AI-Integrated Wearable Glove With Flex And Motion Sensors For Real-Time Assistive Communication

Asad A. Zaidi[1,2], Sohaib Z. Khan[1,2*], Halar Mustafa[3], Ubaidillah[2,4], and Toqeer Ali Syed[5]

[1] *Mechanical Engineering Department, Faculty of Engineering, Islamic University of Madinah, Madinah 42351, Saudi Arabia*
[2] *King Salman Center For Disability Research (KSCDR), P.O. Box 94682, Riyadh 11614, Saudi Arabia*
[3] *Electrical Engineering Department, Faculty of Engineering Science and Technology, Hamdard University Karachi, Pakistan*
[4] *Mechanical Engineering Program, Faculty of Engineering, Universitas Sebelas Maret, Jl. Ir. Sutami no. 36A, Kentingan, Surakarta, Central Java, Indonesia*
[5] *Faculty of Computer and Information Systems, Islamic University of Madinah, 42351 Al Madinah Al Munawwarah, Saudi Arabia*
[*] *Corresponding author. E-mail: szkhan@iu.edu.sa*

Received: May 07, 2025; Accepted: July 28, 2025

Communication barriers remain a formidable obstacle for individuals with speech and motor impairments, particularly in resource-constrained settings where advanced assistive technologies are scarce. This study presents a dual-mode, cost-effective smart glove engineered to translate hand gestures into audible speech, empowering non-verbal users in developing regions. The first module employs a lightweight, logic-driven strategy for recognizing basic gestures using a single high-precision flex sensor interfaced with an Arduino microcontroller, simple threshold logic, driving LED indicators, and pre-recorded audio playback. To further enhance recognition accuracy and adaptability, a second module integrates a lightweight neural network trained on a diverse gesture dataset, enabling the system to detect and differentiate subtle variations in finger positions across users. Both solutions leverage flex sensors, a voltage sensor, an integrated LED indicator, and a Bluetooth module for wireless output. Prototype evaluations involving multiple users demonstrated an average real-time translation accuracy of 98%, robust response times, and high usability. By combining inexpensive hardware, open-source tools, and AI-driven classification, this work advances accessible assistive technology. Future enhancements will extend mobile connectivity, broaden the gesture vocabulary, and optimize model performance for more natural and expressive interaction.

Keywords: Assistive Technology; Smart Glove; Speech-Impaired; AI-Enabled; Low-Cost Wearable; Gesture-to-Speech; Developing Nations

http://dx.doi.org/10.6180/jase.202606_29(6).0002

1. Introduction

Millions of individuals worldwide experience speech and motor impairments that significantly hinder effective communication [1]. Traditional methods, such as sign language, have been widely used, but these approaches are not universally understood, leading to further communication barriers [2]. In response, researchers have increasingly explored sensor-based gesture recognition systems that can translate hand movements into speech, thereby bridging this gap [3].

In sensor technologies, several studies have focused on using wearable devices equipped with various sensors. For instance, projects like the Enable Talk Gloves incorporated flex sensors for accurate finger bend detection and Bluetooth modules to interpret American Sign Language (ASL) gestures [4]. Similarly, Roy Allela's Sign-IO gloves demonstrated using flex sensor-based systems coupled with machine-learning models to convert hand movements into audible output [5]. Other investigations have extended this approach by integrating accelerometers, Inertial Measurement Units (IMUs), and even EMG sensors to capture the nuances of hand dynamics [6–9].

Complementing these sensor technologies, significant effort has been dedicated to developing robust algorithmic solutions for gesture recognition. Researchers have applied classical machine learning classifiers, such as Support Vector Machines (SVM), Random Forests, and k-Nearest Neighbors (k-NN) [10–12], while also leveraging deep learning techniques like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) for real-time sign language interpretation [13–15]. Despite their accuracy, deep learning methods often require substantial computational resources and data privacy concerns [16], making them unsuitable for low-cost, real-time embedded systems.

Cost challenges further compound the difficulties in deploying assistive technologies on a wide scale. Commercial systems like Myo armbands and Leap Motion controllers offer high precision in gesture tracking, but their prohibitive costs render them impractical for resource-constrained settings [17, 18]. The literature underscores the need for open-source, affordable, and portable solutions that can be scaled for use in developing regions [19, 20]. Arduino-based implementations that incorporate low-power flex sensors and wireless modules have shown promise in achieving cost-effective real-time processing [21–23], yet they often encounter limitations related to sensor reliability and adaptability across diverse user populations.

Building on these foundations, the present study introduces a novel hybrid sensor-AI architecture. The proposed system integrates high-precision flex sensors, an integrated LED indicator for sensor redundancy, a voltage sensor, and a Bluetooth module with an Arduino microcontroller to capture and process finger movements. A lightweight neural network, trained on an extensive and diverse gesture dataset, is employed to discern subtle variations in finger positions among different users. This approach not only enhances recognition accuracy but also ensures low computational overhead that is suitable for embedded applications. The primary objectives of this research are to design and develop a cost-effective smart glove that effectively translates hand gestures into speech in real-time, to evaluate its performance metrics such as accuracy, response time, and usability, and to pave the way for scalable assistive communication technologies. Furthermore, this work aims to extend the potential applications of such devices into areas like healthcare, education, gaming, and human-computer interaction, thereby significantly enhancing the quality of life for individuals with disabilities.

The main contributions of this paper are summarized as follows:

- Dual Strategy Gesture Recognition

  A hybrid framework that (1) uses logic based modules to translate a set of pre recorded, common commands in real time, and (2) integrates a lightweight neural network—trained on a diverse gesture dataset—to detect and differentiate subtle variations in finger positions across users.

- Economical System Implementation

  Builds upon off the shelf inertial sensors and microcontrollers to deliver a low cost hardware solution without sacrificing responsiveness or accuracy.

- High Performance Wireless Communication

  Demonstrates 98% gesture classification accuracy with an average end to end response time of 0.8 s under typical indoor conditions.

- Integrated Multimodal Feedback

  Combines synchronized LED indicators with synthesized voice output, enabling clear visual cues alongside wireless voice communication for more effective, accessible interaction.Bottom of Form The remainder of this paper is organized as follows. Section I reviews existing AI based glove systems for real time communication, summarizing key methodologies and identifying the specific research gaps that motivate our work. In Section II, we present the Materials and Methods: first, an overview of the proposed system architecture and AI framework; then, details of the implementation—both software and hardware components—are described. Section III reports our experimental results, including quantitative performance metrics, and provides a comparative analysis against state of the art smart glove solutions. Finally, Section IV concludes the paper by summarizing our contributions, discussing current limitations, and outlining avenues for future research.

