International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-Jun-2026, ISSN (Online): 2395-566X

Design and Implementation of a Wearable Multimodal Hand Gesture Vocalizer for Assistive Communication

Aryan Patel, Era Mane, Tanvi Sonawane, Dr. Vineeta Philip
Department of Electronics and Telecommunication Engineering AISSMS Institute of Information Technology, Pune
Savitribai Phule Pune University, India

Abstract- The persons who have difficulties communicating with society because of hearing and speech problems face challenges that restrict their communication ability with society. While sign language is one possible solution, everyone may not comprehend it. The design and application of a low-cost wearable hand gesture vocalizer that produces both visual and aural outputs from predefined hand gestures are presented in this paper. The suggested system uses flex sensors built into a glove to record finger movements, which are then operated by an Arduino Nano microcontroller. An external audio replay module is used to translate recognized gestures into corresponding audio output and display them on a 16x2 I2C-based LCD at the same time. To improve approachability for users with hearing or vision impairments, the system places a strong emphasis on multimodal feedback, appropriate wearable design and un- complicated hardware. The prototype developed shows reliable gesture identification with minimal latency and offers a extensible platform for future improvements.

Keywords- Hand Gesture Recognition, Assistive Technol- ogy, Flex Sensor Glove, GSM, Multimodal Communication, Wearable Devices.

I. INTRODUCTION

Communication is the most basic human needs; people who have speech and hearing impairment often face remarkable challenges in communicating their needs. Sign language offers a efficient way of communicating, but it has the limitation that both communicating parties must know the signs used. This therefore creates a communication barrier between the differently-abled and the rest of the population, especially in situations that need urgent attention.

Advances in embedded systems and wearable technology have facilitated the development of assistive devices that mitigate this gap by translating hand gestures into intelligible outputs. Although vision-based gesture recognition systems provide high accuracy, they are associated with a number of drawbacks: high computational cost, sensitivity to lighting conditions, and limited portability. On the other hand, wearable sensor-based systems offer a compact, real-time, and cost- effective alternative for continuous daily use.

This paper proposes a wearable hand-gesture vocalizer glove, which converts predefined hand gestures into text and speech outputs. Based on flex sensors integrated on a glove, the system combines with a compact processing unit for portability and ease of use. Multimodal feedback allows visually- and hearing-impaired users to be supported, while a GSM-based communication module allows remote emergency alerts, whereby the device can automatically notify caregivers or an emergency contact when a user uses a user-defined distress gesture. This feature moves the device beyond that of a local communication aid and provides a safety-enabled assistive system for real-world deployment.

II. RELATED WORK

Various approaches to hand gesture identification have been explored in the existing literature. Vision-based systems use cameras and image processing techniques to recognize hand gestures; however, these systems require high computational resources and are sensitive to environmental conditions such as background clutter and light variations.

Sensor-based approaches using flex sensors, inertial measurement units (IMUs), or electromyography (EMG) sensors have gained popularity due to their mobility and lower computational needs. Several glove-based systems using flex sensors and microcontrollers have been proposed to interpret hand gestures into text. However, many such systems either rely on complex hardware configurations, lack audio feedback or do not prioritize appropriate wearable design.

The proposed system differentiates itself by engaging a reduced number of flex sensors, a small sized wearable layout supported by a registered design patent, and dual-mode output in the form of both text and audio, thereby improving functionality and availability.

III. SYSTEM ARCHITECTURE

Four functional layers make up the modular embedded architecture of the suggested hand gesture vocalizer system: sensing, processing, output, and power management. Scalability, simplicity of execution, and future flexibility are all improved by this modular design.

The gesture sensing layer is made up of four flex sensors that are built into a glove that you can wear on your index, middle, ring, and small fingers. These sensors can tell when flexing of a finger through altering the resistance value based on the finger movements. All the sensors will be configured in such a manner that they generate analog voltage values based on the degree of finger bending.

Fig. 1: Design representation of the wearable hand gesture vocalizer glove.

