# Hand Gesture Detection and Classification Accuracy Improvement in NN and Clustering Segmentation Hybrid Methods

**[1]Ratnesh Mishra ,[2] Devendra Kumar Meda**

*[1]ME Scholar, Department of Electronics and Communication, Jabalpur Engineering College, Jabalpur, M.P., India*

*[2]Assistant Professor*

*Department of Electronics and Communication, Jabalpur Engineering College, Jabalpur, M.P., India*

***Abstract*—**When it comes to human-computer interaction, gesture recognition is crucial. In hand gesture recognition, modified convolutional neural network (CNN) models are used for hand posture prediction, hand motion capture, and hand object interaction. We used these models to estimate hand poses on a data set, motion capture accessible postures dataset, and hand object interaction on a dataset. Hand motion capture accuracy is improved by analyzing such components.

***Key words:*** *Gesture Recognition, CNN, RNN, Motion Capture and Object interaction*

## I. **Introduction**

Human-computer interaction has grown to encompass the majority of modern information technology design forms. Hand motion detection and analysis is one of several human computer interface approaches. It can be used to make human-computer interaction more comfortable. The working conditions have an impact on hand motion detection and analysis in the current system. As a result, in this study, we attempt to overcome the congested environment and provide a good outcome, as well as to facilitate human-computer contact [1].

There are numerous ways available for tracking hand motion, such as Haar cascade, sensor-based, edge detection, and so on. To produce faster and more accurate results, this approach employs neural networks rather than the other methodologies stated. When the conditions are favorable to the camera, hand motion detection works properly. Environmental background, light and rotation, translation and size are all aspects that influence how smoothly the system runs. As a result, an effective real-time hand motion detection system is created, which overcomes the limitations encountered in previous systems [2].

## Ii. Research Motivation

Human hand motions provide a natural and effective nonverbal communication mechanism with the computer interface. Hand gestures are major bodily motions made with the hands, arms, or fingers. Hand gesture recognition layers range from static motions with complex backgrounds to dynamic gestures that reflect human feelings and communicate with computers or humans. The hand is used directly as an input to the machine; no intermediate medium is required for communication purposes of gesture detection. Gesture recognition is the process of determining the user's gestures through the computer. [3] Defines how a human can operate a machine using hand gestures.

## Iii. Research Contribution

A hand gesture identification system must be able to analyses and recognize the universal hand gesture in video sequences in a robust and efficient manner. The primary goal of this project is to create an algorithm that recognizes hand gestures under variable lighting and low resolution. The driver and passenger made the signals with their left or right hands. Hand or finger motion is also used to control the motions. Furthermore, the system should be designed in such a way that recognition is done accurately and with enhanced precision. The research's key contributions are divided into four segments. The subsections that follow provide a brief overview of the Research's important contributions [4].

*Problem statement*

Understanding and turning the precise meaning of deaf and dumb people's symbolic gestures into understandable language (Text).

A key deficiency in our society is a social barrier between differently abled people and abled people. Communication is one of the most important characteristics of human beings, who are considered social animals. Communication is also a significant barrier for those with hearing and vocal difficulties. This inability to communicate causes regular complications and impedes a person with hearing and vocal disabilities' daily activities. The fundamental cause of this inequality is that able-bodied people do not acquire or are not taught Sign Language, which is the primary mode of communication for people with hearing and voice difficulties. As a result, abled people are unable to converse regularly with these various segments of society.

Iv. Hand Action

- *Hand poses estimation*

The problem of estimating hand position in Human-computer Interaction can always be done on a system for detecting the presence in specific to the practical significance. This is due to additional parameters and a more prominent articulation. When talking with different objects, self-occlusions and reciprocal occlusions are frequently wide and strong [5].

- *Hand motion capture*
- *Hand object Interaction*

V. Proposed Methodology

*Segmentation techniques*

Segmentation is a critical step in Vision-based Hand Gesture Recognition for Indian Sign Language [6], since it entails splitting the input image into areas of interest (ROIs) that correspond to the hand motions, which are then converted from color to grey scale images. The segmentation method chosen will be determined by the problem's specific requirements and the qualities of the data [7-10].

Convolutional Neural Networks

Before delving deeper into those questions, let's have a look at the proposed classifier. There are two key processes in building a strong classifier in data-driven detection problems: feature extraction and classification. Handcrafted features [10], [11], linear predictive coding [12], empirical mode decomposition [13], principal component analysis [14], and other methods have been used to extract features from raw data. Many other classification algorithms, such as k-nearest-neighbors (KNN) [15], support vector machines [10], [11], dynamic Bayesian networks [7], have been used.

