



Creation of an IoT-Based Gestures-To-Words Sign Language Translator

Yannam Ravi Sankar, Anumala Alekhy, E. Deena Dayalan and Kattiram Lakshmi Prasanna

ABSTRACT: Sign language translator based on Gesture recognition using IoT and arduino. The proposed project will serve to bridge the gap in communication between impaired individuals with ability and people with hearing or speech problems in the general population. This program is a project that will use Media Pipe, a real-time gesture recognition tool that will be used in conjunction with the machine learning algorithms to read sign language gestures and create readable text and speech that can be decrypted. The system uses an Arduino microcontroller to process the gesture data, incorporates the IoT to access the system remotely and monitor it, and uses a DF Mini Player to provide voice feedback when the gesture is recognized. The system also includes a light and fan control as a demonstration module to demonstrate how to interact with a device by providing a gesture. This new solution focuses on inclusivity, access and usability, which is why communication is easy in any setting, whether at home, at work or on the street.

Keywords: Sign language translation, gesture recognition, IoT, Arduino, MediaPipe, machine learning, real-time communication, assistive technology.

Contents

1 Introduction	1
2 Related Work	2
3 The Proposed Methodology	2
3.1 Gesture Recognition	4
3.2 Translation of text and speech	4
3.3 Appliance Control	4
3.4 IoT Integration	4
3.5 The overall process of the work	5
4 Experimental Results	5
4.1 Prototype Implementation	5
4.2 Gesture Recognition Accuracy	5
4.3 Text and Speech Output After the gesture was identified	6
4.4 Overall Performance	6
5 Conclusion	6
6 Future Work	6

1. Introduction

Several scholars have developed sophisticated mechanisms to translate sign language into voice through the use of IoT and AI technology. A smart IoT-based ISL translator that makes it easy to translate gestures in real-time [1], [2] capable of generating a gesture system to detect gestures from people with hearing impairment, [3] made an IoT glove translator that can translate sign language into speech. [4] focused on Indian Sign Language recognition in real-time computer vision, and [5] implemented an ESP32-based setup to have a working sign-to-speech translation.

Other follow-ups are TinyML and AI-based methodologies, which enhance system precision and accuracy in communication. [6] developed a TinyML-based IoT communication device that can be used by the hearing-impaired, while Nagarale et al. [7] made an Arduino-Python-based voice converter that

2020 *Mathematics Subject Classification:* 68T45, 68T07, 94C99.

Submitted November 30, 2025. Published April 29, 2026.

interprets Indian Sign Language. [8] applied a deep learning algorithm to convert sign motions to text and voice, while [9] introduced a model based on CNN, which could translate American Sign Language into Nepali with the text-to-speech features. All these works highlight the growing importance of IoT, embedded systems, and AI in the development of inclusive communication technology. Human touch involves communication, and millions of people with hearing or speech disabilities have to overcome daily difficulties. The main form of communication among the deaf community is sign language, which is not effective because of insufficient knowledge of the whole population. Research indicates that hearing impairment among children has a serious impact on education and socialization at school [10]. In the study of sign language, vision-based gesture recognition is utilized [11], sensor-based gloves are used [12], [13], and accelerometer-based models are used [14]. Vision-based techniques, such as MediaPipe, can be used to recognize items in real-time without additional equipment, which makes them more convenient [15]. The identification of funding sources by neurons is shown by Andrzej Cichocki [16]. In cases of nonlinear models and noisy data, networks are superior to traditional approaches. General and interrogative phrases in Indian Sign Language are categorized with the help of dual-CNN architecture [18] relying on data provided by a wearable sensor [17] to identify signals. The differences in illumination, hand occlusions, and real-time processing should be processed. The application of IoT technology has enabled smart community features such as remote communication and assistive technology [19]. IoT enables a translation system based on sign language to support real-time remote translation and communication. Smart appliances can be controlled by gesture, which makes them easier for disabled people to use [20].

2. Related Work

Sign language recognition has also become an interesting subject, and a number of methodologies working to translate motions into text or voice have been developed. These approaches are generally divided into vision approaches, sensor approaches and hybrid approaches. Early works in gesture recognition relied on vision-based algorithms, such as machine learning and image processing, to recognize hand shapes and motions. A large-scale assessment of the vision-based gesture recognition has been carried out by Wu and Huang (2001) and has shown the importance of robust feature extraction algorithms. Deep learning models, such as convolutional neural networks (CNNs), and frameworks, such as MediaPipe, have been employed recently to significantly increase the recognition accuracy.