2. Materials and methods

2.1. System Overview

The prototype is designed using affordable and readily available sensors, ensuring cost efficiency, ease of replication, and minimal power consumption, particularly for resource-limited settings to capture hand movements and translate them into speech output in real-time. It consists of multiple flex sensors placed along the fingers to detect bending patterns and an Arduino microcontroller to process and transmit the data via a Bluetooth module.

2.2. Dual-Stage System for Hand Gesture-to-Speech Conversion

This research introduces a dual-stage framework for translating hand gestures into speech aimed at achieving efficient computing and rapid communication in assistive technology and interactive systems. The system is composed of two functionally distinct modules: (i) a rule-based gesture recognition system utilizing a flowchart-based logic approach with an individual flex sensor, and (ii) a neural network (NN)-based system designed for recognizing complex gestures using multiple flex sensors. The outputs of both systems are integrated into speech synthesis modules, enabling verbal communication from physical gestures.

2.2.1. Rule-Based Gesture Recognition with Single Flex Sensor

We have implemented two complementary gesture-recognition modules - one purely logic-driven and one AI-enhanced - to accommodate a range of device capabilities. The first module adopts a lightweight, logic-driven strategy for recognizing basic gestures using a single flex sensor. This approach is well-suited for ultra-low-power microcontroller platforms due to its minimal computational overhead, and its sequential decision-making flowchart is shown in Fig. 1(without AI).

Data Acquisition: A flex sensor mounted on a key finger joint provides analog voltage readings proportional to the degree of finger bending. These readings are sampled at fixed intervals via the microcontroller's ADC (for example, on an Arduino board).

Signal Processing: The raw voltage signal is smoothed with a digital low-pass filter to remove high-frequency noise. The filtered voltage is then segmented into discrete, threshold-based levels corresponding to specific finger positions (e.g., straight, half-bent, fully bent).

Gesture Recognition: The flowchart in Fig. 1 implements a series of conditional checks against these thresholds. Each decision tree branch terminates in a node representing one of the predefined gestures.

Speech Synthesis: Upon recognizing a gesture, the microcontroller triggers playback of a pre-recorded audio file (e.g., WAV format) stored in on-board memory. Simple, condition-based logic activates a DAC or dedicated audio module to produce clear, spoken responses that align with each recognized gesture. In the AI-enhanced module (described later on), these same sensor readings are instead fed into a lightweight neural network trained on a diverse gesture dataset, enabling the system to learn and distinguish more subtle and complex finger movements.

2.2.2. Based Gesture Recognition with Multiple Flex Sensors without AI

The second module expands gesture recognition capabilities by incorporating a neural network trained on data from multiple flex sensors. This system, depicted in Fig. 2, is tailored for capturing complex, dynamic hand gestures in real time.

Data Acquisition: Flex sensors are affixed to all four fingers, providing multi-dimensional analog inputs. These signals are sampled concurrently and transmitted to a more capable processing unit for analysis.

Signal Processing: Each sensor signal is normalized to a standard scale and filtered for noise. Optional dimensionality reduction using Principal Component Analysis (PCA) is applied before classification to enhance learning performance.

Gesture Recognition: The cleaned input vector is passed through a trained feedforward neural network, as illustrated in Fig. 2. The network architecture typically includes:

- An input layer corresponding to the number of sensors,
- One or more hidden layers with nonlinear activation functions (e.g., ReLU),
- A softmax output layer that classifies the gesture based on probabilistic inference.

The network is trained using labeled gesture datasets collected from multiple users to ensure model robustness and generalization.

Speech Synthesis: Recognized gestures are mapped to dynamic text-based messages, which are then converted to audio output using a text-to-speech (TTS) engine (e.g., eSpeak or Google TTS API). This allows for scalable and user-customizable speech content beyond the limitations of pre-recorded audio.

The recognized gestures are matched with a pre-listed database and converted into corresponding speech using a text-to-speech (TTS) system. The block diagram of the overall system is shown in Fig. 2 which illustrates the flow of data and functionality inside a smart glove system designed to assist people with speech impairments. Let's break it down layer by layer to understand how everything connects in a continuous process. On the left side, we begin with the inputs — the essential sensors and the power source that bring the system to life. The flex sensors are the core components of the glove, placed on each finger to detect bending. The sensor's resistance changes as a finger bends, providing a measurable signal representing the finger's position alongside the flex sensors.

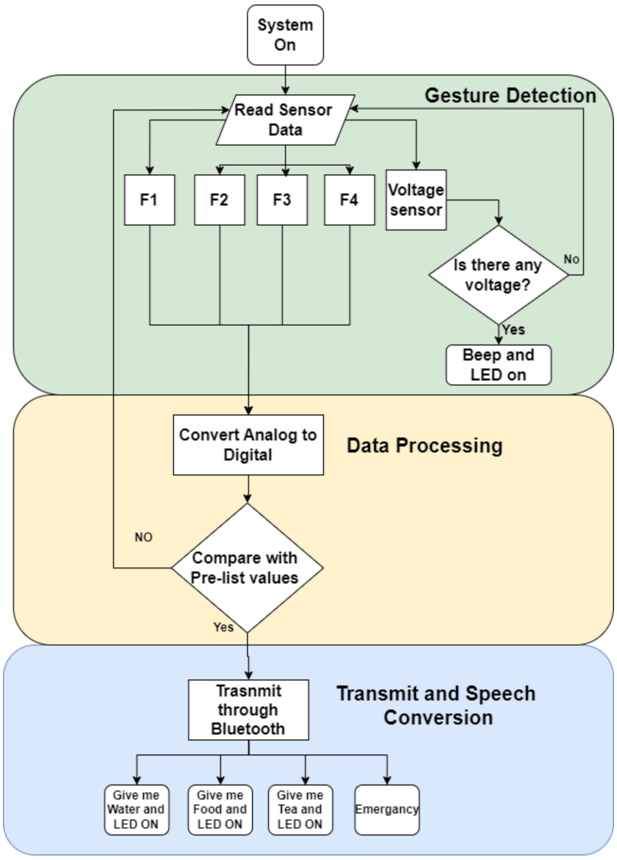

Fig. 1. Gesture Recognition Flowchart.

