iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78009

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 14 Issue III Mar 2026- Available at www.ijraset.com

Sign Language Detection and Text and Word Conversion using AI-ML

Parth Dattatray Khaladkar[1], Parth Santosh Bodke[2], Mayur Govind Chaskar[3], Sagar Laxman Chavan[4], Prof. Thorat S. K.[5]
Computer Engineering Department, Samarth Polytechnic Belhe

Abstract: ***Sign language is very important communication method for people who are hearing or speech disabled. However, many people are not familiar with sign language, which creates a communication gap between them and normal users. The proposed system aims to reduce this gap by detecting hand gestures and converting them into readable text using Artificial Intelligence and Machine Learning techniques. In this project, hand gestures are captured using the built-in laptop webcam. A custom dataset consisting of 10–15 unique gestures is created, where each gesture represents a specific word or message. The collected gesture data is used to train a Recurrent Neural Network (RNN) model for gesture recognition. During real-time operation, the system captures gesture images from the webcam, processes them using image processing techniques, and guess the gesture using the trained model. Once the gesture is recognized, it is automatically converted into associated text and displayed on the screen. The system provides a simple and efficient way to translate sign language gestures into text without requiring any additional hardware devices.***

Keywords: Sign Language Recognition, Gesture Detection, Artificial Intelligence, Machine Learning, Recurrent Neural Network (RNN), Computer Vision, Text Conversion.

I. INTRODUCTION

Communication is an essential part of everyday life. However, individuals with hearing and speech disabilities often face difficulties when interacting with people who do not understand sign language. Sign language uses different hand gestures to represent words, messages, or sentences. Since many people are not familiar with these gestures, communication between hearing-impaired individuals and others becomes challenging. To address this issue, a sign language detection system can be developed using Artificial Intelligence and Machine Learning techniques. Such systems are capable of recognizing hand gestures and converting them into readable text so that the message can be easily understood by others. During real-time operation, the webcam captures the hand gesture, the trained model identifies the gesture, and the system converts it into the respective text output. This approach helps improve communication between hearing-impaired individuals and others.

The proposed system provides the following features:

1) Real-time hand gesture detection using the laptop webcam
2) Recognition of custom sign language gestures using a trained RNN model
3) Conversion of detected gestures into readable text output
4) Ability to assign specific words or sentences to different gestures
5) Capability to extend the system by training additional gesture data in the future

By combining gesture detection, machine learning, and real-time text conversion, the system helps improve communication between hearing-impaired individuals and others, making interaction easier and more accessible.

II. LITERATURE REVIEW

Sign language recognition has become an important research area for improving communication with hearing-disabled and speech-disabled people. Many researchers have developed systems using computer vision and machine learning techniques to detect hand gestures and convert them into alphabets or words. However, most existing systems focus on alphabet-based recognition, which requires multiple gestures to form a complete sentence and makes communication slower.

A. Sign Language Recognition using Machine Learning

Many researchers have developed sign language recognition systems using machine learning and computer vision techniques. These systems detect hand gestures from images or video frames and convert them into text or speech output. Various algorithms are used to train models for recognizing different hand signs.

B. Deep Learning based Gesture Recognition

Some studies have used deep learning models such as neural networks to recognize sign language gestures. These models can learn complex gesture patterns and improve recognition accuracy. However, such systems often require large datasets and high computational resources for training.

C. Sensor-Based Sign Language Systems

Certain systems use special hardware devices such as sensor gloves to detect hand movements. Although these systems provide accurate gesture detection, they require additional hardware which increases cost and reduces convenience for everyday use.

D. Limitation of Existing Systems and Proposed Idea

In many existing sign language recognition systems, gestures are assigned to alphabetical characters. For example, one gesture represents the letter “A”, another represents “B”, and another represents “C”. To form a complete sentence, the user needs to perform multiple gestures corresponding to each letter, which makes the communication process time-consuming.

In the proposed system, instead of assigning individual letters, commonly used words or sentences are directly assigned to specific gestures. This approach reduces the number of gestures required and allows the system to display meaningful text output quickly.

Additionally, the system is designed in a flexible way so that the model can be trained again at any time with new gesture data. This allows new signs or sentences to be added easily, improving the system’s scalability and usability.