Alternatively, sensor-based methods have gloves with sensors that are used to capture hand and finger movement. Flex sensors and accelerators have been applied to the data gloves and demonstrated a high degree of accuracy as far as distinguishing hand motions is concerned (Lokhande, Prajapati, and Pansare, 2015). Li, Chen, Zhang, Wang, and Wang (2012) relied on the data provided by accelerators and electromyographers (sEMG) to improve the accuracy of the identification in Chinese Sign Language. The integration of the IoT and the intelligent community model has made the use of sign language translation even more applicable, making it more accessible to the impaired community. A recent example of the potential effectiveness of gesture control to control appliances is provided by Parveen and Rohitha (2015), who developed a hand gesture system based on a microcontroller to be used by individuals with disabilities.

Despite these advantages, there are still challenges towards achieving real-time processing, scalability as well as linguistic adaptability. The changes in light, hand placements, and the use of gestures still hinder robust performance. The paper contributes to the existing methods by integrating vision-based gesture recognition and IoT-based smart control systems, where the long-term aim is to enhance real-time communication and access among individuals with speech or hearing impairments.

3. The Proposed Methodology

The suggested system is aimed at developing an intelligent sign language translator using gesture recognition, machine learning, and IoT to interpret the sign language in real-time and control smart appliances. There are four key elements of the technique, namely gesture recognition, text and speech translation, appliance control, and IoT integration. The Intelligent Sign Language Translator, as suggested, will consist of four main parts, including gesture recognition, text and speech translation, appliance control, and IoT integration. The entire system flow is depicted in Figure 1.

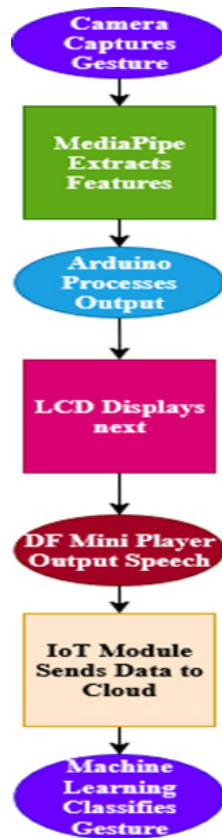


Figure 1: Block Diagram of the Proposed Sign Language Translator System

3.1. Gesture Recognition

Gesture recognition is implemented using MediaPipe, a deep-learning architecture that is used to track hands and classify gestures in real-time [6]. The camera module captures the hand motions, and then the analysis is performed on the motions to detect vital features such as finger positions and palm orientation. The machine learning algorithm is pre-trained to distinguish gestures according to the established representations of sign language. The procedure of gesture recognition includes the following steps:

- Getting live video frames using an external camera module.
- Hand key points extraction with the use of MediaPipe Hand Tracking API.
- Preprocessing the characteristics collected in order to normalize them.
- Gestures classification with the help of a trained machine learning model, e.g. Random Forest or CNN-based classifiers.

3.2. Translation of text and speech

Once a gesture is identified, the system translates it into voice and text to be able to communicate. It is achieved using an Arduino microcontroller and a DF mini player. The process of translation is divided into the following steps:

- The recognized gesture is associated with a corresponding text-based response.
- The recognized gesture is associated with a corresponding text-based response. The microcontroller of the Arduino interprets the text and sends it to an LCD display producing a visual display.
- The DF MiniPlayer is a text-to-voice system that converts typed text into speech with pre-recorded audio snippets.
- This ensures that the hearing-impaired people, as well as people who do not use the sign language, can follow the conversation instantly.

3.3. Appliance Control

The technology also includes gesture-based appliance control and sign language translation, meaning the users can control smart home devices such as lights and fans using predefined gestures [20]. The control method involves the mapping of known motions onto ON/OFF instructions.

The microcontroller called Arduino deciphers commands and activates relay modules. The light or the fan attached responds to a gesture detected. This feature enhances access for the impaired in the sense that they are able to interact with their environment through the automation of gestures.

3.4. IoT Integration

IoT Integration This system will include remote surveillance and management technologies based on the IoT [13]. The system is connected to an IoT platform such as ThingSpeak or Blynk through the ESP8266 NodeMCU module, which enables users to track gesture-based orders remotely. The process of embedding IoT abides by the following process:

- Sensor and gesture data is also transmitted to an IoT platform offered by the cloud. The devices can be controlled and managed through web interface or mobile app by the user.
- Data tracking and analytics, on the other hand, can give an idea of the patterns of use and system effectiveness.
- This enables real-time remote communication, and hence the technology can be used in smart environments.



Figure 2: Prototype of the intelligent sign language translator system

3.5. The overall process of the work

The proposed system is presented in the following way:

- A user undertakes a sign language gesture before the camera.
 - The gesture is captured by the camera and analyzed by MediaPipe.
 - The gesture model uses machine learning to classify the gesture under a preset classification.
 - The result is displayed in a text on an LCD panel and converted into voice.
 - In case the gesture matches an instruction, the Arduino microcontroller triggers the respective appliance.
 - IoT connection helps in the remote oversight and management of gesture-based activities.
- The proposed method will ensure actual-time, accurate sign language translation and smart environment integration, therefore, enhancing accessibility and inclusion.

