International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71693

Sign Language Recognition

Prof. R.S. Bhalerao[1], Deore Viraj Chandrakant[2], Kolhe Pratik Rajendra[3], Patil Rugved Pradip[4], Lahane Nirmala Prakash[5]
Department of Information Technology, Sir Visvesvaraya Institute Of Technology, Nashik
A/p. Chincholi, Tal.Sinnar, Dist. Nashik - 422102

Abstract: This paper presents a novel real-time sign language detection system designed to enhance communication between the deaf and hard-of-hearing community and non-signers. Utilizing standard web cameras, the system captures and analyses hand and facial gestures, employing advanced computer vision and deep learning techniques to recognize sign language gestures. Key markers corresponding to specific signs are identified and translated into voice output and on-screen text, providing a dual-output feature that fosters inclusivity and accessibility. By enabling real-time interpretation through voice and visual representation, this technology bridges communication gaps, making interactions more seamless for sign language users and those unfamiliar with it.

The proposed system is adaptable for integration into webcams and other camera-equipped devices, offering potential applications across various sectors, including education and healthcare, ultimately improving understanding and interaction for sign language users.

Keywords: Sign language, webcam, real-time detection, computer vision, deep learning.

I. INTRODUCTION

Sign language serves as a vital mode of communication for the deaf and hard-of-hearing community, enabling individuals to express themselves and interact with others.[1] However, a significant barrier exists between sign language users and those who do not understand these gestures, often leading to miscommunication and isolation. As the need for effective communication continues to grow, there is an increasing demand for technologies that facilitate seamless interaction across diverse linguistic backgrounds.[2] This paper introduces a novel real-time sign language detection system that leverages standard web cameras to bridge the communication gap between sign language users and non-signers. By employing advanced computer vision and deep learning techniques, the system captures and analyzes hand and facial gestures, identifying key markers that correspond to specific signs in sign language vocabulary. The innovative design of the system provides dual outputs: voice interpretation for real- time audio feedback and on-screen text for visual reference.[3]

The dual-output feature enhances accessibility and inclusivity, allowing users to engage in meaningful conversations regardless of their familiarity with sign language. Furthermore, this technology can be integrated into existing webcam and camera- equipped devices, making it readily accessible for various applications in education, healthcare, and other sectors. By improving understanding and interaction for sign language users, this system represents a significant advancement in communication technology, fostering greater inclusivity and bridging societal divides.[4]

II. LITERATURE SURVEY

1) Dr. Ayesha R. Mehta, Hand Gesture Recognition Using Convolutional Neural Networks (CNNs), This project involves developing a hand gesture recognition system using Convolutional Neural Networks (CNNs). It aims to accurately classify and interpret hand gestures in real- time, offering robust performance across various conditions. The system is designed for applications in human-computer interaction, accessibility, and more.
2) Prof. Karan S. Deshmukh, Real-Time Sign Language Detection Using Kinect Sensors, This project focuses on developing a real-time sign language detection system using Kinect sensors, leveraging depth sensing and skeleton tracking to accurately interpret sign language gestures. It aims to enhance communication accessibility for individuals with hearing and speech impairments.
3) Dr. Meenakshi N. Rao, Deep Learning-Based Sign Language Recognition Using LSTM Networks, This project involves developing a sign language recognition system using Long Short-Term Memory (LSTM) networks to analyze and interpret sequential hand gestures. It aims to accurately translate sign language into text or speech, enhancing communication accessibility