The processing component employs an Arduino Nano microcontroller as its primary controller device. This microcontroller contains an ADC which allows it to continuously read values from the analog sensors. The classification algorithm employed to interpret sensor data will be based on threshold detection, thereby enabling recognition of hand gesture commands.

Feedback is provided in more than one manner by the output layer. There is a 16x2 LCD which is connected to the device using an I2C module. This simplifies the cabling of the device. There is also an external audio output unit which provides feedback by playing pre-recorded voice messages corresponding to gestures made.

The power management layer allows the system to operate independently or to be worn. This is achieved through the use of an external power source which supplies controlled power to each component of the system. The overall design makes sure that communication is easy for users, responses are quick, and gestures are recognized in real time.

IV. HARDWARE DESIGN

Compactness within a wearable form factor, low power consumption, and appropriate integration are given top importance in the hardware execution of the proposed system. The primary sensing elements are four medium-sized flex sensors placed along the dorsal side of the fingers. The thumb is eliminated in order to reduce hardware complexity while maintaining sufficient gesture differentiation.

A voltage divider circuit is designed by connecting each flex sensor in series with a 10 kilo-ohm resistor. The supply voltage is linked to one of the flex sensor’s terminals, and the Arduino Nano’s analog input pin is connected to the junction of the flex sensor and resistor. Ground is connected to the remaining terminal. Accurate measurement of voltage, which is proportional to the bending of the finger, can be achieved using this circuit setup.

Arduino Nano was selected since it is compact in size, consumes low power, and has numerous analog input pins. It processes data from sensors and controls output devices. A 16x2 LCD with an I2C interface is used for visual output to keep wiring simple and keep the digital I/O pins working.

The microcontroller is connected to an external audio playback module in a serial way so that it can give audio feedback. Recognized gestures turn on pre-recorded audio files that are stored on the device. Putting all the electronic parts in the back of the glove makes it more comfortable and balanced for the user. This physical arrangement is backed up by a registered design that shows off the wearable’s unique layout over its internal circuitry.

V. GESTURE RECOGNITION

A prototype version of the suggested hand gesture recog- nition framework is represented by the gesture mapping dis- played in Table I. A 4-bit binary representation, where each bit represents the bending state of a particular finger, is used to define each gesture. A bent finger is represented by a binary value of ”1,” whereas a straight finger is represented by a value of ”0.” Each hand posture can be represented in a condensed and organized form using this binary encoding, such as 1000, 1100, 1110, and 1111.

Four gestures—WATER, HELP, FOOD, and EMERGENCY—are used in the current prototype to show how the recognition algorithm operates.This implementation, however, is not restricted to four commands. A total of 2*4=16 distinct gesture combinations could theoretically be made with just four fingers because the system employs a binary finger- state encoding. Without changing the hardware, more gestures, words, or entire sentences can be incorporated by expanding the software mapping table.

The Arduino Nano samples the analog voltages produced by the flex sensors to start the gesture recognition process. The microcontroller’s ADC is used to convert these voltages, which change according to finger bending, into digital values. Each finger is then classified as either bent or straight using predetermined threshold values. To identify the matching gesture, the resulting binary pattern is compared to a lookup table, as indicated in Table I.

If a valid gesture pattern is determined, the corresponding text message is shown in the LCD and its audio output is activated. This method allows the same input to generate visual and auditory feedback concurrently, providing multi- modal communication. Modularity of the lookup table allows for straightforward addition of vocabulary, while maintaining compatibility for continued future expansion to include multi-word phrases, commands, and full sentences making the system fully scalable beyond simple keyword-based commu- nication.

module.By adding more audio files and updating the gesture-to-track lookup table, this mapping can be expanded to support entire sentences or multi-word phrases. Therefore, without changing the sensing hardware, the suggested audio subsystem can be scaled to phrase-level communication and is not limited to keyword-level output.

