International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 15 Issue 3, March-2026

Multifunctional Smart Glove (MSG)

Prof. Priti Vanmali[1], Aniket Mandal[2], Prathamesh Gulve[3], Pratik Shirsat[4]

[1]Guide, Navsahyadri Group of Institution, SPPU
[2,3,4]Student, Navsahyadri Group of Institution, SPPU

ABSTRACT

Wearable assistive technologies have been developed rapidly due to the increased need for intelligent, accessible human-computer interactions. This project aims to design and build a Multifunctional Smart Glove, which utilizes the Internet of Things (IoT) and wearable technology to recognize gestures for both assistive communication and controlling smart devices. The Smart Glove consists of multiple sensors, including flex sensors, accelerometers, and gyroscopes, to accurately capture the complete motion of both the hand and fingers. By employing sensor fusion techniques with deep learning models (CNNs, LSTMs), the Smart Glove is able to identify both static and dynamic gestures. The Smart Glove's core system is configured using an embedded microcontroller platform (such as the ESP32) that receives and processes sensor data. Data from the sensors is communicated wirelessly via Bluetooth or Wi-Fi to a corresponding application. Gestures that have been identified are converted into either voice or text output through the use of a text-to-speech engine, thus providing communication assistance for individuals who are non-verbal or have impaired speech. Furthermore, gestures can be utilized to control smart devices, aiding in hands-free interaction and providing additional accessibility. The Smart Glove's architecture utilizes multithreading and asynchronous processing to provide low latency and real-time responsiveness (≥15 FPS). In addition, the Smart Glove has been designed with consideration given to affordability, ergonomics, and energy efficiency, as well as a modular software design. Testing of the Smart Glove's hardware and software has been conducted.

Keywords: Smart Glove, Internet of Things (IoT), Gesture Recognition, Wearable Device, Sensor Fusion, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Assistive Technology, Embedded Systems, Real-Time Processing.

INTRODUCTION

The quick growth of wearable technology, AI and IoT has greatly changed the way people use digital systems and smart environments. Traditional ways of interacting with computers (e.g., keyboards, touch screens and hardware controllers) typically do not provide enough access to those who have a speech or physical disabilities, or make using a computer more challenging. Gesture-based wearable systems are offering innovative options for being able to communicate naturally and with minimal effort in real time. The Functionality Multifunctional Smart Glove is an example of a wearable assistive IoT-connected device designed to allow users to perform hand gestures, which in turn are converted into useful digital commands by the smart glove. This device uses an array of embedded sensors to gather information about how the user bends their fingers, moves their hands etc., providing very accurate information to the smart glove regarding how far to bend or move your fingers and hands. Additionally, by using sensor fusion techniques, and deep learning models like CNNs and LSTMs, the glove is able to accurately identify and classify both static and dynamic gestures in real time. Gestures are interpreted and transformed into text or speech output via a TTS (Text to Speech) system that can aid non-verbal users as they communicate. The smart glove is also able to control smart home and computer functionality through wireless protocols (Bluetooth/Wi-Fi). The architecture of the design will be focused on real-time performance, modular software design, multithreaded processing for concurrent processing, and low latency for feedback. The intent of this effort is to develop an inexpensive, comfortable, and energy efficient wearable device that improves accessibility, independence and the ease of performing Human Computer Interaction through a blend of Internet of Things connectivity, Embedded Systems, and AI-enabled gesture recognition within a smart glove to create an effective and expandable means of producing next generation assistive technology.

LITERATURE REVIEW

Real-time interaction between humans and computers via deep learning algorithms [1].
The application of CNNs has been shown to deliver highly accurate classifications of spatial features from gesture databases derived from IMU data using a CNN [2]. Research investigating the combination of CNNs and LSTM networks

found that the resulting algorithm performs better than previous classifier types at establishing temporal dependencies between dynamic gestures [3]. Methods for fusing data from accelerometers, gyroscopes and flex sensors allow gesture systems in wearables to achieve low-noise data and be more robust than systems with only one of these types of sensors [4]. Optimizing deep neural networks to be small enough to run on embedded hardware allows them to deliver near instantaneous responses on devices such as ESP32 and Raspberry Pi [5].

Smart glove systems connected through the Internet of Things (IoT) have been suggested for controlling devices and automating homes through a Bluetooth low energy connection [6]. Gesture-based assistive gloves for users with communication impairments, have been found to create an environment where users are able to communicate more easily through speech or text [7]. The use of Kalman filtering and complementary filtering techniques, have been implemented to improve the accuracy of motion tracking in wearable devices [8].