*Preprocessing*

Input data from the radar are collected and preprocessed before being used for network training and validation

*Feature extraction*

This stage entails obtaining hand gesture information from preprocessed frames. Gestures are typically shown in videos, therefore solving gesture recognition relies solely on feature extraction. The proposed method for hand gesture detection uses deep learning by indirectly automatic learning complex features with a convolutional neural network (CNN).

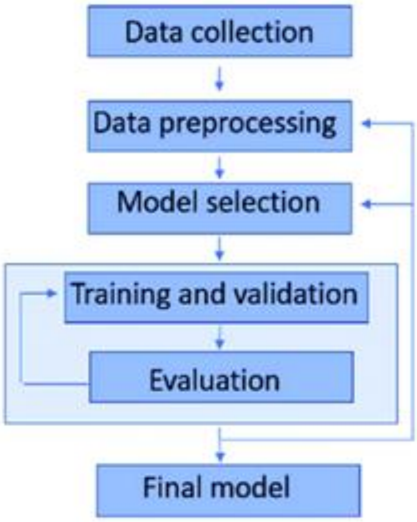


**Fig.1. Flow chart proposed methodology.**

Our suggested system, which records video and then turns it into frames, is a sign language recognition system that recognizes signals using convolutional neural networks.

### Vi. Result And Simulation

Google Colab is used for training model visualisation and debugging. Tensorflow, a machine learning framework developed by Google for designing, constructing, and training deep learning models, is also utilised in Deep CNN model training. TFLearn is a transparent and modular deep learning library built on top of Tensorflow is used to simplify and accelerate experimentation. The number of training steps is given on the xaxis. Testing and validation datasets are included in the dataset. The collection is made up of binary images of hand motions. Almost all hand gestures were recognised and navigation was successfully provided to the car simulation when testing the trained model. The model's identification of hand motions may be influenced by shadows and light source orientation.

*Validation Details*

Validation is accomplished by evaluating our model on the validation dataset using Python's model library's eval function.

*Testing Details*

The testing dataset was created by mixing the datasets and taking 10% of them from the NUS II and datasets.

(https://github.com/SiliconLabs/mltk_assets/raw/master/datasets/rock_paper_scissors_v2.7z).