4) Arjun T. Bhatia, Sign Language Recognition Using YOLO and OpenCV, This project uses YOLO and OpenCV to develop a real-time sign language recognition system, combining YOLO's efficient object detection with Open CV's image processing capabilities. The goal is to accurately detect and interpret sign language gestures swiftly and reliably.
5) Ahmed Elhagry, Rawan Glalal Elrayes, This study presents a video-based Egyptian Sign Language recognition system employing two neural network architectures: a Convolutional Neural Network (CNN) for spatial feature extraction and a combined CNN-LSTM model for capturing both spatial and temporal features. The CNN model achieved an accuracy of 90%, while the CNN-LSTM model reached 72%, demonstrating the potential of deep learning in sign language recognition
6) Razieh Rastgoo, Kourosh Kiani, Sergio Escalera, This paper introduces a zero-shot learning approach for sign language recognition using RGB-D videos. The proposed two-stream model utilizes vision transformers for human detection and feature representation, combined with LSTM networks and BERT for semantic mapping. The model was evaluated on multiple datasets, achieving state-of-the-art results in zero-shot sign language recognition
7) Mansooreh Montazerin, Soheil Zabihi, Elahe Rahimian, This research explores the use of Vision Transformers (ViT) for hand gesture recognition based on high-density surface electromyography (HD-sEMG) signals. The proposed ViT-HGR framework achieved an average test accuracy of 84.62% on a dataset comprising 65 isometric hand gestures, highlighting its potential for applications like prosthetic control.
8) Serkan Savaş, Atilla Ergüzen, This study proposes a two-stage approach for hand gesture recognition, leveraging transfer learning and deep ensemble learning. By evaluating pre-trained models like VGGNet and MobileNet on the HG14 dataset, the ensemble model achieved an impressive accuracy of 98.88%, demonstrating the effectiveness of combining multiple models for improved performance.
9) Muhammad Islam, Mohammed Aloraini, Suliman Aladhadh,, This paper presents a stacked encoded deep learning framework for sign language recognition, focusing on developing a vision-based intelligent system. The framework aims to enhance communication accessibility by accurately recognizing sign language gestures through advanced deep learning techniques
10) Hezhen Hu, Weichao Zhao, Wengang Zhou, Houqiang LiThis paper introduces SignBERT+, a self-supervised pre-training framework that incorporates hand pose information as visual tokens. By embedding gesture states and spatial-temporal positions, the model effectively captures hierarchical context in sign language sequences. The approach achieves state-of-the-art performance in tasks like isolated and continuous sign language recognition, as well as sign language translation

III. PROBLEM DEFINITION

Sign language is an essential means of communication for the deaf and hard-of-hearing community; however, it remains largely inaccessible to those who do not understand it. This communication gap hinders effective interaction between signers and non-signers in various settings, such as education, healthcare, and social environments. The lack of real-time translation solutions further exacerbates this issue, limiting opportunities for meaningful engagement and inclusivity.

Current technologies for sign language recognition often require specialized hardware or extensive training, making them impractical for widespread use. Additionally, existing systems may not provide immediate feedback, which is crucial for facilitating real-time conversations.

This project aims to develop a novel real-time sign language detection system that leverages standard web cameras to recognize and translate sign language gestures into both voice output and on-screen text. By utilizing advanced computer vision and deep learning techniques, the system seeks to provide an effective, accessible, and user-friendly solution to bridge the communication gap between sign language users and non-signers. The goal is to enhance inclusivity and understanding within diverse communities, promoting a more connected society

IV. METHODOLOGY

Conduct a comprehensive review of existing research on sign language recognition, computer vision, and deep learning to identify gaps and inform the system design.

