International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031
www.ijlret.com || Volume 02 - Issue 06 || June 2016 || PP. 01-08

Indian Sign Language Alpha-Numeric Character Classification using Neural Network

[1]Rajesh B. Mapari, [2]Dr. G.U.Kharat
[1]Asst. Prof, [2]Professor
[1]Anuradha Engineering College, Chikhli, India
[2]Sharadchandra Pawar College of Engineering, Otur, India

ABSTRACT: Many approaches for Sign Language recognition have been tried by researchers including Web camera to latest device like Kinect. The Indian Sign language (ISL) consists of signs performed using one hand and two hands as well. In this paper, an Indian Sign Language alpha-numeric character classification approach is presented using RGB camera. DCT of gray scale image and regional properties of black & white images are used to form feature vector of 74 values. Dataset consists of 33 signs (one handed and two handed) performed by 60 signers (students of age 20-22 years) who have given training about how to perform signs which results in total dataset of 1980 signs. Out of this, 90% dataset is used for training and 10% dataset is used for testing/Cross validation. we have trained and tested different Neural Network classifiers like MLP, GFFNN, SVM. We could get maximum classification accuracy as 86.27 % on Cross Validation dataset using MLP Neural Network.
KEYWORDS: ISL, MLP, GFFNN, SVM.

I. INTRODUCTION

Sign language is used as primary tool by deaf-mute to communicate each other. Signs can be manual or Non-manual. Manual signs consist of hand mostly. However Non-manual signs consist of hand, head and facial expressions as well. However for a common person it is very difficult to understand the meaning of signs performed by deaf people. So mostly the interpreter is assisted to acknowledge the meaning. However there are less chances that interpreter may know more than one sign language and the translator may not be available at any time and at any place. The solution to this problem is Human-Compute Interaction system which can be installed at many places like post office, Railway station, Banks etc. So that meaning of sign performed by Deaf-Mute can be understood. There are different sign languages all over the

Many Sign Languages such as American Sign Language (ASL), British Sign Language (BSL), Australian Sign Language, Indian Sign Language (ISL), Chinese Sign Language (CSL) etc. Basically Sign Language recognition approaches can be classified as instrumented glove based and vision based. In first approach equipped sensors measures information related to the shape, orientation, movement, and location of the hand. As it is based on direct coordinate values the segmentation is easily achieved which is difficult compared to bare hand segmentation in vision based system. However due to wearing of many sensors on wrist and arms, it creates difficulty for signer to perform the sign in its natural way.

In contrast, vision based system supports to both manual and non manual signs. However the segmentation is color space based which creates difficulty if background is not uniform and has matching color objects with hand color. Other issue is occlusion handling which may create when hand and face overlap or one hand overlap on another. Recently a new method which considers both the local feature and global feature of gesture is introduced using Kinect sensor. But the problem with this sensor is it’s not support minute details like shape of hand. Another sensor like leap Motion on other hand can handle minute details of hand but doesn't support for global features.

II. RELATED WORK

Most of the research work in sign language recognition system is concern to translation of sign language to text or spoken word. Some systems are as follows.

Karishma Dixit and Anand Singh Jalal [1] have worked on recognition of signs in ISL. First hand is extracted from image. Then after filtering, binary image is used to get Hu invariant moment set and structural shape descriptors which are used as a features. This feature vector is input to multi-class Support Vector Machine (MSVM) for recognition. Adithya v [2] proposed a method to recognize 26 letters of the English alphabet and the numerals from 0-9 of ISL. Database is collected in constrained background. Using distance transform feature set is formed. In [3], J. Rekha et al. have worked on recognition of 23 static and 3 dynamic signs of ISL using different classifier. Using YCbCr skin color model hand region is segmented. To obtain the shape and texture information Principal Curvature Based Region detector & 2-D Wavelet Packet Decomposition methods are used. Set of feature vector given by the PCBR, WPD-2 & finger count for each hand signs are stored in a database. Finally, the generated vector is fed into the multi-class SVM training classifier model that was built in the training stage to classify and recognize the hand gestures.

