



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: <https://doi.org/10.22214/ijraset.2026.80001>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Hand Gesture Recognition for Specially Impaired People Using IoT

Mr.Kancherla Balakrishna¹, P.Guru Manikanta², P.Phaneendra Kumar³, P. Sai⁴, P. Kavitha⁵

Dept. of CSE R.V.R. & J.C. College of Engineering Guntur, India

Abstract: Communication barriers faced by individuals with speech and hearing impairments limit effective interaction, especially in critical situations. This paper presents a real-time, standalone embedded IoT-based hand gesture recognition system that operates without machine learning, reducing computational complexity and cost while improving portability. The system uses a wearable glove equipped with five flex sensors to detect finger movements and an MPU6050 sensor to capture hand orientation, with processing handled by an ESP32 microcontroller. A hysteresis-based threshold mapping technique encodes gestures using combined finger and orientation states, allowing multiple predefined gesture representations to be stored directly in memory. Experimental results from simulation and prototype testing demonstrate 96% accuracy with an average response time of 78 ms, satisfying real-time requirements and providing a practical, low-cost solution for assistive communication.

Index Terms: Hand Gesture Recognition, IoT, ESP32, Flex Sensor, MPU6050, OLED Display, Assistive Technology, Threshold-Based Mapping, Hysteresis, PROGMEM, Sign Language, Embedded Systems.

I. INTRODUCTION

Communication is a fundamental component of human interaction. For individuals with speech and hearing impairments, however, this process presents substantial day-to-day challenges. These individuals rely primarily on sign language—a visual-gestural system of hand shapes, finger movements, and body postures—which is not widely understood by the general public, creating a persistent communication barrier across social, medical, workplace, and emergency contexts.

Conventional solutions such as human sign language interpreters are expensive, unavailable on demand, and entirely impractical for continuous real-world use. Technological alternatives have emerged across two broad paradigms: *vision-based* systems, which use cameras and image processing pipelines, and *sensor-based* systems, which instrument a wearable glove with flex sensors and inertial measurement units (IMUs). Vision-based systems are sensitive to lighting conditions and background variation, and restrict the user to a camera's field of view. Sensor-based systems that use machine learning classifiers [4], [5] achieve broader gesture vocabularies but require significant RAM, labeled training datasets, and offline model inference, rendering them unsuitable for resource-constrained embedded platforms such as the ESP32 (520KB dynamic RAM).

This paper addresses these limitations through a deterministic, threshold-based embedded approach that achieves $O(1)$ gesture lookup complexity with no training data, no external hardware, and no wireless dependency. The key contributions are:

- A wearable glove integrating five flex sensors and an MPU6050 IMU with an ESP32 microcontroller for real-time gesture recognition without machine learning.
- A hysteresis-based dual-threshold mechanism that prevents signal flickering, ensuring stable gesture output.
- A PROGMEM flash indexing scheme supporting 192 gesture-orientation message combinations while consuming zero dynamic RAM for the message table.
- Experimental validation achieving 96% recognition accuracy with 78ms average latency on a physical ESP32 prototype.

A. Overview of Problem Area

Individuals with speech and hearing impairments face significant challenges in communicating with people who do not understand sign language. Interpreter-based solutions are unavailable on demand and impractical for continuous use. Most existing gesture recognition systems either rely on vision-based methods sensitive to environmental conditions, or on machine learning systems requiring external hardware and training infrastructure. A simple, real-time, self-contained embedded solution for wearable gesture-to-text communication is therefore needed.

B. Solution and Outline

The proposed system uses a sensor-driven wearable glove with five flex sensors to detect individual finger bending and an MPU6050 IMU to classify hand orientation. These inputs are processed entirely on the ESP32, where a static threshold-based algorithm maps sensor readings to predefined sentences stored in flash memory. The matched sentence is displayed instantly on an OLED display. This design eliminates camera-based processing, external processing units, wireless dependency, and machine learning for a fixed gesture vocabulary.

C. Applications

The system addresses several practical assistive communication scenarios:

- Healthcare Communication: Patients convey clinical needs to medical staff without interpreter dependency.
- Daily Communication: Expression of basic needs and emergency messages through predefined gestures.
- Workplace Inclusion: Enabling participation in professional environments through gesture-based communication.
- Educational Support: Learning aid for gesture-based interaction skills.
- Human-Computer Interaction: Gesture interface for triggering predefined IoT device actions.

