International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** XI **Month of publication:** November 2025

DOI: https://doi.org/10.22214/ijraset.2025.75103

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

Design and Implementation of an AI Based Real Time Sign Language Recognition

Dr. B N Madhukar[1], Nishanth P[2], Naveen Kumar V[3], Revansidda[4], Sakshi Yeli[5]
Department of Electronics and Communication Engineering AMC Engineering College Bengaluru, India

Abstract: Communication is an essential requirement for human interaction, but it is very difficult for people with speech and hearing disabilities to communicate their thoughts to the non-sign language speakers. To solve this problem, the project AI-Based Real-Time Sign Language Recognition is conceived with the objective of filling this communication gap by developing a system that can translate sign language movements into speech and text, thus facilitating effective and seamless interaction between the deaf, mute, and the general population. The system is deployed on Raspberry Pi 4, with the advantage of the efficiency of TensorFlow Lite for AI-based gesture recognition and OpenCV for real-time image processing. Hand gestures are captured using a camera, and they are processed by a trained AI model that detects and classifies individual signs. Upon detection, they are translated into text, which in turn is transformed into natural spoken output through Google Text-to-Speech (TTS) technology. Apart from speech translation, the detected text is also rendered on an external display device, making the translated message readily available in a broad range of settings, ranging from noisy environments where audio alone might prove inadequate. By combining affordability, portability, and live performance capabilities, this project offers a practical and effective solution that can significantly increase accessibility for the deaf and mute community, ultimately fostering inclusiveness in education, the workplace, and day-to-day social interactions.
Keywords: Sign Language Recognition, Raspberry Pi, TensorFlow Lite, OpenCV, Real-Time Translation, Gesture Recognition, Text-to-Speech (TTS), Accessibility, Human-Computer Interaction, Assistive Technology

I. INTRODUCTION

In today's age of quick technological development,communication is essential for social, educational, and career inclusion. People with hearing and speech disabilities usually face a major hindrance in expressing their thoughts and feelings to the general public because of limited knowledge about sign language among the masses. Sign language is the original and most natural means of communication for this population, but because it has not been adopted universally, it becomes limited in accessibility and creates communication barriers. In an attempt to overcome this problem, artificial intelligence (AI) has proven to be a viable alternative with the potential to facilitate real-time sign language translation into text and speech, thus filling the communication gap between the hearing and speech-impaired populace and the rest of society. Recent breakthroughs in machine learning and computer vision have made it possible to develop systems that can understand gestures with high accuracy. Deep learning technologies, specifically Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have shown great performance in detecting intricate spatio-temporal features like hand shapes, movements, facial expressions, and body postures. Through the use of these models, AI-powered sign language recognition systems can deliver very accurate and consistent translations in real-time. In addition, due to increasing availability of edge computing hardware and optimized frameworks like TensorFlow Lite, deployment of these systems has become increasingly practical and cost-effective, rendering them viable for real-world applications outside of the laboratory. The possible uses of AI-driven sign language recognition spread across diverse fields like education, healthcare, corporate communication, and customer support, where accessibility is the key. By allowing unhampered human–computer and human–human interaction, these systems not only improve accessibility for the deaf and mute population but also social integration and equality of participation. As such, AI-based sign language recognition is increasingly being seen as a game-changing tool, which can change the face of accessibility and push the overall agenda of creating an inclusive society.

II. RELATED WORK

Various researchers have worked on the design of real-time sign language recognition systems based on artificial intelligence, deep learning, and computer vision. Kumar et al. [1] designed an ISL detection system based on pre-trained VGG16 CNN with attention mechanism with 99.8% accuracy on a small database of 702 images, although it was restricted to pre-defined gestures and light-sensitive.

Divya Lakshmi et al. [2] had applied a MobileNet-V2 model on Raspberry Pi 3 for the recognition of English alphabets with 99.52% accuracy, albeit in a system restricted to static signs.

