# A Smart Vision Based Single Handed Gesture Recognition system using deep neural networks

Suguna R[1], Rupavathy N[2], Asmetha Jeyarani R[3]
{drsuguna@veltech.edu.in[1], rupavathy@veltech.edu.in[2], rasmethajeyarani@veltech.edu.in[3]}

[1]Professor, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, (Tamil Nadu), India
[2,3] Assistant Professor, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, (Tamil Nadu), India

**Abstract.**The primary and expressive mode of human communications are gestures. Human can interact with machines using body postures and finger pointing. Advancements in human-computer interaction (HCI) has presented new innovations in technology making the users to communicate with computers in an instinctual manner. Evidences clearly state that future living space will be dominated by sensor-based devices and hence an efficient human-computer interfaces are required to exchange information. Hand gesture interfaces have been employed in multiple domains and has won social acceptance. System requirements for gesture recognition vary with the intended application areas. Responsiveness, Learnability, Cost and Accuracy are major drivers for success of hand gesture recognition systems. This paper suggest a HCI design that requires no wearable markers or gloves. A noninvasive vision based framework has been suggested for human-machine interface. Deep Neural Networks have provided promising results in vision based tasks. Convolutional Neural Networks (CNN) are claimed for image recognition problems as they learn features from images gradually and automatically. An optimal CNN architecture has been proposed to recognize single handed gestures. The images of hand gestures convey a numerical representation of ten digits. Image augmentation has been performed to increase the size of training data for deep learning. Depending on application, the interpretation of gesture can be customized. The classification performance has been analyzed with metrics reported by confusion matrix. The proposed architecture performs well both in training and testing reporting the accuracy of 98.2% and 96.2% respectively. Tuning the hyper parameter has improved test accuracy.

**Keywords:** Human-Computer Interface, Gestures, Hand gesture recognition, Deep Learning, Convolutional Neural Network (CNN).

## 1 Introduction

Gestures are a form of non-verbal communication usually involving body parts. Gesture recognition plays with the goal that human gestures can be interpreted by applying mathematical algorithms. Gestures can be of different types which include hand gesture, face gesture and other body gestures. Any body part contributes to gesture, but most commonly gestures are obtained from face or hand. Nowadays, hand gesture recognition research is gaining more importance as it can interact with machines in a more intuitive way.

I3CAC 2021, June 07-08, Chennai, India

DOI 10.4108/eai.7-6-2021.2308783

Hand gesture recognition can be implemented using vision-based and sensor-based methods. In vision-based method bare hand is used as input whereas in sensor-based method data gloves or sensors are attached to detect the gestures. Vision based hand gesture recognition is classified into two types- 3D hand model-based approach which uses volumetric or skeletal model & appearance-based model which utilizes skin color images. Hand gesture recognition consists of mainly four different stages as shown in Figure 1. The first stage is image capture and pre-processing of the input images. Next stage is the image segmentation, third stage is feature extraction and the last stage is classification. Image capture is usually done using web cameras and Kinect sensor-based cameras. Mainly in pre-processing stage image enhancement and noise filtering process too will happen. Segmentation stage consists of segmenting the hand gesture from the background of the image and the region of interest is segmented as well. Different algorithms contribute in extracting the features from the image and the obtained features are further classified by using different classification algorithms.

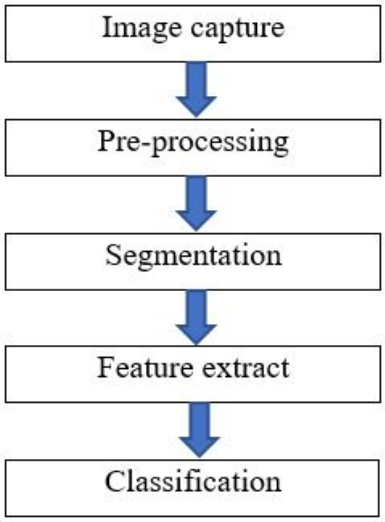


**Fig 1:** Phases in hand gesture recognition system.

Hand gesture recognition system has many applications like medical, gaming controls, home appliances, car driving, sign language representation, interacting with computer and numerous other communication purposes.

*a. Sign Language Recognition*

Sign language is the language used by deaf and dumb people. With the help of sign language, they can communicate with the outer world. American sign language (ASL), Taiwan sign language (TSL), Spanish sign language, Indian sign language, British sign language (BSL) etc. are the well-known sign languages used. Indian people have ISL language and different parts of the country it has different signs, but grammar is same throughout the country. Normal people use different sign language but differently abled persons use this sign language for communication.

