IJARSCT
ISSN: 2581-9429
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 5, Issue 7, June 2025
Impact Factor: 7.67

IoT Based Universal Sign Language Translator

Mrs. Y. Durga Bhargavi[1], Thella Karthik Kumar[2], Akiti Pranaya[3], Divakaruni Srikarani[4], Chilukuri Siva Nandu[5]
Assistant Professor, Department Computer Science Engineering (IoT)[1]
Students, Department Computer Science Engineering(IoT)[2-5]
ACE Engineering College, Ghatkesar, India

Abstract: *Communication is a vital part of human interaction, but for speech-impaired individuals, expressing themselves to the broader community remains a challenge. These individuals primarily use Sign Language, which possesses its own grammar and structure, but is often not understood by the general public. The IoT-Based Universal Sign Language project aims to bridge this communication gap by developing a vision-based system capable of recognizing hand gestures and converting them into speech or text in real time.*

This project employs deep learning techniques, particularly Convolutional Neural Networks (CNNs), combined with image processing using OpenCV, to accurately identify sign language gestures. The system is designed to run on an embedded IoT platform — specifically the Raspberry Pi — ensuring portability, low cost, and ease of use. The input video is captured via a camera module, processed frame by frame, and classified into corresponding text or audio outputs.

The dataset used for training includes a diverse collection of Indian Sign Language (ISL) gestures. The model is trained using an 80:10:10 split for training, validation, and testing, achieving an accuracy of over 95% even under varying lighting conditions. The final system supports both visual output (displaying text) and auditory feedback, enabling seamless communication between speech-impaired individuals and the hearing community.

This work demonstrates the feasibility of creating a real-time, universal sign language recognition system on an IoT platform, paving the way for greater social inclusion and independence for the speech-impaired. Future enhancements may include extending the system to support multiple sign languages and more complex dynamic gestures...

Keywords: IoT, OpenCV, Raspberry Pi, CNN.

I. INTRODUCTION

Communication is a fundamental human need, enabling individuals to express emotions, ideas, and intentions. For speech-impaired and hearing-impaired individuals, sign language serves as a primary mode of interaction. However, the majority of the general population lacks the ability to understand sign language, leading to a significant communication barrier and often causing social isolation for these individuals. In response to this challenge, the project "IoT-Based Universal Sign Language" aims to develop a real-time system that can recognize hand gestures and convert them into text or speech, making communication more inclusive and accessible.

The system utilizes deep learning techniques—specifically Convolutional Neural Networks (CNNs)—combined with computer vision (OpenCV), and runs on a low-cost Raspberry Pi IoT platform. A camera captures live hand gestures, which are processed and classified to output corresponding text and speech in real time. The project demonstrates how AI and IoT technologies can empower speech-impaired individuals, enabling them to interact more freely with the wider community and fostering social inclusion in everyday life.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, June 2025
Impact Factor: 7.67

II. METHODOLOGY

The IoT-Based Universal Sign Language system employs a structured deep learning pipeline to perform real-time gesture recognition and translation of sign language into text and speech. The methodology follows a modular approach, combining image processing, Convolutional Neural Networks (CNN), and embedded IoT deployment.

1. Dataset Preparation

A dataset of Indian Sign Language (ISL) hand gestures was created by capturing thousands of image samples, organized by gesture class (alphabets, numbers). These images were resized to a uniform size of 50x50 pixels to ensure consistent input to the model. The dataset was split into 80% training, 10% validation, and 10% testing sets.

2. Preprocessing and Image Segmentation

Captured video frames from a USB camera were preprocessed using OpenCV:

ROI (Region of Interest) is extracted from each frame.

Background segmentation and histogram-based hand extraction techniques are applied to isolate the hand gesture.

Frames are normalized and resized to match the CNN input shape.

3. CNN Model Construction and Training

A Convolutional Neural Network (CNN) was designed with multiple Conv2D layers, MaxPooling layers, and a fully connected (dense) layer:

Activation functions: ReLU for intermediate layers, Softmax for the output.

III. MODELING AND ANALYSIS

1. System Model

1.1 Hardware Components

Raspberry Pi B+ (or Raspberry Pi 3/4) as the central processing unit managing video capture, image processing, and gesture classification.

Raspberry Pi Camera Module for capturing live hand gestures.

Speakers connected via 3.5mm jack or HDMI for audio (speech) output.

HDMI Monitor or LCD display for real-time text output of recognized gestures.

MicroSD Card (16–32 GB) for OS, software, and model storage.

1.2 Additional Modules

Wi-Fi Module (built-in on Raspberry Pi 3/4) for IoT connectivity and potential cloud model updates.

