Wearable Glove for Sign Language Translation Using Machine Learning

Abdullah Abdulrahman Abdulhak Alqershi

Department of Computer Science

International Islamic University Malaysia

Kuala Lumpur, Malaysia

abdullahsc253@gmail.com

Bawazier Najla Geis Junaid

Department of Computer Science

International Islamic University Malaysia

Kuala Lumpur, Malaysia

jullienajla@gmail.com

Assoc. Prof. Ts. Dr. Amelia Ritahani Bt. Ismail

Department of Computer Science

International Islamic University Malaysia

Kuala Lumpur, Malaysia

amelia@iium.edu.my

***Abstract*— To address communication challenges for the deaf community, this project introduces an innovative wearable glove designed to translate sign language into text. The primary objective is to bridge the communication gap between deaf individuals and those unfamiliar with sign language. Focusing on precision, the glove features flex sensors and a gyroscope to accurately detect sign language gestures. A concise yet impactful data collection phase involved participants performing ten sign language gestures, setting the foundation for a sophisticated K-Nearest Neighbors (KNN) machine learning model. The deployment of this technology was achieved through the Django web framework, resulting in a user-friendly web system for real-time translation. The model showcases an impressive accuracy of 92%, validated through thorough testing, confirming the system's effectiveness and successful operation in real-world scenarios. This achievement emphasizes the solution's technical adeptness and signifies notable progress toward broader inclusivity.**

***Keywords*— Machine Learning, IoT, Sign Language, Flex Sensors, Arduino, KNN (K-Nearest Neighbors),**

I. Introduction

In a world where communication forms the cornerstone of social interaction, the deaf community often faces significant barriers. Traditional sign language, while a primary mode of communication among those with hearing impairments, tends to create misunderstandings and social isolation when interacting with the hearing world. The challenge is further compounded by the inadequacy of existing systems designed to bridge the gap between the deaf and hearing communities, which often fall short in terms of effectiveness and accessibility.

This project introduces an innovative solution to these challenges; a wearable glove that leverages machine learning algorithms and Internet of Things (IoT) technology. The glove's primary function is to detect specific sign language gestures and translate them into text or speech. To achieve this, it is equipped with flex sensors that can accurately recognize finger movements and bends, which is fundamental in sign language. These sensors capture data that is then analyzed by sophisticated machine learning algorithms, enhancing the accuracy and efficiency of the translation process. The glove is primarily focused on interpreting gestures for specific alphabets and is rigorously tested for accuracy and real-time translation capabilities.

The urgency of addressing these communication barriers is underscored by statistics from the World Health Organization, which indicate that over 5% of the world's population requires rehabilitation for disabling hearing loss. This number is projected to exceed 700 million by 2050. Despite the widespread use of sign language within the deaf community, the lack of understanding by the hearing population results in limited social interaction and experiences for those with hearing impairments. Existing solutions, such as sign language interpreters and text-to-speech software, although beneficial, are often not readily available and can be cost-prohibitive.

Developing a wearable glove, integrating flex sensors and IoT technology, is at the heart of this project, focusing on real-time sign language translation. This glove is not just a tool for capturing sign language gestures, but a sophisticated device designed to translate these gestures into spoken language with high accuracy, using carefully selected algorithms. The implications of this technology extend far beyond aiding the deaf community; it promises to enhance inclusivity in diverse settings such as public spaces, educational environments, and businesses. Moreover, the potential applications of this technology extend into areas such as human-robot interaction and virtual assistants, showcasing its versatility. By enhancing the precision and reliability of machine learning algorithms in gesture recognition, the project stands at the forefront of advancing assistive technology and contributing significantly to the field of machine learning. This initiative, therefore, is not just about creating a functional tool; it is about paving the way for a more inclusive and technologically advanced future where communication barriers are a thing of the past.

II. Literature Review

Recent advancements in gesture recognition technology have seen a significant focus on integrating machine learning techniques for enhancing communication aids for individuals with speech and hearing impairments. [1] developed a Bluetooth-enabled glove interfaced with an Android system, which translated hand gestures into artificial speech. This system employed a predefined database of alphabets and words, matching sensor values to create meaningful speech output. It was also proposed to expand the glove's use in virtual reality and IoT interactions, underscoring the potential of machine learning in enhancing gesture recognition.