VII. GSM MODULE INTEGRATION

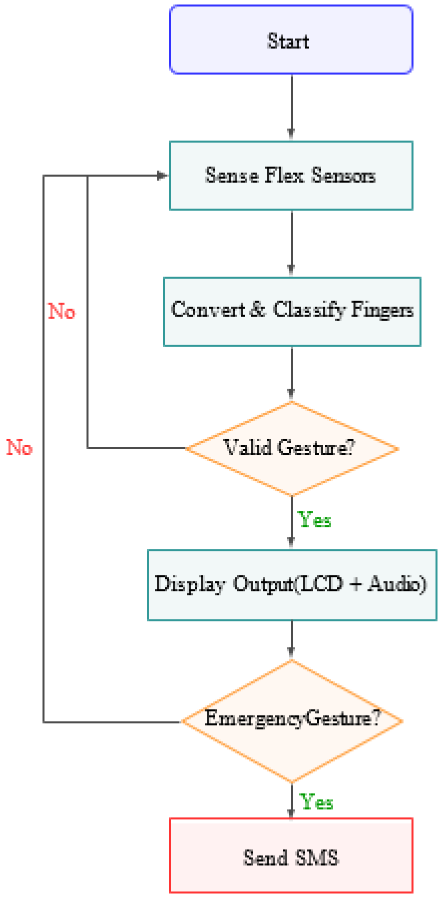

Fig. 2: Gesture recognition system flowchart with emergency SMS alert

VI. AUDIO OUTPUT INTEGRATION

To enhance the accessibility and user-friendliness of the telephony system, it has been built to include anATT Being telecommunications terminals in conjunction with a multi-modal feedback by supplying both visual (textual output) and aural (speech output) []. When a gesture is detected, the I2C-interfaced 16×2 LCD shows the relevant message and an audio output module plays out pre-recorded voice prompt for that corresponding gesture. This two-mode output provides a means of communication in instances where either the user or listener may possess sensory limitations, such as for instance when a hearing-impaired user is able to simply read the LCD output while a visually-impaired listener can listen.

Instead of using cloud-based text-to-speech services, the prototype uses an offline audio playback method for speech output. For embedded wearable assistive systems, offline playback offers three useful benefits: (i) minimal latency and deterministic response; (ii) independence from external networks; and (iii) lower microcontroller computational demands. The microcontroller initiates a track index that corresponds to the recognized gesture label (e.g., WATER, HELP, FOOD, EMERGENCY), and each recognized gesture is mapped to an audio file stored locally on the playback

A GSM-based communication module is added as an additional output layer to enhance the wearable system's usability and safety. The system's primary purpose is to translate hand gestures into speech and text. However, it's crucial to stay in touch with distant family members or caregivers during emergencies. When it recognizes a particular emergency gesture, the GSM module enables the system to send text message alerts.

In this configuration, emergency signals are represented by a specific gesture pattern (e.g., 1111, where all monitored fingers are bent). The Arduino Nano uses serial AT commands to activate the GSM module when the gesture classification algorithm detects this pattern. The GSM module then notifies the caregiver that the user requires urgent assistance by sending an alert message to a preconfigured mobile number.

The user can request assistance using this method without requiring a smartphone, an internet connection, or any other external devices. The GSM-based alert system transforms the gadget from a tool for local communication into a safety-enabled assistive device that can offer remote assistance in dire circumstances. In order to prevent it from interfering with gesture sensing or recognition, the GSM module is also included in the output layer.

The GSM interface is appropriate for both demonstrations and practical use because of its modular design, which enables users to enable or disable it through software settings. The suggested system's dependability and usefulness are significantly increased by this integration.

VIII. USER PERCEPTION AND NEED ANALYSIS

The necessity, usability, and acceptance of the suggested assistive system were evaluated through a user perception survey. The public's awareness of the communication difficulties faced by people with speech and hearing impairments and the perceived value of a glove-based gesture translation device were assessed using an online survey.