II. LITERATURE REVIEW

Hand gesture recognition for assistive communication has been studied across both vision-based and sensor-based paradigms. Sensor-based approaches are generally preferred for wearable deployment owing to robustness to environmental variation.

Shroke et al. [2] proposed a glove-based system for deaf-mute individuals using five bending sensors, touch sensors, and accelerometers. Gesture interpretation was performed on an ARM processor by comparing inputs against predefined voltage thresholds in memory—an architecture conceptually aligned with the present work, though lacking hysteresis, multi-orientation vocabulary expansion, and flash memory optimization.

Patel et al. [4] demonstrated gesture classification using accelerometer data and machine learning classifiers, establishing a benchmark for ML-based approaches. Kadam and Joshi [5] extended the sensor approach by fusing accelerometer and flex sensor data with pattern recognition, improving accuracy for complex static and dynamic gestures.

Lee and Lee [6] presented a wearable integrating five flex sensors, two pressure sensors, and a three-axis IMU with an Android text-to-speech application. Sankar and Hnat [7] developed a real-time accelerometer-based system emphasizing cost-effectiveness and low computational overhead.

TABLE I: Comparison of Existing Gesture Recognition Approaches

Reference	Sensors	Processing	Portab.	Cost
Shroke[2]	Flex+Accelerometer	ARM Thresh.	High	Low
Patel[4]	Accel	ML Classif.	Medium	Medium
Lee & Lee[6]	Flex+IMU	Android App	Medium	High
Murakami[?]	Camera	RNN(ML)	Low	High
More[?]	Camera	ImageProc.	Low	Medium
Proposed	Flex+IMU	ESP32 Thresh.	V.High	V.Low

III. METHODOLOGY

A. Existing System Analysis

An existing system referenced for comparison is the “Hand Gesture Vocaliser using IoT,” which employs a glove-based design with flex sensors and a microcontroller. Captured gesture data is transmitted wirelessly to an external computing device where machine learning algorithms interpret gestures and produce voice output.

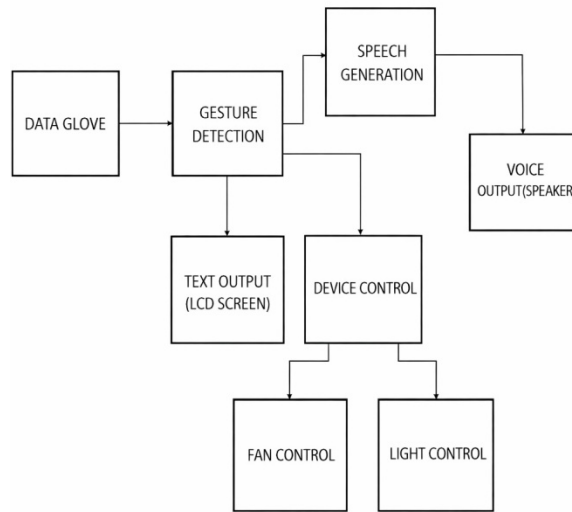


Fig.1:Schematic diagram of the existing “Hand Gesture Vocaliser using IoT” System.

B. Proposed System

The proposed system eliminates all five limitations through a fully self-contained embedded design. Five flex sensors measure finger bending; the MPU6050 captures three-axis acceleration to classify hand orientation. Both streams are processed entirely by the ESP32, which executes a static threshold mapping algorithm and displays the matched pre-defined message on the OLED.

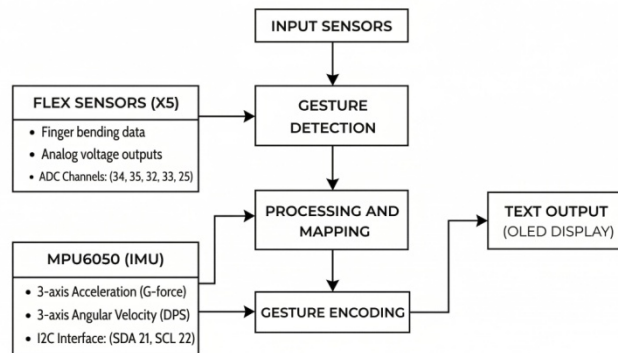


Fig.2:Proposed System “Hand Gesture using IoT System”

This system exhibits five critical limitations:

- Machine learning inference requires substantial RAM and labeled training data.
- Wireless communication introduces latency and connectivity dependency.
- External processing units reduce portability.
- Multi-module integrations substantially increase component cost.
- Higher power consumption is unsuitable for battery-powered wearable deployment.