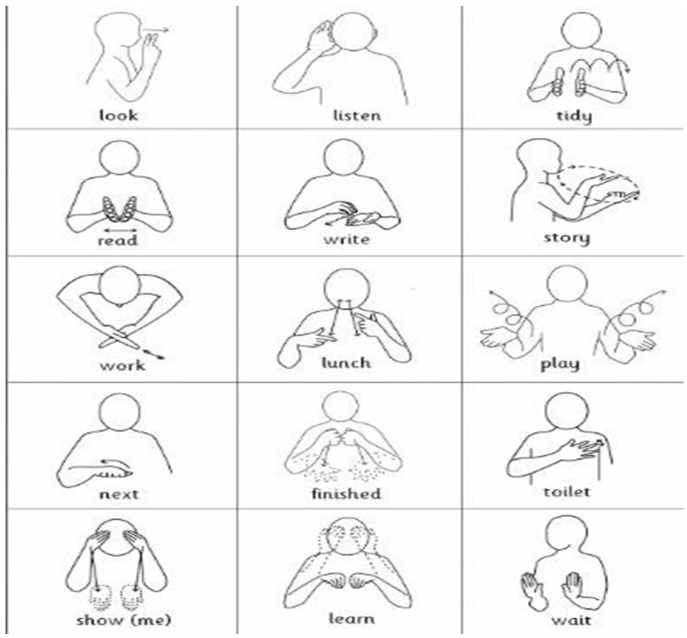
Gather a diverse dataset of sign language gestures, including various signers, lighting conditions, and backgrounds, to ensure robustness and accuracy in detection.

Implement preprocessing techniques to normalize the data, including resizing images, augmenting the dataset, and applying techniques to enhance image quality and visibility.

Utilize deep learning frameworks (e.g., TensorFlow, PyTorch) to develop a gesture recognition model, employing Convolutional Neural Networks (CNNs) to classify and identify sign language gestures from captured images.

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 V May 2025- Available at www.ijraset.com

Integrate facial recognition technology to analyze accompanying facial expressions, which can provide context and enhance the accuracy of gesture recognition.
Identify key markers and features corresponding to specific signs through advanced computer vision techniques, such as optical flow or keypoint detection.
Train the recognition model using the collected dataset, employing techniques like transfer learning to improve performance and reduce training time.



Flow Chart

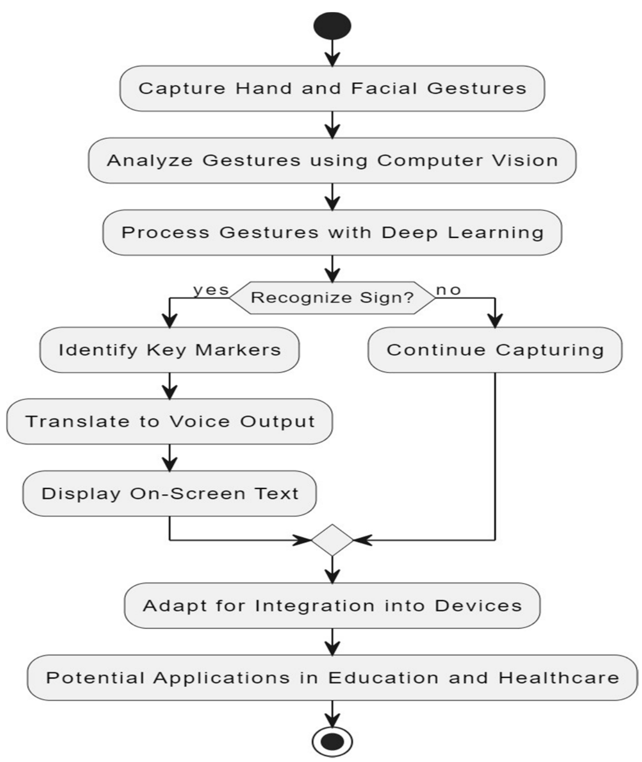


Fig. Flow Chart

V. FUNCTIONAL REQUIREMENTS

- The system must accurately recognize and interpret various sign language gestures using hand and facial movements.
- Upon recognizing a gesture, the system should generate both voice output (speech synthesis) and on-screen text.
- The system must process input from the webcam in real time to facilitate immediate communication without noticeable delays.
- Users should be able to create profiles and customize settings for personalized experiences, such as selecting preferred output modes (voice/text).
- The system must support multiple sign languages, allowing users to select their preferred language for recognition and output.
- Users should have the option to train the system for personalized gestures to improve recognition accuracy.
- The system must include mechanisms to handle the misrecognition of gestures, and provide feedback to users for corrections.
- The system should log interactions and provide analytics to help users track communication patterns and improvements.