In 2015, Ankita Saxena et al.[4] presents principal component analysis which is a fast and efficient technique for recognition of Indian sign language gestures from video stream. In proposed technique 3 frames per second captured from video stream. By comparing three frames static posture is extracted. Static posture image of size 60X80 pixels matched with dataset using PCA and result is image of sign. In year 2006, M.K.Bhuyan & P.K. Bora [5] also worked on ISL recognition for Static signs, Dynamic signs & Sentences. In year 2011 offline approach to recognize sign [6], M.K. Bhuyan et al. have recognized few static hand postures (8 no's) by analyzing texture and key geometrical features of the hand. However due to recent development of inexpensive depth cameras, e.g., the Kinect sensor & Leap Motion, new opportunities opened doors for hand gesture recognition. In 2013 [8], Zhou Ren et al. have used advanced sensors like Kinect to recognize signs from 1 to 10. Using Template matching and Finger-Earth Mover's Distance, A.S.Elons et al. [9] have captured hands and fingers movements in 3D digital format using Leap motion. These temporal and spatial features are fed into a Multi-layer Perceptron Neural Network (MLP). The system was tested on 50 different dynamic signs (distinguishable without non manual features) and the recognition accuracy reached 88% for two different persons. In 2014 ,Giulio Marin et al. [10] proposed a novel hand gesture recognition scheme using Leap motion and Kinect. Feature set of leap Motion consists of Fingertips distances, Fingertips angles and Fingertips elevations. Kinect Feature set consists of Curvature, Correlation. A Multi-class SVM classifier is used for reorganization of sign.

III. EXPERIMENTAL SETUP

A. Data Collection

We have kept Black background using black cloth and Signers have worn black T-shirt while performing sign. This has helped to segment the hand easily from uniform and fixed background. For acquiring image we have used SONY camera of 16.1 mega pixel plus 5x Optical zoom.60 different signers of different age categories performed 33 signs as per the chart mentioned in Fig. 1

Fig. 1. Sample of ISL Signs

B. Feature Extraction

In first phase we read original image as shown in Fig. 2 (a) and cropped it by maintaining height width ratio of hand portion using bounding box technique with L*a*b color space as shown in Fig. 2 (b). This way hand is exactly at the center of image as shown in Fig. 2 (c). Hand image is then converted to 256×256 size RGB image.

Later on image is converted to gray scale image. The gray scale image is divided in to 32×32 block using block processing operation. 2-D DCT of each 32-by-32 block is calculated which results in 64 values.

Filtering operation is carried out by testing various filters but the best result is obtained using Gaussian Filter. Followed by smoothing operation image is converted to black and white image using gray threshold as shown in Fig. 2 (d). However to get proper black and white image to extract regional properties, it must be smooth. So series of morphological operations as shown in Fig 2 (e-i) are performed to get best result. It can be observed from Fig.2 (e) & Fig.2 (i) that jagged edges have been removed.

From the Fig. 2 (i), Regional properties like Area, MajorAxisLength, MinorAxisLength, Eccentricity, Orientation, Convex Area, EquivDiameter, Solidity, Extent & Perimeter are calculated. So feature set consists of 64 DCT values and 10 values of regional properties resulting in feature set of total 74 values.