[2] emphasized the critical role of machine learning algorithms in optimizing IoT systems for gesture classification. This aligns with [3], who demonstrated the efficacy of machine learning models in sign language recognition. These models show considerable promise in bridging the communication gap for the deaf and hard-of-hearing community, indicating a significant shift towards more accessible technology-driven solutions.

Regarding wearable gloves for sign language translation, the focus has predominantly been on two methods of gesture capture: using cameras or sensors. [4] highlighted the necessity of cameras in capturing hand images for gesture interpretation. Conversely, [5] and [6] focused on sensor-based solutions, where flex sensors and other IoT devices played a crucial role in detecting and translating hand movements. These efforts collectively illustrate the evolution of gesture recognition systems from visual to sensor-based technologies, offering enhanced mobility and user comfort.

[7] explored the effectiveness of flex sensors and MPU6050 sensors in a smart glove designed for sign language translation. They emphasized the precision of these sensors in measuring hand movements, integral to accurate gesture interpretation. Complementing this, the "Smart Hand Glove" project [8] and the framework by [9] demonstrated accessible and cost-effective approaches to gesture recognition, employing similar sensor technologies and highlighting the potential for future advancements in complex gesture interpretation.

Finally, [10] introduced a dynamic gesture recognition approach using Hidden Markov Models and D-S evidence theory, focusing on human-computer interaction within the IoT context. Along with [11], who proposed an intelligent electronic glove system, these developments represent a significant contribution to the field of gesture recognition technology. They underscore the importance of advanced algorithms and prototype selection methods in creating efficient and reliable communication systems for the deaf and speech-impaired, thereby paving the way for more inclusive and technologically advanced solutions in human-computer interaction.

III. METHODOLOGY

A. *Circuit and Glove Design*

The circuit for the wearable sign language translation glove incorporates various components to enable precise gesture detection and communication. As illustrated in Fig. 1, the primary components are as follows:

1. *Flex Sensors (4.5 and 2.2 inches):* These sensors detect finger bending and flexing for sign language gestures. Smaller 2.2-inch sensors ensure a precise fit for the pinky and thumb, enhancing accuracy.
2. *ESP32 Microcontroller:* With ample analog inputs and integrated Wi-Fi, the ESP32 reads, processes, and manages data from flex sensors and MPU6050. It enables real-time translation of sign language into text or speech via Wi-Fi, suitable for IoT applications.
3. *MPU6050:* The MPU6050 provides orientation and motion data for improved gesture recognition.
4. *Wire Jumpers:* Essential for connecting components on the breadboard, wire jumpers ensure smooth signal, power, and data flow among the microcontroller, flex sensors, and MPU6050.

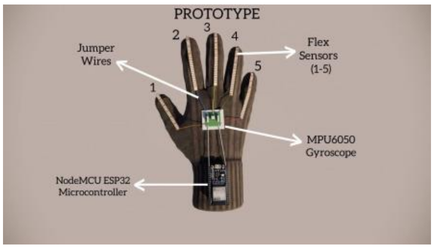

Fig. 1. The wearable glove design

The illustration in Fig. 2 depicts the circuit setup, the specific connections are outlined as follows:

- VCC (Power) of the flex sensors is connected to ESP32's 3.3V supply.
- The signal wires from each flex sensor are connected to specific ESP32 pins as follows:
  - Pinky (2.2 inches) to ESP32 Pin 39 (VN)
  - Ring (4.5 inches) to ESP32 Pin 34
  - Middle (4.5 inches) to ESP32 Pin 35
  - Index (4.5 inches) to ESP32 Pin 32
  - Thumb (2.2 inches) to ESP32 Pin 33
- A resistor is used to connect the signal wires of the flex sensors to the ESP32's GND (Ground) for proper signal referencing.