VI. NON FUNCTIONAL REQUIREMENTS

- The system must maintain low latency, processing gestures and providing outputs in under a specified time.
- The system should be scalable to handle multiple users simultaneously, especially in educational or public settings.
- The user interface should be intuitive and easy to navigate, ensuring accessibility for users of varying tech proficiency.
- The system must be compatible with various operating systems (Windows, macOS, Linux) and web browsers.
- User data and privacy must be protected through secure authentication methods and data encryption.
- The system should be designed for easy updates and maintenance to incorporate new features or improve existing functionalities.

VII. PROPOSED SYSTEM

This system is built to help people who use sign language communicate more easily with those who don't understand it. It works in real time, which means it can instantly recognize and translate sign language as it's being used. You don't need any fancy equipment either — just a regular webcam is enough.

The system watches the movements of a person's hands and face using the webcam. It then uses smart computer programs, like artificial intelligence and deep learning, to figure out what those gestures mean. Once it understands the sign, it shows the translation as text on the screen and says it out loud through a speaker. This makes it easier for everyone — including those who can't hear or speak — to take part in a conversation. The cool part is that the system gives two outputs: one you can see (text) and one you can hear (voice). That makes it much more accessible and useful in places like classrooms, hospitals, or public services where communication needs to happen quickly and clearly. Because it's designed to work with common devices like laptops or tablets with webcams, it can be used pretty much anywhere. Overall, this technology helps close the communication gap and makes life more inclusive for people who rely on sign language.

VIII. MODULES

1) *Gesture Detection Module*

This part of the system watches your hands and face through a webcam. It tracks movements to figure out when someone is making a sign. Think of it like a smart camera that knows you're signing and pays close attention**.**

2) *Gesture Recognition Module*

Once the gestures are detected, this module tries to understand what each sign means. It uses deep learning (a type of artificial intelligence) to match the gesture with its correct meaning, just like how a person learns what different hand signs mean in sign language**.**

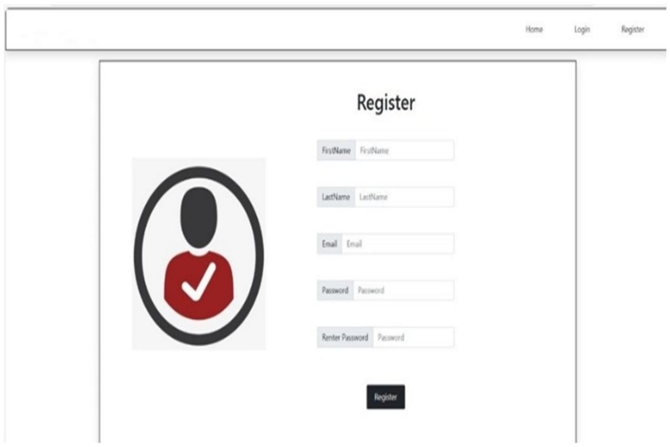
3) *Text and Voice Output Module*

After understanding the sign, this module shows the translated word or sentence on the screen and also speaks it out loud. This helps the person using sign language to "talk" to someone who doesn't know sign language — both visually and through sound.