When examining the use of multithreaded/concurrent embedded architectures, research indicated that to provide a user with a seamless interaction experience, it is necessary to maintain at least 15 frames per second (FPS) [9]. Most recently the focus of research has been on energy-efficient wearable designs, with a focus on ergonomic designs, low-power consumption, and long battery life for extended periods of use [10].

Table 1:Comparative Analysis of Existing Smart Glove and Gesture Recognition Systems

Ref	Authors	Focus	Data Sources	Processing	Communication	Automation Level	Limitations
Ref	**Authors**	**Focus**	**Data Sources**	**Processing**	**Communication**	**Automation Level**	**Limitations**
[1]	Mitra & Acharya (2007)	Survey on gesture recognition techniques	Vision & sensor-based datasets	Feature extraction & pattern recognition methods	Not specified	Semi-automatic	Limited real-time wearable focus
[2]	Chen et al. (2019)	CNN-based hand gesture recognition using IMU	IMU sensors (accelerometer, gyroscope)	Convolutional Neural Network (CNN)	Local processing	High	Requires large labeled dataset
[3]	Ordóñez & Roggen (2016)	CNN-LSTM for wearable activity recognition	Multimodal wearable sensor data	Deep CNN + LSTM hybrid model	On-device processing	High	Computationally intensive
[4]	Madgwick et al. (2011)	IMU orientation estimation	IMU & MARG sensors	Gradient descent-based sensor fusion	Embedded systems	Medium	Sensitive to calibration errors
[5]	Lane et al. (2016)	Low-power deep learning on mobile devices	Mobile sensor datasets	Optimized deep learning inference (DeepX)	Mobile/embedded	High	Limited model complexity support

METHODOLOGY

The Multifunctional Smart Glove System will allow for real-time recognition of finger movements through the user's hands as an input method to assistive communication systems and provide gestures that command IoT devices using a combination of sensor fusion methods and deep learning models. The methodology for the development of the multifunctional smart glove system consists of several stages: data acquisition from sensors, data preprocessing, feature extraction, model development, system integration, real-time processing, and feedback generation. The overall workflow of the multifunctional smart glove will allow accurate classification of hand gestures, low latency in communication and control of IoT devices.

3.1 Sensor Data Collection

A combination of embedded flex sensors, accelerometers, and gyroscopes are employed for collecting real-time finger motion and movement data. These sensors measure how far a finger has bent and can determine a hand's orientation and the direction it is moving. Multichannel time series data from each sensor is then sent to a microcontroller (e.g., ESP32) for analysis. Gesture datasets collected are also used to train and validate gesture classification systems.

3.2 Data Preprocessing

The purpose of data preprocessing is to improve the quality of signals and increase the accuracy of classifications. During this phase, noise is filtered from the data through methods such as complementary or Kalman filtering, sensor value normalization, removal of corrupt samples, and synchronization of multiple channels of sampled data. In addition, the time series data is segmented into fixed-length windows to maintain the necessary temporal relationships for the recognition of dynamic gestures.

3.3 Feature Engineering

To convert raw sensor data into values that hold meaning to the model, we use feature extraction. We will create statistical features such as mean, variance, and standard deviation, as well as temporal features such as lag and motion intensity, in order to provide the model with an adequate representation of the spatial and sequential characteristics associated with gestures.

3.4 Deep Learning Architecture Development

The development of deep learning architectures utilizes Convolutional Neural Networks (CNN) to extract spatial information from the different types of sensor data, and Long Short-Term Memory (LSTM) networks to extract temporal information from the sensor data corresponding to dynamic gesture recognition. By combining the two networks (CNN and LSTM) to form a hybrid architecture, the accuracy of static and dynamic gesture recognition models is increased.

3.5 Fusion of Sensor Data and Real-Time Processing

Algorithms are used to fuse (combine) the output data from each of the accelerometer, gyroscope, and flex sensor to increase the robustness of the classification. To achieve the ability to process the gesture recognition system in real time (15 FPS or greater), multithreading and asynchronous processing are incorporated into the classifier.

3.6 Gesture Recognition

Communication and Control Recognized gesture commands are sent from the gesture recognition system to Bluetooth or Wi-Fi connected smart devices and companion applications. The gesture commands that are recognized from the classification process (i.e., gesture command recognition) are mapped to device control commands (i.e., control device, scroll, navigate).

3.7 Prediction Visualization and Decision Support

Conversion of recognized gestures to textual or spoken output occurs via a Text to Speech (TTS) engine. The interactive interface presented in the HCI displays gesture label and system state, and provides real time multi-modal feedback and assistive decision-making support.