III. PROBLEM STATEMENT

Investors face multiple challenges:

1) People who are hearing-disabled or speech-disabled mainly communicate using sign language, but most people are not familiar with sign language gestures.
2) This creates a communication gap between disabled individuals and the general public in daily interactions.
3) In many existing systems, each hand gesture represents a single alphabet (for example A, B, C), and sentences are formed by combining multiple gestures
4) This process becomes time-consuming and inefficient for real-time communication.
5) Therefore, there is a need to develop a system that can detect gestures and directly convert them into commonly used sentences to make communication faster and easier.
6) The proposed system uses a webcam-based gesture detection model that can recognize signs and convert them into readable text output.

IV. PROPOSED SYSTEM ARCHITECTURE

A. System Architecture

The proposed system is designed to recognize sign language gestures and convert them into text and words using Artificial Intelligence and Machine Learning techniques. The system is developed using the Python programming language and various libraries for image processing and prediction.

1) Input Module: The system captures hand gestures using a webcam or camera. The captured images or video frames act as input for the system.
2) Image Processing Module: The captured frames are processed using image processing techniques. Python libraries such as OpenCV and NumPy are used to detect and extract hand gestures from the video frames.
3) Feature Extraction: Important features from the detected hand gestures are extracted. These features help the machine learning model understand the pattern of the gesture.
4) Machine Learning Model: The extracted features are passed to the trained machine learning model. The model identifies the gesture and predicts the corresponding sign language sentence.
5) Text and Word Conversion : After prediction, the detected sign language gesture is converted into readable **text**.
6) Output Display: The final result is displayed on the screen so that users can see the translated text from the sign language gesture.

B. Workflow

1) Webcam Capture: The system opens the laptop webcam and captures real-time video frames of the hand gesture.

2) Hand Key point Detection: Each frame is analyzed to identify the hand joints and their coordinates, which represent the position of fingers and palm in the frame.
3) Frame Preprocessing & Feature Extraction: The frame is preprocessed (resizing, noise removal, segmentation), and important features like joint positions, angles, or distances between joints are extracted.
4) Gesture Recognition using RNN: The extracted features are fed into a trained RNN model, which predicts the performed gesture.
5) Result Selection: The predicted gesture with the highest confidence is selected.
6) Text Conversion: The recognized gesture is converted into the assigned sentence and can also be converted into speech output.

V. LIMITATIONS

1) The system currently supports a limited number of trained gestures, so it cannot recognize gestures that are not included in the dataset.
2) Gesture recognition accuracy may depend on lighting conditions and camera quality.Sentiment analysis optional, requiring external data collection
3) The model needs retraining to add new gestures, which requires additional data collection and training time.

VI. ADVANTAGES

1) The system reduces the communication gap between disabled individuals and normal users by converting hand gestures into understandable text.Multi-stock and multi-timeframe analysis
2) The system works using a standard laptop webcam, so no additional hardware or sensors are required.Portfolio simulation for investment strategy planning
3) The model can be trained again to add new gestures anytime, which increases the flexibility of the system.
4) Unlike traditional systems, the proposed system maps gestures directly to commonly used sentences, which makes communication faster.

