

# Hand Sign Detection Using Deep Learning

**R. Arun Deepika**

*PG Scholar, Department of Computer Science and Engineering  
Mohamed Sathak Engineering College, Kilakarai, Tamil Nadu, India*

**S. Ummul Hyrul Fathima**

*Assistant Professor, Department of Computer Science and Engineering  
Mohamed Sathak Engineering College, Kilakarai, Tamil Nadu, India*

**K. Seeni Pulavar Pitchai**

*Assistant Professor, Department of Computer Science and Engineering  
Mohamed Sathak Engineering College, Kilakarai, Tamil Nadu, India*

## Abstract

The hand sign detection is an artificial intelligence system that detects and interprets sign language. Computer vision-based machine learning-based gestures. The system captures hand language gestures using computer vision and machine learning techniques. The system captures hand processes the visual data with deep learning models like and moves through cameras. Models such as Convolutional Neural Networks (CNN) and other AI algorithms. The core idea of this project is to close the gap in communication between deaf or mute people, and non-signers, by turning their hand gestures into text or speech in real-time. The system enhances access and inclusion in education, health care, and customer services. The proposed system incorporates gesture recognition, computer vision, and machine learning algorithms to make sure that the sign language is accurately detected and efficiently. This technology supports real-time communication and promotes equal opportunities for hearing-impaired individuals. The main objective of this project is to bridge the communication gap between deaf or mute individuals and non-signers by converting hand gestures into text or speech in real time. The system improves accessibility and inclusivity in areas such as education, healthcare, and customer service. By integrating gesture recognition, computer vision, and machine learning algorithms, the proposed system ensures accurate and efficient sign language detection. This technology supports real-time communication and promotes equal opportunities for hearing-impaired individuals.

**Keywords:** Hand Sign Detection, Convolutional Neural Network (CNN), OpenCV, Computer Vision and Machine Learning, Recurrent Neural Network (RNN).

## OPEN ACCESS

Volume: 1

Issue: 2

January 2026 to June 2026

E-ISSN: 3108-3420



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

## Introduction

The development of communication technologies has proceeded very quickly, yet there are still important problems of individuals with hearing and speech impairments in their daily interactions. Sign language is among the key channels of communication of such people, but it is not common knowledge among the general population. This generates a communication barrier which in many cases results in social isolation and inaccessibility

of services. Technological communication has improved at a very fast rate, yet people with hearing and speech difficulties continue to experience serious difficulties in all daily interactions. One of the main ways of communication of such people is Sign Language, which is not popularly known by the general population. This causes a gap in communication that in most cases translates to social isolation and inaccessibility to services.

To resolve this problem, the creation of smarter systems, including Hand Sign Detection with the help of Artificial Intelligence (AI) and Machine Learning (ML) has become relevant. THEIR objective is to identify hand gestures and translate them into either text or speech. With the help of such techniques as the computer vision, the convolutional neuron networks (CNN) and deep learning, one can create an effective and precise sign recognition models.

The current project is dedicated to the design and the implementation of a Hand Sign Detection system, able to read gestures and give meaningful output. The system works through a webcam or camera as input to get the hand motions and process them through trained models. The Idea is to develop a user-friendly and effective solution that will bridge the communication gap between the users of the sign language and users of non-sign language..

### **Problem Statement**

Communication is an essential aspect of human interaction. Nonetheless, people who have hearing and speech impairments use Sign language as the major means of communication and it is not well understood by the rest of the population. This poses a great communication gap between users and non-users of Sign language, resulting in challenges in everyday activities, education and social integration

Current approaches to Sign language recognition frequently require specialized hardware like data gloves or sensors, which are expensive, cumbersome and not convenient enough to use in everyday life. There are also other challenges associated with Vision-based systems that are based on conventional image processing methods, including sensitivity to lighting situations, intricate backgrounds, and variations in hand movements.