*b. Gaming Control*

Gaming control uses augmented reality with 3D modeling in which 3D characteristics are controlled by hand gestures. It uses virtual controlled hand movement which provides efficient and user-friendly interface.

*c. Medical Applications*

Wearable hand gesture recognition system is used in the field of medical application and healthcare. Hand gestures are used to enable remote control of medical devices, navigation of contactless MRI and X- ray devices. In medical data visualization environment vision-based hand gesture recognition system is used to interpret user gestures in real time.

*d. Controlling home appliances*

Hand gesture recognition system is used for controlling home appliances like TV, music player etc. Television can be controlled by using hand gestures. Hand gestures are used to turn ON and OFF control of TV, increasing the volume and changing the channels. It is also used to control music player operations like rewind, fast forward etc.

*e. Robot Control*

One of the most interesting application of gestures is to control the robot action. Gestures are used to control robot by giving instructions to move forward, stop, do actions etc. Different finger authoring gestures are used for controlling them.

*f. Graphic editor control*

Graphic editor control system uses hand gesture recognition system for tracking and locating preprocessing operations for drawing and editing graphic system. For drawing the shapes like rectangle, oval, circle etc. and for drawing lines like vertical and horizontal, for editing commands like copy, delete, move, swap etc. gestures are used.

## 2 Related Works

(Ozcan&Basturk, 2019) used 165 different finger images for gesture recognition. The representation of finger print images have been performed by different classifiers like support vector machine, k-nearest neighbors, Naive-Bayes, decision tree learning, and deep neural networks. They obtained recognition rate of 96.73%, 95.77%, 61.94%, 52.73% and 98.31% respectively. (Wang et al., 2017) in his work, used Kinect- based algorithms for hand gesture analysis and evaluated CSG- EMD based hand gesture recognition system. They used KNN classification and obtained an accuracy of 99.7%.

(Jiang et al., 2020) used sensor-based hand gesture recognition system with EMG and FMG waveform and obtained accuracy of 81.5% for EMG, 80.6% for FMG and 91.6% for EMG-FMG. (Dong et al., 2018) used CNN and DCCNN for hand gesture recognition. DCCNN extracts richer features for training and obtained higher accuracy than CNN. For the number of iterations of 20000 they have obtained 71.41% accuracy for CNN and 76.38% for DCCNN. When the number of iterations is reduced from 20000 to 10000 the accuracy improved to 76.56% for CNN and 81.25% for DCCNN. (Abdal&Rasel, 2019) proposed HOG and SVM in two different datasets. For dataset 1 they obtain an accuracy of 97.5% and for dataset 2 the accuracy was 97.4%. (Han et al., 2016) proposed CNN based hand gesture recognition system. They used a dataset which consist of 76000 frames and obtained an accuracy of 95.8% and used joint tracking framerate of 11 fps. They used convolutional pose machines (CPM) hand tracking system.

(Hu & Wang, 2020) used CNN based hand gesture recognition system and obtained an accuracy of 90% when the batch size is bigger and obtained an average accuracy of 93%. They used five fully connected layers with raw data and scaled data and obtained accuracy of about 97%. (Yang et al., 2016) used hand gesture recognition for smart glass application in IoMTW and they compared it with dynamic gesture recognition system and obtained good

classification rate. (Mahmood &Abdulazeez, 2019) created a dataset by digital camera that uses threshold to extract the feature of hand gesture and applies to neural network and obtained accuracy of 90%. They enhanced the filter to extract 50 features and successfully obtained accuracy of 98%. (Alom et al., 2019) worked on ASL dataset and used deep CNN architecture. They have used combination of CNN and SVM and obtained and accuracy of 98.2%. (Kalbhor& Deshpande, 2018) proposed CNN based hand gesture recognition for two different sign languages ASL and SLD and obtained an accuracy of about 100% and 98.3% respectively. (Zhang et al., 2020) proposed hand pose classifier based on FGMM fussy gaussian mixture model and obtained 91.11% accuracy with SVM, 88.33% with MLP and 98.06% with CPM.