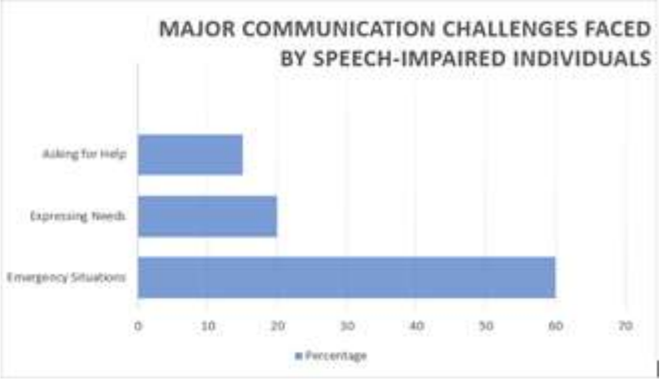

Fig. 3: Major communication challenges faced by speech-impaired individuals based on user perception survey.

Survey results showed that many respondents had interacted with speech- or hearing-impaired individuals, and about 90% agreed that communication with them is difficult for the general public, confirming the real-world relevance of the problem. The most significant challenge was found to be emergency situations, which were followed by challenges in communicating needs and asking for assistance. Additionally, the majority of respondents mentioned that sign language is not commonly understood, highlighting the necessity of a translation-based tool.

Over 95% of participants agreed that a gesture-to-text and voice glove would be very helpful, indicating a strong acceptance of the suggested solution. Choice of multimodal communication system was justified as a result of preferences exhibited by the majority regarding usage of both the visual display screen and audio. In addition, a large proportion was in favor of the emergency SMS system based on GSM technology.

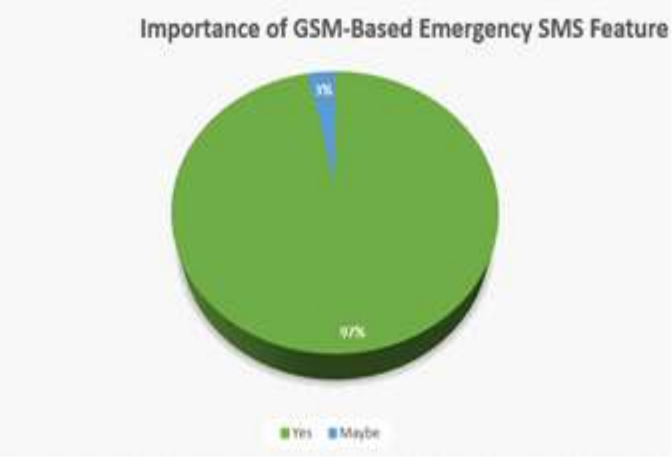

Fig. 4: Survey results showing user acceptance of the GSM-based emergency SMS alert functionality.

The survey focused on awareness, utility, and emergency functionality and was administered to professionals, students, and general users using anonymous multiple-choice questions. Overall, the findings show a high level of user acceptance, societal need, and potential for practical adoption. The results successfully support the system’s design goals and support additional development with target users, despite being constrained by a small and non-targeted sample.

IX. RESULTS AND DISCUSSION

A. Prototype Gesture Set

The gesture mapping utilized in the current prototype is shown in Table I. A proof-of-implementation of the recogni- tion pipeline is represented by the four implemented gesture codes (1000, 1100, 1110, and 1111). The system architecture allows for expansion even though the current prototype only demonstrates four gestures. With four binary finger states, up to 24 = 16 unique combinations are possible, and per- user calibration and software updates can incorporate more gestures.

B. Evaluation Methodology

The user wearing the glove conducts repeated trials to assess the prototype. For every class of gesture, (i) Raw sensor values from the ADC, (ii) classified finger states, (iii) recognized output label, and (iv) LCD and audio trigger events are all recorded during N trials. Metrics for response time and recognition accuracy are used to report performance.

C. Recognition Accuracy

The ratio of correctly identified gestures to all gesture attempts is used to calculate recognition accuracy.