4. Experimental Results

The proposed Intelligent Sign Language Translator has been tested in terms of the accuracy of gesture recognition, text-to-speech performance, appliance control via IoT, and prototype testing. The findings can show that the system is effective at real-time sign language translation.

4.1. Prototype Implementation

An Arduino microcontroller, camera module, LCD display, DF Mini Player, and IoT module (ESP8266 NodeMCU) were used to create a hardware prototype. The prototype is lastly depicted in Figure 2. The system detects hand gestures through a camera and processes them through MediaPipe and a machine learning model. The output is translated and shown on an LCD screen and turned into speech by the DF Mini Player.

4.2. Gesture Recognition Accuracy

The machine learning model was trained on 50 common sign language gestures and tested with 1000 gesture samples. The recognition accuracy was evaluated under different lighting conditions, as shown in Table 1. It has 50 common sign language gestures that it used to train the machine learning model, and it tested the machine with 1000 gesture samples. These findings show that MediaPipe recognition with machine learning is very accurate and is a little sensitive to extreme light conditions, [15].

Table 1: Gesture Recognition Accuracy Under Different Conditions

Lighting Condition	Accuracy (%)
Normal Indoor Light	94.2
Low Light	89.6
Bright Outdoor Light	85.3

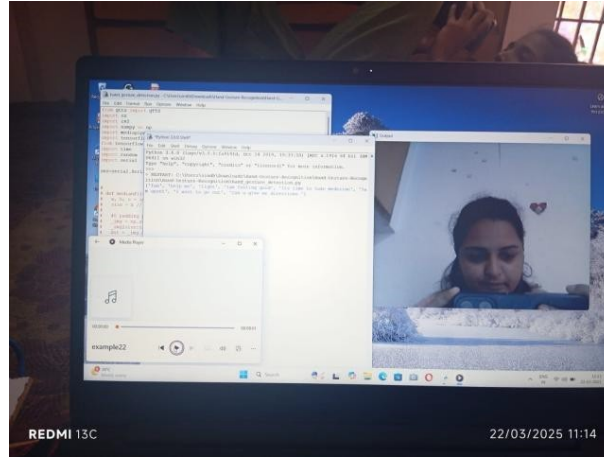


Figure 3: Display Showing Recognized Gesture Output

4.3. Text and Speech Output After the gesture was identified

The system would show the matching text on an LCD display and create a speech output, with the DF Mini Player indicating the sample outputs. After the gesture was identified, the system would show the matching text on an LCD display and create speech output with the DF Mini Player. Figure 3 indicates the sample outputs.

4.4. Overall Performance

The obtained results validate the statement that the suggested system is efficient in converting sign language gestures into both texts and speech and allows controlling smart homes with gestures.

5. Conclusion

The paper introduced an intelligent sign language translator with the application of gesture recognition, machine learning, and IoT. The system was able to convert sign language movements into text and speech, as well as show how gesture-based smart home control could work. Key findings include that the gesture recognition model demonstrated a high accuracy of 94.2 percent. The text-to-speech system offered satisfactory and prompt audio feedback. The IoT module allowed remote monitoring and controlling of devices by gestures. Such findings demonstrate the convenience of the suggested system and its usability among deaf and mute people: it provides them with an opportunity to communicate in real time and also connects to a smart home.

6. Future Work

Future research will focus on:

- Enhancing gesture recognition accuracy with deep learning models (e.g., CNNs, Transformers).
- Extending the support of a variety of sign languages (e.g., ASL, BSL, ISL).
- Combining with AI voice helpers with automated communication support

- Making it portable and battery operated so that it can be moved more. The combination of complex AI and IoT technologies can change the face of assistive communication and make regular communication more inclusive and accessible.

