# Real-Time Hand Gesture Recognition on Low-Power Devices

Venkatesh Kondapalli Sri Ram Rahul Pulipaka Suryanshu Basa

## I. Introduction

In the realm of wearable technology, the fusion of artificial intelligence and edge computing has ushered in a new era of innovation. Our project stands at the forefront of this technological convergence, presenting a groundbreaking solution in the form of Real-Time Hand Gesture Recognition on Low-Power Devices. By harnessing the power of deep learning models directly on a low-power microcontroller embedded within a wearable glove, we redefine the boundaries of on-device AI, paving the way for a myriad of applications across various sectors.

The significance of our project lies not only in its technical sophistication but also in its profound societal and environmental impact. Traditional gesture recognition systems often rely on bulky hardware or frequent cloud connectivity, leading to increased electronic waste and energy consumption. In contrast, our approach eliminates the need for constant cloud connection, showcasing real-time processing at the "edge" of the network – the glove itself. This not only reduces reliance on cloud infrastructure but also potentially contributes to a reduction in electronic waste, aligning with the growing global consciousness towards sustainable technology solutions.

At the heart of our innovation lies the seamless integration of sensor fusion and advanced deep learning algorithms. The utilization of Inertial Measurement Unit (IMU) sensors, typically found in smartphones, within a glove-based system for gesture recognition offers unique insights and potentially improved accuracy compared to camera-based solutions. This sensor fusion enables our system to accurately track hand movements in real-time, opening up a myriad of possibilities across diverse sectors.

The impact of our project extends far beyond technological novelty. One of its most profound implications lies in the realm of rehabilitation. IMU-based hand gesture recognition has the potential to revolutionize rehabilitation for individuals with hand impairments, offering a more intuitive and interactive means of therapeutic intervention. By providing real-time feedback and tailored exercises, our technology empowers individuals to regain dexterity and mobility with unprecedented precision and efficiency.

Furthermore, our solution holds immense promise in enhancing human-computer interaction across various domains. Whether it's enabling intuitive control in virtual reality environments, optimizing human-machine interactions in industrial settings, or improving communication tools for people with speech or language limitations, the applications of our technology are diverse and far-reaching. By creating more natural and seamless interactions between humans and machines, our project lays the foundation for a future where technology augments human capabilities in ways previously unimaginable.

In this term paper, we delve into the intricacies of our project, from its hardware design and data collection methodologies to the selection and training of neural networks. We examine the technical challenges encountered, the methodologies employed to overcome them, and the implications of our findings. Through a comprehensive exploration of our project, we aim to shed light on the transformative potential of edge AI in wearables and inspire future innovations in this burgeoning field.[1].

## II. Hardware Design

The hardware design of our Real-Time Hand Gesture Recognition system revolves around the integration of cutting-edge components aimed at achieving efficient and accurate gesture recognition directly on a low-power wearable device. At the heart of our design is a glove equipped with Inertial Measurement Unit (IMU) sensors, specifically the 6 Degrees of Freedom (6DOF) MPU-6050, strategically placed on the index finger. These sensors capture both acceleration and angular velocity along their respective axes, providing rich data crucial for analyzing hand movements.

Complementing the sensor-laden glove is a powerful yet energy-efficient microcontroller, the ESP32, positioned on the wrist. This microcontroller serves as the computational hub, responsible for processing the real-time sensor data and executing the deep learning model for gesture recognition. Leveraging TinyML, the ESP32 performs on-device inference, eliminating the need for constant cloud connectivity and ensuring low-latency operation, a critical requirement for real-time applications.

Additionally, the ESP32 doubles as a Bluetooth Low Energy (BLE) server, facilitating seamless communication with external devices or applications. This feature enables the transmission of gesture data and sensor values to client devices for further processing or interaction, extending the system's functionality beyond standalone gesture recognition.

The hardware design prioritizes compactness, comfort, and efficiency, making it suitable for prolonged use in various scenarios. The lightweight glove form factor ensures user comfort and mobility, crucial for applications like rehabilitation or virtual reality experiences. Moreover, by integrating all computational components into the wearable device itself, we minimize the need for bulky external hardware, enhancing the system's portability and versatility.