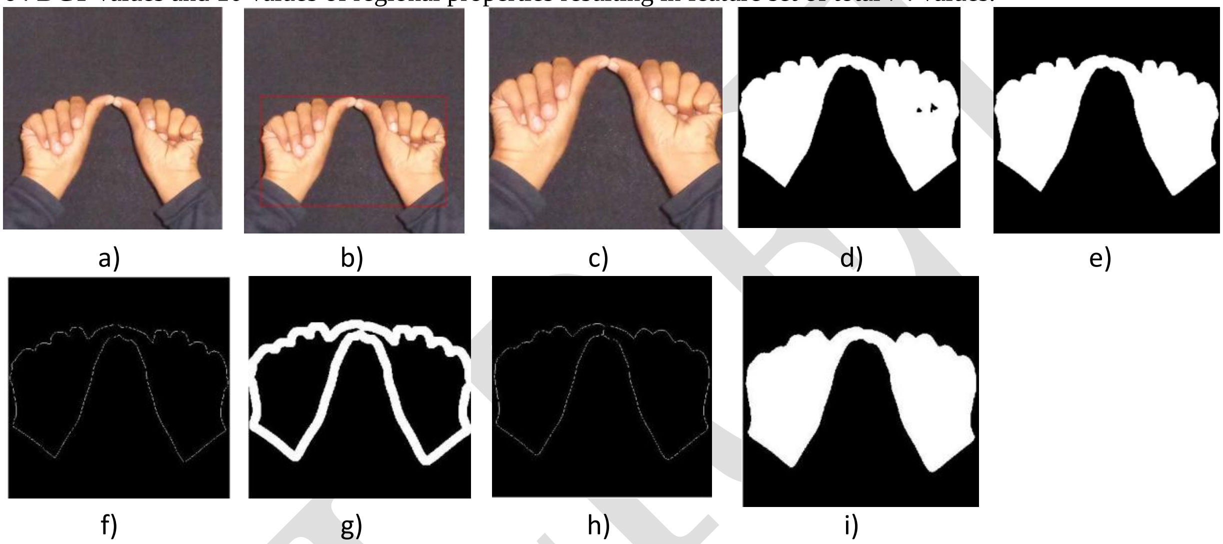

Figure 2 a) original RGB image b) bounding box c) hand at the center of image d) morphological closing operation e) holes filling operation f) morphological remove operation g) dilation operation h) thinning operation i) filling of holes

IV. CLASSIFICATION

A. Multilayer Perceptron Neural Network (MLPNN)

Following trials have been performed on Multilayer Perceptron (MLP) Neural Network to get optimal parameters for minimum MSE and maximum percentage Average Classification Accuracy. Feature vectors are divided into two parts as 90 % for training (TR) and 10% for Cross validation (CV).

By keeping only one hidden layer, first network is tested to search number of Processing Element (PE) required in Hidden Layer which gives minimum Mean Square Error (MSE) on training dataset. Fig. 3 shows that the minimum MSE for Cross Validation is given by processing element (PE) number 39.

Different transfer function like Tanh, LinearTanh, Sigmoid, LinearSigmoid, Softmax and Learning rules like Step, Momentum, Conjugate Gradient, Quick Propagation, Delta Bar Delta are varied in hidden Layer to get maximum percentage classification accuracy as shown in Fig. 4.

MLP Neural Network with the following parameter setting gives maximum Percentage classification accuracy of 94.63 % on training and 86.28 % on CV dataset. Table 1 & 2 shows confusion matrix and correct classification accuracy for CV data respectively. Tagging of Data: 90% for Training & 10% Cross validation.

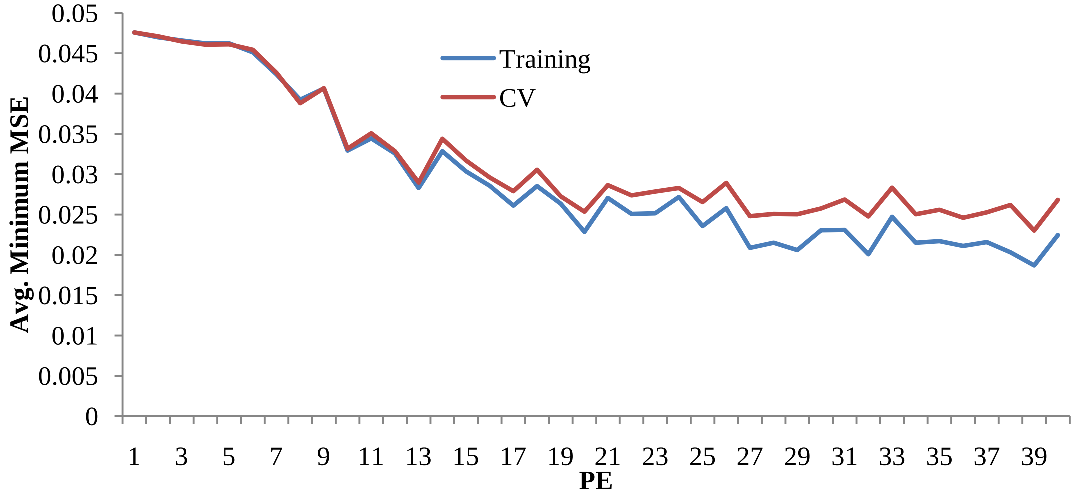

Fig. 3. Processing Element (PE) Vs Minimum MSE

Input Layer: Input PE - 74 Exemplars - 1782
Hidden Layer: PE - 39 Transfer Function - Tanh
Learning Rule - Momentum Momentum - 0.7 Step Size - 0.1
Output Layer: Output PE's - 33 Transfer Function - Tanh
Learning Rule - Momentum Momentum - 0.7 Step Size - 0.1