Beyond this, most of the current systems have constraints in their capability to recognize gestures accurately in real-time and tend to fail with dynamic gestures or with varying user hand shapes and orientations. The Absence of big and diversified data further diminishes the efficacy and generalization of such systems.

Hence, it is necessary to create an effective, cost-efficient, and real-time hand sign detection system based on the latest technologies, including computer vision and deep learning. The system should be able

to accurately detect hand gestures and translate them into readable text or speech, which will close the communication gap and enhance accessibility by hearing and speech-impaired people.

### **Objectives of the Study**

The key aims of this study are:

- To gather Rough Visual Data of the HAND SIGN.
- To Ease the image to a more convenient and more accurate analysis.
- To identify and locate the hand(s) in the input image or video frame.
- To Extract important features of the hand region to facilitate the classification of gestures.
- To Categorise the Gesture, using extracted features.
- To give me meaningful feedback based on gestures detected.

### **Literature Review**

#### **Existing Work**

Sign Language recognition has been a vibrant research field over the past few decades. Primarily, systems created in the early days were mainly based on hardware-oriented solutions, including data gloves and sensors to record movements of the hands. Precise, but still cumbersome and expensive, These systems were able to detect finger bending and hand orientation, but this required users to wear special equipment.

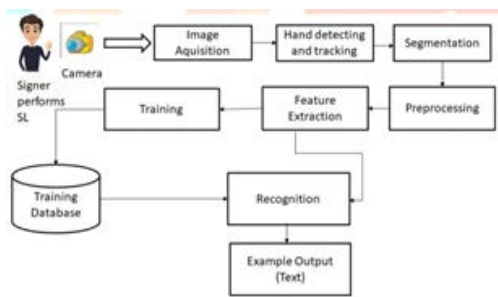
The Hidden Markov Model (HMM) was used by researchers such as Starner and Pentland (1995) to recognize the signs of the Sign language. Their system used wearable sensors in conjunction with input with cameras to recognize dynamic gestures in real time. Despite being an innovative method, it needed controlled environments and dedicated hardware..

Subsequently, VISION-Based Systems appeared, as an alternative to Hardware-Dependent methods. These systems involved cameras to record hand gestureures and implement image processing methods like segmentation, edge detection, and feature extraction. Classification was done using traditional machine learning algorithms such as the Support Vectors Machines (SVM) and the K-Nearest

Neighbors (KNN). These methods minimized the need to rely on hardware, but had problems with lighting, background, and hand shape variations..

As deep learning progressed, the widespread use of convolutional neural networks (CNNs) in gesture recognition based on images was achieved. CNN MODELS were able to increase accuracy significantly because features are automatically extracted on images. Researchers also came to co-train CNN with recurrent neuron net (RNN) and long short-term memory (LSTM) networks to identify dynamic gestures and sequences.

Recent systems are concerned with real-time sign language recognition with the use of WEBCAMS and mobile devices. The purpose of these systems is to translate hand gestures into text or speech which will make communication more user-friendly to individuals with hearing impairments or speech problems. Some Applications, as well, combine Natural Linguistic Processing (NLP) to enhance Sentence Construction.



**Figure 1 Existing Architecture**

### Related Work

A number of researchers have been involved in the development of SILRS through a variety of methods.

One of the first real-time sign recognition systems based on Hidden Markov models was introduced by Starner and Pentland (1995). Theirs was more of a strategy to identify continuous gestures, but it needed devices that could be worn and controlled environments.

A detailed survey of Sign Langore cognition systems was aided by Ong and Ranganath (2005). They divided the methods into vision-based and sensor-based methods and discussed their benefits and drawbacks.

Koller et al. (2015) suggests to apply deep learning techniques to continuous recognition of sign language. Their Work used CNNs in combination with sequence models to enhance the recognition accuracy of dynamic gestures.

A CNN-based system of gesture recognition based on large datasets was developed by Pigou et al. (2018). Their model had better performance in the recognition of stational hand signs in motion but had to be highly computational.