For the MPU6050 gyroscope:

- VCC (Power) is connected to ESP32's 3.3V supply.
- GND (Ground) is connected to ESP32's Ground.
- SCL (Serial Clock) signal is connected to ESP32 Pin 22.
- SDA (Serial Data) signal is connected to ESP32 Pin 21.

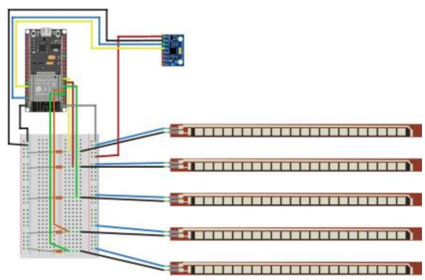

Fig.2. Circuit design

This circuit setup allows the ESP32 microcontroller to receive data from the flex sensors and MPU6050, process it, and utilize Wi-Fi connectivity for real-time sign language translation, making it a versatile and powerful assistive technology device.

B. System Architecture

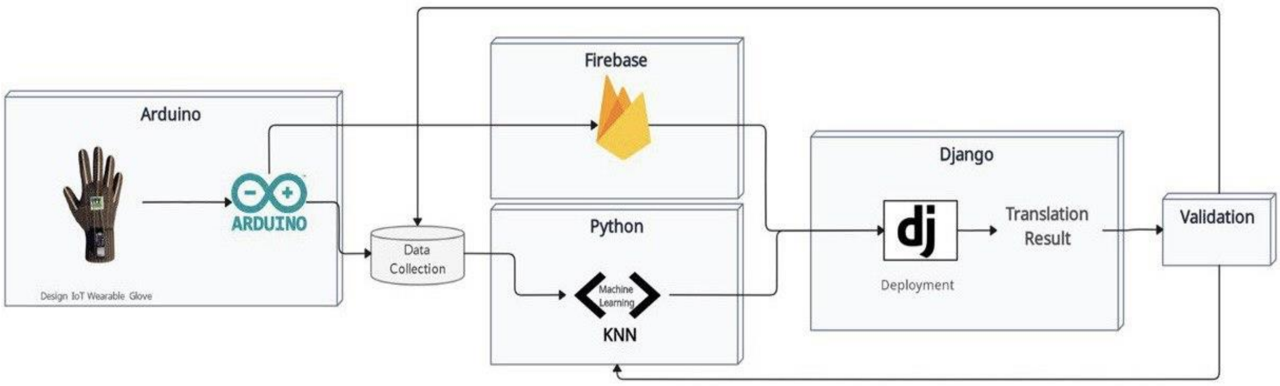

Fig.3. System Architecture

The wearable glove, central to the system architecture, is equipped with sensors like flex sensors and a gyroscope, connected to an Arduino microcontroller. This setup captures detailed hand movements and gestures, storing the data in CSV format. The Arduino setup plays a dual role; while it captures and saves the gesture data, it also transmits this data in real-time to a Firebase database. This cloud-based database is instrumental in synchronizing the sensor data with the user's gestures, ensuring that the collected information is up-to-date and accurately represents the user's actions.

The machine learning component, a K-Nearest Neighbors (KNN) classifier, is trained using the collected gesture data. This training process enables the model to interpret various sign language gestures, turning complex hand movements into understandable sign language elements. Once the model is adequately trained and validated for its accuracy and reliability, it is ready for deployment.

The deployment of the trained model is facilitated through the Django web framework. In the deployed system, the Django application continuously communicates with the Firebase database, fetching the latest gesture data in real-time. As new gesture data arrives, the application uses the pre-trained KNN model to interpret these gestures and generate predictions. The predictions, which translate the gestures into text or speech, are then presented to the user. This process ensures users receive immediate and accurate sign language translations, enhancing the communication experience for the deaf and hard-of-hearing community. The integration of Arduino with Firebase and the subsequent utilization of Django for deploying the machine learning model exemplify a harmonious blend of hardware, cloud computing, and web technologies, making the system a noteworthy achievement in assistive technology.