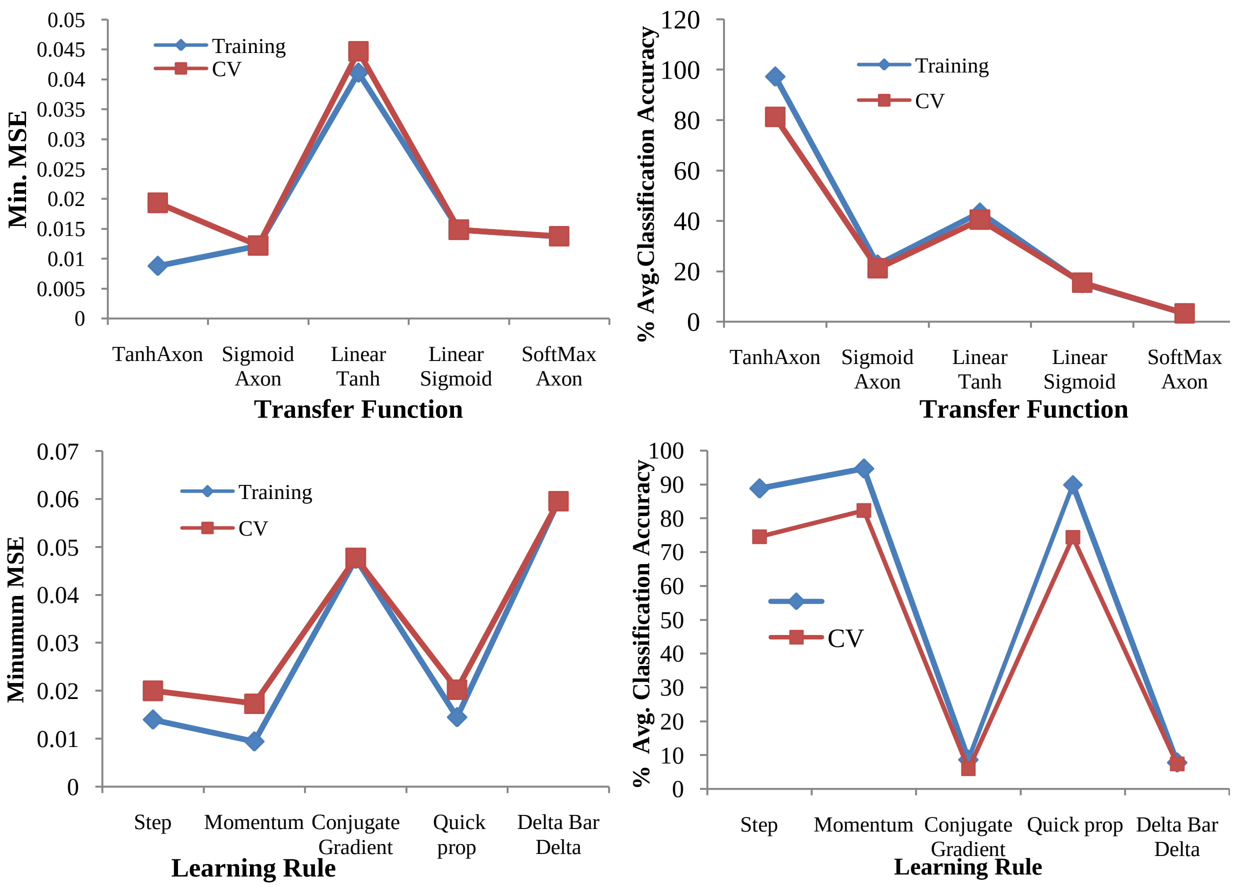

Fig. 4. Variation of Minimum MSE and Percentage average classification accuracy with different transfer functions and learning rules

International Journal of Latest Research in Engineering and Technology (IJLRET)
ISSN: 2454-5031
www.ijlret.com || Volume 02 - Issue 06 || June 2016 || PP. 01-08

Table 1: Performance Matrix for Cross Validation (CV) data set using GFF Neural Network