D. Response Time and Output Latency

The time between the formation of a gesture (stable finger position) and the generation of an output is known as the system response time. (i) ADC sampling; (ii) threshold classi- fication; (iii) lookup-based gesture mapping; (iv) LCD refresh; and (v) audio trigger comprise the total end-to-end latency. The prototype’s audio playback trigger delay was [XX ms], and the average response latency from gesture stabilization to LCD update was [XX ms]. These numbers show that the system is suitable for assistive communication in real time when prompt feedback is needed.

E. Discussion and Practical Observations

Threshold-based binary encrypting can be applied to low-power microcontroller platforms due to its computationally light nature. I2C LCD communication is yet another positive aspect of the system since it reduces the pin usage and makes connection easier. However, this approach should be altered due to the fact that the baseline values obtained from the sensor are dependent on many factors such as the glove dimensions, finger length, placement of the sensor, and positioning of the hand. Adaptive recognition algorithms (lightweight ML algorithms or user-specific threshold calibration at initialization) would be a possible solution.

X. FUTURE SCOPE

Accuracy, effectiveness, and applications of the Hand Ges- ture Vocalizer can be improved. Future iterations may use state-of-the-art technology, such as machine learning and artificial intelligence, to detect a variety of different gestures,
thereby allowing the device to provide assistance in a wide variety of human interactions.

The introduction of wireless technologies, including Bluetooth, Wi-Fi, and GSM modules, enables the device to communicate with other smart phones and gadgets. The inclusion will facilitate the user’s ability to run mobile applications and interact from a distance. Moreover, communication can take place between users in various languages via the multilingual speech output module.

Additionally, you can add to its comfort and portability by designing a complete wearable glove that incorporates both sensors and improved battery control technology. These modifications will help improve the smartness, organization, and practicality of the device, which will be helpful for people to interact socially.

XI. CONCLUSION

In this research, the design and implementation of a hand gesture vocalizer have been discussed that is intended to assist people who cannot speak or hear. It has been proposed that the system be implemented using an Arduino Nano board for processing, flex sensors for gesture recognition, and a compact glove design that can produce multiple modes of output including voice and text. Through this system, there is a bridge that connects people with disabilities and other normal people because it translates hand gestures into sounds and visuals.

The sensor-driven approach makes the technology user-friendly on a daily basis due to low energy consumption, portability, and real-time responsiveness. The binary coding and look-up table driven gesture recognition method is scalable since it allows updating software in order to incorporate new gestures in the vocabulary without having to update the hardware. In the study, the design and implementation of a hand gesture vocalizer have been addressed, which would help people that lack the ability to communicate verbally or through hearing. It is suggested that the technology could be implemented using Arduino Nano processor, flex sensors for gesture recognition, and small-size glove design that could provide various means of information transfer including voice and text messages.

By introducing this device, it is possible to create a link between people with disabilities and others since they would be able to transform hand gestures into sound and visual signals. The multimodal feedback mechanism helps to make the communication more convenient for people with diverse abilities since it benefits both users and recipients. The introduction of the GSM-based alarm mechanism enhances the safety of the users since they could receive assistance during emergencies without using mobile phones and the Internet. Conclusions drawn from experimental testing and perception surveys indicate that the suggested approach is widely acceptable, workable, and socially relevant. The suc- cessful implementation of the prototype proves its viability as a communication device due to reliable gesture detection and minimal lag. At the moment, the system uses static gestures for classification based on a threshold mechanism; however, it is an excellent foundation for further developments in the realm of adaptability, dynamic gesture recognition, and even machine learning. Having everything taken into account, the wearable vocalizer suggested in this paper can be classified as user-friendly, scalable, and relatively inexpensive assistive technology.