ElTokhy and Abas [3] had developed a user-initiated recognition system on Raspberry Pi 4 through segmentation, feature extraction, and deep learning classification with 92% accuracy and portability but restricted by pre-defined gestures and hardware limitations. Kumar et al. [4] used Mediapipe for landmark extraction and CNN-based classification for ASL alphabets (A–Z) with 99.95% accuracy using a big dataset of 166,000 images, although only static gestures and ASL were used. Srivastava et al. [5] used TensorFlow Object Detection API with SSD MobileNet-V2 to recognize ISL, which was achieved at 85.45% confidence on 650 images, although limited to alphabets and affected by small dataset. Appalanaidu et al. [6] developed a CNN and YOLO-based system on Raspberry Pi 4 with the integration of TTS, providing portability and real-time capability but suffering from visually similar gestures and environmental changes. Although these works provide a substantial advance in gesture recognition, they are mostly constrained by pre-defined datasets, static gesture management, and sensitivity to environments, indicating the necessity of an enhanced, scalable, and real-time Indian Sign Language recognition system.

III. PROBLEM STATEMENT

Speech and hearing-impaired persons use sign language as their natural form of communication. The fact that the general population fails to adopt this form of communication as a common mode exposes them to severe communication challenges. The use of human interpreters or written messages is not always possible, especially where meetings are held in real-time and there is no time for translation or interpretation. Over the last decade, there have been efforts to create automated recognition systems, with most current techniques based on wearable technology such as gloves, which are cumbersome, or being confined to a controlled setting with limited datasets. Additionally, many systems are only applicable to static gestures or particular alphabets, thereby restricting their use to natural, continuous sign language. These limitations prevent strong, real-time, and economical solutions from being deployed in real-world settings like classes, offices, hospitals, and public services.To overcome these challenges, one needs a portable, cost-effective, and efficient system that can accurately recognize sign language gestures in real time without the need for specialized hardware. This gives rise to creating an AI-based real-time sign language recognition system based on deep learning and computer vision on an embedded platform such as Raspberry Pi. By integrating TensorFlow Lite for light inference and OpenCV/Mediapipe for real-time image processing, such a system has the ability to bridge the communication gap and enhance inclusivity for the hearing and speech-impaired populace.

IV. SYSTEM DESIGN OVERVIEW

The suggested real-time sign language recognition system using AI combines hardware and software aspects to achieve affordability, portability, and real-time execution. The overall configuration is segregated into hardware architecture, software architecture, and functional workflow, represented through a block diagram (Fig. 1) and A flow chart (Fig. 2).

A. Hardware Architecture

The hardware platform focuses on the Raspberry Pi 4 (4 GB RAM) as the central processing unit for executing the trained deep learning model and processing real-time video. A USB webcam is utilized to capture a live video of hand gestures as the main input device. The output of the recognized text is shown on an external display, with a speaker giving auditory feedback through speech synthesis of recognized text. Other peripherals including a mouse, keyboard, HDMI cable, and 32 GB microSD card enable system functionality, storage, and user interaction. The system operates using a 5V, 3A power supply, providing stable and continuous performance. This small configuration guarantees that the system is cost-effective and lightweight, allowing it to be comfortably deployed in educational, workplace, and public service settings.

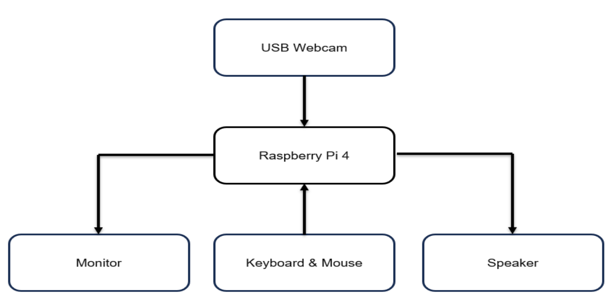


Fig. 1. Block diagram of the proposed AI-based real-time sign language recognition system.

B. Software Architecture

The Raspberry Pi operates on Raspberry Pi OS, which is the execution environment for the project. The project is implemented using Python 3.x, with various libraries incorporated to undertake certain tasks. OpenCV is utilized for real-time video capture and preprocessing of the frames. Mediapipe is utilized for hand tracking and the extraction of 21 landmarks from the identified hand, which are normalized to remove variance in hand size and position. These characteristics are then input to the trained deep neural network, constructed utilizing TensorFlow and optimized to TensorFlow Lite for efficient running on the Raspberry Pi. Lastly, the output gesture label is displayed on the screen, and pyttsx3, a local text-to-speech library, translates the identified gesture into spoken words. This two-mode output increases accessibility in any environment.