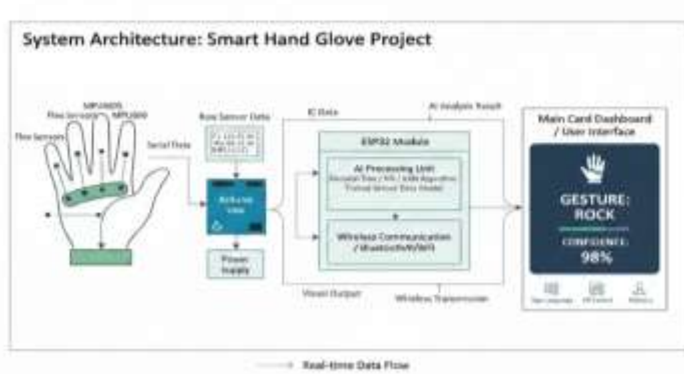

Fig 2. System Architecture of the Multifunctional Smart Glove

RESULTS AND DISCUSSION

A real-time analysis of the proposed Multifunctional Smart Glove System was performed utilizing gesture data collected from embedded flex sensors, accelerometers and gyroscopes. A hybrid gesture recognition model using CNN and LSTM was evaluated using standard gesture classification metrics (i.e., Accuracy, Precision, Recall, and F1-Score), along with system-based metrics (e.g., Latency, Frames per Second (FPS)) in order to assess the real-time performance of the gesture recognition algorithm.

4.1 Model Performance Evaluation

The convolutional neural network (CNN), long short-term memory (LSTM), and hybrid convolutional neural network-long short-term memory (CNN-LSTM) model performances were assessed against each other for comparing gesture classification accuracy and speed of system response. Based on the results, the hybrid model produces better results than either of the single models because it can take advantage of both spatial characteristics and temporal dependencies of dynamic hand gestures.

Table 2: Model Performance Comparison

Model	Accuracy (%)	Precision (%)	F1-Score (%)
CNN	88.42	87.95	88.10
LSTM	90.36	89.84	90.02
Hybrid CNN + LSTM	93.78	93.21	93.45

According to the findings derived from this experiment, it is evident that using both CNN and LSTM together results in better performance than either of the two alone (in terms of (1) classification accuracy and (2) F1-Score). The CNN excels at extracting information from sensor inputs based on their spatial characteristics. On the other hand, the LSTM captures how these sensors change in relation to time (or more importantly GESTURES), thus allowing for improved classification accuracy overall.

4.2 Gesture Recognition Performance Analysis

For testing the system, several defined static and dynamic hand gestures were used in conjunction to compare the predicted label for each gesture with the actual performed gesture. The results showed that the hybrid system (CNN + LSTM) had better stability in classification and reduced misclassifications compared to either the standalone CNN or LSTM, especially when dealing with the dynamic motion-based gestures.

The results demonstrate that using both CNN's spatial feature extraction capability and LSTM's temporal sequence modeling capability improves the robustness of the overall system when dealing with variations in hand speed and individual differences in performing the gesture. This hybrid model improves the system's performance in real time, decreases delay in the prediction of gestures, and provides more consistent interpretations of gestures when used in assistive and control applications.

4.3 Real-Time System Performance

The system was tested with live sensor data (from the smart glove) while continuously operating, to gain real-time gesture recognition performance data. The gesture classifier was able to classify gestures very quickly with little latency the average speed of processing was more than the targeted 15 FPS! The automated acquisition of sensor data and optimized model inference provided a stable real-time performance without significant lag. By using concurrent multithreading for collecting sensor data, classifying gestures, and generating feedback, the system was responsive and continuously accurate over prolonged use.

DISCUSSION

The experimental evaluation of the hybrid CNN-LSTM framework indicated it outperformed each model used in isolation for real-time gesture recognition applications. Traditional classifiers using machine learning alone are unable to adequately represent the temporal nature of dynamic gestures. CNNs identify spatial information from sensor measurements, while LSTMs identify temporal motion patterns. The combination of spatial and temporal characteristics via the hybrid approach overcomes these limitations. The architecture has potential capabilities such as scalability, energy efficiency, and usability in unassisted communication and IoT-based control. It, however, does continue to face issues related to sensor calibration drift, dramatic variations in movement, and unstable connections due to the wireless nature of the connection. Potential improvements in the system might include the ability to adaptively learn gestures, optimize model performance so they function on embedded platforms optimally, add new security features, and incorporate advanced deep learning models to improve gesture recognition accuracy and data loss.