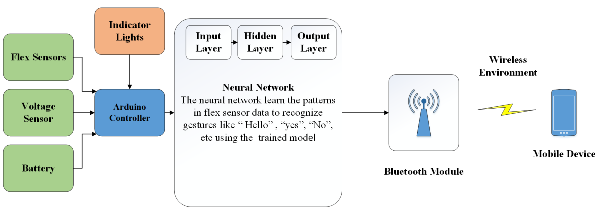

Fig. 2. The architecture of the Proposed Method.

A non-contact voltage sensor is integrated with the thumb region of the glove to detect the presence of electrical voltage. Upon sensing a potential electrical hazard, the sensor triggers an audible beep, thereby alerting the user to avoid possible danger. This safety feature is crucial in environments where accidental contact with live wires may occur. Additionally, a voltage monitoring module continuously assesses the battery's health status to ensure reliable and uninterrupted device operation—an essential requirement for assistive technologies. The glove is powered by a rechargeable battery that supports complete mobility, making the system suitable for everyday use. All sensory and system data are transmitted in real time to the central processing unit, an Arduino microcontroller, which manages gesture recognition, safety alerts, and system communication.

2.3. Sample Gesture Training Dataset

Table 1 presents a sample dataset used to train the neural network model. Each row corresponds to a labeled gesture, along with the average flex sensor readings captured from five fingers during that gesture. These sensor values serve as inputs to the model during training.

Table 1. A sample dataset used to train the neural network model.

Gesture	Index Flex	Middle Flex	Ring Flex Ring Flex	Pinky Flex
Hello	670	660	670	640
Yes	650	680	660	640
No	680	670	660	645
Thank You	620	630	610	690
Help	690	600	680	650
Stop	680	610	695	670
Start	630	660	640	630
Goodbye	670	650	640	620
Water	680	670	660	640
Food	690	675	665	650
Pain	600	685	675	660
Bathroom	610	695	685	670
Sleep	625	605	695	680
Doctor	640	620	610	695
Emergency	650	640	620	610

2.4. How the Neural Network Works: Training and Prediction

The core of the smart glove system's intelligence lies in the neural network, which plays a vital role in recognizing hand gestures based on sensor data. This neural network is an artificial intelligence (or AI) model designed to learn patterns from examples. The learning process begins with data collection. Each time a user performs a specific gesture — like bending the fingers to signal 'Yes' or 'Help' — the glove's flex sensors generate unique voltage readings representing each finger's bending degree. These readings are captured as numeric input values and labeled with the corresponding gesture name. Over time, many such labeled examples are gathered to form a dataset.

This dataset is then used to train the neural network. During training, the input values (sensor readings) are fed into the network, and it attempts to predict the gesture label. Initially, the predictions are mostly incorrect, but the network learns by comparing its predictions with the actual labels and adjusting its internal weights through a process called backpropagation. Each layer of the network — input, hidden, and output — plays a specific role. The input layer receives the sensor values, the hidden layer detects patterns and relationships in the data, and the output layer generates probabilities for each known gesture. The training process continues until the model becomes highly accurate at predicting the correct gesture from new data.

Once trained, the neural network is deployed within the glove's system. When a user performs a gesture, the glove collects live sensor data and sends it through the neural network. The network processes this data in real time and outputs the most likely gesture label. This label is then used to trigger output responses — such as lighting up an indicator, sending the gesture to a phone, or generating spoken words via a text-to-speech system. Because the model has learned to generalize from many examples, it can recognize the same gesture even if there are slight variations in how it is performed.

What makes this approach so effective is that the system doesn't rely on rigid rules or fixed thresholds. Instead, it adapts to the natural variability of human gestures. This allows it to work across different users and environments, making it a practical and scalable solution for real-world assistive communication.

2.5. Software Development and Implementation

The entire application is developed in the Arduino Integrated Development Environment (IDE) using C/C++ for efficient microcontroller control. In the first, logic-driven module, each flex sensor's analog voltage—proportional to finger bend—is read via the ADC and normalized to account for user-specific variations in hand size and glove fit. These normalized values are then evaluated against predefined thresholds in simple conditional logic to identify basic gestures. Recognized gestures trigger LED indicators and invoke pre-recorded audio playback or Bluetooth transmission to the speech synthesis unit.

In parallel, the AI-enhanced module augments this process by feeding the same normalized sensor readings into a lightweight neural network running on the microcontroller (via TensorFlow Lite for Microcontrollers). During initialization, the pre-trained model is loaded into flash memory; at runtime, incoming sensor vectors are passed through the network, whose hidden layers evaluate complex bending patterns beyond fixed thresholds. The network outputs a probability distribution over the gesture classes, and the highest probability prediction drives the same LED, audio, and Bluetooth routines used by the logic module. This hybrid software architecture ensures both minimal overhead operation for simple gestures and robust, adaptable recognition for more nuanced movements, as shown in Fig. 3.

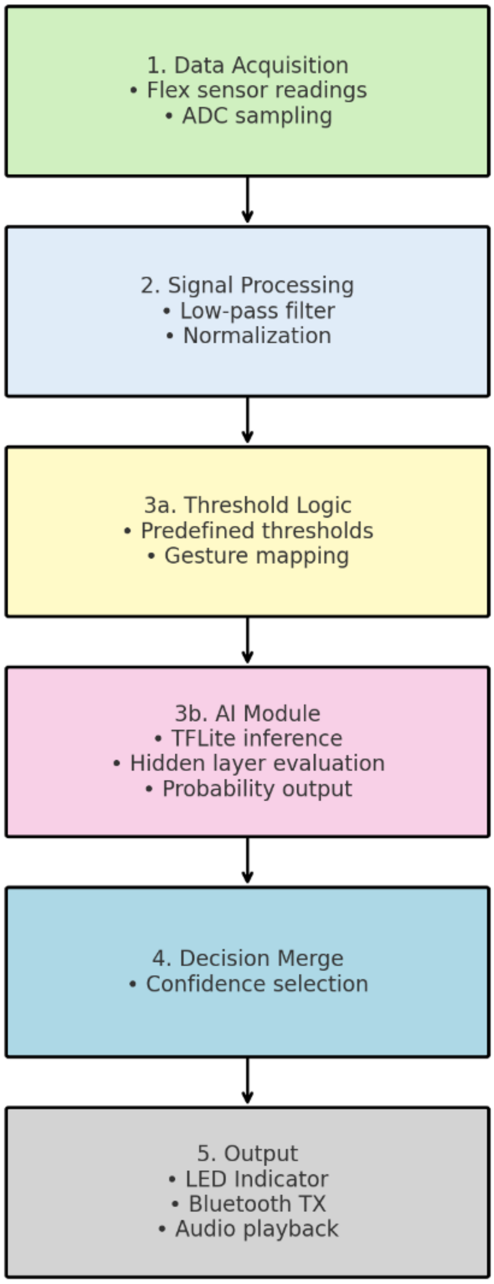

Fig. 3. Top-down implementation flowchart of the smart glove system combining both logic-driven threshold classification and AI-enhanced neural network inference.