C. Functional Workflow

The operation flow starts with video recording from the webcam, followed by preprocessing with OpenCV and Mediapipe for recognizing and obtaining hand landmarks. Preprocessed data is then fed into the TensorFlow Lite model, which determines the gesture and produces the output accordingly. The identified output is shown as text and also translated into voice, thereby making it accessible in both silent and noisy rooms. The entire sequential procedure is illustrated in the flow chart presented in Fig. 2.

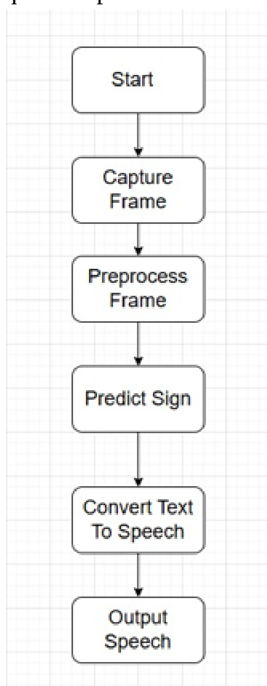


Fig. 2. Flow chart for sign language recognition.

This system design combined makes hardware and software parts work in harmony to provide real-time gesture recognition, filling gaps in communications and fostering inclusivity in the hearing and speech-impaired society.

V. IMPLEMENTATION

The proposed AI-based real-time sign language recognition system is implemented in four stages. In the data acquisition stage, gesture images are captured through a USB webcam attached to the Raspberry Pi. The camera is started using a Python script that also puts the dataset in order by making individual directories for each gesture and thus storing the captured images systematically. At the creation of the dataset phase, the original image data is preprocessed through the use of Mediapipe Hands library, and 21 landmarks per hand are extracted. They are normalized by scale and position, and the data processed is saved in a structured format (data.pickle file) for training. At the model training phase, a deep learning model is created through the utilization of TensorFlow. The model uses a hybrid strategy involving CNN and LSTM layers to identify spatial and temporal gesture characteristics. After being trained, the model is saved in TensorFlow Lite format to ensure efficient execution on the Raspberry Pi. In the prediction phase, the webcam records a live video stream, and every frame is fed through Mediapipe to detect landmarks, which are then sent through the trained TensorFlow Lite model for classification.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

The identified gesture is output as text to the screen and also translated into speech in real-time by the pyttsx3 text-to-speech engine, which closes the recognition loop.

A. *Experimental Parameters*

Test and validation system configuration used for experimentation is shown in Table 1.

Theseparameters specify the experimental setup and hardware requirements for performance testing.

Table 1: Experimental parameters used for gesture recognition in the proposed system.

Parameter Name	Description / Value
Gesture Names	Hi,Yes, Good Job, No, Victory
Model Used	MediaPipe+ TensorFlow Lite
Number of Keypoints Detected	21 hand landmarks
Camera Resolution	640 × 480 px
Frame Rate (FPS)	~30 frames per second
Confidence Score Range	92–97%
Environment Condition	Indoor lighting, plain background
Processing Device	Raspberry Pi 4 (Python 3 environment)

B. *Performance Testing*

The system's performance was analyzed considering the gesture recognition accuracy based on the following formula:

$$Accuracy\ (\%) = \left(\frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions}\right) \times 100$$

For multi-class gesture classification, accuracy can also be represented using confusion matrix parameters as:

$$Accuracy(\%) = \left\{\frac{TP + TN}{TP + TN + FP + FN}\right\} \times 100$$

where TP and TN stand for true positives and true negatives, while FP and FN refer to false positives and false negatives, respectively. Based on experimental results, the proposed system achieved a recognition accuracy of 96.44%. This confirms its reliability and efficiency in real time.