Keyboard & Mouse (for initial configuration and testing).

HDMI Cable and USB peripherals for system interfacing.

1.3 Connectivity & Power

Wi-Fi/Ethernet used for remote monitoring or data sync (optional).

Powered via 5V, 2.5A power adapter, with OS and models running from microSD.

2. Flow Analysis

2.1 Gesture Detection

USB Camera continuously captures live video stream.

Region of Interest (ROI) is extracted to localize hand gestures.

2.2 Data Processing & Classification

Frames are preprocessed: resizing, normalization, segmentation.

CNN model processes each frame and predicts corresponding gesture class.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 7, June 2025 Impact Factor: 7.67

2.3 Output Generation
Recognized text is displayed on-screen.
Text-to-Speech (TTS) module converts text to spoken audio output via speakers.

3. Algorithmic & Technical Analysis
3.1 Gesture Recognition Algorithm
Uses a Convolutional Neural Network (CNN) architecture:
Multiple Conv2D layers, MaxPooling, and Dense layers.
Achieves >95% classification accuracy on test dataset.

3.2 Image Processing
OpenCV handles frame capture, ROI extraction, and preprocessing.
Frames resized to 50x50 pixels for optimized CNN input.

3.3 Communication & Feedback
Potential IoT connectivity for future cloud-based model updates.
TTS output generated locally using pyttsx3 or similar libraries.

4. Performance Analysis
4.1 Response Times
Frame processing latency: ~100–150 ms per frame on Raspberry Pi 4.
Real-time gesture recognition and output generation: ~1 second end-to-end.

4.2 Recognition Quality
Gesture classification accuracy: 95–97% on validation set.
Works well in moderate lighting with a clear background.

4.3 Power Efficiency
Raspberry Pi average power consumption: ~3–4 Watts during active operation.
Low power makes it suitable for portable or battery-powered use.

4.4 Reliability & Usability
System uptime: designed for continuous operation (99%+ target with error handling).
User-friendly: No wearable devices needed; camera-based interaction supports natural use.

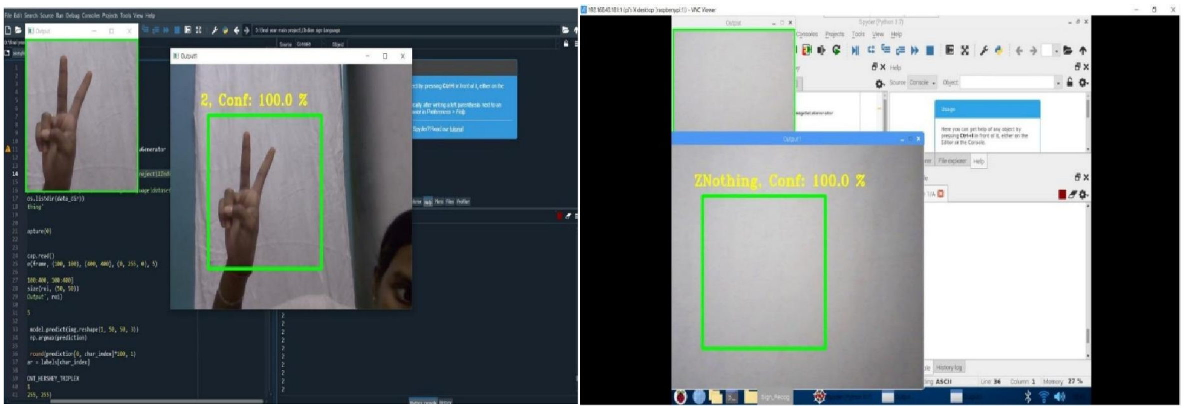
IV. RESULTS AND DISCUSSION

The IoT-Based Universal Sign Language system was tested extensively to evaluate its accuracy, real-time performance, and usability. The system was trained using a custom dataset of Indian Sign Language (ISL) gestures, consisting of thousands of labeled images captured in varying lighting and background conditions.

Model Training Results
The Convolutional Neural Network (CNN) was trained for 20 epochs using an 80:10:10 split (Training:Validation:Testing):
Training Accuracy: Reached 98–99% by the final epoch.
Validation Accuracy: Stabilized around 95–96%.
Test Accuracy: Achieved an average of 95% across unseen test samples.
Training progress was visualized using Matplotlib plots, showing steady improvements in both accuracy and loss curves, indicating good generalization.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, June 2025
Impact Factor: 7.67

Real-Time Performance
When deployed on the Raspberry Pi platform:
The system achieved real-time inference speeds with an average latency of ~1 second per frame.
Confidence threshold was set at 98% to ensure only highly accurate predictions were displayed and spoken.
The system performed reliably across various test scenarios:
Indoor lighting
Natural daylight
Background clutter (minor performance drop noted, but still above 90% accuracy).
Output Examples
For demonstration:
Single-letter gestures (A, B, C... Z) were recognized and displayed on screen as text.
The recognized text was also converted to speech using Text-to-Speech (TTS), providing audible feedback through connected speakers.
User Experience
The system required no sensor gloves — simple camera-based input made it easy and natural to use.