References

1. C. A. Kumar, V. S. Reddy, N. Kandarpa and P. Sharma, *Smart IoT system for Indian sign language translator*, Proc. 2nd Int. Conf. Intelligent Cyber Physical Systems and Internet of Things (ICoICI), Coimbatore, India, 254–258, (2024). doi:10.1109/ICoICI62503.2024.10696760.
2. D. Ganeshkumar, T. Subasri, R. Kaleeshwari, R. Raksha and D. Muthukumar, *An IoT based gesture recognition system for vocally and hearing impaired individuals*, Proc. 9th Int. Conf. Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 685–689, (2023). doi:10.1109/ICACCS57279.2023.10112835.
3. S. Mathupriya, D. Roopa and M. Subashini, *IoT translator for sign language based on glove*, Proc. Intelligent Computing and Control for Engineering and Business Systems (ICCEBS), Chennai, India, 1–7, (2023). doi:10.1109/ICCEBS58601.2023.10448829.
4. H. Muthu Mariappan and V. Gomathi, *Real-time recognition of Indian sign language*, Proc. Int. Conf. Computational Intelligence in Data Science (ICCIDS), Chennai, India, 1–6, (2019). doi:10.1109/ICCIDS.2019.8862125.
5. S. Yalabaka, S. S. Siddam Shetty, C. Rajendra Prasad, P. Sadula, L. P. K. and P. Sothuku, *Sign language to speech conversion using ESP32*, Proc. 4th Asian Conf. Innovation in Technology (ASIANCON), Pimpri Chinchwad, India, 1–4, (2024). doi:10.1109/ASIANCON62057.2024.10838176.
6. R. Sharma, A. Gupta and A. Kumar, *A TinyML solution for an IoT-based communication device for hearing impaired*, Expert Systems with Applications, 246, 123147, (2024). doi:10.1016/j.eswa.2024.123147.
7. D. P. Nagarale, S. B. Sangale, A. J. Rukade, D. R. Wadd and S. S. Halunde, *IoT-based sign to speech converter system*, Int. J. Res. Applied Science and Engineering Technology, (2024). doi:10.22214/ijraset.2024.64413.
8. P. Duraisamy, A. Abinayasrijanani, M. Amrit Candida and P. Dinesh Babu, *Transforming sign language into text and speech through deep learning technologies*, Indian Journal of Science and Technology, 16(45), 4177–4185, (2023). doi:10.17485/IJST/v16i45.2583.
9. B. Paneru, B. Paneru and K. N. Poudyal, *AI-driven translation of American sign language to Nepali using CNN and text-to-speech conversion*, Systems and Soft Computing, 6, 200165, (2024). doi:10.1016/j.sasc.2024.200165.
10. D. Mumba, *Perceptions of teachers and pupils on factors affecting academic performance of pupils with hearing impairment*, Ph.D. Thesis, University of Zambia, (2022).
11. Y. Wu, *Vision-based gesture recognition: A review*, IEEE Signal Processing Magazine, 18(3), 51–60, (2001). doi:10.1109/79.924888.
12. P. Lokhande, R. Prajapati and S. Pansare, *Data gloves for sign language recognition system*, Int. J. Computer Applications, (2015).
13. K. Ellis and J. C. Barca, *Exploring sensor gloves for teaching children sign language*, Advances in Human-Computer Interaction, 2012, 210507, (2012). doi:10.1155/2012/210507.
14. Y. Li, X. Chen, X. Zhang, K. Wang and Z. J. Wang, *A sign-component-based framework for Chinese sign language recognition using accelerometer and sEMG signals*, IEEE Trans. Biomedical Engineering, 59(10), 2695–2704, (2012). doi:10.1109/TBME.2012.2208750.
15. R. Akmeliawati, M. P.-L. Ooi and Y. C. Kuang, *Real-time Malaysian sign language translation using colour segmentation and neural networks*, Proc. IEEE Instrumentation & Measurement Technology Conf. (IMTC), 1–6, (2007). doi:10.1109/IMTC.2007.379019.
16. A. Cochocki and R. Unbehauen, *Neural networks for optimization and signal processing*, John Wiley & Sons, (1993).
17. G. Saggio, P. Cavallo, M. Ricci, V. Errico, J. Zea and M. E. Benalcázar, *Sign language recognition using wearable electronics with k-NN and CNN algorithms*, Sensors, 20(14), 3879, (2020). doi:10.3390/s20143879.
18. K. Suri and R. Gupta, *Convolutional neural network array for sign language recognition using wearable IMUs*, arXiv preprint arXiv:2004.11836, (2020).
19. X. Li, R. Lu, X. Liang, X. Shen, J. Chen and X. Lin, *Smart community: An Internet of Things application*, IEEE Communications Magazine, 49(11), 68–75, (2011). doi:10.1109/MCOM.2011.6069710.
20. A. Parveen and U. Rohitha, *Microcontroller-based hand gesture recognition system using flex sensor for disabled people*, Int. J. Computer Applications, (2015).

Yannam Ravi Sankar,
Department of Electrical and Electronics Engineering,
QIS College of Engineering and Technology, Ongole, India.
E-mail address: ravi.yannam2@gmail.com

and

Anumala Alekhya,
Department of Electrical and Electronics Engineering,
QIS College of Engineering and Technology, Ongole, India.
E-mail address: alekhya.anumala@qiscet.edu.in

and

E. Deena Dayalan,
Department of Electronics and Communication Engineering,
QIS College of Engineering and Technology, Ongole, India.
E-mail address: deenaieee@yahoo.com

and

Kattiram Lakshmi Prasanna,
Department of Computer Science Engineering,
QIS College of Engineering and Technology, Ongole, India.
E-mail address: katharamprasannareddy@gmail.com