VI. RESULTS AND DISCUSSION

The performance of the proposed system was measured in terms of recognition accuracy, real-time responsiveness, and usability. The trained model generated high accuracy on the test dataset, exhibiting efficient recognition of pre-defined Indian Sign Language gestures. In real-time testing, the system could successfully capture gestures using the webcam and process them with minimal latency, providing a smooth experience. Screenshots of the identified gestures on the screen confirm the system's real-world applicability. Additionally, the system presented here is cost-effective and transportable compared to other works presented in the literature review while achieving competitive accuracy. Nevertheless, performance was found to be light-sensitive and sensitive to hand occlusion, which remain open issues for real-world usage.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

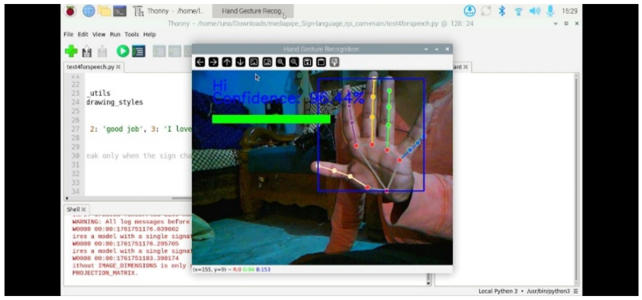


Fig. 3 Hand Gesture Recognition Output.

The system effectively identifies and classifies the "Hi" gesture in real-time with MediaPipe keypoint tracking. The model has a confidence rate of 96.44%, visibly showing hand landmarks and bounding box for gesture confirmation.

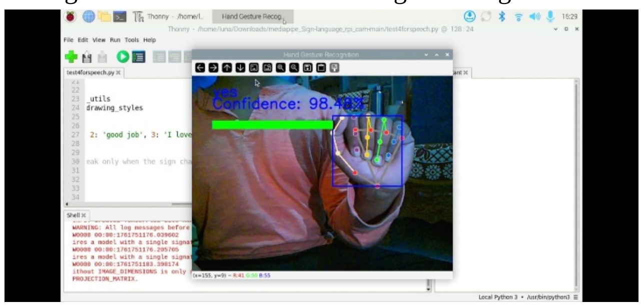


Fig. 4. Hand Gesture Recognition Output

The system effectively identifies and classifies the "yes" gesture in real-time with MediaPipe keypoint tracking. The model has a confidence rate of 98.48%, visibly showing hand landmarks and bounding box for gesture confirmation.

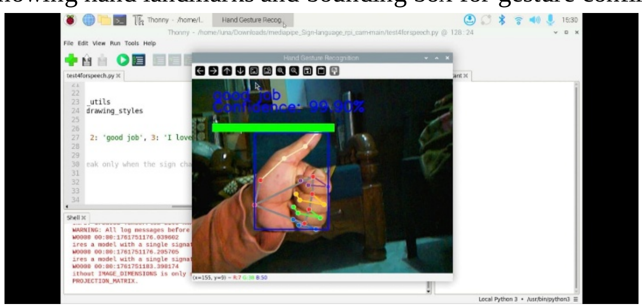


Fig. 5 Hand Gesture Recognition Output.

The system effectively identifies and classifies the "good job" gesture in real-time with MediaPipe key point tracking. The model has a confidence rate of 99.90%, visibly showing hand landmarks and bounding box for gesture confirmation.

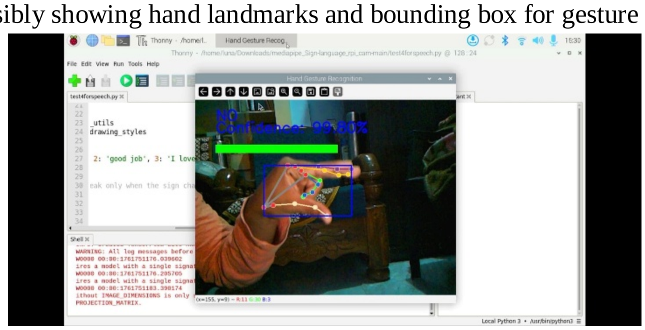


Fig.6. Hand Gesture Recognition Output.

The system effectively identifies and classifies the "No" gesture in real-time with MediaPipe keypoint tracking. The model has a confidence rate of 99.80%, visibly showing hand landmarks and bounding box for gesture confirmation.