C. Data Collection

The project's data collection involved gathering inputs from 13 participants, resulting in a comprehensive dataset of 3087 entries. Each data set comprised readings from flex sensors attached to the thumb, index, middle, ring, and pinky fingers, along with X and Y axis values from the gyroscope. This approach ensured a rich and diverse range of hand movements and gestures, capturing the intricate nuances of sign language. Specifically, the data collection focused on accurately capturing 10 key American Sign Language (ASL) gestures: "Hello," "Goodbye," "You," "OK," "I Hate You," "I am/I," "Fine," "Yes," "No," and "I Love You."

Participants were instructed to perform these gestures while the glove, equipped with sensors, tracked the precise movements and orientations of their hands in real-time. The Arduino setup facilitated the seamless recording of these movements and enabled the immediate storage of data in a CSV or table format, ensuring efficiency and organization in data handling.

After the data collection phase, the dataset underwent cleaning and preprocessing, which refined the total number of usable datasets to 2895. This step, while crucial in enhancing the dataset's quality, involved standardizing the data to ensure consistency across all measurements, particularly for the selected ASL gestures. This process was key in preparing the dataset for the next stage of building a machine learning model, which would be responsible for the accurate interpretation and translation of sign language gestures based on the collected data.

D. KNN Model

The K-Nearest Neighbors (KNN) algorithm plays a central and pivotal role in the successful execution of this project. KNN is a versatile and fundamental machine-learning technique that is particularly well-suited for classification tasks. Its core principle involves classifying new data points by evaluating the majority class among their closest neighbors in a given feature space, as shown in Fig. 4. This proximity-based classification hinges on the calculation of a distance metric, often employing the Euclidean distance measure, which quantifies the dissimilarity or similarity between data points. A crucial parameter in the KNN algorithm is 'k,' representing the number of neighboring data points considered during classification. The selection of an appropriate 'k' value is paramount, as it significantly influences the model's sensitivity to the local structure of the data.

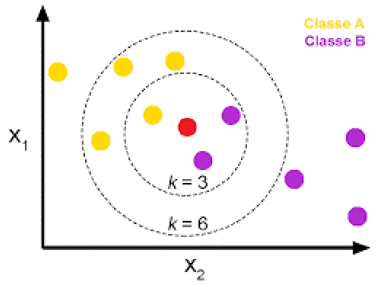

Fig. 4. KNN Model

In this project, the KNN algorithm played a central role as a fundamental machine-learning method for classification tasks. It functioned by classifying new data points based on the majority class among their nearest neighbors, relying on a distance metric, commonly the Euclidean distance. The parameter 'k,' representing the number of neighbors considered, was set to 8 for this project. The dataset, comprising hand gesture data collected from flex sensors and gyroscopes, underwent preprocessing steps that included the removal of outliers using Z-scores and the standard division into training and test sets. Subsequently, the KNN classifier was trained on the prepared dataset, enabling it to accurately classify hand gestures based on sensor readings.

E. System Deployment

The deployment of the system was approached with a focus on real-time functionality and user experience. Using Django, a Python-based web framework, a web application that could efficiently handle real-time data processing and user interactions was created. Django's inherent strengths in managing database operations and user interfaces provided a solid foundation for integrating complex functionalities. Furthermore, Django's compatibility with Python enabled seamless integration of the machine learning model, facilitating efficient data processing and real-time analytics. The core functionality of the system revolves around the K-Nearest Neighbors (KNN) model, which was seamlessly integrated into the Django framework. This integration allowed for direct communication with Firebase, a cloud database known for its real-time data handling capabilities.

In practice, as users interacted with the glove, equipped with sensors for gesture recognition, the data was instantly transmitted to Firebase. From there, the Django application continuously retrieved this data, feeding it into the machine learning model for real-time translation of gestures into text or speech. The immediacy of this process was crucial for the system's effectiveness, ensuring that users experienced minimal delay between their gestures and the corresponding translations. To complement this functionality, the system's front end was crafted with HTML and CSS, prioritizing ease of use and visual appeal. This design choice not only made the application more engaging for users but also ensured that the interface was accessible, enhancing the overall user experience. Through this combination of Django's robust back-end capabilities and a carefully designed front-end, the system achieved a balance of technical efficiency and user-centered design.