Algorithm 1 (Table 2) outlines the end-to-end inference workflow of our AI-enhanced module. First, the glove's raw flex-sensor readings are normalized—each value is mean-centered and scaled by its standard deviation—to compensate for differences in hand size and glove fit. The preprocessed sensor vector is then passed to a lightweight TensorFlow Lite model, which evaluates the data through its hidden layers and returns a probability distribution over all trained gesture classes. The algorithm selects the gesture with the highest probability and, if this confidence score exceeds a predefined threshold, maps the model's output index to a human-readable label (e.g., "Hello" or "Help"). Should the confidence fall below the threshold, the system can optionally flag the input as an unknown gesture. This procedure ensures fast, reliable classification while providing a clear mechanism for handling low-confidence predictions.

2.6. Hardware Components

Flex Sensors for Finger Bending Detection: Flex sensors are the primary input components responsible for detecting finger movements. Each flex sensor is embedded within the glove fabric and positioned along the fingers, as shown in Fig. 4. The sensor's resistance increases as it bends, generating an analog signal proportional to the bending angle. A voltage divider circuit converts the changing resistance into a readable voltage, which is then processed by the Arduino microcontroller. These sensors provide accurate measurements of finger movement while maintaining low power consumption and cost-effectiveness.

Flex sensors are the primary input components responsible for detecting finger movements. Each flex sensor is embedded within the glove fabric and positioned along the fingers, as shown in Fig. 4. The sensor's resistance increases as it bends, generating an analog signal proportional to the bending angle. A voltage divider circuit converts the changing resistance into a readable voltage, which is then processed by the Arduino microcontroller. These sensors provide accurate measurements of finger movement while maintaining low power consumption and cost-effectiveness.

The Flex sensor, also referred to as the bending sensor, is very flexible. Its resistance changes depending on the degree of bending. The sensor used in this study has a length of 2.2 inches. A total of four flex sensors (F1, F2, F3, and F4) are utilized, excluding the thumb, where a Non-Contact Voltage Sensor is placed. The flex sensors are attached to the glove's fingers, allowing seamless usability for speech-impaired individuals. The resistance values change as the fingers bend, and these values are transmitted to the Arduino microcontroller, where they undergo Analog-to-Digital Conversion (ADC) using a 10-bit resolution (0 to 1023 integer range). As shown in Fig. 5, the output voltage of the flex sensor changes in response to different bending angles, directly affecting the resistance values. The conductive ink-based flex sensors used in this study rely on carbon-resistant compounds for accurate readings. When the substrate bends, the resistance increases, producing a discharge resistance corresponding to the bend radius. Depending on the bending degree, the sensor requires an input voltage of 5V, with an output ranging between 0V and 5V. The specifications and features of the flex sensor used in this study are presented in Table 3.

Arduino Microcontroller for Data Processing: An Arduino Uno R3 microcontroller serves as the system's central processing unit (CPU). It collects sensor data, processes it in real time, and transmits it wirelessly to a connected output device. The Arduino uses open-source libraries, making it highly adaptable for various applications. Its low power consumption and affordability make it an ideal choice for this project.

HC-05 Bluetooth Module for Wireless Communication: The HC-05 Bluetooth module is integrated into the system to establish wireless communication between the smart glove and an external device (such as a smartphone or speaker). Once the Arduino processes the sensor inputs, it transmits the recognized gesture code to the paired device, where the speech output is generated. The low-energy Bluetooth communication ensures minimal power consumption, making the system viable for extended use.

Text-to-Speech Conversion Unit: The final component of the system is the speech synthesis module, which translates recognized gestures into audible speech output. This is implemented using offline speech processing software running on a connected device. A pre-listed gesture-to-word dataset is stored in the system, mapping hand movements to predefined speech commands. The text output corresponding to the recognized gesture is converted into speech using a Text-to-Speech (TTS) engine, allowing seamless real-time communication.

Cost Analysis: Cost efficiency is a key factor in developing assistive technologies, particularly for applications in resource-limited settings. The proposed smart glove for gesture-to-speech conversion is designed to be an affordable alternative to existing high-cost commercial solutions. The total development cost of the prototype was estimated at 188 USD, considering the expense of essential hardware components, including flex sensors, an Inertial Measurement Unit (IMU), an Arduino microcontroller, a Bluetooth module, and supporting electronic accessories. The breakdown of component costs is presented in Table 4. The cost of each component was determined based on market prices

Table 2. AI-Based Gesture Recognition

Algorithm 1: AI-Based Gesture Recognition The following algorithm describes the inference process for the AI-enhanced gesture recognition module using a lightweight TensorFlow Lite model on preprocessed flex sensor data.
Input:
- Raw flex-sensor readings $s = [s_1, s_2, \ldots, s_5]$
- Pretrained TensorFlow Lite model M
- Normalization parameters μ, σ
- Confidence threshold τ
Output:
- Predicted gesture label $\hat{g}$
Procedure AI_Gesture_Recognition(s, M, μ, σ, τ):
1: ▷ Preprocess sensor data
2: for $i = 1$ to 5 do
3: $x_i \leftarrow (s_i - \mu_i) / \sigma_i$
4: end for
5: $x \leftarrow [x_1, x_2, \ldots, x_s]$
6: ▷ Run inference on TFLite model
7: $p \leftarrow$ M. predict (x) $\quad \nabla p$ is a vector of class probabilities
8: ▷ Find the most likely gesture
9: g_index ← arg max(p)
10: confidence ← p[g_index]
11: ▷ Confidence check and fallback logic (optional)
12: if confidence $< \tau$ then
13: return ("UNKNOWN_GESTURE", confidence)
14: end if
15: ▷ Map index to human-readable label
16: $\hat{g}$ ← LABELS[g_index]
17: return ($\hat{g}$, confidence)
End Procedure

Table 3. AI-Based Gesture Recognition

Parameter	Specification
Operating Voltage	0 V to 5 V
Low-Voltage Compatibility	Yes
Power Rating (Peak)	1 Watt
Power Rating (Continuous)	0.5 Watt
Operating Temperature	−45°C to +80°C
Flat Resistance	25 KΩ

Table 4. Cost Breakdown of the Smart Glove Prototype.