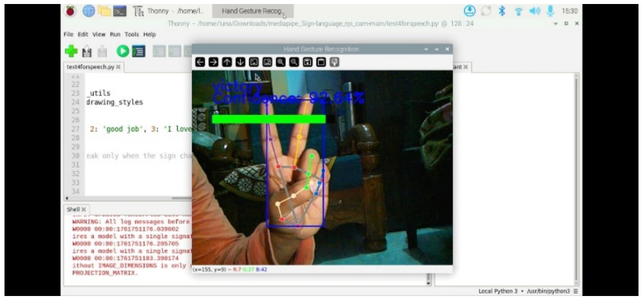


Fig. 7. Hand Gesture Recognition Output

The system effectively identifies and classifies the "Victory" gesture in real-time with MediaPipe keypoint tracking. The model has a confidence rate of 92.64%, visibly showing hand landmarks and bounding box for gesture confirmation.

VII.BENEFITS AND DRAWBACKS

The intended system has various benefits. It is cost-effective and portable since it uses Raspberry Pi and a webcam rather than costly GPUs or wearable devices. It provides real-time functionality, permitting smooth communication. The system is designed specifically for Indian Sign Language (ISL), directly addressing accessibility for the domestic hearing and speech-impaired population. In addition, it does not use gloves or special sensors, so it is comfortable and convenient to use. Even with these advantages, the system has its drawbacks. Its performance reduces when illuminated lightly or gestures are partially obstructed. Training data is of limited size, which keeps the variety of gestures recognized in check. Moreover, the system as of now supports only predetermined static gestures and hasn't yet accommodated continuous or dynamic sign sequences, restricting its scalability for sophisticated communication.

VIII. CONCLUSION

The current contribution is a low-cost and portable AI-driven sign language recognition system in real-time implemented on Raspberry Pi. Computer vision and deep learning come together to make the system efficiently recognize pre-defined ISL gestures and map them to text and speech output. The system with dual-mode feedback guarantees accessibility across diverse environments, even noisy ones. Experimental results validate that the system works well in real time, thus closing the communication gap between sign language users and non-signers. In general, the project showcases a practical solution for inclusivity and accessibility for the hearing and speech-impaired society.

IX. FUTURE SCOPE

The system can be extended in a number of ways. Increasing the dataset to cover a more extensive vocabulary of gestures will increase scalability and precision. Dynamic and ongoing sign recognition will be supported in future iterations, allowing the recognition of entire sentences instead of single words. Multi-language support will enable the system to be portable across regions and sign language variations. Further, integration with wearable or mobile devices can make it applicable outside of stationary environments. As edge AI optimization improves, real-time recognition accuracy can be maximized with less latency and power usage.

REFERENCES

[1] S. Srivastava, A. Gangwar, R. Mishra, and S. Singh, “Sign Language Recognition System using TensorFlow Object Detection API,” International Conference on Advanced Network Technologies and Intelligent Computing (ANTIC-2021), Springer, 2021.

[2] R. Kumar, S.K. Singh, A. Bajpai, and A. Sinha, “Mediapipe and CNNs for Real-Time ASL Gesture Recognition,” ArXiv Preprint arXiv:2305.05296v3, 2023.