Molchanov et al. (2016) investigated the application of 3D convolutional neuronal networks to gesture recognition. THEIR Methodology IS BOTH Spatial and Temporal and therefore it is appropriate to dynamic gestures.

More Recent Work revolves around the use of lightweight models which can be operated in mobile devices. Researchers are also looking into application of transfer learning and pre-trained models to both save training time and enhance accuracy.

Furthermore, there are studies that have incorporated a sign language recognition system with a speech synthesist system to offer voice output in real-time. This improves the communication between the users and the non-users of sign languages.

### Limitations of Existing Systems

Nevertheless, even with the considerable progress, the current Sign language recognition systems still have a number of problems.

Reliance on Controlled Environments is one of the significant constraints. Many Systems work well in ideal lighting conditions and simple backdrops but fail to work in real scenarios where there is variable lighting and varying backgrounds.

The other difficulty is the difference in Hands gesture among various users. Variations in hand size, shape and style of movement can influence the accuracy of recognition. The vast majority of systems are trained on small datasets, and this decreases their generalization capabilities.

Hardware-Basd Systems, though precise, cannot be considered user-friendly because they require wearable sensors like gloves and sensors. These systems are also costly, and cannot be utilized on a daily basis.

The vision-based systems encounter challenges in accurately detecting the hand regions and this is particularly hard when the background is intricate or when there are multiple objects in the scene. There is also the Occlusion (When a part of the hand blocks another) which influences performance.

Another tricky issue is Dynamic Gesture Recognition. Although the recognition of static gestures can be relatively easily achieved, continual signs language involves sequences of gestures, which cannot be recognized without more sophisticated models, such as LSTM or transformers. Even at this time, it is difficult to attain high accuracy.

The other limitation is that it does not have standard data sets. Various researchers apply different datasets, and it is challenging to compare outcomes and enhance models on a comparable basis..

Performance is also a concern as it can be real-time. Deep Learning models are characterized by the need to have high computational resources, which may restrain their application in low-end devices, including smartphones.

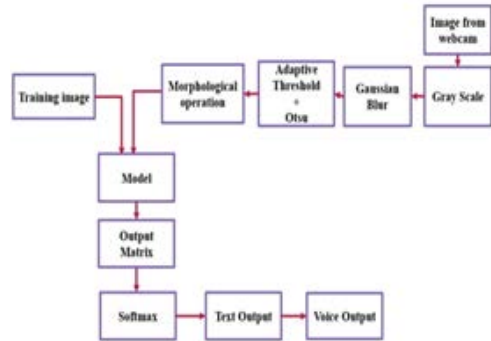
Also, the majority of systems typically only recognize gestures, and do not take facial expressions or body language, which are also significant aspects of sign language.

**Proposed System**

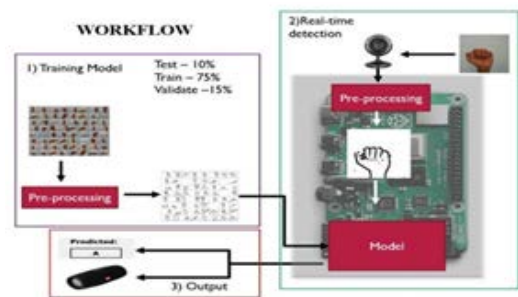
The System Proposed proposes a system that builds a real-time hand sign detection system based on computer vision and Deep learning. IT is able to identify the use of the hand gesture and translate it into a text or speech in order to minimize the communication gap among hearing and speech-impaired people. Unlike other current systems, IT does not use sensors but a webcam; hence it is cost-efficient and user-friendly. The System consists of such steps as image capture, preprocessing, hand detection, feature extraction, and classification with the help of a CNN model..

Lastly, the recognized gesture is shown in the form of text or speech. It is efficient and easy to use as it operates in real-time and manages various lighting conditions.

