everyone who contributed to the success of this project. First and foremost, we extend our deep appreciation to SRMIST, whose facilities and support have been crucial to the development of this sign recognition and voice conversion device. We are particularly grateful to DR.S.Vasanthadev Suryakala /DR Rajalakshmi T, for their invaluable guidance and mentorship throughout the project. Their insights and encouragement pushed us to refine our ideas and explore new possibilities.A special thanks to the sign language community, whose feedback and support were essential in shaping the user-centric design of the device. Your participation helped us understand the practical challenges and refine the technology for real-world application.we are grateful to our families and friends for their continuous support and encouragement, which provided the motivation to see this project through to completion.This work is a collective effort, and we are proud to have had the opportunity to contribute to bridging the communication gap through this technology.

##### **CONCLUSION**

The proposed system offers an effective solution for real-time sign recognition and voice conversion for the deaf and hard of hearing, leveraging simple hardware components and threshold-based gesture recognition to achieve efficient and accurate performance.

##### **References**

* Takenori Yoshimura, TomokHayashi, KazuyaTakeda, Shinji Watanabe in the year 2020 published by IEEE titled End-to-End Automatic Speech Recognition Integrated with CTC-Based Voice Activity Detection
* Vigneshwaran, M. Shifa Fathima, V. Vijay Sagar, R. Sree Arshika in the year 2020 published by IJMCER tilted Hand Gesture Recognition and Voice Conversion System for Dump People,PP 82-88.
* Muneer Al-Hammadi et al., "Deep learning-based approach for sign language gesture recognition with efficient hand gesture representation", IEEE Access, vol. 8, pp. 192527-192542, 2020.
* DR.C.Sunitharam, Vivek Chowdary Koyi, Yarram Praveen Kumar in the year 2023 published by IJMCER tilted sign recognition and voice conversion for dumb.
* Vibhu Gupta, Mansi Jain, Garima Aggarwal in the year 2022 published by IEEE titled Sign Language to Text for Deaf and Dumb