IV. Result and Discussion

A. System Evaluation

As illustrated in Fig. 5, the wearable glove, equipped with flex sensors and a gyroscope, stands as a pivotal component of the system. These sensors are strategically placed on the glove to capture intricate hand movements and gestures accurately. This configuration ensures that the glove effectively records the sign language gestures made by the user, laying the foundation for the subsequent translation process. The physical appearance of the glove, as depicted in Figure 5, showcases its ergonomic design, making it user-friendly and comfortable for individuals with hearing impairments.

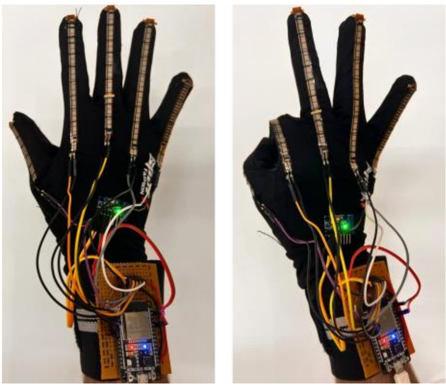

Fig. 5. Sensor-equipped wearable glove

On the other hand, Fig. 6 provides a visual representation of the web application interface, which serves as the user's gateway to real-time gesture translation. The web application is thoughtfully designed, offering an intuitive and accessible platform for users to interact with the system. Its user-friendly layout and design prioritize ease of use and visual appeal. Through the seamless integration of the pre-trained machine learning model, the web application fetches sensor data from Firebase, generates instant translations, and presents them to the user in real-time. This visual depiction in Figure 6 highlights the practicality and effectiveness of the web application in bridging the communication gap for individuals with hearing impairments.

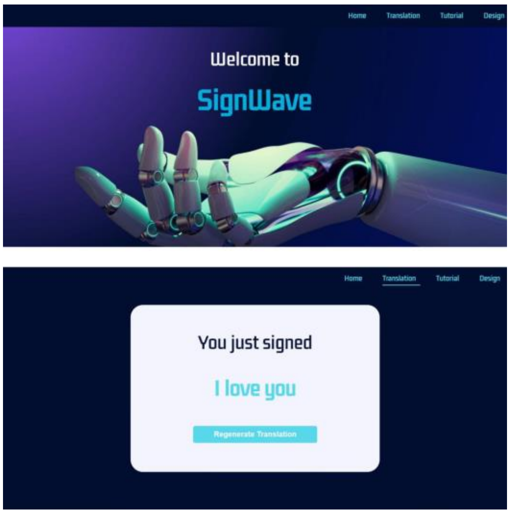

Fig. 6. Web Application Interface

B. *Model Performance*

The classification report provided in Table I is a crucial component in evaluating the performance of the K-Nearest Neighbors (KNN) model, particularly for a classification task involving distinct categories or gestures. This report is meticulously structured, presenting key metrics like Precision, Recall, and F1-score, alongside the Support for each category. These metrics are fundamental in understanding how well the model performs for each specific class.

TABLE 1. MODEL CLASSIFICATION REPORT

	Precision	Recall	F1-score	Support
Hello	0.87	0.89	0.93	55
Goodbye	0.91	0.84	0.88	51
You	0.93	0.98	0.96	56
Ok	0.96	0.96	0.96	50
I Hate You	0.98	0.89	0.88	51
I am	0.92	0.92	0.92	60
Fine	0.91	0.94	0.92	32
Yes	0.93	0.88	0.90	65
No	0.85	0.98	0.91	45
I Love You	0.93	0.95	0.94	56
Model Accuracy			0.92	516

In the KNN model's classification report, the gestures 'Hello' and 'Goodbye' exhibit lower precision (0.87 and 0.91, respectively) and recall (0.89 and 0.84, respectively), likely due to their physical resemblance, leading to frequent misclassifications. This illustrates the challenge of distinguishing visually or contextually similar gestures. Conversely, 'No' stands out with a high recall of 0.98 but a lower precision of 0.85, indicating its effective identification but also a tendency to be over-identified, possibly due to it having a motion. Similarly, 'I Hate You' and 'I Love You', despite their contrasting meanings, might share physical similarities, as reflected in their precision and recall, leading to some confusion in the model's predictions. This pattern in the data underscores the importance of nuanced training in gesture recognition models, especially when dealing with gestures that have subtle differences or similarities.