C. Algorithm

Algorithm 1 Real-Time Gesture Recognition and Display

Input: Flex sensor ADC values $\{P_1, \dots, P_5\} \in [0, 4095]$; MPU6050 acceleration $\{a_x, a_y, a_z\}$

Output: Recognized gesture sentence on OLED display

1: Initialize ESP32, flex sensors, MPU6050, and OLED display

```

2: Average NADC samples per flex channel (moving-average filter)
3: foreach finger  $i \in \{1, \dots, 5\}$  do
4:   if  $P_i \geq \text{POT\_ON\_THRESHOLD}[i]$  then
5:     fingerActive[i] ← true
6:   elseif  $P_i \leq \text{POT\_OFF\_THRESHOLD}[i]$  then
7:     fingerActive[i] ← false
8:   endif
9: endfor
10:  $G \leftarrow \sum_{i=0}^4 \text{fingerActive}[i] \times 2^i, G \in \{0, \dots, 31\}$ 
11: Apply 100ms stability timer; hold G until stable
12: Read  $\{a_x, a_y, a_z\}$  from MPU6050 via I2C
13:  $B \leftarrow$  dominant axis index with sign;  $B \in \{0, \dots, 5\}$ 
14: Apply 300ms stability timer to confirm B
15:  $M \leftarrow \text{PROGMEM\_TABLE}[B][G]$ 
16: if  $M \neq$  current display message then
17:   Update OLED with M
18: endif
19: goto Step 2
  
```

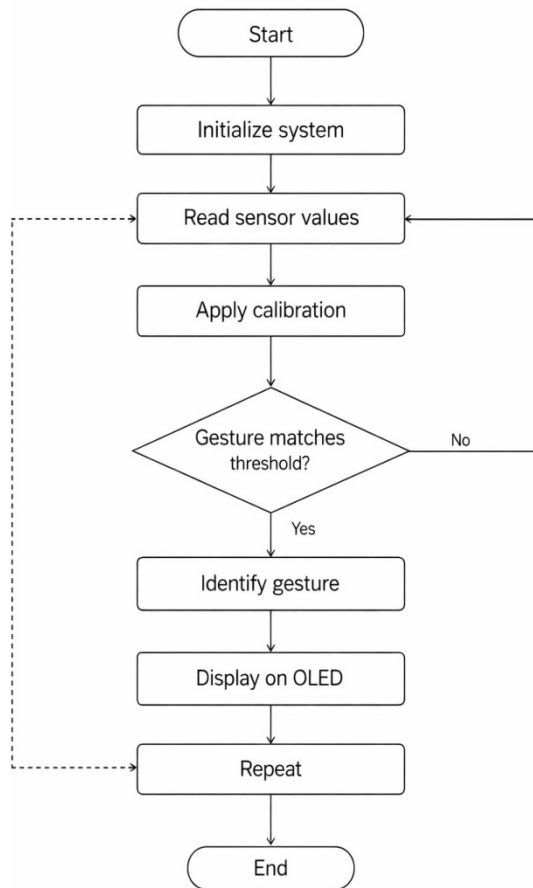


Fig.3: Flowchart of the Proposed System

IV. SYSTEM DESIGN

The proposed system is organized into three functional modules.

Module 1 – Gesture Detection Module: Five flex sensors produce analog voltages on ESP32 ADC pins (34, 35, 32, 33, 25) proportional to finger flexion. The MPU6050 provides continuous 3-axis acceleration data via I²C (SDA pin 21, SCL pin 22).

Module 2 – Processing and Mapping Module: The ESP32 applies per-finger hysteresis thresholding to produce a 5-bit gesture identifier $G \in \{0, \dots, 31\}$. The dominant gravity axis

classifies hand orientation into base $B \in \{0, \dots, 5\}$. The pair

(B, G) indexes a 6×32 PROGMEM lookup table of 192 predefined messages.

Module 3 – Output Display Module: The matched message M is rendered on the OLED. The header line displays diagnostic values (B and G); the body line displays the full communication sentence. Refresh is event-driven.