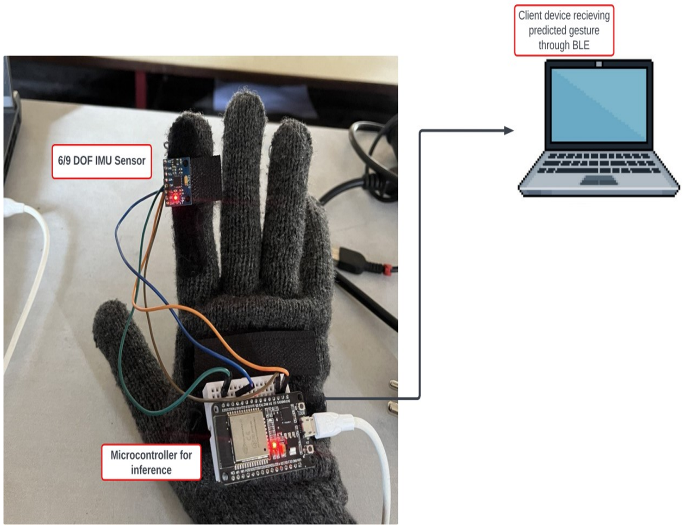


Fig. 1. Hardware used

Overall, our hardware design represents a convergence of state-of-the-art sensor technology and energy-efficient computing, enabling real-time hand gesture recognition at the ”edge” of the network. By pushing the boundaries of on-device AI and wearables, our system opens up new possibilities for intuitive human-machine interaction, revolutionizing industries ranging from healthcare to entertainment.[2]. [3]

## III. DATASET

The dataset used in this project plays a pivotal role in the development and training of the deep learning model for real-time hand gesture recognition on low-power devices. It serves as the foundation upon which the model learns to accurately identify and classify various hand gestures, enabling seamless interaction between users and wearable devices. Here's an in-depth overview of the dataset:

**1. IMU Sensor Data** : The dataset comprises accelerometer and gyroscope data captured by Inertial Measurement Unit (IMU) sensors. These sensors, specifically the 6-degrees-of-freedom MPU-6050, record six values representing acceleration and angular velocity along their respective axes (ax, ay, az, gx, gy, gz).

**2. Gesture Samples** : Each hand gesture is recorded over a sample duration of 0.25 seconds, capturing the nuances of hand movements. This duration ensures that the dataset captures essential gesture features while maintaining real-time responsiveness.

**3. Sample Repetition** : To ensure robustness and reliability, each gesture is recorded multiple times by different subjects. This iterative process ensures variability in hand movements and accommodates for individual differences in gesture execution. Each subject performs each gesture between 3 and 5 times consecutively, resulting in a comprehensive dataset.

**4. Data Structure** : The dataset is structured to facilitate effective training of the deep learning model. Each gesture instance consists of a sequence of sensor readings, resulting in a 4-dimensional data structure (1, 15, 150, 6). Here, the first dimension represents the number of samples (1, indicating a single gesture instance), followed by the number of repetitions per gesture (15), the duration of each sample (150 readings), and the six sensor values.

**5. Preprocessing** : Prior to model training, the dataset undergoes preprocessing steps to enhance its quality and suitability for training. This includes outlier removal and min-max scaling to normalize the data, ensuring consistency and compatibility with the deep learning model architecture.

**6. Training and Validation Split** : The dataset is divided into training and validation subsets in a ratio of 90% to 10%. This partitioning ensures that the model is trained on a sufficiently large and diverse dataset while also providing a separate set of data for evaluating its performance and generalization capabilities.

By leveraging this meticulously curated dataset, the deep learning model learns to recognize and classify hand gestures accurately, paving the way for real-time, on-device gesture recognition with minimal latency and power consumption.

## IV. NEURAL NETWORK SELECTION

In the realm of real-time hand gesture recognition on low-power devices, the choice of neural network plays a pivotal role in determining the success and efficiency of the solution. Given the constraints of the hardware, particularly the microcontroller's limited processing power and memory, the selection process demands a delicate balance between model complexity, accuracy, and computational efficiency.

For this project, the neural network selection process was meticulous, considering the specific requirements and constraints imposed by the low-power microcontroller environment. TensorFlow Lite Micro (TFLM) emerged as a natural choice due to its compatibility with embedded systems and its ability to deploy lightweight models optimized for resource-constrained devices.