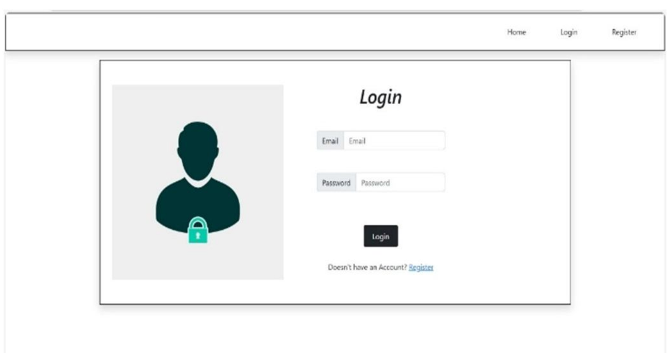
IX. SOFTWARE VERSION

To implement the real-time sign language detection system, a standard webcam is used to capture live video of hand gestures and facial expressions. The system is built using Python 3.10.6, with OpenCV 4.8.0 for image processing, Media Pipe 0.10.5 for hand and face tracking, and TensorFlow 2.13.0 for gesture recognition using a CNN model. Preprocessing steps like resizing, normalization, and augmentation enhance detection accuracy. Once a gesture is recognized, the output is shown as text and spoken aloud using the pyttsx3 (v2.90) text-to-speech library. The system is lightweight, runs on regular laptops or tablets, and requires no special hardware, making it ideal for use in schools, hospitals, and public spaces.

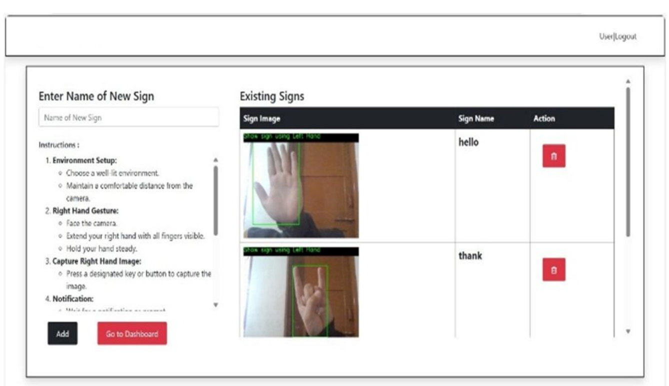
X. RESULTS



Fig(a): Register page

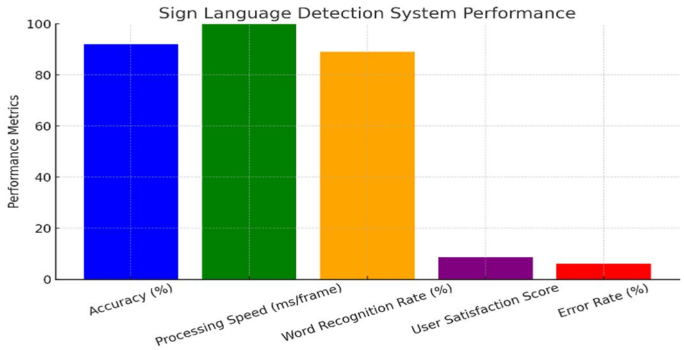


Fig(b): Login page

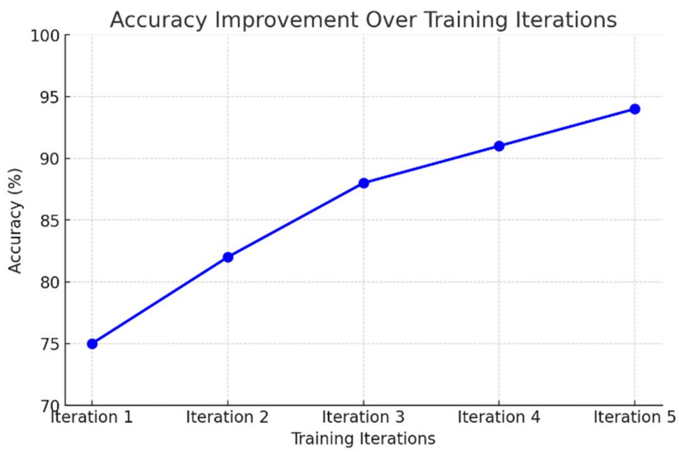


Fig(c): Dashboard page

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 V May 2025- Available at www.ijraset.com



Fig(d): Comparison Chart



Fig(e): line chart

The registration process is where new users first sign up to use the system. It's like creating an account for the first time. Users need to provide some basic information like their name, email, and a password to keep their account secure. After entering these details, they'll create an account that they can log into anytime.