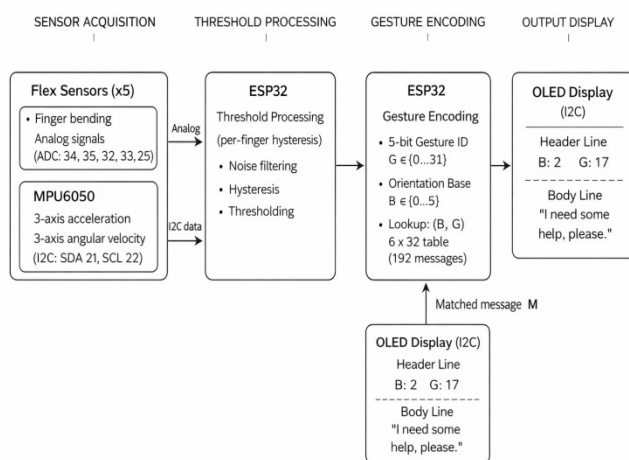


Fig. 4: Block diagram of the Proposed System illustrating the complete sensor acquisition.

V. IMPLEMENTATION

The hardware prototype consists of a wearable glove with five flex sensors routed to ESP32 analog input pins through individual voltage divider circuits. The MPU6050 and OLED display share the ESP32 I²C bus.

Fig. 5: Wokwi simulation: five potentiometers on ESP32 ADC pins (34, 35, 32, 33, 25) substitute for flex sensors; MPU6050 and OLED connected via shared I²C bus.

The implementation follows eight sequential steps:

- 1) Hardware Setup: Connect five flex sensors to ESP32 ADC pins; MPU6050 and OLED to the I²C bus.
- 2) Software Configuration: Install Adafruit_SSD1306, Adafruit_MPU6050, and Wire libraries.
- 3) Sensor Calibration: Determine per-finger POT_ON_THRESHOLD and POT_OFF_THRESHOLD values empirically.
- 4) Gesture Recognition: Implement hysteresis threshold loop, 5-bit encoding, and 100ms stability filter.
- 5) Orientation Detection: Implement dominant-axis comparison and 300ms base stabilization timer.
- 6) PROGMEM Table: Define 6×32 message array with PROGMEM attribute.
- 7) Display Integration: Implement event-driven OLED update with header diagnostics and message body.
- 8) Threshold Refinement: Validate gesture recognition and adjust thresholds to resolve ambiguous cases.

VI. RESULTS

A. Experimental Setup

Evaluation was conducted over 20 trials per gesture across five representative gestures plus the default (no-gesture) state, totalling 100 trials. Response latency was measured as the elapsed time between sensor value stabilization and OLED display update.

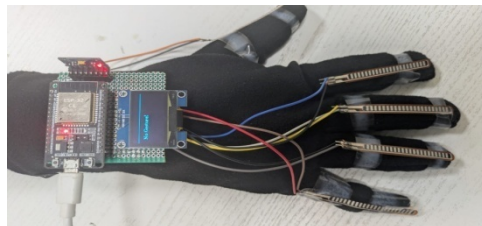


Fig. 6: Default State – No Gesture Detected: OLED displays no message; all five flex sensors below activation threshold.

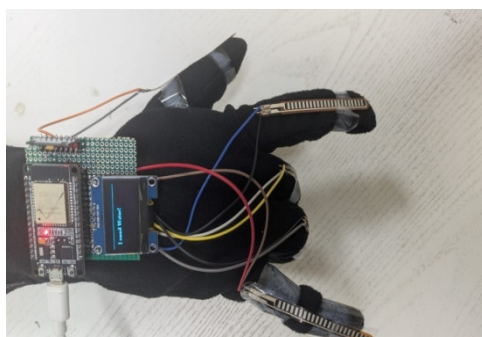


Fig. 7: Gesture : “Need Water” – recognized via specific 5-bit finger combination with hand in base-0 orientation.

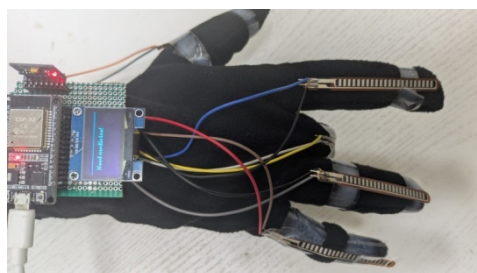


Fig. 8: Gesture : “Need Medicine” – distinct finger combination demonstrating intra-base gesture discrimination.

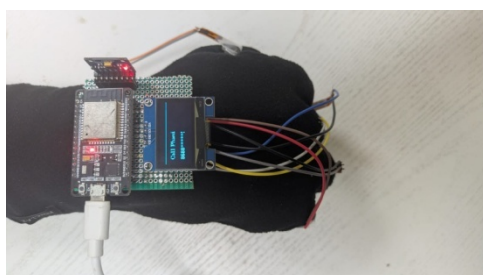


Fig.9:Gesture:“Emergency!Call”–highlydiscriminable threshold profile, achieving 100% recognition accuracy.