VII. RESULTS AND PROJECT OUTPUT

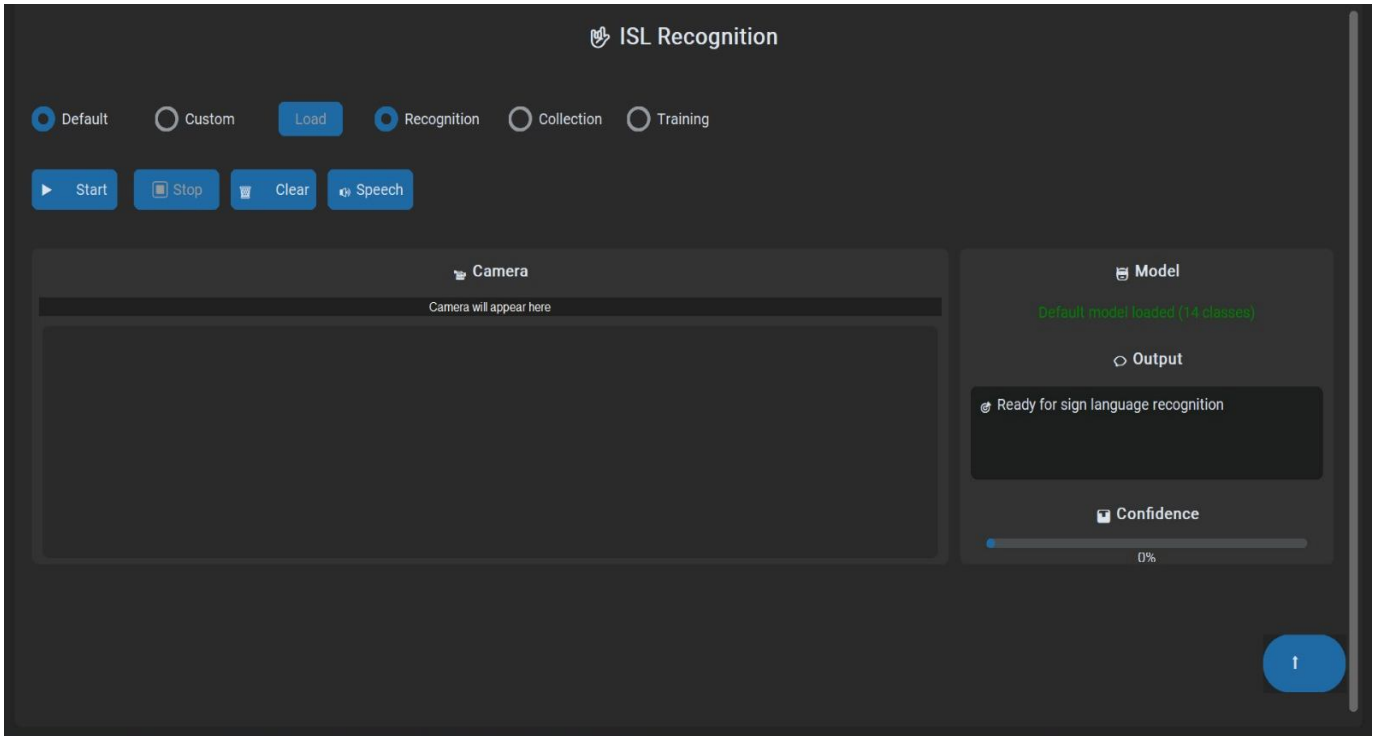


Fig.1-Home screen

A. Model Training for Sign 1

In this step, the system is trained to recognize the first sign gesture using the dataset. The machine learning model learns the pattern of the hand gesture during training.

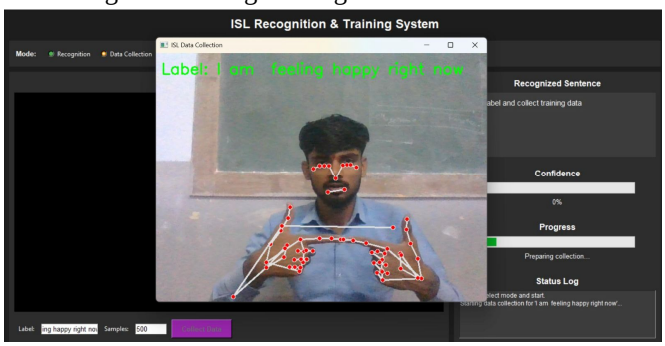


Fig.2- Model Training

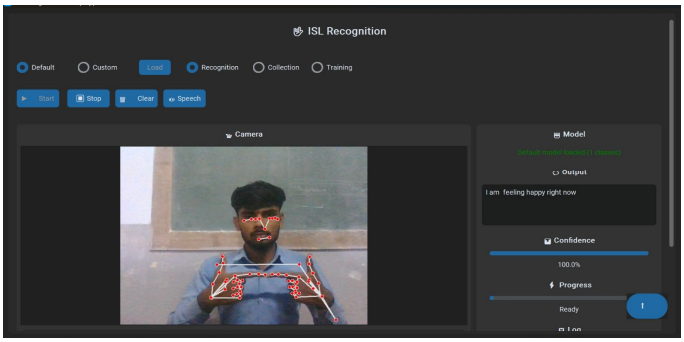


Fig.3-Output

B. Model Training for Sign 2

The system is trained with the dataset for the second sign gesture. The model analyzes the hand gesture features to identify the correct sign.

Fig.3- Model Training

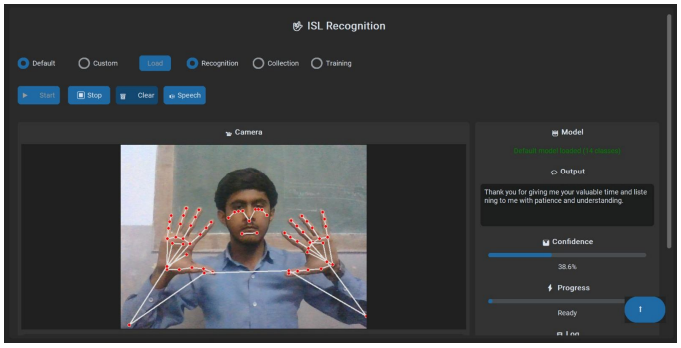


Fig.5- Output

C. Model Training for Sign 3

In this step, the system is trained to recognize the third sign gesture using machine learning techniques.