(Murugeswari&Veluchamy, 2015) proposed SIFT algorithm to extract key points from each hand gesture image and used different classification algorithms like HMM, ANN and SVM. They have obtained 97% accuracy with SVM, 86% with ANN and 79% with HMM model. (Nyirarugira et al., 2016) proposed particle swarm optimization method with three different classifiers HMM, LCS and PSO. Obtained 93.3% accuracy with HMM model, 94% with LCS and 94.2% with PSO. (Sahoo et al., 2019) proposed PCA based reduced deep CNN feature for static hand gesture recognition. PCA dimension reduction technique is used to reduce the redundant features in their feature vector. They used ASL digits and alphabets and obtained an accuracy of about 95% with ASL digit and 92.6% with ASL alphabets.

(Mirehi et al., 2019) used hand gesture descriptor based GNG graph method for 2170 images with 31 gestures and classified by using LDA and obtained an accuracy of about 90%. (Neethu et al., 2020) proposed CNN based hand gesture recognition system with connected component analysis algorithm and obtained an accuracy of about 96.2%. (Khan et al., 2017) proposed robust algorithm for hand gesture recognition used in electronic equipment inside vehicles and obtained detection accuracy of about 100%. (Suguna& S, 2017) used k-means clustering algorithm to classify hand gesture recognition by extracting shape features.

(Han et al., 2016) used skin model and background subtraction algorithm and given this to two stage CNN classifiers. They have taken 10 gestures with 10-fold cross validation on the system, given 10000 trained images and 2000 testing images and obtained an accuracy of about 93.8%. (Bheda&Radpour, 2017) used deep learning convolutional neural network for American sign language (ASL) and obtained an accuracy of about 82.5%. (Ozcan&Basturk, 2019) used CNN to recognize human actions and tested on sign language digits for two different datasets like sign language digit dataset and Thomas Moeslund's gesture recognition dataset and obtained and accuracy of about 98.09% and 94.33% respectively. N S (Sreekanth& Narayanan, 2017) proposed convex hull algorithm for American Sign Language (ASL) with different digits from 0-9 and obtained an accuracy of 89% to 98%. (Rady et al., 2019) proposed enhanced automatic model for hand gesture recognition using CNN method. They used both depth and color information with Kinect sensor and applied to three different datasets and obtained an accuracy of 84.67%, 99.5% and 99.85%.

(Kalam et al., 2019) proposed two-layer CNN using residual learning for 7000 rotated images and obtained an accuracy of about 97.28%. (Bhavsar& Trivedi, 2017) conducted review on sign language recognition system. They considered different feature extraction methods and taken accuracy which is done on different signs, numbers, alphabets and words. (Ghosh & Ari, 2016) proposed LCS feature set with block-based feature and is applied for 24 static ASL Hand Alphabet and obtained 82% accuracy. (Sharma &Verma, 2015) used hand gesture shapes and positions recognition system and obtained an accuracy of 95.44%.

## 3 Materials and Methods

The dataset for the proposed work has been downloaded from web resource *https://github.com/ardamavi/Sign-Language-Digits-Dataset.git*. The dataset consists of 2089 images with 10 classes, each image is single handed representing a digit ranging from 0 to 9. Sample images from the dataset are shown in Figure 2.

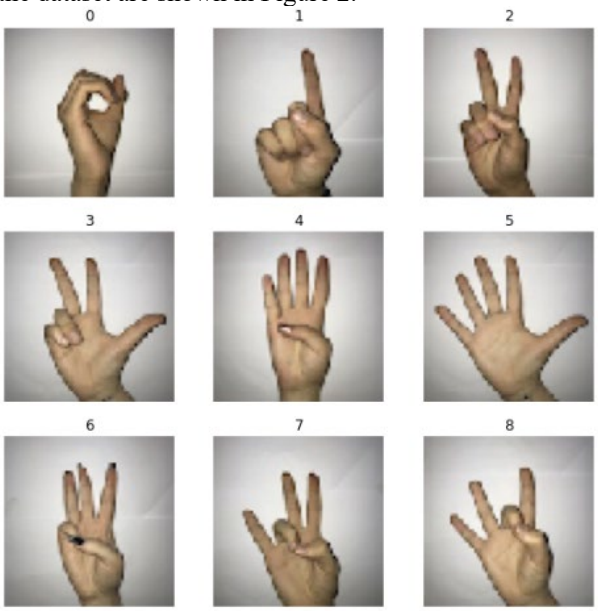


**Fig 2:** Sample hand gesture images

**Data Augmentation**

Deep Neural Networks require a huge training data to attain good output results. Image augmentation is a technique that helps in building efficient image classifier with minimum training data and enhances the performance of deep neural networks. Image augmentation artificially generates images by applying different processing techniques such as shifting, sheering, flips and rotations. Hence it increases the number of training samples to make the deep network to perform learning efficiently. The training sample size has been improved by performing data augmentation.