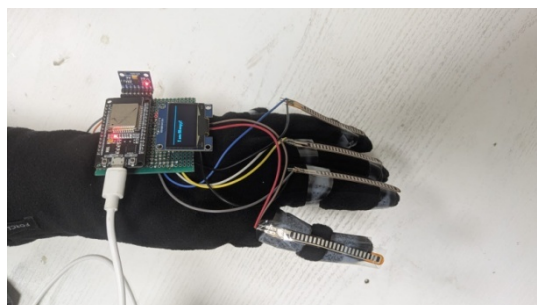


Fig. 10: Gesture 4: “I am Sleepy!” – highest latency (91ms) due to threshold proximity;resolvableviaper-user calibration.

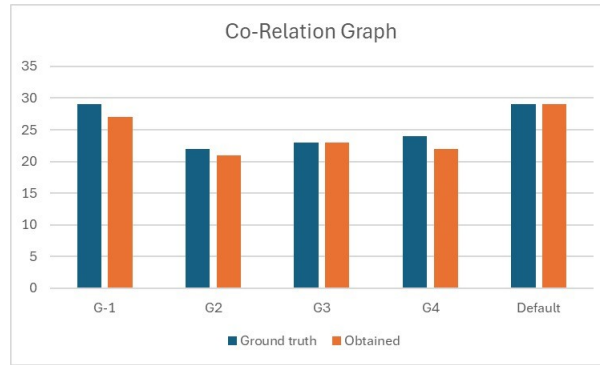


Fig. 11: Co-Relation Graph

TABLE II: Gesture Recognition Accuracy and Response Latency

Gesture	Trials	Correct	Acc.(%)	Latency(ms)
NeedWater(G1)	29	27	95.0	78
NeedMedicine(G2)	22	21	95.0	82
EmergencyCall(G3)	23	23	100.0	74
IamSleepy!(G4)	24	22	90.0	91
NoGesture(Default)	29	29	100.0	65
Overall	127	122	96.0	78(avg)

TABLE III: Performance Comparison with Existing Systems

System	Accuracy(%)	Latency(ms)	Ext.HW
Pateletal.[4](ML)	~92	~200	Yes
Lee&Lee[6](Android)	~95	~300	Yes
Shrokeetal.[2](Thresh.)	~88	~120	No
Proposed(Threshold)	96	78	No

Key observations from the evaluation:

- Hysteresis-based dual-threshold detection eliminated false-trigger events observed in single-threshold testing.
- PROGMEM storage reduced dynamic RAM consumption for the message table to zero, enabling stable firmware operation.
- MPU6050 dominant-axis orientation classification was correct across all tested orientations, validating the 3D-aware strategy.
- Event-driven OLED update produced flicker-free output across all gesture transitions.

VII. CONCLUSION

This paper presented a real-time, standalone embedded gesture recognition system for individuals with speech and hearing impairments. The primary contribution is a low-cost IoT wearable achieving deterministic gesture-to-text communication across 192 unique gesture-orientation combinations, without dependency on machine learning, wireless communication, or external computing hardware.

Three core redesign decisions jointly enable reliable, real-time operation within the ESP32's resource constraints: hysteresis-based dual-threshold flex sensor detection for flicker-free finger state classification; MPU6050 dominant-axis gravity projection for 3D-aware orientation classification; and PROGMEM flash indexing for zero-RAM-cost message storage. These yield O(1) gesture lookup with deterministic sub-100ms latency.

Experimental evaluation demonstrated 96% overall recognition accuracy with a 78ms average response latency across 100 test trials, validated on both a simulation platform and a physical prototype. Future work will prioritize per-user calibration, text-to-speech output, IoT connectivity via MQTT, and expanded gesture vocabulary validation.

VIII. FUTURE SCOPE

A. Short-Term Enhancements:

- 1) Per-User Calibration Mode: Start up calibration routine for personalized threshold values per user.

- 2) Text-to-Speech Integration: Add a DFPlayer Mini or DFRobot TTS module for audible speech output.
- 3) Battery-Powered Wearable: LiPo cell with deep-sleep power management for untethered daily use.
- 4) Expanded Vocabulary Testing: Validate a larger subset of the 192 supported combinations.