REFERENCES

1. Implementation of sign language recognition with TinyML on Arduino (conference paper),” AIP Conference Proceedings, 2024.
2. Z. Zhou et al., “A Portable Sign Language Collection and Translation Framework,” [PMC indexed journal article], 2022.
3. H. Kim et al., “Smart gloves-based triboelectric nanogenerator for sign language detection,” Human-Centric Intelligent Systems (Springer), 2025.
4. T. Kanwal et al., “A Smart Glove-Based System for Dynamic Sign Language Translation Using LSTM Networks,” MDPI Proceedings, 2025.
5. A. Filipowska et al., “Machine Learning-Based Gesture Recognition Glove: Design and Implementation,” Sensors, vol. 24, 2024..
6. M. A. Alzubaidi et al., “A Novel Assistive Glove to Convert Arabic Sign Language into Text and Speech,” ACM Transactions (ACM Digital Library), 2023. doi: 10.1145/3545113.
7. R. Wu et al., “Full-Fiber Auxetic-Interlaced Yarn Sensor for Sign- Language Translation Glove Assisted by Artificial Neural Network,” Nano-Micro Letters, 2022.
8. Y. Liu et al., “A wearable system for sign language recognition enabled by a convolutional neural network,” Nano Energy, 2023.
9. C. Yan et al., “A wearable sign language translation device utilizing silicone-hydrogel hybrid triboelectric sensor arrays and machine learn- ing,” Nano Energy, 2024. doi: 10.1016/j.nanoen.2024.110425.
10. P. Hrishikesh, M. George, and S. Krishna, “Vision Based Gesture Recognition: Real-Time Translation Application for American Sign Language,” Procedia Computer Science, vol. 244, pp. 1038–1045, 2024.
11. Recognizing hand gestures with pressure sensor-based motion sensing,” IEEE Transactions on Biomedical Circuits and Systems, 2019.
12. K. Kudrinko, E. Flavin, X. Zhu, and Q. Li, “Wearable sensor-based sign language recognition: A comprehensive

review," IEEE Reviews in Biomedical Engineering, vol. 14, pp. 82–97, 2020.

13. K. Sangeethalakshmi, K. G. Shanthi, A. M. Raj, S. Muthuselvan, P. M. Taha, and S. M. Shoaib, "Hand gesture vocalizer for deaf and dumb people," Materials Today: Proceedings, vol. 80, pp. 3589–3593, 2023.
14. Mittal, P. Kumar, P. P. Roy, R. Balasubramanian and B. B. Chaudhuri, "A Modified LSTM Model for Continuous Sign Language Recognition Using Leap Motion," in IEEE Sensors Journal, vol. 19, no. 16, pp. 7056 7063, 15 Aug.15, 2019.
15. H. Hu, J. Zhou, Z. Yuan, "SignBERT+: Hand-aware Self-supervised Pre-training for Sign Language Understanding", arXiv preprint, arXiv:2305.04868, 2023.
16. K. John, A. Sreekumar, and L. R. Mathew, "Wearable gesture detection glove for mute people," in Proc. 2021 2nd Int. Conf. Adv. Comput., Commun., Embedded Secure Syst. (ACCESS), pp. 97–101, IEEE, 2021.
 Springer Nature Singapore, pp. 345–353, 2020
17. E. Shanthini, N. Akshaya, and J. Haritha, "Hand Gesture Vocalizer using MobileNetV2 for Deaf Mutes," in Proc. 2023 3rd Int. Conf. Pervasive Comput. Social Netw. (ICPCSN), pp. 1481–1485, IEEE, 2023.
18. A. Gul, B. Zehra, S. Shah, N. Javed, and M. I. Saleem, "Two-way smart communication system for deaf and dumb and normal people," in Proc. 2020 Int. Conf. Inf. Sci. Commun. Technol. (ICISCT), pp. 1–4, IEEE, 2020.
19. M. I. Saleem, P. Otero, S. Noor, and R. Aftab, "Full duplex smart system for Deaf and Dumb and normal people," in Proc. 2020 Global Conf. Wireless Opt. Technol. (GCWOT), pp. 1–7, IEEE, 2020.
20. A. G. Wijaya, E. M. Nawangwulan, and N. P. Lestari, "Development of hand gesture vocalizer system for deaf-dumb people using IoT and machine learning," in Proceedings of the 2021 8th International Con- ference on Electrical and Electronics Engineering (ICEEE), Surakarta, Indonesia, 2021