**System Architecture**



**Figure 2 Proposed System**



**Figure 3 Workflow Architecture**

System architecture will be comprised of the following elements:

- User Interface
- Gaussian Blur
- Morphological Operation
- Training Images
- Output Matrix

**Methodology**

- Image/Video capture: Gather visual raw information of the hand signature with a camera and equipment such as openCV and FFMPEG.
- Preprocessing: PREP the Image to be analyzed by making it grayscale (where necessary), normalizing, background-subtraction to isolate the hand, and resizing/cropping the image. Pandas and OpenCV are some of the tools used..
- Hand Detection: Detect and Locate The Hand (S) in the Image/Video using Pre-Trained Models (Mediapipe, Yolo, Opencv Haar Cascade) to Get 21 Key Landmarks.

- Feature Extraction: Extraction Key Features Of The Hand Region (21 Key Hand Landmarks Using Mediapipe), compute distances, angles and other geometrical characteristics. Tools Mediapipe, OpenCV, Tensorflow/Keras/Pytorch.
- Gesture Recognition (Classification): Classify The Gesture based on extracted Features using machine learning (SVM, random forest, k-nn using sci-kit-learn) or deep learning (CNNs, RNNS/LSTMS using tensorflow/keras/pytorch).
- Output And Visualization: GIVE Feedback based on detected gestures, e.g. outputting text or speech in a Sign language, controlling devices, or showing the recognized gesture.

### Module-description

Capture Module: OpenCV and FFMPEG are used to capture images or videos.

- Preprocessing module: use openCV and Panda to clean and prepare images.
- Detection Module: Utilizes Mediapipe, YOLO, OPENCV to detect hands.
- Feature Extraction Module: relies on MEDIapipe on landmarks, opencv on geometric features and tensorflow/keras/pytorch on deep learning based extraction.
- Classification Module: Uses Scikit-Learn as a traditional ML and Tensorflow/Keras/Pytorch as a deep learning model.
- Output Module: Processes Visualisation and Feedback based on Gestures identified.

### Implementation

The implementation would include coding these steps with the help of the specified tools and libraries.

- Capture Video and Preprocess images with OpenCV.
- Integrate Mediapipe to Detection of Hand and Landmarks in Real Time.
- Develop a Model Using Tensorflow/Keras/Pytorch For Gesture Classification Based On Extracted Features.
- Apply Application-Specific Feedback.

### Front-End Development for AI Applications

This project entailed the development of interactive front-end interfaces with python-based web application frameworks like streamlit, flask,

and pyqt5 to develop user-friendly dashboard to AI-based apps.

### Tasks Performed

In-Place Image And Video Upload options to process in the mode of deep learning model.

- Created dynamic charts and visualization to display the outcomes of AI models.
- Optimized frontend responsiveness to seamless interactions and real-time performance.
- Developed and deployed real-time AI applications with components of UI.

### Technologies Used

- Streamlit, Flask, Pyqt5 - Front-End Development.
- HTML, CSS, JAVASCRIPT - To improve the customization of UI.
- MATPLOTLIB, SEABORN, PLOTLY - in order to visualize data.

### Image and Video Processing Using Opencv

Considering that the project was related to AI-based applications, image and video processing workflows were processed with OpenCV.

### Tasks Performed

- Developed Real-Time Object Detection and Tracking Pipe-lines.
- Applied Image Preprocessing methods, e.g. resizing, denoising, edge detection.
- AI and computer vision model selection tools on developed ROI (region of interest) selection.
- Incorporate webcam and video streaming capabilities into applications.
- Adopted Frame extraction and Image normalization to input machine learner model.
- Developed and tested real time gesture recognition workflows based on computer vision techniques.

### Technologies Used

- The technologies utilized in this project include a number of software libraries and tools which facilitate image processing, machine learning, and application development.
- Image and video processing tasks, including the capture of frames in the Webcam, image scaling, filtering and the detection of hand regions are done using the openCV.