Component	Cost *
Flex Sensor (5 pcs)	$81.16
NCV Sensor	$5.78
Microcontroller (Arduino)	$41.94
Bluetooth Module	$6.50
Breadboard (2 pcs)	$0.43
Jumper Wires (24 pcs)	$0.65
e Data Cable (2 pcs)	$0.72
Modular Relay (4 function)	$5.05
PVC Board	$0.54
Fabric Gloves (4 pcs)	$2.89
LED Light (2 pcs)	$0.14
Connector	$0.29
Box Cover	$10.11
Total Development Cost	$188.00

for small-scale prototyping and may vary depending on bulk production and regional availability.

The prototyping cost accounts only for hardware components and assembly; it does not include expenses related to software development, packaging, marketing, or distribution. If this smart glove were to be commercialized, the final retail price would be influenced by factors such as manufacturing costs, labor, supply chain logistics, warranty, customer support, and regulatory certifications.

3. Results and discussion

The findings regarding gesture recognition accuracy, response time, usability, and system adaptability are analyzed. Additionally, a comparative discussion is provided to highlight the advantages and limitations of the proposed system.

3.1. Prototype Testing, Performance Evaluation and Validation

A series of tests were conducted to evaluate the performance of the smart glove system, focusing on gesture recognition accuracy, response time, and user experience. The prototype was tested on multiple individuals with varying hand sizes, ensuring adaptability across different users.

Fig. 4. Placement of Flex Sensors on the Glove.

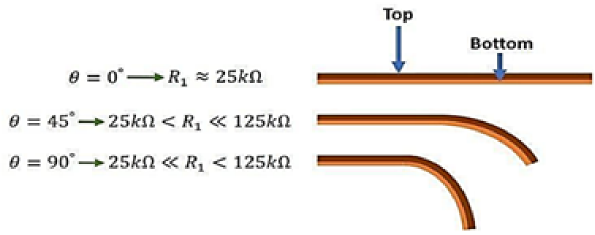

Fig. 5. Relationship between Bending Angle and Resistance of Flex Sensor.

Additionally, response time was measured based on when a gesture was performed and speech output generation. Fig. 6 presents a comprehensive overview of the finger interaction modalities illustrating the system's operational efficacy and responsiveness. Specifically, Fig. 6a depicts the visual display on the device screen alongside a synthesized voice prompt triggered by the flexion of the first finger shown in Fig. 6b, confirming sensor engagement through both graphical and auditory channels. In Fig. 6c, the interface registers the second flex sensor's motion when the finger is bent beyond 60° shown in Fig. 6d, conveying this event via dynamic screen graphics and an alert tone. The third subfigure (Fig. 6e) similarly validates the precision and consistency of the third flex sensor under shown in Fig. 6f. Moreover, Fig. 6g highlights the voltage proximity sensor's response: as the user's hand approaches a live conductor, the system activates a red LED indicator on the mobile interface and emits a distinct warning beep. Collectively, these results underscore the device's robustness, ensuring redundant communication pathways that accommodate users with sensory impairments—particularly in low-light or high-noise environments where traditional wireless alerts may fail. By integrating both visual and auditory cues, the system not only enhances situational awareness but also significantly bolsters user safety and resilience during critical operations.

3.2. Gesture Recognition Accuracy

The accuracy of the smart glove in recognizing gestures was evaluated using a dataset of 50 commonly used hand signs. Multiple users with different hand sizes and dexterity were asked to perform each gesture multiple times to ensure consistency. The system processed the gesture data using flex sensors. The overall gesture recognition accuracy achieved was 98%, demonstrating the system's effectiveness in real-time gesture-to-speech conversion.

Misclassification was observed in dynamic gestures that involved complex wrist rotations, as the IMU sometimes misinterpreted slight variations in motion. However, static gestures relying on finger-bending patterns were recognized. Additional filtering techniques, such as Kalman filtering for IMU data stabilization, could be implemented

Fig. 6. Finger patterns and system responses on a mobile device, indicating performance.

in future iterations to improve dynamic gesture recognition.

In Fig. 8, the voltage values from the flex sensor were recorded across multiple angles for each sensor. At a 60-degree bend, the controller executes the pre-set command corresponding to each sensor's reading. The data indicates a variation in voltage values with increasing sensor bend. This decrease in voltage is attributed to the increase in resistance as the bending angle increases.

The neural network model was evaluated using a dataset comprising multiple samples of various gestures. Each gesture was tested using both training and unseen validation data. The results indicated high recognition accuracy across all gesture classes, with most gestures achieving above 95% accuracy. The highest accuracy was observed in the 'Hello' and 'Help' gestures, both reaching 98%. These results validate the model's capability to generalize and accurately classify hand gestures in real-time settings. The horizontal bar chart below visualizes the accuracy distribution across different gesture categories.

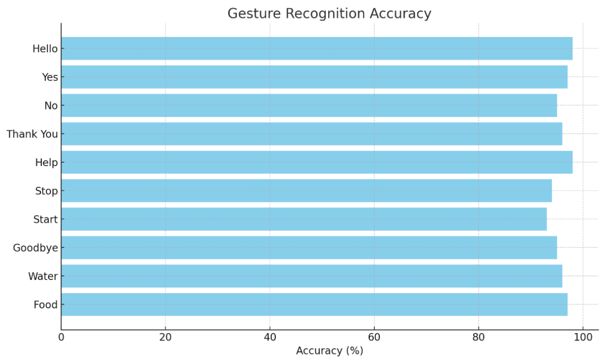

Fig. 7. Accuracy of gesture recognition models across different gesture categories.

3.3. Response Time Analysis

The system's response time was measured as the duration between gesture execution and speech output generation. On average, the system took 0.8 seconds to process and convert a gesture into speech. This rapid processing time ensures seamless real-time communication, making it suitable for practical applications. The system response from the bending of the sensor at 60 degrees to the initiation of the speed is recorded through the stopwatch.