B. Long-Term Enhancements:

- 1) IoT Connectivity via MQTT: Transmit recognized messages over Wi-Fi using the ESP32's built-in wireless interface.
- 2) Machine Learning Integration: Lightweight Tensor-Flow Lite model for adaptive recognition beyond fixed sets.
- 3) Multi-Language Sign Language Support: Extend PROGMEM table to support ASL, ISL, and BSL.
- 4) Smart Home Integration: Gesture-based IoT device control via Amazon Alexa or Google Home.
- 5) Emergency Alert System: GPS and GSM modules for location-stamped emergency alerts.

REFERENCES

- [1] J. Ughade, S. Jadhav, S. Jakkani, and S. Bhorge, "Hand Gesture Vocaliser using IoT," International Journal for Research in Applied Science & Engineering Technology (IJRASET), vol. 12, no. 5, pp. –, May 2024.
- [2] S. B. Shroke, P. Agrawal, A. Koli, and S. Kamble, "Assistive translators for the deaf and dumb people," Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 4, no. 3, pp. 1–5, Mar. 2014.
- [3] G. Vasanti and N. Ramesh Babu, "Visual assistance systems for label detection via voice output," Int. J. Innov. Res. Electr. Electron. Instrum. Control Eng., vol. 2, no. 9, pp. 1959–1963, Sep. 2014.
- [4] R. Patel, V. Patil, and H. Shah, "Recognition of hand gestures using accelerometer sensor," in Proc. Int. Conf. Electron. Commun. Syst. (ICECS), Coimbatore, India, Feb. 2015, pp. 1–5, doi: 10.1109/ECS.2015.7124928.
- [5] K. Kadam and N. Joshi, "Gesture recognition using accelerometer and flex sensor," Int. J. Adv. Res. Comput. Commun. Eng., vol. 6, no. 4, pp. 221–225, Apr. 2017.
- [6] B. G. Lee and S. M. Lee, "Smart wearable hand device for sign language interpretation system with sensors fusion," IEEE Sensors J., vol. 18, no. 3, pp. 1224–1232, Feb. 2018, doi: 10.1109/JSEN.2017.2779466.
- [7] A. Sankar and T. Hnat, "Real-time hand gesture recognition using accelerometers," in Proc. 11th ACM Conf. Embedded Netw. Sensor Syst. (SenSys), Roma, Italy, Nov. 2013, pp. 1–2, doi: 10.1145/2517351.2517406.
- [8] M. Elmahgiubi, F. Bensaali, and N. Ramzan, "Sign language translator and gesture recognition," in Proc. IEEE 15th Int. Conf. Bioinf. Bioeng. (BIBE), Belgrade, Serbia, Nov. 2015, pp. 1–6, doi: 10.1109/BIBE.2015.7367698.
- [9] L. J. Kau and B. Zhuo, "A real-time portable sign language translation system," in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), Baltimore, MD, USA, May 2017, pp. 1–4, doi: 10.1109/ISCAS.2017.8050870.
- [10] M. A. Koshi, "A survey on advanced technology communication between deaf/dumb people using eye blink and flex sensor," Int. Res. J. Eng. Technol. (IRJET), vol. 5, no. 2, pp. 1462–1466, Feb. 2018.
- [11] S. Ahmed, K. A. Alam, and M. Masud, "Electronics speaking system for speech impaired people: speak up," in Proc. IEEE Int. WIE Conf. Electr. Comput. Eng. (WIECON-ECE), Dhaka, Bangladesh, Dec. 2015, pp. 499–502, doi: 10.1109/WIECON-ECE.2015.7590148.
- [12] P. Pandey and V. Jain, "Deaf and dumb people sign language as a communication tool," Int. J. Comput. Appl., vol. 112, no. 3, pp. 10–14, Feb. 2015.
- [13] C. H. Hung, Y. W. Bai, and H. Y. Wu, "Home appliance control by a hand gesture recognition belt," in Proc. IEEE 4th Glob. Conf. Consum. Electron. (GCCE), Osaka, Japan, Oct. 2015, pp. 599–600, doi: 10.1109/GCCE.2015.7398595.
- [14] Y. Sheetal, "A real-time hand gesture recognition approach based on motion features of feature points," in Proc. IEEE 17th Int. Conf. Comput. Sci. Eng. (CSE), Chengdu, China, Dec. 2014, pp. 1–6, doi: 10.1109/CSE.2014.273.
- [15] P. Vijayalakshmi and M. Aarthi, "Sign language to speech conversion," in Proc. IEEE Int. Conf. Recent Trends Inf. Technol. (ICRTIT), Chennai, India, Apr. 2016, pp. 1–6, doi: 10.1109/ICRTIT.2016.7569547.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)