**Convolution Neural Network**

Computer Vision has been perceiving immense growth with the advent of deep learning. One of the algorithm, Convolutional Neural Network have been producing promising results in image based classification tasks.

Convolutional Neural Network (CNN) is deep learning algorithm that accepts image as input, learns various aspects of image through filters and able to discriminate its class with other. It is a multilayer neural network capable of analyzing visual features in the given input image. CNN is able to capture the spatial and temporal dependencies in the image through application of appropriate filters. CNN comprises of two main parts:

**Feature Learning :** Convolution tool that recognizes various features in images

**Classification / Prediction :** A fully connected layer that collects input from convolutional layer to predict image description.

The architecture of CNN is composed of different kinds of layers

**Convolutional layer:** This layer convolves the image with filters and creates a feature map by examining few pixels at a time. Convolution operation helps in extracting features such as edges, color and gradient orientation. With additional convolutional layers high level features of image are identified. The convolved output layer results in reduced dimensionality with valid padding or dimensionality is increased / retained by applying same padding.

**Pooling layer:** The role of this layer is to down sample image into a form that are easy for processing while preserving the significant information for better prediction. Pooling layer reduces the spatial size of the convolved layer. This leads to reduction in computational power to process the data. Particularly it facilitates the extraction of dominant features that are invariant to position and rotation. There are two types of pooling: Max pooling and Average pooling. Both differ in type of computation made over the portion of image prescribed by the kernel. Max pooling retrieves the maximum value of the scanned region while average pooling return the mean of the values in the region. Max pooling acts as a noise suppressor. It removes noise activations along with dimensionality reduction.

**Fully connected input layer:** This layer flattens the outputs of the preceding layer and converts them into single dimension vector.

**Fully connected layer:** This layer comprises of feed forward neural network and back propagation algorithm is employed to train the model. With relevant weights and activation function, this layer learns the nonlinear combination of high level feature represented by previous layer. Over a series of epochs, the model is able to map the input to target output.

**Fully connected output layer:** Uses soft max classification technique to generate probabilities for determining the class of the image.

**CNN Architecture**

The architecture used for hand gesture recognition is shown in Figure 3 with number of parameters used during each layer is listed in Table 1.