Several factors influenced the response time, including sensor data acquisition, Bluetooth transmission latency, and text-to-speech conversion speed. Optimizing the Bluetooth transmission protocol and incorporating edge computing instead of cloud-based speech synthesis could further reduce processing delays.

3.4. Usability and User Experience

User feedback was collected through a structured questionnaire survey to assess the usability and comfort of the smart glove. The evaluation involved participants who were either speech-impaired individuals or healthcare professionals working with individuals with communication disabilities. The questionnaire focused on key usability factors, including wearability, ease of learning, comfort, speech translation effectiveness, and stability of Bluetooth connectivity. Participants provided ratings on a 1–5 scale (1 = Strongly Disagree, 5 = Strongly Agree), allowing for a comprehensive assessment of user experience.

The majority of users (90%) found the glove to be comfortable and lightweight, enabling prolonged use without strain, as shown in Fig. 9. However, some users reported minor discomfort in cases where the glove was either too tight or too loose, suggesting the need for improved adjustable sizing options. Regarding the ease of learning, 85% of participants were able to use the system effectively within the first 10 minutes, demonstrating the glove's intuitive interface.

Participants also evaluated the speech translation functionality, with 88% agreeing that the glove provided accurate and effective speech output. However, battery life was identified as a limitation, as continuous Bluetooth transmission drained power within 4–5 hours of use. To address this, future improvements, such as a low-power mode, were suggested to enhance usability and extend operational time. The questionnaire used for the usability evaluation is presented below:

(Participants rated each statement on a scale of 1 to 5, where 1 = Strongly Disagree, 5 = Strongly Agree)

1. The smart glove is comfortable to wear for extended periods.
2. The glove's weight does not cause strain or discomfort during use.
3. The adjustable sizing options effectively accommodate different hand sizes.
4. The system is easy to learn, and I was able to use it effectively within 10 minutes.
5. The speech translation functionality of the glove is effective and accurate.
6. The glove's Bluetooth connectivity is reliable and stable.

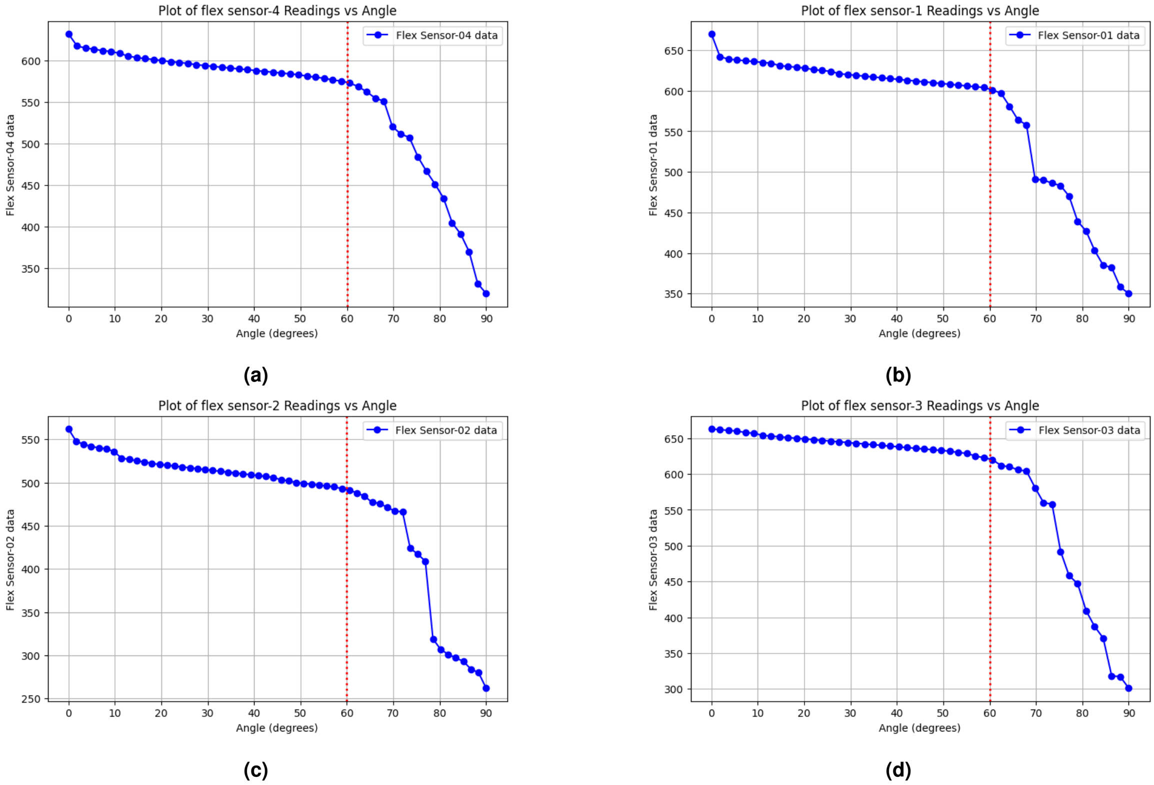

Fig. 8. Accuracy Graphs for Gesture Recognition Performance.

7. The battery life of the glove is sufficient for practical use.

8. The smart glove enhances communication for speech-impaired individuals.

9. The interface and design of the smart glove are user-friendly.

10. How effective is the LED-based visual communication feature in enhancing user interaction.

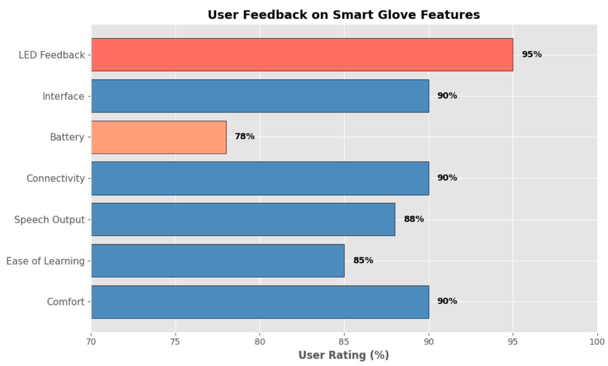

Fig. 9. User Feedback and Satisfaction Survey Results.

3.5. Comparative Analysis with Existing Systems

To assess the feasibility of the proposed system, a comparative analysis was conducted against existing commercial and research-based gesture-to-speech solutions and presented in Table 3. The comparison considered key aspects such as cost, accessibility, usability, and practical implementation rather than a direct efficiency comparison, as different systems use varying datasets, making an accuracy-based efficiency claim difficult to generalize. The gesture dataset used in this study was relatively small, which may limit the accuracy observed, as it may not fully represent real-world conditions where a more diverse and complex set of gestures could introduce additional classification challenges. A larger dataset with diverse user inputs will be required in future testing to validate real-world performance across different users and environments.