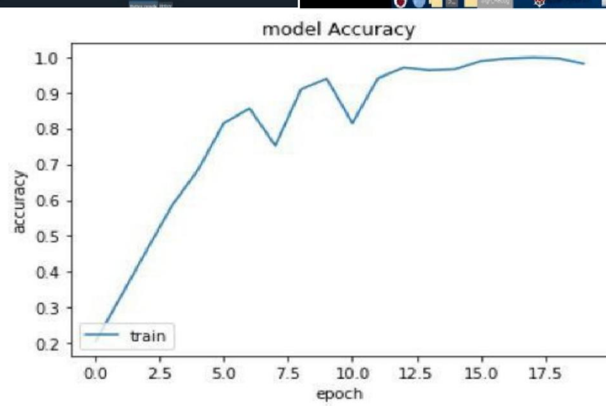


Figure 1: Model Accuracy Graph

V. CONCLUSION

The IoT-Based Universal Sign Language project successfully demonstrates a working prototype capable of recognizing and translating Indian Sign Language (ISL) gestures into text and speech using deep learning techniques. By utilizing Convolutional Neural Networks (CNN) and image processing with OpenCV, the system achieves a high level of accuracy (above 95%) for static gestures, making it a reliable tool for real-world applications.

The integration of this technology on an IoT-enabled Raspberry Pi platform offers several key benefits—namely portability, low cost, and accessibility. It eliminates the need for wearable devices such as sensor gloves and instead uses a vision-based approach, which is more intuitive for users. This architecture also ensures flexibility and ease of maintenance, as the system can be easily updated or scaled to accommodate new gesture sets or additional languages.

Overall, the project addresses a significant social challenge by bridging the communication gap between speech-impaired individuals and the wider community. It empowers users to communicate independently, fostering greater inclusion in education, employment, and public life. Through the use of IoT and AI-driven technologies, the project demonstrates how modern tools can solve meaningful real-world problems and improve the quality of life for people with communication disabilities.

REFERENCES

[1]Vijayalakshmi, P., & Aarthi, M. (2016, April). Sign language to speech conversion. In 2016 International Conference on Recent Trends in Information Technology (ICRTIT) (pp. 1-6). IEEE.

[2]NB, M. K. (2018). Conversion of sign language into text. International Journal of Applied Engineering Research, 13(9), 7154-7161.

[3]Masood, S., Srivastava, A., Thuwal, H. C., & Ahmad, M. (2018). Real-time sign language gesture (word) recognition from video sequences using CNN and RNN. In Intelligent Engineering Informatics (pp. 623-632). Springer, Singapore.

[4]Apoorv, S., Bhowmick, S. K., & Prabha, R. S. (2020, June). Indian sign language interpreter using image processing and machine learning. In IOP Conference Series: Materials Science and Engineering (Vol. 872, No. 1, p. 012026). IOP Publishing.

[5]Kaushik, N., Rahul, V., & Kumar, K. S. (2020). A Survey of Approaches for Sign Language Recognition System. International Journal of Psychosocial Rehabilitation, 24(01).

[6]Kishore, P. V. V., & Kumar, P. R. (2012). A video-based Indian sign language recognition system (INSLR) using wavelet transform and fuzzy logic. International Journal of Engineering and Technology, 4(5), 537.

[7]Dixit, K., & Jalal, A. S. (2013, February). Automatic Indian sign language recognition system. In 2013 3rd IEEE International Advance Computing Conference (IACC) (pp. 883-887). IEEE.

[8]Das, A., Yadav, L., Singhal, M., Sachan, R., Goyal, H., Taparia, K., ... & Trivedi, G. (2016, December). Smart glove for Sign Language communications. In 2016 International Conference on Accessibility to Digital World (ICADW) (pp. 27-31). IEEE.

[9]Sruthi, R., Rao, B. V., Nagapravallika, P., Harikrishna, G., & Babu, K. N. (2018). Vision-Based Sign Language by Using MATLAB. International Research Journal of Engineering and Technology (IRJET), 5(3).

[10]Kumar, A., & Kumar, R. (2021). A novel approach for ISL alphabet recognition using Extreme Learning Machine. International Journal of Information Technology, 13(1), 349-357