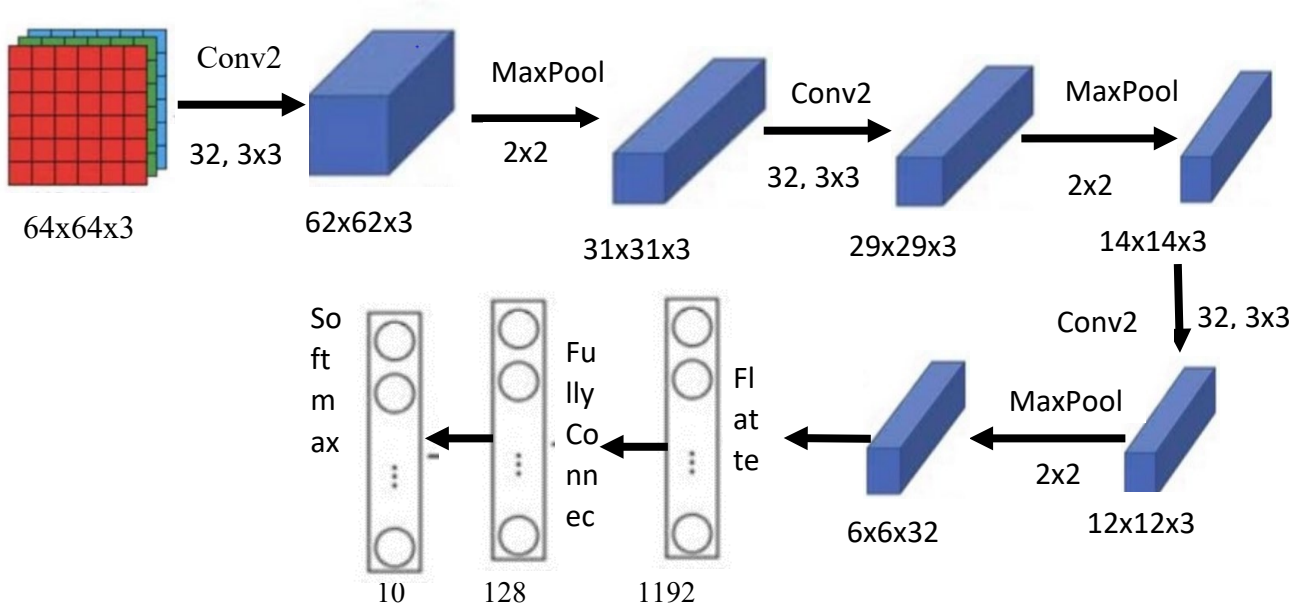


**Figure 3 :** Proposed CNN architecture for Sign Digit Prediction

**Table 1:** CNN Layers Design and parameter details

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 62, 62, 32)        896
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 31, 31, 32)        0
_________________________________________________________________
dropout (Dropout)            (None, 31, 31, 32)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 29, 29, 32)        9248
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 14, 14, 32)        0
_________________________________________________________________
dropout_1 (Dropout)          (None, 14, 14, 32)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 12, 12, 32)        9248
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 6, 6, 32)          0
_________________________________________________________________
dropout_2 (Dropout)          (None, 6, 6, 32)          0
_________________________________________________________________
flatten (Flatten)            (None, 1152)              0
_________________________________________________________________
dense (Dense)                (None, 128)               147584
_________________________________________________________________
dense_1 (Dense)              (None, 10)                1290
=================================================================
Total params: 168,266
Trainable params: 168,266
Non-trainable params: 0
```

RELU activation function was applied to every output of convolution and fully connected layer except the output soft max layer. After building the model the performance on test data was evaluated and the accuracy score of 96.17%.

## 4 Results and discussion

The classifier was trained with 100 epochs and the performance of the model with training and validation datasets are shown in Figure 4. Similarly as the epochs increased the loss has considerably reduced during training phase and validation test as seen in Figure 5.

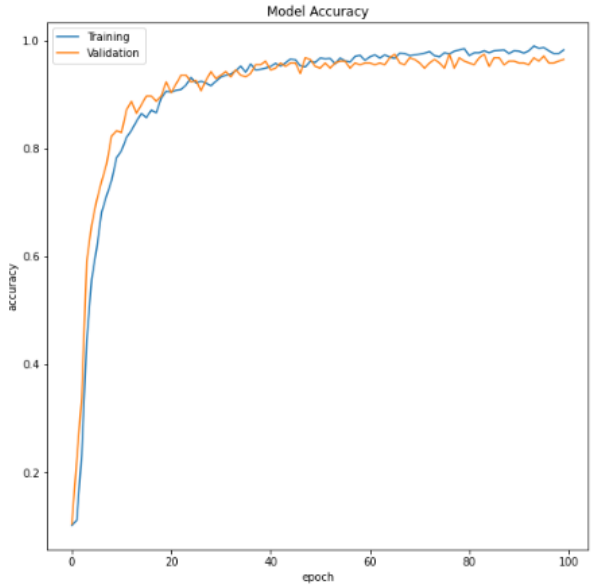


**Fig 4:** Model Accuracy on Training and Validation dataset

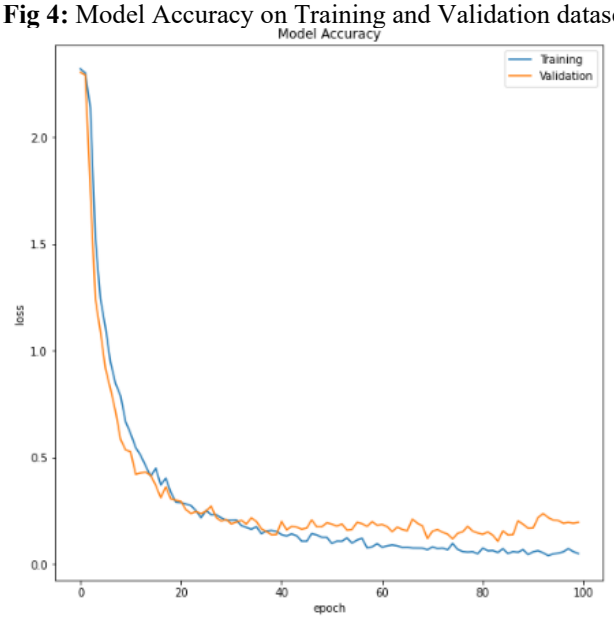


**Fig 5 :** Loss incurred during training and validation phase

To assess the prediction rate of each class confusion matrix on test data are examined and the classification reports are depicted in Table 2.

**Table 2** : Confusion Matrix showing the misclassified class in test data

| | Predicted Class | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Actual Class | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| | 0 | 31 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 1 | 0 | 31 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 0 | 2 | 28 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| | 3 | 0 | 0 | 0 | 29 | 0 | 0 | 0 | 0 | 1 | 1 |