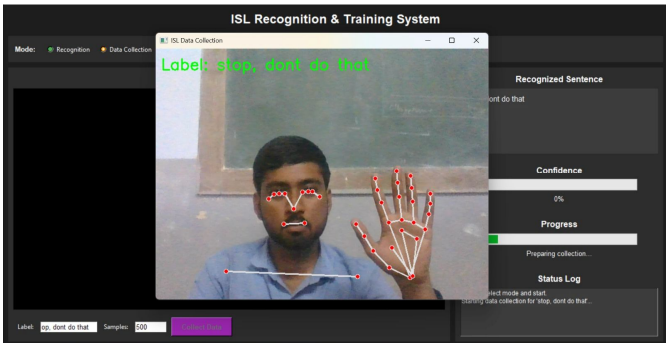


Fig.6- Model Training

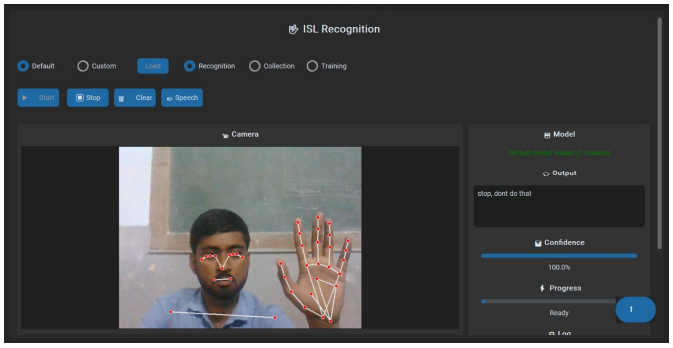


Fig.7- Output

VIII. CONCLUSION

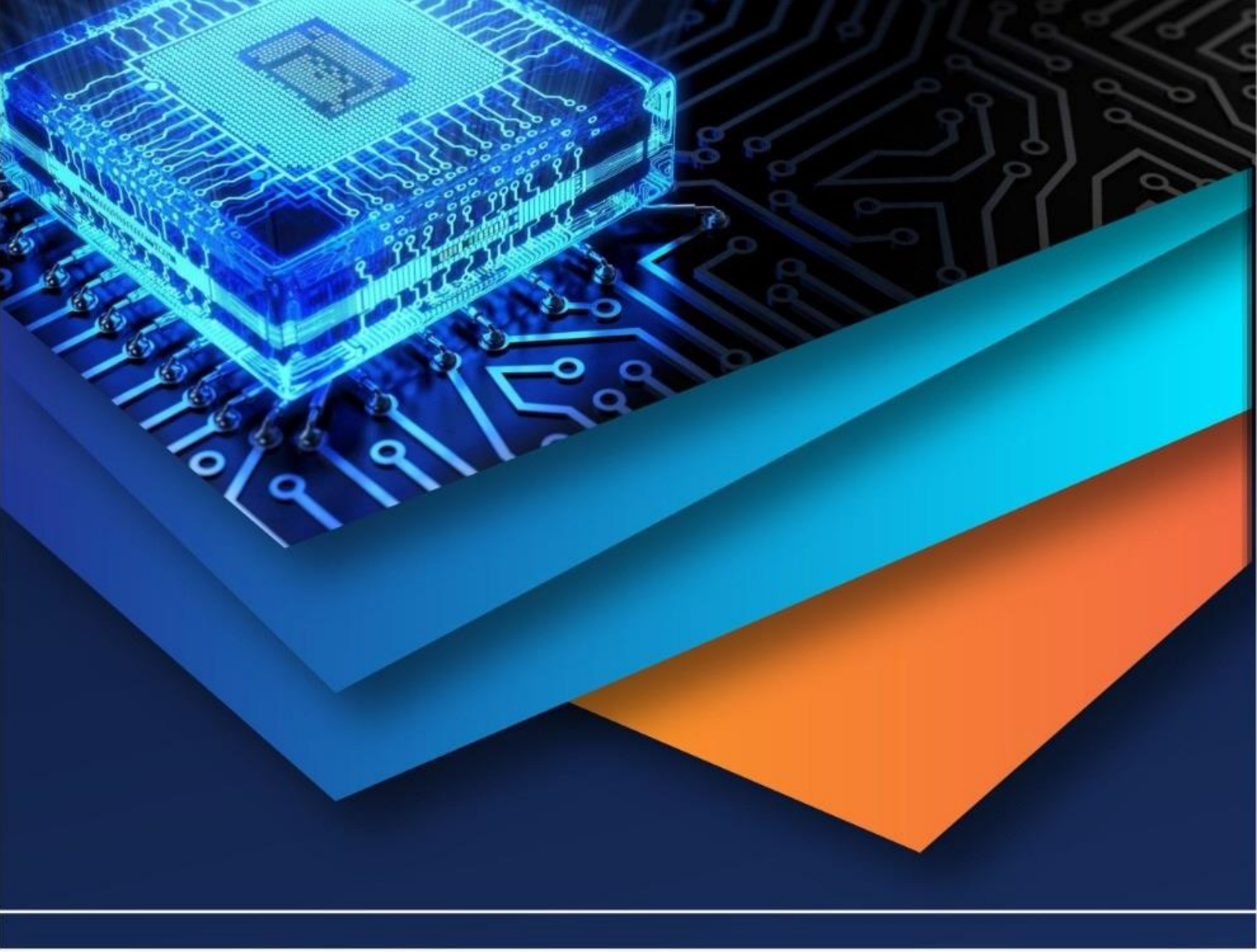
The Sign Language Detection and Text and Word Conversion system was successfully developed using Python and AI/ML techniques. The system captures hand gestures through a webcam and converts them into readable text or words.In this project, a limited number of sign language gestures were trained in the model. Instead of training separate gestures for each alphabet, specific gestures were assigned to commonly used sentences. This approach was used because training gestures for every alphabet would require a large dataset and more time for model training. The trained model was able to recognize the defined gestures and display the corresponding text output accurately.