- NUMPY is applied in numerical operations, array manipulation and matrix calculations needed during image preprocessing and model input preparation.
- Tensorflow and Keras are applied to build and train deep learning models to classify hand signs and recognize gestures.
- The WEB-based user interface and real-time application dashboard to visualize the camera feed and prediction result is developed using the Streamlit.
- The WEB-based user interface and real-time application dashboard to visualize the camera feed and prediction result is developed using the STREAMLIT.
- Python is the primary programming language to be used to combine all the parts of the system.
- Also Matplotlib and Scikit-Learn are used to visualize data, preprocess data and assess model performance.

**Data Handling & Preprocessing Using Pandas and Numpy**

- Extensive work was done on data handling, manipulation, and preprocessing to make sure that AI models have high-quality inputs.
- This consisted of image resizing, image normalization, image noise reduction, and image data augmentation methods to enhance model performance and accuracy.
- Adequate preprocessing contributed to the minimization of the irrelevant background data and the concentration on the features of interest like the shape of hands and gestures patterns.
- This eventually enhanced the overall efficiency, accuracy and consistency of the HAND SIGN recognition system in real time environments.

**Tasks Performed**

- Data Cleaning And Preprocesses (Missing Value Handling, Normalization, Feature extraction).
- CSV And Excel File Handeling In Inputs of structured data..
- Live Data Processing of AI-Predictions.

**Technologies Used**

- Pandas - To handle structured data.
- Numpy -High-speed array computations.

**Output Module and Interpretation**

The Output Module Converts the raw prediction of the model into meaningful feedback that can be used by the user:

- Prediction and classification: The Predicting the last softmax layer produces a probability distribution over all possible sign classes. THE Sign Class that has the highest probability is chosen as the final prediction.
- Text Output: Showing The Recognized Sign in text (e.g. Hello or A).
- Speech Output: Vocalizing the recognized sign using a text-to-speech (TTS) engine, which is particularly useful with sign language recognition systems.
- Action Command: In Controlling applications the recognized sign is mapped on a particular action (e.g. a swipe left sign causes a slide change or scroll).
- Real-Time Feedback: Overlap the predicted sign and/or detected hand binding box onto the real-time video feed to provide the user with immediate feedback.