| | 4 | 0 | 0 | 0 | 0 | 30 | 1 | 0 | 0 | 0 | 1 |
| | 5 | 0 | 0 | 0 | 0 | 1 | 31 | 0 | 0 | 0 | 0 |
| | 6 | 1 | 0 | 1 | 1 | 4 | 0 | 25 | 0 | 0 | 0 |
| | 7 | 1 | 1 | 1 | 6 | 0 | 0 | 0 | 20 | 2 | 0 |
| | 8 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 29 | 0 |
| | 9 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |

The accuracy rate of each class has been analyzed and shown in Table 3.

**Table 3:** Prediction rate of classifier on Test data

| Class | Accuracy in % |
|---|---|
| 0 | 100 |
| 1 | 100 |
| 2 | 90.3 |
| 3 | 93.75 |
| 4 | 93.75 |
| 5 | 96.9 |
| 6 | 78.1 |
| 7 | 64.5 |
| 8 | 93.5 |
| 9 | 96.77 |

Class 6 and class 7 has lower prediction rates compared to other classes. Class 6 and class 7 are erroneously recognized as class 4 and class 3 respectively. The comparison of performance metrics such as precision, recall and f1-score are shown in Figure 6.

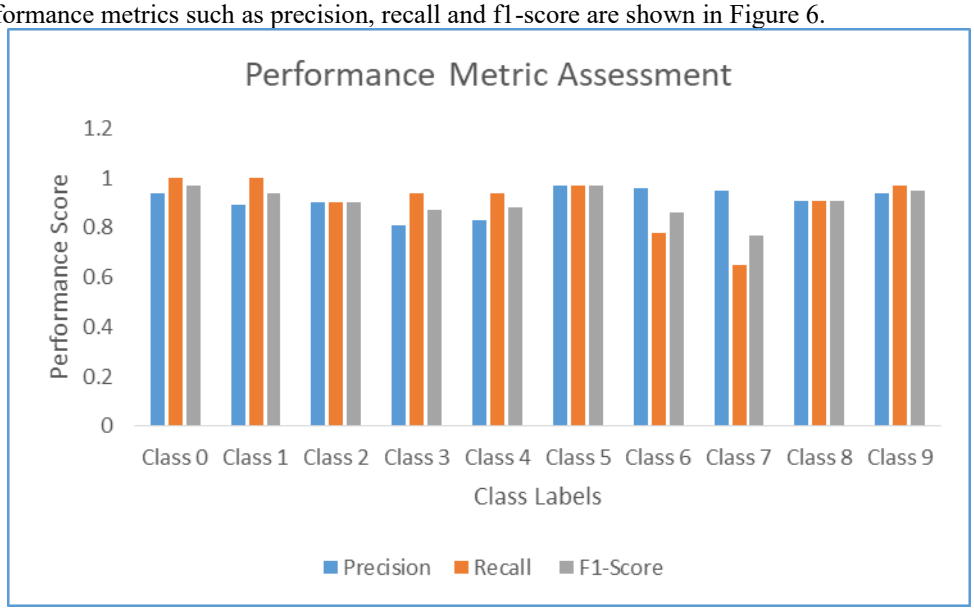


**Fig 6:** Comparison of performance metrics among classes

Sample results of test data predictions are shown in Figure 7


**Fig 7:** Sample output of test data predictions

**Tuning of Hyper parameters**

To improve the performance number of filers used in Convolution layers are increased and tested. The increase in filters have shown considerable improvements in both training and testing phases. Figure 8 shows the test accuracy with varied number of filters.

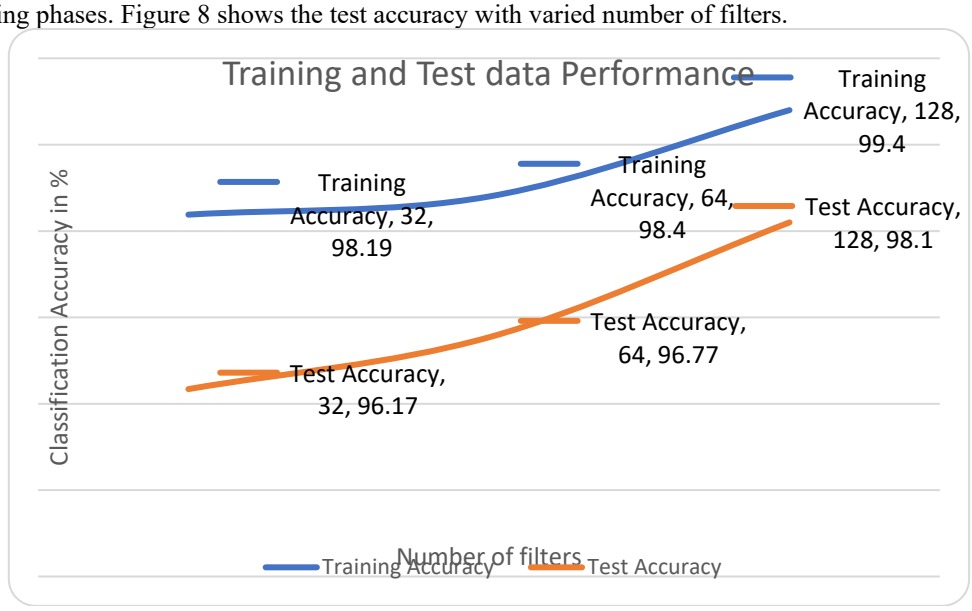


**Fig** 8: Performance Analysis with tuning hyper parameter

## Conclusion

A vision based noninvasive and cost effective hand gesture recognition system has been proposed in this work. Several research works have been carried out on different datasets. The proposed CNN architecture performs well with minimum filters and layers. The simplicity in design reduces the response time during testing. Analyzing the misclassified samples and improving the training size may enhance the recognition rate. By tuning the hyper parameter the recognition rate has improved yielding an accuracy score of 98.1% in test dataset

## References


[1] Abdal, A., &Rasel, S. (2019). An Efficient Framework for Hand Gesture Recognition based on Histogram of Oriented Gradients and Support Vector Machine. December. https://doi.org/10.5815/ijitcs.2019.12.05

[2] Alom, M. S., Hasan, M. J., & Wahid, M. F. (2019). Digit recognition in sign language based on convolutional neural network and support vector machine. 2019 International Conference on Sustainable Technologies for Industry 4.0, STI 2019, 0, 24–25. https://doi.org/10.1109/STI47673.2019.9067999