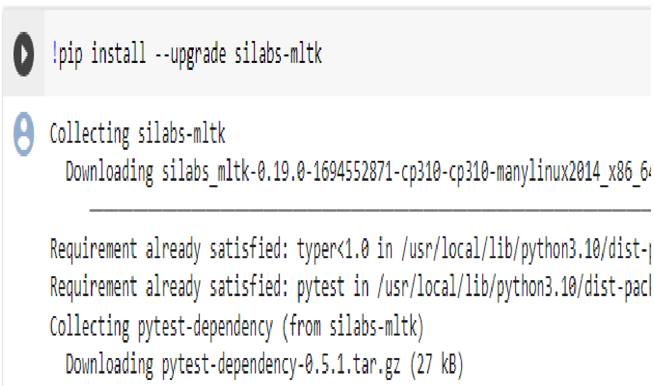


**Fig.2. PIP python.**

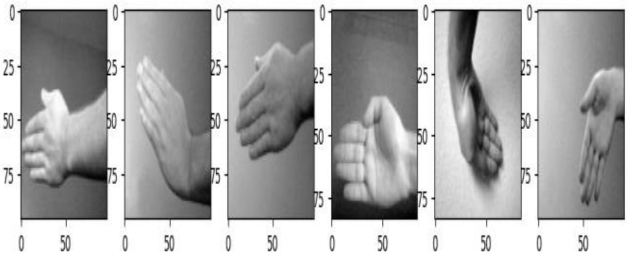

**Fig.3. PaperGesture Image**

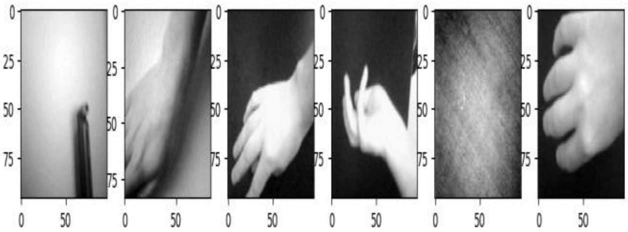

**Fig.4. UnknownGesture Image**

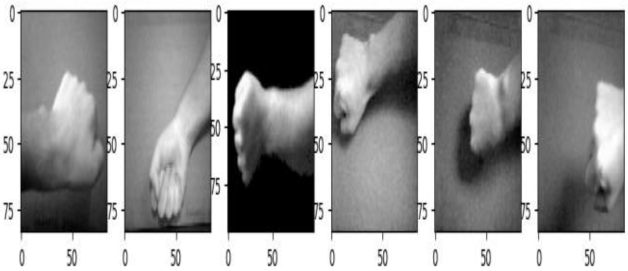

**Fig.5. Rock Gesture Image**

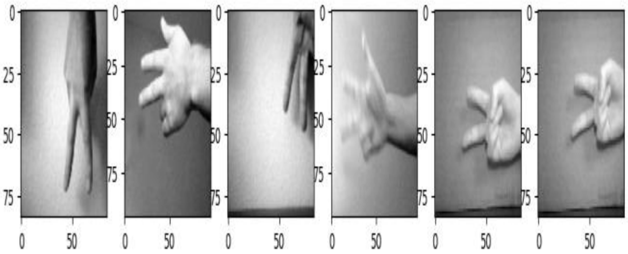

**Fig.6. ScissorGesture Image**

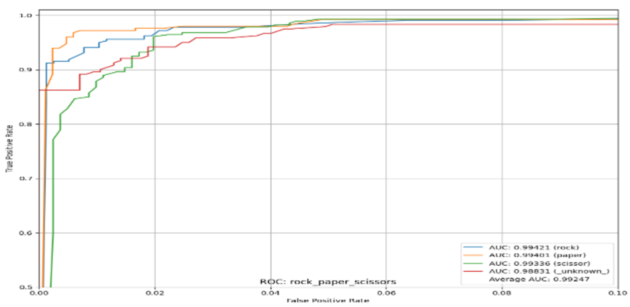


**Fig.7. ROC of all Gesture.**

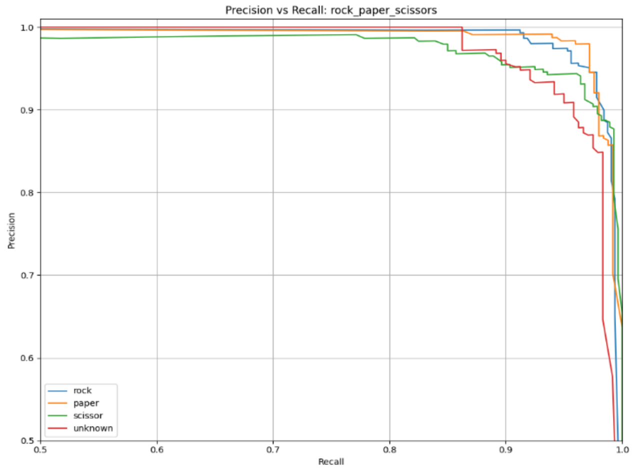


**Fig.8. Precision and recall curve.**

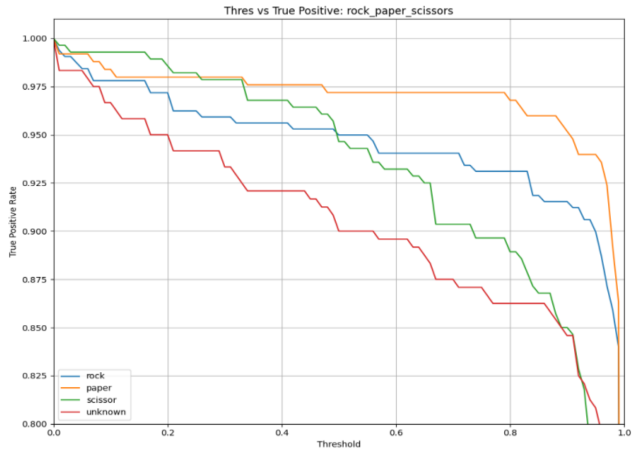


**Fig.9. True positive rate curve.**

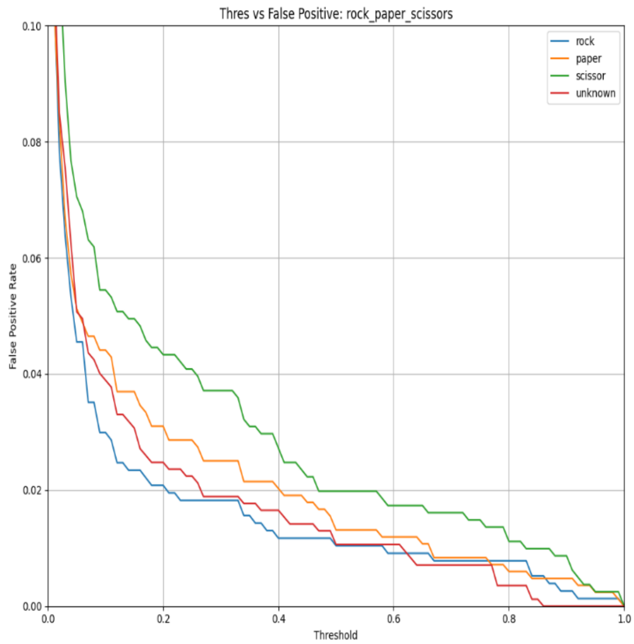


**Fig.10. False positive rate.**

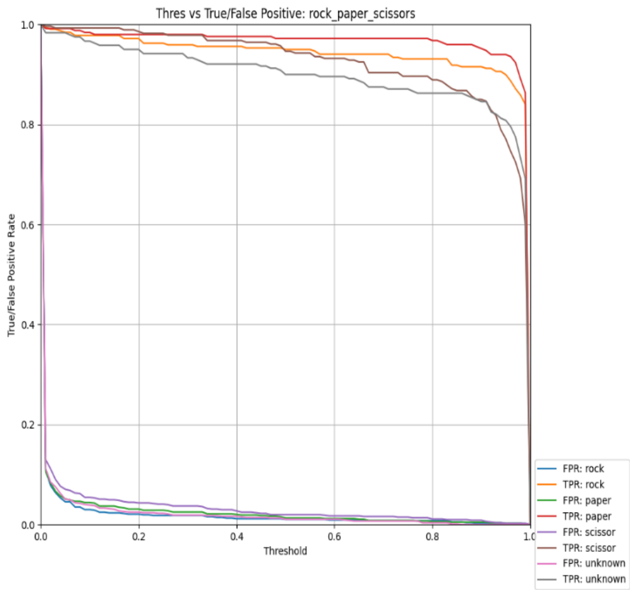


**Fig.11. ratio of True and False positive rate.**

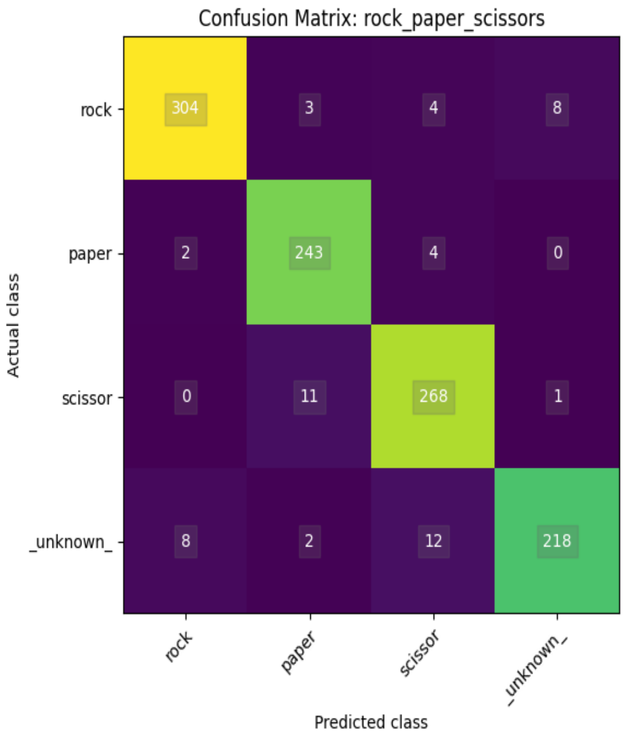


**Fig.12. Confusion matrix.**

**TABLE 1. Class Accuracies**

| **Class** | **AUC %** |
|---|---|
| paper | 97.590 |
| scissor | 95.714 |

| | |
|---|---|
| unknown | 95.298 |
| Average ROC | 99.247 |

**TABLE 2. Class ROC AUC.**

| **Class** | **AUC %** |
|---|---|
| rock | 99.421 |
| paper | 99.401 |
| scissor | 99.336 |
| unknown | 98.831 |

The graph depicts the model's training and validation losses. Both losses are reducing progressively. Although this model may not always operate completely, there is much we can do to improve its performance. We are using photos of collected hand motions to test the model. The model uses these photographs as input parameters to determine the class to which they belong. Before transmitting the image, we must confirm that the proportions are the same as they were during the training phase. The model properly predicted the following test photos. However, it is not always correct if the image contains too many other elements in the background.