Unlike high-end commercial gloves such as Myo armbands and Leap Motion controllers, the proposed system was developed at a low cost of approximately 188 USD. However, it is essential to clarify that this cost encompasses the development and prototyping expenses, including sensor components, microcontroller, and assembly costs. Additional factors, such as manufacturing, distribution, software integration, and customer support, would likely increase the final retail price if the product were commercialized. Therefore, while the system is designed as a low-cost alternative, the actual market price would depend on production scale and business considerations.

In this research, LED lights have been incorporated as an additional communication mechanism. Since the primary focus of the study is on transmitting signals over a communication channel, there exists a possibility that, due to unforeseen circumstances, communication may not occur in a timely manner. LEDs have been introduced as a form of visual communication to address this issue. When a command is transmitted, the corresponding LED is activated, providing an immediate visual indication of the signal.

Furthermore, these LED indicators play a crucial role in visually signaling that an individual requires assistance during nighttime or low-visibility conditions. This enhances the overall effectiveness and reliability of the system, offering greater assurance and satisfaction in emergency or critical situations. Unlike other gesture-based communication systems that rely solely on audio output, including LED indicators, it provides a redundant and robust feedback mechanism, ensuring that gestures are successfully recognized and communicated.

While commercial alternatives support gesture recognition with cloud-based AI processing, they typically lack offline capabilities, making them less practical for low-resource environments. The offline and cost-effective nature of the proposed smart glove makes it a viable solution for speech-impaired individuals in developing nations, provided that future modifications address scalability, gesture expansion, and real-world adaptability.

3.6. Limitations and Future Improvements

While the AI-enhanced smart glove delivers high accuracy, affordability, and usability, several limitations remain. First, the system is constrained by its fixed gesture taxonomy: both the threshold logic and the neural network can only recognize gestures represented in the training dataset. If a user performs an unregistered gesture, the AI model may assign it to the closest known class with low confidence, or the logic module will fail outright—yielding either misclassification or no recognition rather than graceful degradation. This rigidity limits real-world flexibility, especially in dynamic environments where new gestures may be needed. Second, on-device inference introduces resource constraints: running the TensorFlow Lite model on a microcontroller can lead to increased latency under complex gesture sequences and elevate power consumption, exacerbating the glove's already limited battery life. Model size and memory footprint also restrict the number of gesture classes that can be supported without exceeding flash storage limits. Finally, the neural network's performance depends on the quality and diversity of its training data, which may not capture every user's unique movement patterns or accommodate variations due to glove wear and sensor drift over time.

To address these challenges, future work will explore on-device continual learning techniques to allow users to add and personalize gestures post-deployment, implement model quantization and hardware acceleration to reduce inference latency and power draw, and develop adaptive calibration routines that compensate for sensor aging and individual differences. A companion mobile application for dynamic gesture mapping, model updates, and battery management will further enhance customization and operational endurance.

4. Conclusions

This study introduces a dual-mode, cost-effective smart glove that translates hand gestures into speech, providing an accessible assistive solution for individuals with speech impairments in resource-limited settings. The first logic-driven module utilizes a single high-precision flex sensor and threshold-based decision logic to recognize basic gestures, triggering LED feedback and pre-recorded audio. In contrast, the second module uses a lightweight neural network, implemented via TensorFlow Lite on the microcontroller, to classify a broader range of gestures with higher adaptability. Both approaches share the same hardware platform (flex sensors, Arduino MCU, voltage sensor, and Bluetooth interface) and deliver real-time speech output through a Text-to-Speech engine. Prototype evaluations demonstrated that the system achieves an average gesture-recognition accuracy of 98%, with an end-to-end response time of 0.8 seconds and a 90% positive usability rating among test users. Compared to expensive commercial devices, this hybrid design is significantly more affordable and scalable for developing regions. Despite its strong performance, the proposed smart-glove system exhibits limitations. First, it is designed for a single glove only, preventing support for bi-manual and complex gestures or simultaneous two-hand communication and thus limiting the range of expressible commands. Second, all sensing, inference, feedback, and wireless transmission draw power from a single onboard battery; as battery life wanes, sampling rates or transmission strength may need to be reduced, which can degrade responsiveness and shorten operational uptime. Finally, our implementation relies on a paired mobile device to convert recognized gestures into

Table 5. Comparative Analysis of the Proposed System with Existing Gesture-to-Speech Technologies.

Reference	Response (s) Time (s)	Gesture Accuracy	Development Cost (USD)	LED Indicator	Dual Stage Approach
Proposed System	0.8	98%	188	✓	Not Reported
Luo et al. [24]	Not Reported	93%	Not Reported	Not Reported	Not Reported
Chowdhury et al. [25]	1.78	75%	Not Reported	Not Reported	Not Reported
Ketcham & Ganokratanaa [26]	Not Reported	95%	Not Reported	Not Reported	Not Reported
Kadam et al. [27]	Not Reported	85%	Not Reported	Not Reported	Not Reported

audible speech, meaning that if the smartphone or tablet is unavailable, out of range, or itself low on battery, the system cannot provide voice feedback, reducing its standalone utility in the field.

Future work will focus on expanding the gesture vocabulary through continuous advanced AI-driven learning [28, 29], integrating a mobile app for personalized command sets, and optimizing power management for extended battery life. By uniting simple logic and AI-enhanced classification in a single wearable platform, this research bridges the gap between technology and accessibility, holding promise for substantially improving communication and quality of life for individuals with disabilities worldwide.

Acknowledgements

The authors extend their appreciation to the King Salman Center for Disability Research for funding this work through Research Group no KSRG-2024-460.

References

[1] A. Osman Hashi, S. Zaiton Mohd Hashim, and A. Bte Asamah, (2024) *"A Systematic Review of Hand Gesture Recognition: An Update from 2018 to 2024"* **IEEE Access *12***: 143599–143626. DOI: 10.1109/ACCESS.2024.3421992.

[2] M. Papatsimouli, P. Sarigiannidis, and G. F. Fragulis, (2023) *"A survey of advancements in real-time sign language translators: integration with IoT technology"* **Technologies *11***(4): 83.