[3] Bhavsar, H., & Trivedi, J. (2017). Review on Feature Extraction methods of Image based Sign Language Recognition system. Indian Journal of Computer Science and Engineering (IJCSE), 8(3), 249–259. https://doi.org/10.1016/j.clinph.2005.11.003

[4] Bheda, V., &Radpour, D. (2017). Using Deep Convolutional Networks for Gesture Recognition in American Sign Language. http://arxiv.org/abs/1710.06836

[5] Dong, X., Xu, Y., Xu, Z., Huang, J., Lu, J., Zhang, C., & Lu, L. (2018). A static hand gesture recognition model based on the improved centroid watershed algorithm and a dual-channel CNN. ICAC 2018 - 2018 24th IEEE International Conference on Automation and Computing: Improving Productivity through Automation and Computing. https://doi.org/10.23919/IConAC.2018.8749063

[6] Ghosh, D. K., & Ari, S. (2016). On an algorithm for Vision-based hand gesture recognition. Signal, Image and Video Processing, 10(4), 655–662. https://doi.org/10.1007/s11760-015-0790-4

[7] Han, M., Chen, J., Li, L., & Chang, Y. (2016). Visual hand gesture recognition with convolution neural network. 2016 IEEE/ACIS 17th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, SNPD 2016, 287–291. https://doi.org/10.1109/SNPD.2016.7515915

[8] Hu, B., & Wang, J. (2020). Deep Learning Based Hand Gesture Recognition and UAV Flight Controls. International Journal of Automation and Computing, 17(1), 17–29. https://doi.org/10.1007/s11633-019-1194-7

[9] Jiang, S., Gao, Q., Liu, H., & Shull, P. B. (2020). A novel, co-located EMG-FMG-sensing wearable armband for hand gesture recognition. Sensors and Actuators, A: Physical, 301, 111738. https://doi.org/10.1016/j.sna.2019.111738

[10] Kalam, M. A., Mondal, M. N. I., & Ahmed, B. (2019). Rotation Independent Digit Recognition in Sign Language. 2nd International Conference on Electrical, Computer and Communication Engineering, ECCE 2019, 1–5. https://doi.org/10.1109/ECACE.2019.8679172

[11] Kalbhor, S. R., & Deshpande, A. M. (2018). Digit Recognition Using Machine Learning and Convolutional Neural Network. Proceedings of the 2nd International Conference on Trends in Electronics and Informatics, ICOEI 2018, 604–609. https://doi.org/10.1109/ICOEI.2018.8553954

[12] Khan, F., Leem, S., & Cho, S. H. (2017). Hand-based gesture recognition for vehicular applications using IR-UWB radar. Sensors (Switzerland), 17(4). https://doi.org/10.3390/s17040833

[13] Mahmood, M. R., &Abdulazeez, A. M. (2019). Different Model for Hand Gesture Recognition with a Novel Line Feature Extraction. 2019 International Conference on Advanced Science and Engineering, ICOASE 2019, 2018, 52–57 . https://doi.org/10.1109/ ICOASE.2019. 8723731

[14] Mirehi, N., Tahmasbi, M., &Targhi, A. T. (2019). Hand gesture recognition using topological features. Multimedia Tools and Applications, January. https://doi.org/10.1007/s11042-019-7269-1