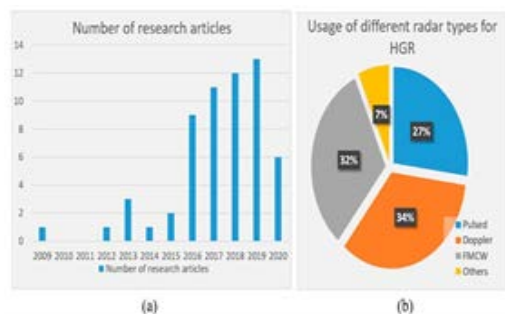


Figure 4 EDA

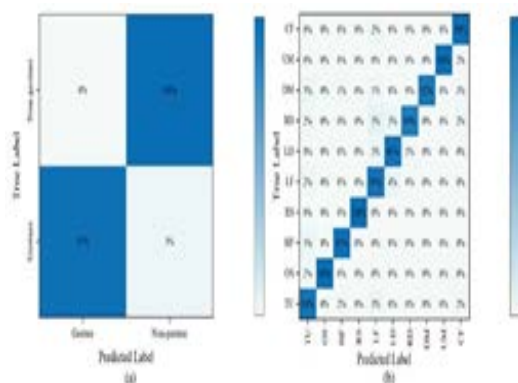


Figure 5 Confusion Matrix

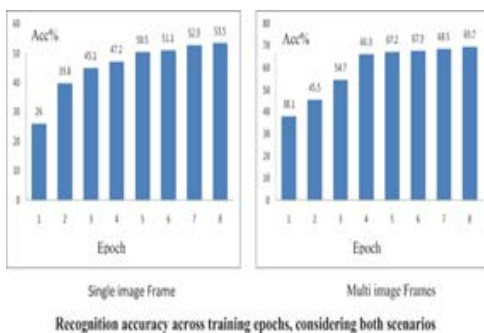
## Results and Discussion

The Hand Gesture recognition system was deployed according to the six-step pipeline: Image/Video Capture, Prepro processing, Hand Detection, Feature Extraction, Gesture Recognition (Classification) and Output and Visualization. THE System Making use of tools such as OpenCV to capture and preprocess, Mediapipe to detect hands and extract features (21 key landmarks), and machine learning/deep learning classifiers (SVM, random forest, k-nn via scikit-learn; cnns, rnns/LSTMS via tensorflow/keras/pytorch) to classify. The evaluation of the model performance is typically part of the discussion and is based on the metrics calculated based on the confusion matrix (ACCURACY, PRECISION, RECALL, F1-SCORE). The results show the effectiveness of the system in identifying hand gestures to be used in applications such as a Sign language translator or controlling a device.

### Result Observed

The System obtained a level of accuracy in identifying gestures depending on the extracted features.

- Specific gestures were accurately categorized with high level of precision and recall.
- Misclassifications Wer Werd Between Similar gestures (E.g., Some Hand poses could be confused with others).
- The confusion matrix was used to calculate performance metrics (like accuracy).

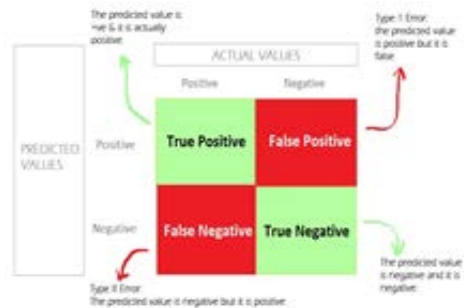


**Figure 6 Matching Performance Graph**

## Key Observations

The Model Achieved a high true positive rate and a high consistency of predictive reliability, which affirms the strength of the deep learning model.

- Hi True Positive Value
- Low False Negative Value
- Reliable Prediction Consistency



**Figure 7 Confusion Matrix of Hand Detection Model**

## Accessibility Effectiveness Analysis

USER testing was carried out to individuals who had varying DISABILITIES To assess the usability of the system.

User type	Accessibility Feature	Outcome
Blind users	Text-To-Speech	High us ability
Deaf users	Speech-To-Text	Accurate transcription
Mute users	Gesture recognition	Effective communication

The Finding confirms that the system has a significant positive impact on the accessibility and independent use.

## Future Work

The existing system can be improved in a number of ways to provide it with better accuracy, efficiency, and usability. Enhancing the size of the dataset to capture a larger range of variations of the HGs is one of the key aspects that they can do to improve the model in the real-world context and make it better generalise.

Future Work can also be devoted to supporting dynamically gesture (continuous sentence of Sign language) rather than simply some static use of hands. This will bring the system closer to real

human communicating and make it more practical. It can be assisted by Integrating with Natural language Processing (NLP) to construct meaningful sentences out of the recognized gestures.

Another Enhancement is to streamline the system to work with mobile devices to enable users to access it conveniently using smartphones. It can also be expanded to multilingual support to make its application in various regions and languages possible.

In addition, voice output and text-to-speech features will be implemented, thus improving accessibility. Broadening the use to IOT devices and smart assistants can also be integrated with the system.

This project can be developed in the future into a full-fledged communication tool that will help close the divide between the users of Sign Language and the general population and make the process of communication more inclusive and effective.

## Conclusion

In conclusion, the Hand Sign Detection System is an effective solution that can be used to overcome the communication barrier that hearing and speech-impaired individuals experience. The system can identify gestures made by hands and frame them into intelligible text or speech with a combination of computer vision and machine learning algorithms.