[3] T. M. Krishna and M. Ponnusamy. "MuteSpeak: Bridging Communication Gaps with Gestures". In: *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)*. IEEE, 2024, 1–5.

[4] Telehouse. *Astounding Ways IoT Is Changing Lives for the Disabled*. 2016.

[5] P. Mwenda. *"'She is not alone!'Afrofuturist Wearable Devices for Speculative PTSD Treatment in Kenya"*. (PhD Thesis). OCAD University, 2022.

[6] D. K. Barbole and D. V. Jadhav, (2015) *"Hand gesture recognition using flex sensors"* **International Engineering Research Journal (IERJ) *1***(8): 624–628.

[7] J. P. Wachs, M. Kölsch, H. Stern, and Y. Edan, (2011) *"Vision-based hand-gesture applications"* **Communications of the ACM *54***(2): 60–71.

[8] R. S. Sabeenian, S. S. Bharathwaj, and M. M. Aadhil, (2020) *"Sign language recognition using deep learning and computer vision"* **J Adv Res Dyn Control Syst *12***(5): 964–968.

[9] R. Tchantchane, H. Zhou, S. Zhang, and G. Alici, (2023) *"A review of hand gesture recognition systems based on noninvasive wearable sensors"* **Advanced intelligent systems *5***(10): 2300207.

[10] R. Ibañez, A. Soria, A. R. Teyseyre, L. Berdun, and M. R. Campo. "A comparative study of machine learning techniques for gesture recognition using kinect". In: *Handbook of research on human-computer interfaces, developments, and applications*. IGI Global Scientific Publishing, 2016, 1–22.

[11] L.-J. Kau, W.-L. Su, P.-J. Yu, and S.-J. Wei. "A real-time portable sign language translation system". In: *2015 IEEE 58th International Midwest Symposium on Circuits and Systems (MWSCAS)*. IEEE, 2015, 1–4.

[12] S. Bhushan, M. Alshehri, I. Keshta, A. K. Chakraverti, J. Rajpurohit, and A. Abugabah, (2022) *"An experimental analysis of various machine learning algorithms for hand gesture recognition"* **Electronics *11***(6): 968.

[13] M. Al-Qurishi, T. Khalid, and R. Souissi, (2021) *"Deep learning for sign language recognition: Current techniques, benchmarks, and open issues"* **IEEE Access *9***: 126917–126951.

[14] S. Katoch, V. Singh, and U. S. Tiwary, (2022) *"Indian Sign Language recognition system using SURF with SVM and CNN"* **Array *14***: 100141.

[15] M. F. Wahid, R. Tafreshi, M. Al-Sowaidi, and R. Langari, (2018) *"Subject-independent hand gesture recognition using normalization and machine learning algorithms"* **Journal of computational science *27***: 69–76.

[16] S. Yin, H. Li, A. A. Laghari, L. Teng, T. Reddy Gadekallu, and A. Almadhor, (2025) *"FLSN-MVO: Edge Computing and Privacy Protection Based on Federated Learning Siamese Network With Multi-Verse Optimization Algorithm for Industry 5.0"* **IEEE Open Journal of the Communications Society *6***: 3443–3458. DOI: 10.1109/OJCOMS.2024.3520562.

[17] H. A. Javaid, M. I. Tiwana, A. Alsanad, J. Iqbal, M. T. Riaz, S. Ahmad, and F. A. Almisned, (2021) *"Classification of hand movements using MYO armband on an embedded platform"* **Electronics *10***(11): 1322.

[18] M. Abduo and M. Galster, (2015) *"Myo gesture control armband for medical applications"* **University of Canterbury**:

[19] A. Bhowmick and S. M. Hazarika, (2017) *"An insight into assistive technology for the visually impaired and blind people: state-of-the-art and future trends"* **Journal on Multimodal User Interfaces *11***: 149–172.

[20] J. Á. Ariza and J. M. Pearce, (2022) *"Low-cost assistive technologies for disabled people using open-source hardware and software: a systematic literature review"* **IEEe Access *10***: 124894–124927.

[21] N. Saleh, M. Farghaly, E. Elshaaer, and A. Mousa. "Smart glove-based gestures recognition system for Arabic sign language". In: *2020 international conference on innovative trends in communication and computer engineering (ITCE)*. IEEE, 2020, 303–307.

[22] S. S. Rautaray and A. Agrawal, (2015) *"Vision based hand gesture recognition for human computer interaction: a survey"* **Artificial intelligence review *43***: 1–54.

[23] H. Shaheen and T. Mehmood. "Talking gloves: Low-cost gesture recognition system for sign language translation". In: *2018 IEEE Region Ten Symposium (Tensymp)*. IEEE, 2018, 219–224.

[24] Y. Luo, Z. Wang, J. Wang, X. Xiao, Q. Li, W. Ding, and H. Y. Fu, (2021) *"Triboelectric bending sensor based smart glove towards intuitive multi-dimensional human-machine interfaces"* **Nano Energy *89***: 106330.

[25] M. F. F. I. Chowdhury, A. S. Adittya, S. Ahnaf, S. T. S. Rafid, S. Islam, and A. S. N. Huda. "ReVive Grip: Charting a Path for Affordable Post-Stroke Hand Rehabilitation". In: *2023 26th International Conference on Computer and Information Technology (ICCIT)*. IEEE, 2023, 1–6.

[26] M. Ketcham and T. Ganokratanaa, (2023) *"Intelligent Gloves for Assisting Individuals with Visual Impairment."* **International Journal of Online & Biomedical Engineering *19***(13):

[27] K. Kadam, R. Ganu, A. Bhosekar, and S. D. Joshi. "American sign language interpreter". In: *2012 IEEE Fourth International Conference on Technology for Education*. IEEE, 2012, 157–159.

[28] X. Meng, X. Wang, S. Yin, and H. Li, (2023) *"Few-shot image classification algorithm based on attention mechanism and weight fusion"* **Journal of Engineering and Applied Science *70***(1): 14. DOI: 10.1186/s44147-023-00186-9.

[29] S. Yin, H. Li, A. A. Laghari, T. R. Gadekallu, G. A. Sampedro, and A. Almadhor, (2024) *"An Anomaly Detection Model Based on Deep Auto-Encoder and Capsule Graph Convolution via Sparrow Search Algorithm in 6G Internet of Everything"* **IEEE Internet of Things Journal *11***(18): 29402–29411. DOI: 10.1109/JIOT.2024.3353337.