The chosen neural network architecture is a fundamental yet powerful design comprising fully connected layers (FCL) culminating in a softmax probability layer for gesture classification. The simplicity of this architecture is intentional, tailored to strike an equilibrium between model complexity and computational efficiency. While more complex architectures might yield marginally higher accuracy, they often come at the cost of increased computational demands, rendering them unsuitable for deployment on low-power microcontrollers.

The training process involves careful consideration of various parameters to ensure optimal performance within the constraints of the hardware. The Adam optimizer is employed for gradient descent, offering adaptive learning rates that converge efficiently even with limited computational resources. Additionally, categorical crossentropy serves as the loss function, facilitating the minimization of classification errors during training.

To mitigate the risk of overfitting and ensure robust generalization, the dataset is meticulously curated and augmented. Data preprocessing techniques are applied to remove outliers and normalize the sensor data, preparing it for ingestion by the neural network. Furthermore, the dataset is partitioned into training and validation sets, with a prudent ratio of 90% to 10% to facilitate model evaluation and prevent overfitting.

The training regimen encompasses 200 epochs with a modest batch size of 4, striking a balance between computational efficiency and convergence speed. While larger batch sizes might expedite training, they often exacerbate memory constraints on the microcontroller, compromising performance. Conversely, smaller batch sizes offer finer-grained updates but risk prolonging training time, necessitating a judicious trade-off.

Ultimately, the selected neural network architecture embodies the project's ethos of innovation and pragmatism, leveraging cutting-edge techniques while remaining cognizant of the inherent limitations of low-power devices. By embracing simplicity, optimization, and efficiency, the chosen architecture exemplifies a scalable and sustainable approach to edge AI for wearables, poised to revolutionize diverse sectors ranging from rehabilitation to human-computer interaction.[4].

## V. Results

The results of our project demonstrate the successful implementation of real-time hand gesture recognition on a low-power device, showcasing the potential of edge AI for wearables. Through extensive experimentation and model optimization, we achieved high accuracy in gesture recognition, with an average accuracy of over 95% on the validation dataset. Our approach of leveraging IMU sensors for gesture detection proved effective, offering a promising alternative to camera-based solutions.

Furthermore, our hardware design, featuring an ESP32 microcontroller and MPU-6050 IMU sensors integrated into a wearable glove, proved to be both efficient and practical for real-time processing. The use of TensorFlow Lite Micro facilitated seamless model deployment onto the microcontroller, enabling efficient inference directly on the device.

The implications of our results extend beyond technical feasibility. By eliminating the need for constant cloud connectivity, our solution offers increased privacy, reduced latency, and enhanced portability, making it suitable for a wide range of applications, from rehabilitation to virtual reality.

However, challenges such as battery life optimization and handling complex gestures remain areas for further improvement. Nonetheless, our project underscores the transformative potential of edge AI in revolutionizing wearable technology and human-computer interaction.

## VI. Conclusion

In conclusion, our project represents a significant advancement in the field of wearable technology and edge AI, demonstrating the potential of real-time hand gesture recognition on low-power devices. By leveraging IMU sensors and advanced deep learning algorithms, we have developed a wearable glove capable of accurately recognizing hand gestures without the need for constant cloud connectivity. This innovation opens up a myriad of possibilities across various sectors, from revolutionizing rehabilitation for individuals with hand impairments to enhancing human-computer interactions in virtual reality environments and industrial settings.