The Project shows the way in which current technologies can be applied to develop inclusive solutions to real-world issues. BETter accuracy in gesture recognition is provided by the implementation of deep learning models like CNN, and real-time processing makes them more usable.

Despite the satisfactory performance of the system, it still has some limitation which includes sensitivity to the lighting conditions, noise in the background, and differences in hand gestures. Despite these struggles, the SYSTEM provides a good foundation on which more improvements and advancements can be made..

On the whole, this project helps develop Assistive Technologies and enhances the availability of better communication. IT underscores the significance of applying AI to social good and enhancing individual with disability quality of life.

## References

- Arachchi, S. K., Hakim, N. L., Hsu, H.-H., Klimenko, S. V., & Shih, T. K. (2018). Real-time static and dynamic gesture recognition using mixed space features for 3D virtual world interactions. In *Proceedings of the International Conference on Advanced Information Networking and Applications Workshops*. pp. 627-632.
- Chen, Q., Georganas, N. D., & Petriu, E. M. (2007). Real-time vision-based hand gesture recognition using Haar-like features. In *Proceedings of the IEEE Instrumentation & Measurement Technology Conference*, pp. 1-6.
- Cheng, H., Yang, L., & Liu, Z. (2015). Survey on 3D hand gesture recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 26, pp. 1659-1673.
- Cui, R., Liu, H., & Zhang, C. (2017). Recurrent convolutional neural networks for continuous sign language recognition by staged optimization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 7361-7369.
- Dawar, N., & Kehtarnavaz, N. (2018). Real-time continuous detection and recognition of subject-specific smart TV gestures via fusion of depth and inertial sensing. *IEEE Access*, 6, pp. 7019-7028.
- Dong, H., Danesh, A., Figueroa, N., & El Saddik, A. (2015). An elicitation study on gesture preferences and memorability toward a practical hand gesture vocabulary for smart televisions. *IEEE Access*, 3, pp. 543-555.
- Escobedo, E., & Camara, G. (2016). A new approach for dynamic gesture recognition using skeleton trajectory representation and histograms of cumulative magnitudes. In *Proceedings of the SIBGRAPI Conference*, pp. 209-216.
- Feichtenhofer, C., Pinz, A., & Wildes, R. (2016). Spatiotemporal residual networks for video action recognition. In *Advances in Neural Information Processing Systems*, pp. 3468-3476.
- Georgi, M., Amma, C., & Schultz, T. (2015). Recognizing hand and finger gestures with IMU-based motion and EMG-based muscle

