© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716012

Real-Time Sign Language Detection System Using Mediapipe and Machine Learning

GLADIS KEZIAH A[1], MONISHA V[2], VINNIAMMAL B[3]
[1, 2, 3]*Department of Computer Science and Engineering, Velammal Engineering College*

Abstract- ***It's usually difficult for people with hearing impairments to communicate with those who don't know sign language. Sign language recognition systems are attempts to fix this problem, and they do so by using computer vision and machine learning. This particular paper describes a way to detect sign language in real time, using MediaPipe and machine learning to understand hand movements and turn them into text you can read. The system uses a webcam to 'see' the hand signs and MediaPipe's hand tracking to find specific points on the hand. The positions of these points on the hand become the data used to train a machine learning model to recognize many different sign language gestures. The system can work out what's being signed almost as it happens, it is very accurate, and it doesn't need a super powerful computer. Therefore, it's a cheap and effective method to help deaf and mute people communicate, and could be used in teaching or as a tool to help with daily life.***

Index Terms- ***Sign Language Recognition, MediaPipe, Machine Learning, Computer Vision, Hand Gesture Detection, Real-Time Detection***

I. INTRODUCTION

Deaf and mute people commonly use sign language to communicate. Meaning is expressed through how the hands move, expressions on the face, and how the body is positioned. Unfortunately, because very few of us know sign language, this creates a problem getting messages between those with hearing difficulties and everyone else.

Luckily, because of improvements in machine learning and computer vision, we can now make computer programs that 'see' and understand signs as they are being made. Earlier versions of these programs needed expensive, awkward data gloves and equipment with sensors to track the hands.

Now, though, things like MediaPipe developed by Google can follow hands and pinpoint important points on them, all using a normal camera, and very quickly. And by pairing MediaPipe with machine learning that categorizes signs, we can make sign language recognition systems that are quite precise.

This paper describes a sign language recognition system which works in real time, utilizes a standard webcam to detect hand gestures, MediaPipe to extract hand landmark features, and a trained machine learning model to classify gestures.

II . LITERATURE REVIEW

Numerous studies have investigated various methods for recognizing sign language. The first systems used data gloves and sensors to follow hand movement. While these systems were very accurate, they were very expensive and required additional proprietary equipment.

More recent studies have focused on methods using vision systems, specifically systems utilizing cameras coupled with computer vision methods. Classification of images of hand gestures using deep learning techniques has led to the popularity of Convolutional Neural Networks (CNNs) for this.

In hand tracking, MediaPipe has proven to be a successful rapid prototyping tool, successfully identifying 21 unique hand landmarks. Many researchers have integrated the MediaPipe framework with classifiers based on various machine learning (ML) techniques, including Support Vector Machines (SVM), Random Forests, and Neural Networks.

Even with these advancements, the majority of current systems either demand a substantial amount of computation or are not able to function in real time. The main goal of this system is to provide a lightweight and efficient solution to the problem of

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716012

real time recognition of sign language gestures using a combination of MediaPipe and various machine learning techniques.

III. EXISTING SYSTEM

Most recent sign language detection systems rely on either sensor-based gloves or deep learning image-based models. Sensor-based systems need users to wear special gloves that have sensors on them to track finger movements. But while these systems are accurate, they are too costly and are impractical to use in everyday life.

Vision-based systems that utilize deep learning have big data and machine learning hardware requirements for training and implementation. In addition, some systems have slow processing speeds, high computational costs, and they lack gesture recognition. Many systems have no real-time gesture recognition which decreases their real-world usability.

IV. PROPOSED SOLUTION

The proposed system uses MediaPipe hand tracking and machine learning techniques to detect sign language movements in real time.

The system is made up of four important elements:
1. Video Capture Module--The webcam captures real-time video frames of the user's hand gestures.
2. Hand Landmark Detection--MediaPipe identifies 21 critical points of the hand, including fingertips, joints, and wrist positions.
3. Feature Extraction--The coordinates of these landmarks are retrieved and normalized to form a feature vector representing the hand gesture.
4. Gesture Classification--A machine learning classifier trained on labeled gesture data predicts the matching sign language symbol or phrase.

The predicted output is then displayed as text on the screen, enabling communication between sign language users and others.

V. WORKING

The working process of the recommended system is:
1. The webcam is used mainly to capture the real-time video input.
2. Each video frame is processed using the MediaPipe framework.
3. MediaPipe detects the hand and extracts 21 landmark points.
4. The landmark coordinates are converted into feature vectors.
5. The features are analyzed using the trained machine learning model.
6. The system predicts the corresponding sign language gesture.
7. The recognized gesture is displayed as text on the screen.

This process occurs continuously, allowing real-time sign language recognition.

VI. CONCLUSION

This paper presented a Real-Time Sign Language Detection System using MediaPipe and Machine Learning. The system utilizes MediaPipe for accurate hand landmark detection and machine learning algorithms for gesture classification. The proposed approach provides a lightweight and efficient solution that can recognize sign language gestures in real time without requiring expensive hardware.

This system can help improve communication between hearing-impaired individuals and the general public. Future developments can further improve the system by supporting a larger vocabulary of gestures and integrating speech output to enhance accessibility.

VII. FUTURE WORK

The proposed system can be enhanced in several ways to improve its usability and accuracy.

Future work may include expanding the dataset to support complete sign language sentences rather than individual gestures Convolutional techniques of deep learning include Convolutional. Neural Networks

(CNNs) and Recurrent Neural Networks. Gesture may also be enhanced through integration of (RNNs). Additionally, the system can be implemented as a mobile application so that users can access the sign language recognition system using smartphones. Integration with speech synthesis can also allow the system to convert detected gestures into spoken language.

REFERENCES

[1] Z. Zhang, Z. Cao, T. Simon, S. Wei, and Y. Sheikh, "MediaPipe Hands: On-device Real-time Hand Tracking," *arXiv preprint arXiv:2006.10214*, pp. 1–9, Jun. 2020. doi: 10.48550/arXiv.2006.10214.

[2] S. C. W. Ong and S. Ranganath, "Automatic Sign Language Analysis: A Survey and the Future Beyond Lexical Meaning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 6, pp. 873–891, Jun. 2005. doi: 10.1109/TPAMI.2005.112.

[3] O. Koller, H. Ney, and R. Bowden, "Deep Learning of Mouth Shapes for Sign Language Recognition," in *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW)*, Santiago, Chile, 2015, pp. 85–91. doi: 10.1109/ICCVW.2015.20.

[4] Google AI, *MediaPipe Framework Documentation*

[5] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.