Overall, the table suggests that while the model performs well overall (as indicated by the overall accuracy of 0.92, there are specific areas where gesture similarity impacts its precision and recall. This highlights the importance of training the model with a diverse set of examples for each gesture and possibly fine-tuning it to better distinguish between gestures that are similar in appearance or context.

C. *Translation Validation*

TABLE II. VALIDATION TEST RESULT

Person	A	B	C	D	E	F	G	H	I	J	Accuracy
Hand Size	3	3	3	3	2	1	1	1	1	2	
Hello	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	100%
Goodbye	✓	✓	✓	✓	✓	✓	×	×	✓	✓	80%
You	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	100%
Ok	✓	✓	✓	✓	✓		✓	✓	✓	✓	100%
I Hate You	×	✓	✓	×	✓	✓	×	✓	✓	×	60%
I am/ I	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	100%
Fine	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	100%
Yes	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	100%
No	✓	✓	✓	✓	✓	×	✓	✓	×	×	70%
I Love You	✓	✓	✓	✓	×	✓	✓	✓	✓	✓	90%

The performance evaluation of the translation system, conducted with 10 individuals, demonstrates its effectiveness in real-world conditions. This diverse group of participants, comprising both males and females, exhibited a range of hand sizes, which is a crucial factor in the system's functionality. Hand size significantly impacts the flex sensor values in the glove, as different hand sizes lead to variations in the degree of bending. This variation is essential to consider, as it directly affects the accuracy of gesture recognition and translation.

Table II presents an insightful breakdown of the system's performance across three distinct hand sizes: large (denoted as '3'), medium ('2'), and small ('1'). The variation in hand sizes among participants provides a robust test of the system's adaptability and accuracy. The differences in flex sensor readings due to hand size variations challenge the system's ability to consistently interpret gestures.

However, certain gestures with close similarities presented more of a challenge. For instance, the phrases "I hate you" and "I love you," which have subtle differences in the gestures, resulted in a lower accuracy rate of 60% for "I hate you." Similarly, the gesture for "No" faced a reduction in accuracy, achieving 70%, primarily due to its resemblance to the gesture for "You." The differences in accuracy levels can be linked to factors like the hand size of the users, which affects how the flex sensors read the gestures and the natural likeness found in some sign language gestures. By gathering more diverse data, especially focusing on signs that share similarities, the system can be better trained to distinguish between closely resembling gestures, leading to improved overall performance.

V. CONCLUSION

In conclusion, the development of the wearable glove for sign language translation using machine learning represents a significant breakthrough in addressing communication barriers faced by the deaf and hard-of-hearing community. The system's remarkable accuracy rate of 92.05% in translating sign language gestures into text or speech underscores its effectiveness and potential to revolutionize the way individuals with hearing impairments interact with the world. Through the integration of flex sensors, gyroscope technology, and the K-Nearest Neighbors (KNN) machine learning model, this innovative solution offers a practical and efficient means of bridging the communication gap between the deaf community and those unfamiliar with sign language.

Notably, the system's real-world performance evaluations, involving a diverse group of participants, have demonstrated its capability to accurately interpret a range of sign language gestures. While challenges exist in distinguishing between closely resembling gestures, particularly "I hate you" and "I love you," the system's overall performance is highly promising. With further data collection and training, it is expected that the system will continue to improve in recognizing subtleties and nuances in sign language. In essence, this project signifies a significant step towards inclusivity and accessibility, offering a potential lifeline to over 5% of the global population with disabling hearing loss and paving the way for a more inclusive future where communication barriers are a thing of the past.