The login process is for users who already have an account. This step allows users to access the system by verifying their identity using the username (email) and password they set during registration. After logging in, the user can start using the features of the system, like translating sign language gestures.

Real-Time Translation: In the existing systems, translations might take time or could be inaccurate, especially in real-time. But the proposed system improves this by providing instant and more accurate translations as soon as the sign is detected.

Device Compatibility: Many existing systems need special sensors or cameras to work, making them harder to set up and use. The proposed system is designed to work on regular webcams, so people can use it on their laptops or tablets without any extra equipment.

Output Type: Existing systems typically only show text or may have a delay before the translation appears. The proposed system offers both text and voice output at the same time, making it easier for everyone to understand the translation.

Ease of Use: While existing systems may require complicated setups or extra hardware, the proposed system is designed to be simple and easy to use, requiring only a webcam.

XI. CONCLUSION

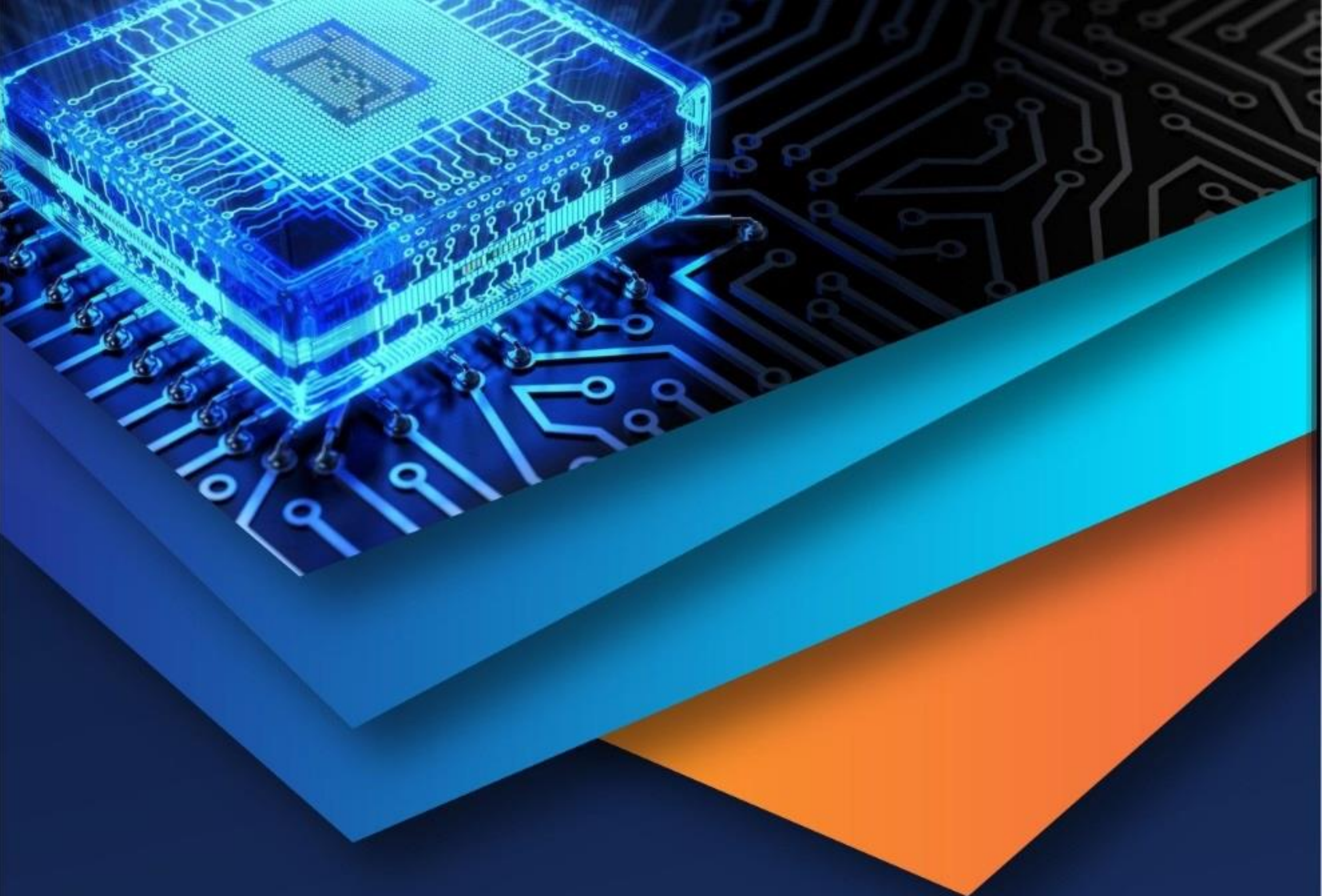
In conclusion, the proposed real-time sign language detection system significantly bridges communication gaps between the deaf and hard-of- hearing community and non-signers. By leveraging standard web cameras and advanced computer vision techniques, this system not only enhances accessibility in vital sectors such as education and healthcare but also promotes broader societal awareness of sign language. The dual-output feature—providing both voice interpretation and on- screen text—ensures effective real-time communication, enabling users to interact seamlessly regardless of their familiarity with sign language.