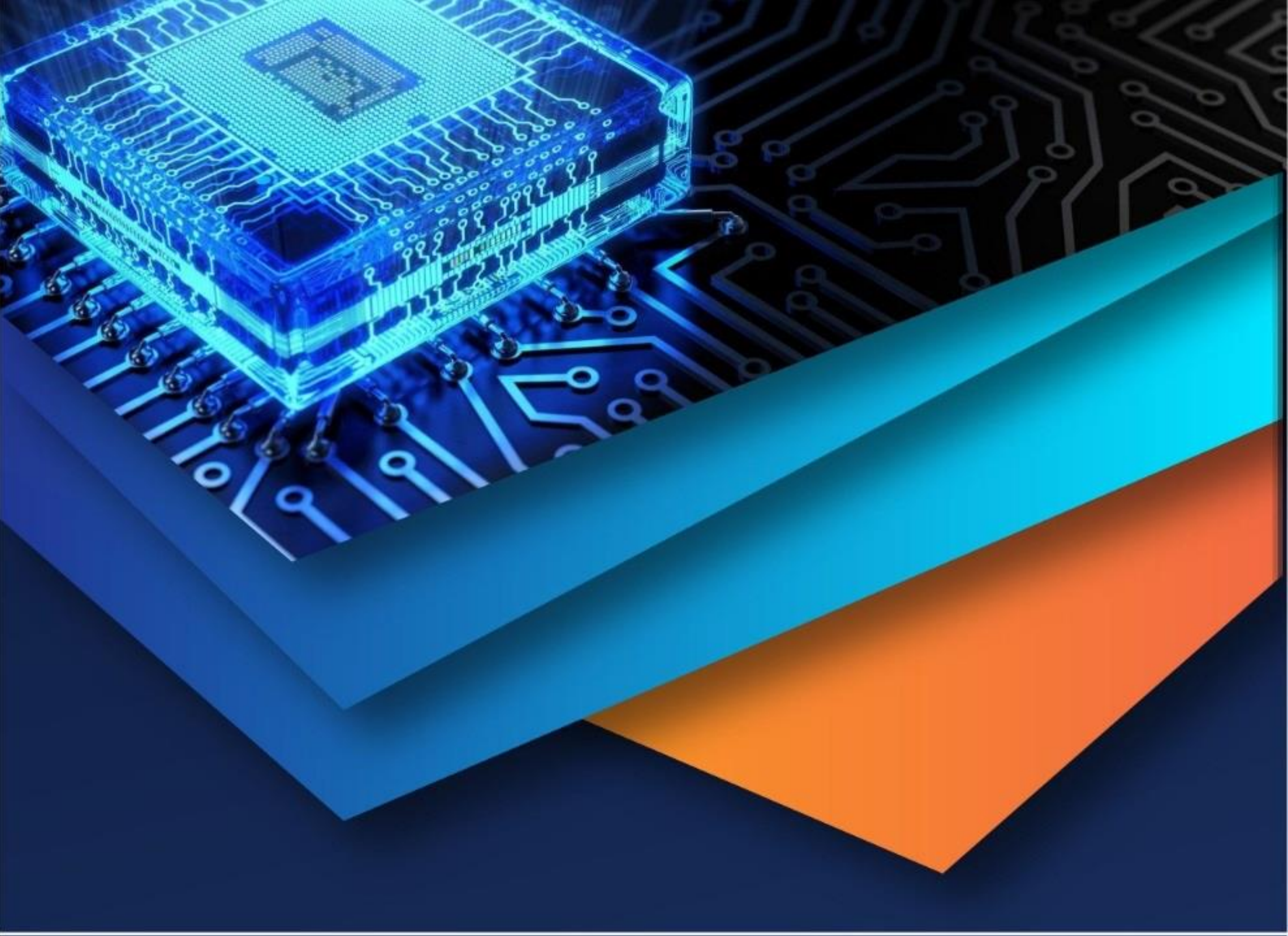
[3] M. A. R. ElTokhy and S. S. Abas, “User-Driven Sign Language Recognition Using AI: AI Tool for Enhanced Communication in Closed Communities,” Nanotechnology Perceptions, vol. 20, no. S14, pp. 3644–3652, 2024.

[4] S. Kumar, R. Rani, and U. Chaudhari, “Real-time sign language detection: Empowering the disabled community,” MethodsX, vol. 13, 2024.

[5] A. M. A., D. L. P., S. K. S., and S. V. N., “A Comprehensive Deep Learning-Based System for Real-Time Sign Language Recognition and Translation Using Raspberry Pi,” International Journal of Computer Trends and Technology (IJCTT), vol. 72, no. 12, pp. 8–16, Dec. 2024.

[6] R. A. Ch., N. S. Ramya, K. Sumanjali, K. V. Lakshmi, and K. Gayatri, “Sign Language Recognition and Speech Conversion Using Raspberry Pi,” International Journal of Creative Research Thoughts (IJCRT), vol. 8, no. 5, pp. 2103-2104, May 2020

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)