[15] Murugeswari, M., &Veluchamy, S. (2015). Hand gesture recognition system for real-time application. Proceedings of 2014 IEEE International Conference on Advanced Communication, Control and Computing Technologies, ICACCCT 2014, 978, 1220–1225. https://doi.org/10.1109/ICACCCT.2014.7019293
[16] Neethu, P. S., Suguna, R., &Sathish, D. (2020). An efficient method for human hand gesture detection and recognition using deep learning convolutional neural networks. Soft Computing. https://doi.org/10.1007/s00500-020-04860-5
[17] Nyirarugira, C., Choi, H. R., & Kim, T. (2016). Hand gesture recognition using particle swarm movement. Mathematical Problems in Engineering, 2016. https://doi.org/10.1155/2016/1919824
[18] Ozcan, T., &Basturk, A. (2019). Transfer learning-based convolutional neural networks with heuristic optimization for hand gesture recognition. Neural Computing and Applications, 31(12), 8955–8970. https://doi.org/10.1007/s00521-019-04427-y
[19] Rady, M. A., Youssef, S. M., & Fayed, S. F. (2019). Smart gesture-based control in human computer interaction applications for special-need people. NILES 2019 - Novel Intelligent and Leading Emerging Sciences Conference, 1, 244–248. https://doi.org/10.1109/NILES.2019.8909324
[20] Sahoo, J. P., Ari, S., & Patra, S. K. (2019). Hand gesture recognition using PCA based deep CNN reduced features and SVM classifier. Proceedings - 2019 IEEE International Symposium on Smart Electronic Systems, ISES 2019, 221–224. https://doi.org/10.1109/iSES47678.2019.00056
[21] Sharma, R. P., &Verma, G. K. (2015). Human Computer Interaction using Hand Gesture. Procedia Computer Science, 54, 721–727. https://doi.org/10.1016/j.procs.2015.06.085
[22] Sreekanth, N. S., & Narayanan, N. K. (2017). Static hand gesture recognition using mon-vision based techniques. International Journal of Innovative Computer Science & Engineering, 4(2), 33–41. www.ijicse.in
[23] Suguna, R., & S, N. P. (2017). Hand Gesture Recognition Using Shape Features. 117(8), 51–54. https://doi.org/10.12732/ijpam.v117i8.11
[24] Wang, C., Liu, Z., Zhu, M., Zhao, J., & Chan, S. C. (2017). A hand gesture recognition system based on canonical superpixel-graph. Signal Processing: Image Communication, 58, 87–98. https://doi.org/10.1016/j.image.2017.06.015
[25] Yang, A., Park, D., Chun, S., Kim, J., Yang, A., Park, D., Chun, S., & Kim, J. (2016). Detection of Hand Gesture and its Description for Wearable Applications in IoMTW International Conference on Advanced Communication Technology, ICACT 598–601.
[26] Zhang, T., Lin, H., Ju, Z., & Yang, C. (2020). Hand Gesture Recognition in Complex Background Based on Convolutional Pose Machine and Fuzzy Gaussian Mixture Models. International Journal of Fuzzy Systems, 22(4), 1330–1341. https://doi.org/10.1007/s40815-020-00825-w