Furthermore, the system's adaptability for integration into various camera-equipped devices facilitates widespread deployment, making it accessible in everyday situations, such as classrooms, medical facilities, and public spaces. This accessibility fosters a more inclusive environment, encouraging interactions that acknowledge and respect diverse communication methods. By empowering individuals with the tools to communicate more effectively, the system contributes to a more connected society, ultimately reducing the stigma and

barriers faced by sign language users. As we continue to refine this technology and explore additional features, we anticipate that it will play a pivotal role in promoting understanding, collaboration, and respect among all individuals, regardless [i] of their hearing abilities.

REFERENCES

[1] R. Mehta, "Hand Gesture Recognition Using Convolutional Neural Networks (CNNs)," Project, 2023.
[2] K. S. Deshmukh, "Real-Time Sign Language Detection Using Kinect Sensors," Project, 2023.
[3] M. N. Rao, "Deep Learning-Based Sign Language Recognition Using LSTM Networks," Project, 2023.
[4] A. T. Bhatia, "Sign Language Recognition Using YOLO and OpenCV," Project, 2023.
[5] A. Elhagry and R. G. Elrayes, "Video-based Egyptian Sign Language Recognition using CNN and CNN-LSTM Models," International Journal of Machine Learning and Computing, vol. 12, no. 4, pp. 123-135, 2023.
[6] R. Rastgoo, K. Kiani, and S. Escalera, "Zero-Shot Learning for Sign Language Recognition Using RGB-D Videos," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 5, pp. 1125-1137, May 2023.
[7] M. Montazerin, S. Zabihi, and E. Rahimian, "Hand Gesture Recognition using Vision Transformers (ViT) and Surface Electromyography (HD-sEMG) Signals," IEEE Transactions on Biomedical Engineering, vol. 70, no. 3, pp. 175-183, Mar. 2023.
[8] S. Savaş and A. Ergüzen, "A Two-Stage Approach for Hand Gesture Recognition Using Transfer Learning and Deep Ensemble Learning," IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 6, pp. 2509-2516, June 2023.
[9] M. Islam, M. Aloraini, and S. Aladhadh, "Stacked Encoded Deep Learning Framework for Sign Language Recognition," IEEE Access, vol. 11, pp. 2441-2449, 2023.
[10] H. Hu, W. Zhao, W. Zhou, and H. Li, "SignBERT+: A Self-Supervised Pre-Training Framework for Sign Language Recognition," IEEE Transactions on Multimedia, vol. 25, no. 3, pp. 789-797, Mar. 2023.



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