VI. FUTURE WORK

Upgrading the web application to enable user-contributed sign language gestures is a critical enhancement. This improvement aims to build a larger and more varied database, significantly boosting the system's accuracy and scope. By embracing a broader spectrum of sign language expressions, the system will become more inclusive and adaptable to diverse user needs.

The personalization of the system is a major focus for future development. This feature will enable users to have their real-time database, which will be particularly useful for capturing and reading sensor values from the wearable glove. This level of customization not only enhances the accuracy of translations but also ensures greater usability and convenience in various environments.

Another area of future work involves improving the system's adaptability to different hand sizes and shapes. This enhancement will focus on refining the glove's sensor sensitivity and the model's ability to interpret a wider range of hand movements. By doing so, the system will become more flexible and accommodating to individual physical differences, which is crucial for accurately translating sign language gestures for a diverse user base. This improvement will ensure that the system remains effective and reliable, regardless of variations in users' hand characteristics.

The final aspect of future enhancements involves upgrading the database architecture to support more complex functionalities. A robust and advanced database is critical for efficiently managing larger data volumes and enabling the integration of sophisticated machine-learning algorithms. This improvement is vital for scaling the system, enhancing its performance, and expanding its capabilities to meet future technological demands and broader application areas.

REFERENCES

[1] S. Y. Heera, M. K. Murthy, V. S. Sravanti, and S. Salvi, "Talking hands — An Indian sign language to speech translating gloves," IEEE Xplore, Feb. 01, 2017. https://ieeexplore.ieee.org/abstract/document/7975564.

[2] D. K. Choudhary, R. Singh, and D. Kamthania, "Sign Language Recognition System," SSRN Electronic Journal, 2021, doi: https://doi.org/10.2139/ssrn.3832151.

[3] B. Rajalingam, R. Kumar, P. Deepan, P. Patra, and S. Bavankumar, "A Smart System for Sign Language Recognition using Machine Learning Models," Dec. 2022, doi: https://doi.org/10.1109/icac3n56670.2022.10074007.

[4] M. A. Ahmed, B. B. Zaidan, A. A. Zaidan, M. M. Salih, and M. M. bin Lakulu, "A Review on Systems-Based Sensory Gloves for Sign Language Recognition State of the Art between 2007 and 2017," Sensors, vol. 18, no. 7, p. 2208, Jul. 2018, doi: https://doi.org/10.3390/s18072208.

[5] Y.-F. Fu and C.-S. Ho, "Development of a programmable digital glove," Smart Materials and Structures, vol. 17, no. 2, p. 025031, Mar. 2008, doi: https://doi.org/10.1088/0964-1726/17/2/025031.

[6] Yigremachew Eshetu, "A Real-Time Ethiopian Sign Language to Audio Converter," International Journal of Engineering Research and Technology (IJERT), vol. V8, no. 08, Sep. 2019, doi: https://doi.org/10.17577/ijertv8is080255.

[7] A. Yudhana, J. Rahmawan, and C. U. P. Negara, "Flex sensors and MPU6050 sensors responses on smart glove for sign language translation," IOP Conference Series: Materials Science and Engineering, vol. 403, p. 012032, Oct. 2018, doi: https://doi.org/10.1088/1757-899x/403/1/012032.

[8] P. SABU, "smart-hand-glove," GitHub, Jan. 24, 2023. https://github.com/pscretn/smart-hand-glove.

[9] A. Suri, S. K. Singh, R. Sharma, P. Sharma, N. Garg, and R. Upadhyaya, "Development of Sign Language using Flex Sensors," IEEE Xplore, Sep. 01, 2020. https://ieeexplore.ieee.org/document/9215392.

[10] G. Li, H. Wu, G. Jiang, S. Xu, and H. Liu, "Dynamic Gesture Recognition in the Internet of Things," IEEE Access, vol. 7, pp. 23713–23724, 2019, doi: https://doi.org/10.1109/access.2018.2887223.

[11] P. D. Rosero-Montalvo et al., "Sign Language Recognition Based on Intelligent Glove Using Machine Learning Techniques," IEEE Xplore, Oct. 01, 2018. https://ieeexplore.ieee.org/document/8580268.