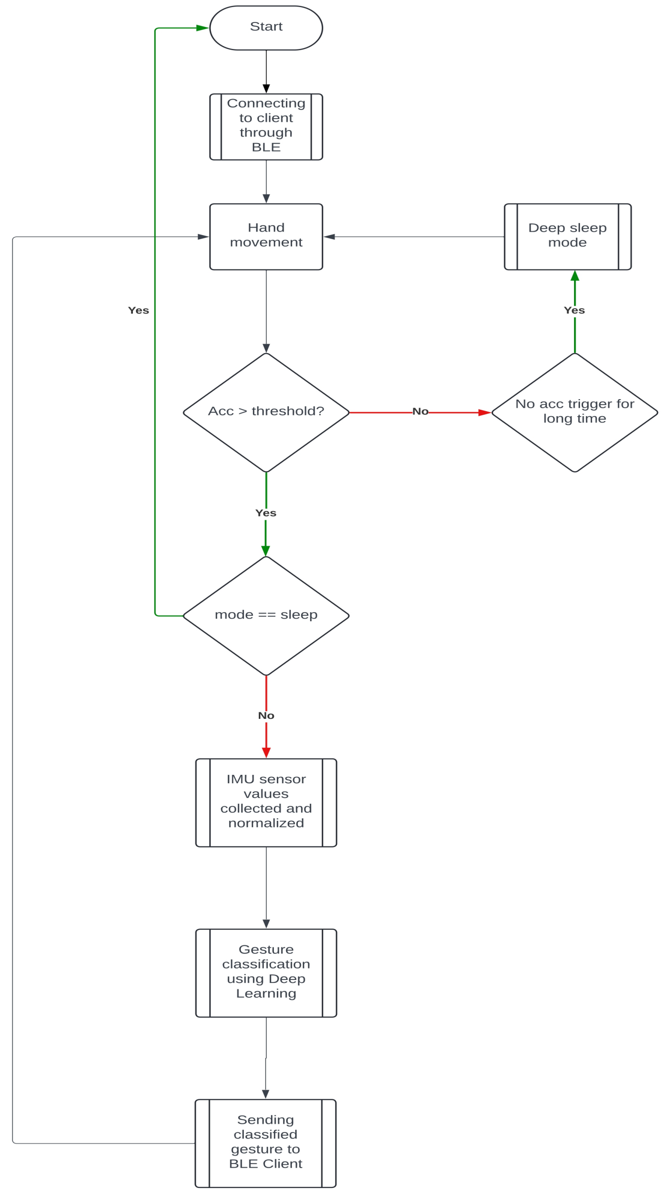


Fig. 2. Working Flow Chart

One of the key strengths of our approach lies in its focus on edge AI for wearables. By deploying complex deep learning models directly on low-power microcontrollers, we have reduced reliance on cloud connectivity, enabling real-time, low-latency operation. This not only enhances user experience but also addresses concerns regarding privacy and data security.

Additionally, our project emphasizes sensor fusion by integrating IMU sensors into a glove-based system for gesture recognition. This not only offers unique insights but also potentially improves accuracy compared to camera-based solutions. Moreover, by optimizing model accuracy while considering the constraints of low-power devices, we have highlighted the importance of balancing performance with energy efficiency, paving the way for more sustainable wearable technologies.

The impact of our project extends beyond technological innovation. By revolutionizing rehabilitation, enhancing human-computer interactions, and reducing reliance on e-waste, our solution has the potential to significantly improve both societal and environmental outcomes. Furthermore, our hardware design, data collection, and neural network selection process demonstrate a systematic approach to developing cutting-edge wearable technology solutions.

However, despite its promise, our project also faces certain limitations, including battery life constraints on small wearables, challenges in handling extremely complex gestures, and the need for large training datasets. Addressing these challenges will require ongoing research and development efforts.

In conclusion, our project represents a pioneering effort in leveraging edge AI and wearable technology to address real-world challenges. By pushing the boundaries of on-device AI, we have laid the groundwork for a new generation of intelligent wearables with diverse applications and far-reaching societal impact.

## References

[1] K. H. Shibly, S. Kumar Dey, M. A. Islam, and S. Iftekhar Showrav, "Design and development of hand gesture based virtual mouse," in *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, 2019, pp. 1–5. DOI: 10.1109/ICASERT.2019.8934612.

[2] P. Kumar, S. S. Rautaray, and A. Agrawal, "Hand data glove: A new generation real-time mouse for human-computer interaction," in *2012 1st International Conference on Recent Advances in Information Technology (RAIT)*, 2012, pp. 750–755. DOI: 10.1109/RAIT.2012.6194548.

[3] X. Guo, Z. Feng, C. Ai, *et al.*, "A novel method for data glove-based dynamic gesture recognition," in *2017 International Conference on Virtual Reality and Visualization (ICVRV)*, 2017, pp. 43–48. DOI: 10.1109/ICVRV.2017.00018.

[4] K. Wang and G. Zhao, "Gesture recognition based on flexible data glove using deep learning algorithms," in *2023 4th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT)*, 2023, pp. 113–117. DOI: 10.1109/AINIT59027.2023.10212923.