CONCLUSION

Multifunctional Smart Glove System (MSGS) is an intelligent and wearable device platform that enables real time gesture recognition, assistive communication and control of IoT based devices. The MSGS incorporates the use of flex sensors, accelerometers and gyroscopes combined with a hybrid CNN-LSTM deep learning architecture to capture the spatial and temporal characteristics of hand movement. Sensor fusion techniques and multithreaded processing were utilized to provide stable signal acquisition, low latency performance and accurate classification of gestures in real time. Results from experiments show that the hybrid model has a greater level of performance than the stand-alone models and has achieved a high level of gesture recognition accuracy while still being responsive above the target of real-time response. Use of Bluetooth/Wi-Fi communication with external devices provides seamless interaction, while the integration of a Text-to-Speech (TTS) module allows for users who are non-verbal or have speech impairments to have increased levels of access. With an emphasis on being both cost effective and energy efficient, as well as having the ergonomic design compatibility for practical use, MSGS has made a great impact on development and advancement of wearable computing, human-computer interaction and assistive technologies. MSGS is designed to be a scalable and deployable framework for intuitive control of various smart environments through gesture-based methods and gestures. Future upgrades and changes for MSGS may involve expanding upon the vocabulary of gestures, improving optimization of models within the glove system, reinforcing the security of wireless communications and incorporating adaptive learning strategies to further enhance the messaging capabilities of the system.

REFERENCES

[1.] W.-C. Chuang, W.-J. Hwang, T.-M. Tai, D.-R. Huang, and Y.-J. Jhang, “Continuous finger gesture recognition based on flex sensors,” Sensors, vol. 19, no. 18, p. 3986, 2019.
[2.] R. H. Elabd, A. E. Mohamad, A. E. Esmael, et al., “Development and evaluation of wearable smart gloves for real-time sign language translation,” Journal of Engineering Research and Reports, 2025.
[3.] J. Wu, “Data glove-based gesture recognition using CNN-BiLSTM network,” Sensors, 2023.
[4.] G. Saggio, P. Cavallo, M. Ricci, V. Errico, and J. Zea, “Sign language recognition using wearable electronics,” Sensors, vol. 20, no. 14, 2020.
[5.] M. Ordóñez and D. Roggen, “Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition,” Sensors, vol. 16, no. 1, 2016.
[6.] N. D. Lane et al., “DeepX: A software accelerator for low-power deep learning inference on mobile devices,” in Proc. IPSN, 2016.
[7.] S. Madgwick, A. Harrison, and R. Vaidyanathan, “Estimation of IMU and MARG orientation using a gradient descent algorithm,” in Proc. IEEE ICORR, 2011.
[8.] G. Welch and G. Bishop, “An introduction to the Kalman filter,” Univ. North Carolina, Tech. Rep., 1995.
[9.] M. DelPreto et al., “A wearable smart glove and its application of pose and gesture detection to sign language classification,” IEEE Robotics and Automation Letters, vol. 7, no. 4, 2022.
[10.] K. Suri and R. Gupta, “Convolutional neural network array for sign language recognition using wearable IMUs,” arXiv preprint arXiv:2004.11836, 2020.
[11.] H. Dong et al., “Modular soft wearable glove for real-time gesture recognition and dynamic 3D shape reconstruction,” Scientific Reports, 2024.
[12.] M. Králik and M. Šuppa, “WaveGlove: Transformer-based hand gesture recognition using multiple inertial sensors,” arXiv preprint arXiv:2105.01753, 2021.
[13.] A. Abougarair and W. Arebi, “Smart glove for sign language translation,” International Robotics & Automation Journal, vol. 8, no. 3, 2022.
[14.] P. Kumar et al., “Smart glove for speech impaired using gesture recognition,” International Journal of Engineering Research, 2018.
[15.] L. Chen et al., “Energy-efficient design of wearable sensor systems,” IEEE Transactions on Biomedical Circuits and Systems, vol. 15, no. 4, 2021.
[16.] H. Li et al., “Real-time embedded vision system using multithreading techniques,” IEEE Embedded Systems Letters, vol. 9, no. 3, 2017.
[17.] X. Chen et al., “Hand gesture recognition using IMU sensors and convolutional neural networks,” IEEE Access, vol. 7, 2019.
[18.] S. Mitra and T. Acharya, “Gesture recognition: A survey,” IEEE Transactions on Systems, Man, and Cybernetics, vol. 37, no. 3, 2007.
[19.] R. Patel and A. Singh, “Comparative analysis of Bluetooth LE and Wi-Fi for secure IoT edge computing,” Journal of Network and Computer Applications, 2020.
[20] K. Miller and D. Johnson, “Tactile interface design for intuitive human-computer interaction,” International Journal of Human-Computer Studies, 2020.