This system demonstrates how artificial intelligence and image processing can help reduce the communication gap between deaf or mute individuals and others.In the future, the system can be improved by including more gestures and training the model for complete sign language alphabets to make the system more advanced and accurate. In this system, a limited number of sign language gestures were trained and assigned to commonly used sentences. Training separate gestures for each alphabet was not implemented because it requires a large dataset and more training time.The system demonstrates how artificial intelligence can be used to assist communication for deaf and mute individuals. It provides a simple way to translate hand gestures into text so that others can easily understand the message.

The project also shows that machine learning and computer vision technologies can be effectively applied in real-world communication systems. In the future, the system can be expanded by adding more sign language gestures, improving model accuracy, and converting the text output into voice for better communication.

REFERENCES

[1] OpenCV Contributors, "OpenCV-Python Tutorials: Video Capture and Processing," OpenCV Documentation, https://docs.opencv.org/4.x/dd/d43/tutorial_py_video_display.html, Accessed: March 2025
[2] Scikit-learn Developers, "Scikit-learn 1.3.0: Machine Learning Utilities for Evaluation," https://scikit-learn.org/1.3/, Accessed: March 2026.
[3] Matplotlib Development Team, "Matplotlib 3.7.2: Visualization of Training Metrics and Results," https://matplotlib.org/3.7.2/, Accessed: March 2026.
[4] Google AI Edge, "MediaPipe Solutions: Holistic Landmark Detection (v0.10.7)," MediaPipe Documentation, https://ai.google.dev/edge/mediapipe/solutions/vision/holistic, Accessed: March 2026.
[5] Pandas Development Team, "Pandas 2.0.3 Documentation: Data Handling and Preprocessing," https://pandas.pydata.org/docs/, Accessed: March 2026.
[6] Abdulrahman-tech, "ASL Recognition: Real-time Sign Language using PyTorch and MediaPipe," GitHub Repository, 2024–2025. [Online]. Available: https://github.com/Abdulrahman-tech/asl-recognition
[7] Abdulrahman-tech, "ASL Recognition: Real-time Sign Language using PyTorch and MediaPipe," GitHub Repository, 2024–2025. [Online]. Available: https://github.com/Abdulrahman-tech/asl-recognition
[8] N. M. Bhat, "pyttsx3: Offline Text-to-Speech Conversion Library (v2.90)," PyPI and Documentation, https://pypi.org/project/pyttsx3/, Accessed: March 2026.
[9] Google AI Edge, "MediaPipe Solutions: Holistic Landmark Detection (v0.10.7)," MediaPipe Documentation, https://ai.google.dev/edge/mediapipe/solutions/vision/holistic, Accessed: March 2026.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY