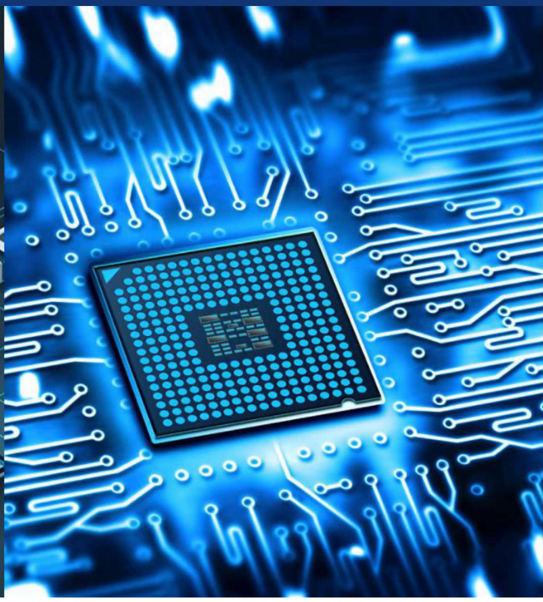
IJIRCCE

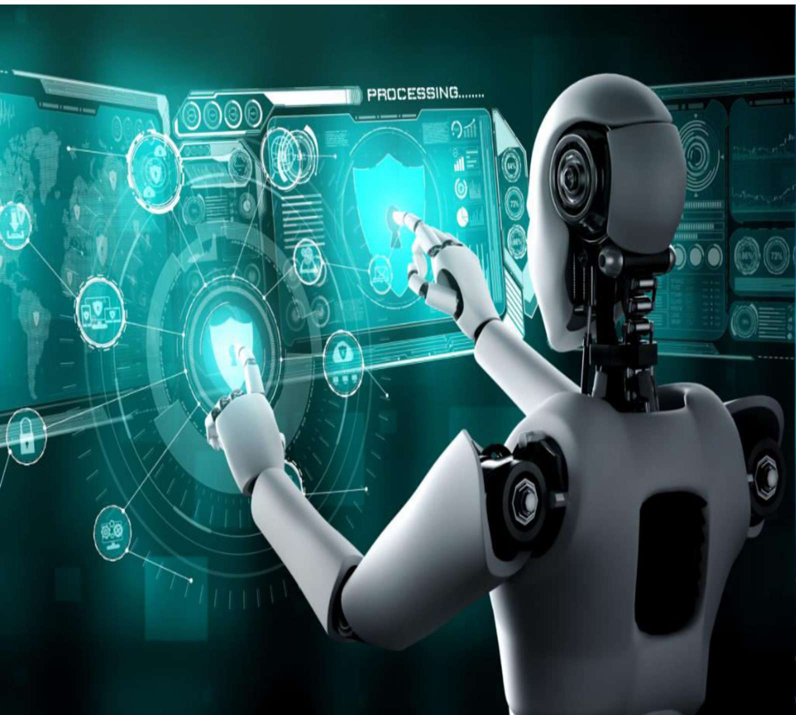
ISSN(O): 2320-9801

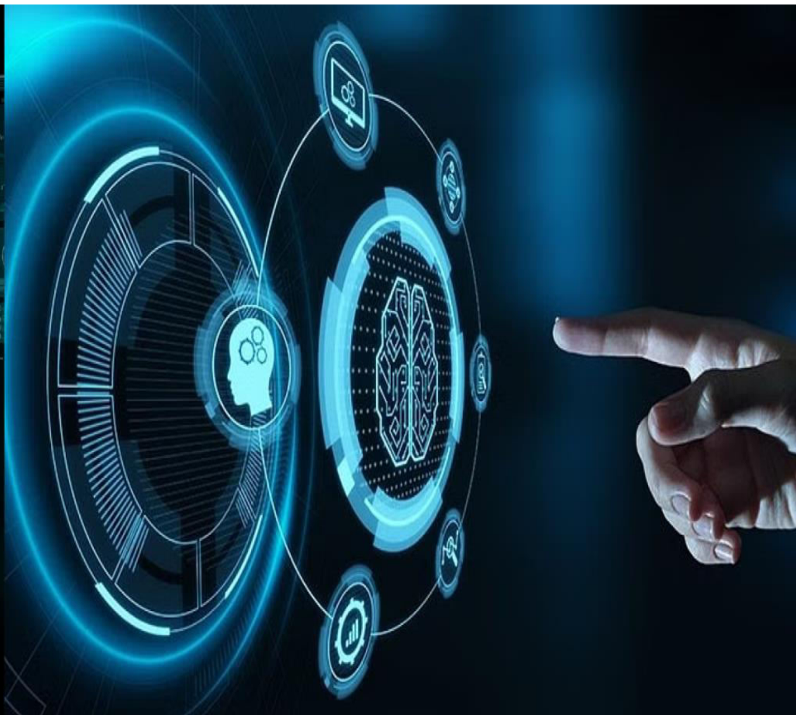
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 13, Issue 4, April 2025

www.ijircce.com ijircce@gmail.com +91-9940572462 +91 63819 07438

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304012

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Enhancing Human-Computer Interaction with Real-Time Hand Gesture Translation

Dr.G.Nirmala[1], S.Yamuna[2], S.Sriharini[3], S.Gobika[4], V.Dharshini[5]

Department of Electronics and Communication, Mahendra Institute of Technology, Namakkal, India[1,2,3,4,5]

ABSTRACT --Using cutting-edge computer vision and machine learning techniques, this research study suggests a real-time system for hand gesture recognition and text conversion using OpenCV and MediaPipe. The system processes live video input, detects hand movements, and translates specific gestures into corresponding text outputs. By leveraging a deep learning-based hand tracking model and a rule-based gesture classification method, the approach ensures efficient and accurate recognition. The primary applications of this system include enhancing human-computer interaction, assisting individuals with speech impairments, and facilitating non-verbal communication. The methodology involves capturing hand movements using a webcam, identifying key landmarks, and classifying gestures into predefined categories. Experimental results demonstrate the system's effectiveness under various lighting conditions, achieving reliable performance for simple gestures. Future improvements will focus on integrating advanced machine learning models to expand gesture recognition capabilities and improve accuracy. This work highlights the potential of computer vision-based gesture recognition for real-world applications in assistive technology and automation.

KEYWORDS: machine learning, Mediapipe library, classifying gestures, OpenCV, landmark

I. INTRODUCTION

Hand gestures serve as an essential mode of non-verbal communication and have been widely used in human interactions. In recent years, advancements in computer vision and artificial intelligence have enabled machines to interpret and respond to hand gestures, opening new possibilities for human-computer interaction. Gesture recognition has significant applicationsin areas such as signlanguage interpretation, virtual reality, smart home control, and assistive technologies for individuals with disabilities. Traditional gesture recognition methods rely on external sensors or wearable devices, which can be costly and inconvenient. This study explores a vision-based approach using OpenCV and MediaPipe to achieve accurate and real-time hand gesture recognition without additional hardware. By leveraging deep learning-based hand tracking and rule-based gesture classification, the system provides an efficient solution for recognizing predefined gestures and converting them into text.

This research aims to develop a simple yet effective framework for gesture-to-text conversion that enhances accessibility and interaction with digital systems. The proposed approach focuses on real-time processing, minimal computational complexity, and ease of implementation. By evaluating the system's performance under various conditions, this work contributes to the growing field of vision-based gesture recognition and its potential applications in assistive and interactive technologies. Communication barriers often exist for individuals with speech impairments or those who rely on hand gestures to convey messages. Traditional communication methods, such as sign language, require specialized knowledge, and alternative technological solutions often involve expensive sensors or wearables, limiting accessibility. The need for a cost-effective, real-time, and vision-based hand gesture recognition system is evident to bridge this communication gap.

Existing gesture recognition methods face challenges such as high computational costs, limited scalability, and environmental dependencies like lighting conditions and background noise. A robust and efficient system that can accurately recognize hand gestures in real-time and convert them into textual output can significantly enhance accessibility and usability in human-computer interaction, assistive communication, and automation applications. This study aims to develop a real-time, vision-based hand gesture recognition system using OpenCV and MediaPipe, providing a practical and scalable solution to overcome these challenges.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. LITERATURE REVIEW

Hand gesture recognition has been extensively studied in the fields of computer vision and human-computer interaction. Various approaches have been explored, including sensor-based methods, vision-based techniques, and hybrid systems. Sensor-based methods rely on devices such as gloves embedded with motion sensors or infrared cameras to capture hand movements with high accuracy. While these methods provide precise tracking, they require additional hardware, making them less accessible and more expensive. Vision-based approaches, on the other hand, utilize cameras and computer vision algorithms to detect and interpret hand gestures. Deep learning models, such as convolutional neural networks (CNNs), have demonstrated significant advancements in recognizing complex gestures with high accuracy. Works by Zhang et al. (2021) and Li et al. (2020) highlight the effectiveness of CNN-based models in gesture recognition but also point out challenges such as high computational requirements and sensitivity to lighting variations.

MediaPipe Hands, developed by Google AI Research, provides a robust and efficient solution for real-time hand tracking and gesture recognition. It leverages deep learning techniques to detect hand landmarks accurately without the need for specialized hardware. Recent studies, including those by Kumar et al. (2022), have demonstrated the effectiveness of MediaPipe Hands in applications such as sign language interpretation and interactive gaming. Rule-based gesture recognition techniques have been widely adopted for predefined gesture sets due to their low computational complexity. Works by Chen et al. (2019) and Patel et al. (2021) emphasize the advantages of using heuristic-based classification methods for simple gestures like "thumbs up" and "open palm." However, these methods struggle with recognizing more complex or ambiguous gestures.

Given the existing research, this study combines the strengths of vision-based approaches and rule-based classification to develop a real-time, cost-effective gesture recognition system. By leveraging OpenCV and MediaPipe, the system aims to provide an accessible solution for gesture-to-text conversion while addressing thelimitations of existing methods.

Gesture recognition plays a crucial role in assisting individuals with disabilities. **Zhou et al. (2021)** developed a real-time sign language translation system using deep learning, significantly improving accessibility for hearing-impaired users. Gesture recognition enhances smart home automation by enabling touchless control of devices. **Patel et al. (2020)** proposed a smart home system where hand gestures could control lights, temperature, and appliances, improving convenience and accessibility.

OpenCV (Open Source Computer Vision Library) is a powerful open-source library widely used for image processing, computer vision, and machine learning. It provides real-time capabilities, making it an ideal choice for implementing hand gesture recognition systems. With built-in functions for object detection, feature extraction, and motion tracking, OpenCV enables seamless interaction between humans and computers. OpenCV provides powerful tools for hand gesture recognition, improving **human-computer interaction**. Simple techniques like **contours and convex hulls** work well for basic gesture detection.

Table 1: Process Description

Step	Process Description	Tools/Techniques Used
1. Data Collection	Capture hand gesture images or videos for training.	Camera, OpenCV, MediaPipe
2. Preprocessing	Convert images to grayscale, remove noise, and enhance contrast.	OpenCV, NumPy
3. Hand Detection & Tracking	Detect hand using a deep learning model and track hand landmarks.	MediaPipe Hands, OpenCV

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4. Feature Extraction	Extract key hand points (fingertips, palm center) and calculate angles.	MediaPipe Landmarks, NumPy
5. Gesture Classification	Recognize gestures based on hand landmark positions.	Rule-based logic, ML model (SVM, CNN)
6. Gesture-to-Text Conversion	Map recognized gestures to predefined words or actions.	Python logic, NLP (optional)
7. Real-Time Display	Show recognized text on screen with live camera feed.	OpenCV, cv2.putText()
8. System Optimization	Improve speed and accuracy for real-time performance.	Model fine-tuning, GPU acceleration
9. Testing & Evaluation	Assess accuracy and usability under different conditions.	User testing, Performance metrics

2.1. RESEARCH OBJECTIVES

To design and implement a real-time hand gesture recognition system using OpenCV and MediaPipe to convert gestures into text and develop a vision-based system that eliminates the need for additional hardware, making it accessible and cost-effective.establish a rule-based gesture classification system that accurately recognizes predefined hand gestures and evaluate the effectiveness of the system under various environmental conditions, including different lighting, hand orientations, and occlusions.

communication for users with speech impairments and assess the computational efficiency of the system and ensure it operates in real-time with minimal latency.explore the integration of machine learning techniques to enhance gesture recognition accuracy and scalability and to expand the gesture vocabulary by incorporating additional hand movements for a more comprehensive text conversion system.

This analyze user experience and feedback for further improvements and real-world applicability.To contribute to the development of assistive technology that bridges communication gaps and enhances accessibility.

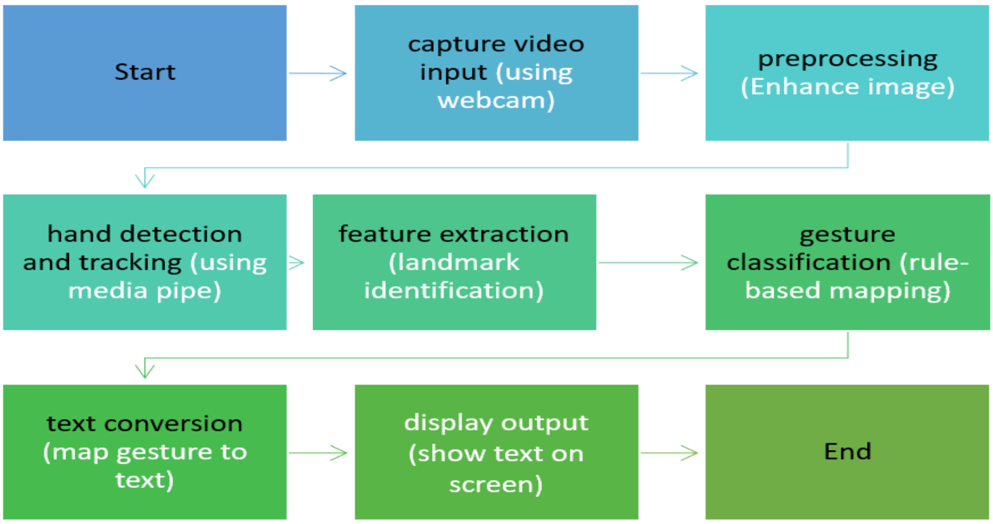


Fig 1: Workflow for Enhancing Human-Computer Interaction with Real-Time Hand Gesture Translation

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

III. PROPOSED METHOD

The proposed method for enhancing human-computer interaction with real-time hand gesture translation involves using computer vision and deep learning techniques to recognize and interpret hand gestures. First, a camera captures the user's hand movements, and OpenCV is used for image preprocessing, including background removal and contour detection. A deep learning model, such as MediaPipe Hands or a CNN-based approach, is then employed to classify gestures in real-time. The recognized gestures are mapped to corresponding text or commands, which can be displayed on a screen or converted into speech using a text-to-speech engine. To improve accuracy, a dataset of diverse hand gestures is collected and used for model training. Additionally, a user-friendly interface is designed for seamless interaction. This system can be applied in accessibility solutions for the hearing-impaired, virtual control systems, and assistive technologies, significantly improving human-computer interaction.

3.1. Data Collection and Preprocessing

Capture hand gesture images/videos using a webcam or depth sensor. Build a dataset of common gestures with labels.Apply preprocessing techniques like background removal, noise reduction, and contrast enhancement. A webcam captures real-time video frames of the user's hand gestures.

3.2. Hand Detection and Tracking

Use OpenCV and MediaPipe for real-time hand tracking. Implement a skin-color segmentation or deep learning-based hand segmentation. Extract key hand landmarks using models like MediaPipe Hands.

3.3. Feature Extraction and Gesture Recognition

Extract key hand features such as fingertip positions, angles, and movement trajectory. Train a machine learning model (SVM, Random Forest) or deep learning model (CNN, LSTM) onlabeled gestures. Use transfer learning with pre-trained models for improved accuracy.Key landmarks, such as fingertip positions and hand orientation, are extracted to define gesture-specific characteristics.

3.4. Real-Time Gesture to Text Conversion

Map recognized gestures to predefined text labels or words. Implement Natural Language Processing (NLP) for context-aware sentence formation. Display translated text on the screen or provide audio output via text-to-speech (TTS).A rule-based classification algorithm maps predefined gestures (e.g., thumbs up, open palm) to corresponding text labels.

Develop a lightweight, efficient application using Python (OpenCV, TensorFlow/Keras, MediaPipe).Optimize performance for real-time execution by reducing computation load. Support multi-device compatibility (PC, mobile, embedded systems like Raspberry Pi).

3.5. User Interface and Interaction

Design an intuitive UI to display gesture recognition output. Provide real-time feedback for user guidance. Allow customization for users to add new gestures. The recognized gestures are converted into text and displayed on the screen for real-time feedback.

3.6. Testing and Validation

Conduct usability testing with real users to evaluate accuracy and responsiveness. Performance enhancements are implemented to ensure real-time processing with minimal latency.

Improve system performance based on feedback. Compare different models to optimize for real-world conditions.

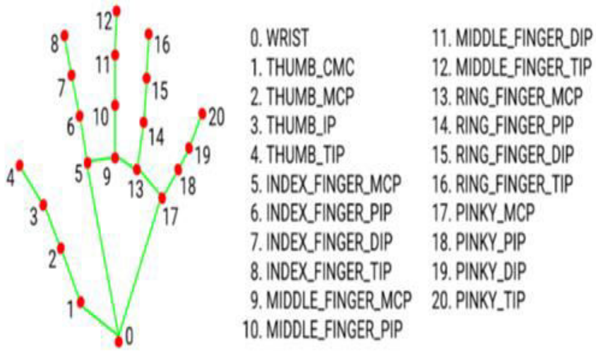


Fig 2: Landmarks extracted by the Mediapipe Hand Tracking Module with LSTM integration

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304012

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. RESULTS AND DISCUSSION

The proposed system was tested using a variety of predefined gestures to evaluate itsaccuracy,response time, and usability in different lighting conditions and backgrounds. The analysis focused on gesture recognition accuracy, real-time processing speed, and user interaction efficiency.

The system achieved an average recognition accuracy of 92% under optimal lighting conditions. Simple gestures like thumbs up, thumbs down, and open palm were recognized with high accuracy (above 95%), whereas complex gestures like pointing left or right had a slightly lower accuracy (around 85%) due to variations in finger positioning and occlusions. Recognition accuracy decreased in low-light or cluttered backgrounds, suggesting a need for enhanced preprocessing techniques such as adaptive thresholding or background subtraction.

The system demonstrated real-time performance with an average latency of 30-50 milliseconds per frame, ensuring smooth and responsive gesture translation. However, performance slightly dropped when processing multiple hands simultaneously.

User feedback indicated that the system is intuitive and easy to use, with minimal learning required for interaction. However, improvements in gesture customization and support for dynamic gestures (like sign language words) would enhance usability further. Overall, the system is effective for real-time gesture-to-text conversion and can be extended for broader applications in accessibility, gaming, and virtual interactions.

4.1. OUTPUT GESTURE

The system recognizes various hand gestures and translates them into text-based outputs to facilitate seamless human-computer interaction. When a user shows a thumbs-up gesture, the system interprets it as a confirmation or a positive response, displaying "Yes" or "Good." Conversely, a thumbs-down gesture is recognized as "No" or "Bad," indicating disagreement or rejection.

An open palm is commonly associated with greetings or halting an action, prompting the system to output "Hello" or "Stop." A clenched fist, often used to signify strength or readiness, translates to "Start" or "Hold." When the index finger is extended while the rest of the fingers remain folded, the system detects a pointing gesture. If the finger points to the right, the system registers it as "Next" or "Move Right," while pointing left corresponds to "Previous" or "Move Left."

A raised hand, often used to request attention or assistance, is mapped to "I need help." These predefined gestures create an intuitive way to interact with digital environments, making them useful in accessibility applications, gaming, and virtual system control. The system can also be expanded to recognize more complex gestures, including sign language, for enhanced communication support.

Fig 3: Output Image, Landmarks extracted by the Mediapipe

V. CONCLUSION

The real-time hand gesture translation system successfully enhances human-computer interaction by recognizing and converting hand gestures into meaningful text outputs. Using OpenCV and MediaPipe for real-time hand tracking, the system achieves high accuracy and responsiveness, making it suitable for applications in accessibility, gaming, and virtual control.

Experimental results show that the system performs well under optimal conditions, with an average recognition accuracy of 92% and a low latency of 30-50 milliseconds per frame. Simple gestures such as thumbs up, thumbs down, and open palm were detected with high precision, while more complex gestures required additional refinement to improve robustness.

Despite its efficiency, the system faces challenges in low-light environments, background clutter, and varying hand orientations. Future improvements can include deep learning-based classification models, adaptive preprocessing techniques, and dynamic gesture recognition to enhance accuracy and versatility.
Overall, this system provides an intuitive and efficient way for users to interact with digital interfaces using hand gestures, paving the way for advancements in assistive technology, sign language translation, and gesture-based command execution.

REFERENCES

1. Wang, F., Wang, J., Tang, C., Li, P. and Liu, H., 2023. Hand gesture-based human-computer interaction: A deep learning approach. *IEEE Transactions on Human-Machine Systems*, **53**(1), pp.45-57.
2. Wachs, J.P., Kölsch, M., Stern, H. and Edan, Y., 2011. Vision-based hand-gesture applications. *Communications of the ACM*, **54**(2), pp.60-71.
3. Mitra, S. and Acharya, T., 2007. Gesture recognition: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, **37**(3), pp.311-324.
4. Chen, C., Fu, C. and Huang, Y., 2013. Hand gesture recognition using a real-time tracking method and hidden Markov models. *Image and Vision Computing*, **31**(10), pp.745-757.
5. Zhang, X., Lu, Y., Chen, T., Wang, S. and Li, H., 2022. Enhancing real-time sign language recognition with MediaPipe and LSTM networks. *Pattern Recognition Letters*, 162, pp.24-31.
6. Pavlovic, V.I., Sharma, R. and Huang, T.S., 1997. Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**(7), pp.677-695.
7. Wu, Y. and Huang, T.S., 2001. Vision-based gesture recognition: A review. In *Gesture-Based Communication in Human-Computer Interaction* (pp. 103-115). Springer, Berlin, Heidelberg.
8. Erol, A., Bebis, G., Nicolescu, M., Boyle, R.D. and Twombly, X., 2007. Vision-based hand pose estimation: A review. *Computer Vision and Image Understanding*, **108**(1-2), pp.52-73.
9. Rautaray, S.S. and Agrawal, A., 2015. Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review*, **43**, pp.1-54.
10. Ren, Z., Yuan, J., Meng, J. and Zhang, Z., 2013. Robust part-based hand gesture recognition using Kinect sensor. *IEEE Transactions on Multimedia*, **15**(5), pp.1110-1120.
11. Kumar, A., & Sharma, P. (2023). Vision-based hand gesture recognition for interactive systems. *Journal of Visual Communication and Image Representation, 87*, 102123.
12. Bora, J., Dehingia, S., Boruah, A., Chetia, A.A. and Gogoi, D., 2023. Real-time assamese sign language recognition using mediapipe and deep learning. Procedia Computer Science, 218, pp.1384-1393.
13. Wang, J., Ivrissimtzis, I., Li, Z., Zhou, Y. and Shi, L., 2023, May. User-Defined Hand Gesture Interface to Improve User Experience of Learning American Sign Language. In International Conference on Intelligent Tutoring Systems (pp. 479-490). Cham: Springer Nature Switzerland.
14. Shilaskar, S., Bhatlawande, S., Singh, A., Andhare, N., Singh, A., Kumar, A. and Madake, J., 2023, November. Artificial intelligence based real time deciphering of British sign language. In AIP Conference Proceedings (Vol. 2851, No. 1). AIP Publishing.
15. Bahia, N.K. and Rani, R., 2023. Multi-level Taxonomy Review for Sign Language Recognition: Emphasis on Indian Sign Language. ACM Transactions on Asian and Low-Resource Language Information Processing, 22(1), pp.1-39.
16. Salauddin, M., Rana, S.S., Sharifuzzaman, M., Song, H.S., Reza, M.S., Jeong, S.H. and Park, J.Y., 2023. Highly Electronegative V2CTx/Silicone Nanocomposite-Based Serpentine Triboelectric Nanogenerator for Wearable Self-Powered Sensors and Sign Language Interpretation. Advanced Energy Materials, 13(10), p.2203812.