- activity sensing. In *Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies*, pp. 99-108.
- Grant, H., & Lai, C.K. (1998). Simulation modeling with artificial reality technology (SMART): An integration of virtual reality and simulation modeling. In *Proceedings of the 1998 Winter Simulation Conference*, pp. 437-441.
- Guo, Z. (2011). Research of hand positioning and gesture recognition based on binocular vision. In *Proceedings of the IEEE International Symposium on VR Innovation*, pp. 311-315.
- Hakim, N. L., Sun, S.-W., Hsu, M.-H., Shih, T. K., & Wu, S.-J. (2019). Virtual guitar: Using real-time finger tracking for musical instruments. *International Journal of Computer Science and Engineering*, 18, pp. 438-450.
- Huang, J., Zhou, W., Li, H., & Li, W. (2015). Sign language recognition using 3D convolutional neural networks. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1-6.
- Kılıboz, N. Ç., & Güdükbay, U. (2015). A hand gesture recognition technique for human-computer interaction. *Journal of Visual Communication and Image Representation*, 28, 97-104.
- Lee, S.-H., Sohn, M.-K., Kim, D.-J., Kim, B., & Kim, H. (2013). Smart TV interaction system using face and hand gesture recognition. In *Proceedings of the IEEE International Conference on Consumer Electronics (ICCE)*, pp. 173-174.
- Li, Y. (2012). Hand gesture recognition using Kinect. In *Proceedings of the IEEE International Conference on Computer Science and Automation Engineering*, pp. 196-199.
- Liu, H., & Wang, L. (2018). Gesture recognition for human-robot collaboration: A review. *International Journal of Industrial Ergonomics*, 68, pp. 355-367.
- Melax, S., Keselman, L., & Orsten, S. (2013). Dynamics based 3D skeletal hand tracking. In *Graphics Interface Conference*, pp. 63-70.
- Molchanov, P., Gupta, S., Kim, K., & Kautz, J. (2015). Hand gesture recognition with 3D convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1-7.
- Mummadi, C., Leo, F., Verma, K., Kasireddy, S., Scholl, P., Kempfle, J., & Laerhoven, K. (2018). Real-time and embedded detection of hand gestures with an IMU-based glove. *Informatics*, 5, Article 28.
- Oyedotun, O. K., & Khashman, A. (2017). Deep learning in vision-based static hand gesture recognition. *Neural Computing and Applications*, 28, pp. 3941-3951.
- Plouffe, G., & Cretu, A.-M. (2015). Static and dynamic hand gesture recognition in depth data using dynamic time warping. *IEEE Transactions on Instrumentation and Measurement*, 65, pp. 305-316.
- Ren, Z., Meng, J., & Yuan, J. (2011). Depth camera-based hand gesture recognition and its applications in human-computer interaction. In *Proceedings of the International Conference on Information, Communications & Signal Processing*, pp. 1-5.
- Ren, Z., Meng, J., Yuan, J., & Zhang, Z. (2011). Robust hand gesture recognition with Kinect sensor. In *Proceedings of the ACM International Conference on Multimedia*, pp. 759-760.
- Ren, Z., Yuan, J., Meng, J., & Zhang, Z. (2013). Robust part-based hand gesture recognition using Kinect sensor. *IEEE Transactions on Multimedia*, 15, pp. 1110-1120.
- Shahroudy, A., Liu, J., Ng, T.-T., & Wang, G. (2016). NTU RGB+D: A large-scale dataset for 3D human activity analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1010-1019.
- Simonyan, K., & Zisserman, A. (2014). Two-stream convolutional networks for action recognition in videos. In *Advances in Neural Information Processing Systems*, pp. 568-576.
- Sturman, D. J., & Zeltzer, D. (1994). A survey of glove-based input. *IEEE Computer Graphics and Applications*, 14, pp. 30-39
- Wachs, J. P., Kölsch, M., Stern, H., & Edan, Y. (2011). Vision-based hand-gesture applications. *Communications of the ACM*, 54, pp. 60-71.

- Wang, L., Qiao, Y., & Tang, X. (2015). Action recognition with trajectory-pooled deep-convolutional descriptors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4305-4314.
- Wang, R. Y., & Popović, J. (2009). Real-time hand-tracking with a color glove. *ACM Transactions on Graphics*, 28, Article 63.
- Wu, Y.-K., Wang, H.-C., Chang, L.-C., & Li, K.-C. (2013). Using HMMs and depth information for signer-independent sign language recognition. In *Multi-Disciplinary Trends in Artificial Intelligence*, pp. 79-86.
- Xu, D., Wu, X., Chen, Y.-L., & Xu, Y. (2015). Online dynamic gesture recognition for human-robot interaction. *Journal of Intelligent & Robotic Systems*, 77, pp. 583-596.
- Yang, C., Jang, Y., Beh, J., Han, D., & Ko, H. (2012). Gesture recognition using depth-based hand tracking for contactless controller application. In *Proceedings of the IEEE International Conference on Consumer Electronics*, pp. 297-298.
- Zhang, C., & Tian, Y. (2015). Histogram of 3D facets: A depth descriptor for human action and hand gesture recognition. *Computer Vision and Image Understanding*, 139, pp. 29-39.