**Vii. Conclusions**

The proposed hand recognition system is extremely valuable because it may be used as a human-computer interface as well as to assist persons who are paralyzed. In the future, equipment can be simply controlled by hand gestures. Sensor techniques were used in prior designs, but no more sensors are required in this system. Deep convolutional neural network algorithms can achieve this. Deep learning works with large amounts of data. The convolutional network is a multilayer algorithm broken into two parts. The first is feature extraction using a convolutional neural network, and the second is classification using fully connected layers.

The neural network is trained on a large set of data, and the aim is met. The proposed method correctly predicts the signs of sign and certain frequent words under varied lighting conditions and speeds. The photos accurately mask by offering a range of values that can dynamically recognise a human hand. CNN is used in the suggested method for image training and classification. The pictures' more informative features are finely isolated and used for categorization and training. To ensure accuracy, a total of static pictures are used in the training process for each sign.

As future work and major development prospects it is suggested:

• Investigate and compare the findings of alternative machine learning techniques applied to the problem of hand gesture classification.

• Include not only the ability to recognize 3D movements, but also the ability to operate with many cameras to gain a full 3D environment and accomplish view-independent recognition, thereby eliminating some restrictions of the current system.

**References**

[1] Alam, M. M., Islam, M. T., & Rahman, S. M. (2022). Unified learning approach for egocentric hand gesture recognition and fingertip detection. Pattern Recognition, 121, 108200.

[2] Zhang, G., Wang, L., Wang, L., & Chen, Z. (2023). Handraising gesture detection in classroom with spatial context augmentation and dilated convolution. Computers & Graphics, 110, 151-161.

[3] Caputo, A., Giachetti, A., Giannini, F., Lupinetti, K., Monti, M., Pegoraro, M., &Ranieri, A. (2020). Sfinge 3d: A novel benchmark for online detection and recognition of heterogeneous hand gestures from 3d fingers' trajectories. Computers & Graphics, 91, 232-242.

[4] Mazhar, O., Navarro, B., Ramdani, S., Passama, R., & Cherubini, A. (2019). A real-time human-robot interaction framework with robust background invariant hand gesture detection. Robotics and Computer-Integrated Manufacturing, 60, 34-48.

[5] Avola, D., Cinque, L., De Marsico, M., Fagioli, A., &Foresti, G. L. (2020). LieToMe: Preliminary study on hand gestures for deception detection via FisherLSTM. Pattern Recognition Letters, 138, 455-461.

[6] Si, J., Lin, J., Jiang, F., & Shen, R. (2019). Hand-raising gesture detection in real classrooms using improved RFCN. Neurocomputing, 359, 69-76.

[7] Salim, U. T., &Dawwd, S. A. (2019). Systolic hand gesture recognition/detection system based on FPGA with multiport BRAMs. Alexandria Engineering Journal, 58(3), 841848.

[8] Simao, M., Neto, P.&Gibaru, O. (2019). Improving novelty detection with generative adversarial networks on hand gesture data. Neurocomputing, 358, 437-445. 9. Raheja, J. L., Chaudhary, A., &Maheshwari, S. (2014). Hand gesture pointing location detection. Optik, 125(3), 993-996.

[9] Bhiri, N. M., Ameur, S., Alouani, I., Mahjoub, M. A., &Khalifa, A. B. (2023). Hand gesture recognition with focus on leap motion: An overview, real world challenges and future directions. Expert Systemswith Applications, 120125.

[10] Wang, X., & Zhu, Z. (2021). Vision–based framework for automatic interpretation of construction workers' hand gestures. Automation in Construction, 130, 103872.

[11] Anand, K., Urolagin, S., & Mishra, R. K. (2021). How does hand gestures in videos impact social media engagementInsights based on deep learning. International Journal of Information Management Data Insights, 1(2), 100036.

[12] Caputo, A., Giachetti, A., Soso, S., Pintani, D., D'Eusanio, A., Pini, S., & Tran, M. T. (2021). SHREC 2021: Skeletonbased hand gesture recognition in the wild. Computers & Graphics, 99, 201-211.

[13] Sadeddine, K., Chelali, F. Z., Djeradi, R., Djeradi, A., &Benabderrahmane, S. (2021). Recognition of userdependent and independent static hand gestures: Application to sign language. Journal of Visual Communication and Image Representation, 79, 103193.

[14] Sharma, S., & Singh, S. (2021). Vision-based hand gesture recognition using deep learning for the interpretation of sign language. Expert Systems with Applications, 182, 115657.

[15] Lazarou, M., Li, B., &Stathaki, T. (2021). A novel shape matching descriptor for real-time static hand gesture recognition. Computer Vision and Image Understanding, 210, 103241