Performance	O3	O4	O5	O6	O7	O8	O9	OA	OB	OC	OD	OE	OF	OG	OH	OI	IJ	OK	OL	OM	ON	OO	OP	OQ	OR	OS	OT	OU	OV	OW	OX	OY	OZ
Percent Correct	100	80	100	60	100	60	100	88	100	83	83	100	100	86	100	100	86	67	100	100	86	100	100	86	100	20	33	100	71	100	100	60	100

Table 2: Confusion Matrix for Cross Validation (CV) data set using GFF Neural Network

Output / Desired	O3	O4	O5	O6	O7	O8	O9	OA	OB	OC	OD	OE	OF	OG	OH	OI	IJ	OK	OL	OM	ON	OO	OP	OQ	OR	OS	OT	OU	OV	OW	OX	OY	OZ
O3	**6**	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O4	0	**4**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O5	0	0	**7**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O6	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0
O7	0	0	0	0	**5**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O8	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
O9	0	0	0	0	0	1	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OA	0	0	0	0	0	0	0	**7**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OB	0	0	0	0	0	0	0	1	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OC	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OD	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OE	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
OF	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OG	0	0	0	0	0	0	0	0	0	0	0	0	0	**6**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**6**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**5**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IJ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**6**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OK	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	**6**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	**7**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**5**	1	0	0	0	0	0	0	0	0	0	0	0	0
ON	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**6**	0	0	0	0	0	0	0	0	0	0	0	0
OO	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	**10**	0	0	0	0	0	0	0	0	0	0	0
OP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	**5**	0	0	0	0	0	0	0	0	0	0
OQ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**6**	0	0	0	0	0	0	0	0	0
OR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**6**	0	0	0	0	0	0	0	0
OS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	2	0	0	0	0	0	0
OT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	**1**	0	0	0	0	0	0
OU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**8**	0	0	0	0	0
OV	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**5**	0	0	0	0
OW	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**9**	0	0	0
OX	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**8**	2	0
OY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**	0
OZ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**9**

B. Generalized Feed Forward Neural Network (GFFNN)

Like MLP Neural Network we have performed similar trials using GFFNN as shown in Fig. 5 and 6.

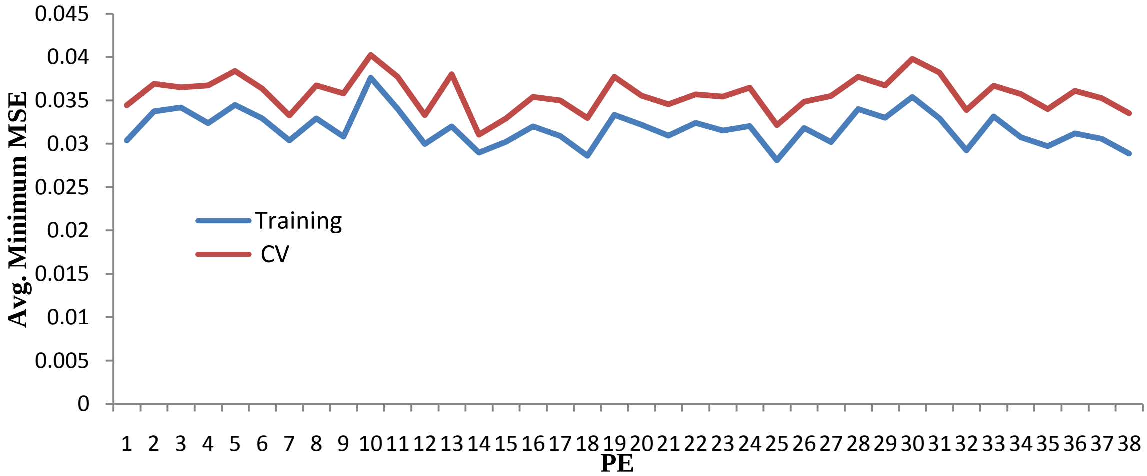

Fig. 5. Processing Element (PE) Vs Minimum MSE

With the following parameter setting we have got maximum Percentage classification accuracy of 93.52 % on training and 76.98 % on CV dataset.

Tagging of Data: 90% for Training & 10% for Cross Validation (CV).

Input Layer: Input PE - 74 Exemplars - 1782

Hidden Layer: PE - 14 Transfer Function - Tanh Learning Rule - Momentum
Momentum - 0.7 Step Size - 0.1

Output Layer: Output PE's - 33 Transfer Function - Tanh Learning Rule - Momentum
Momentum - 0.7 Step Size - 0.1

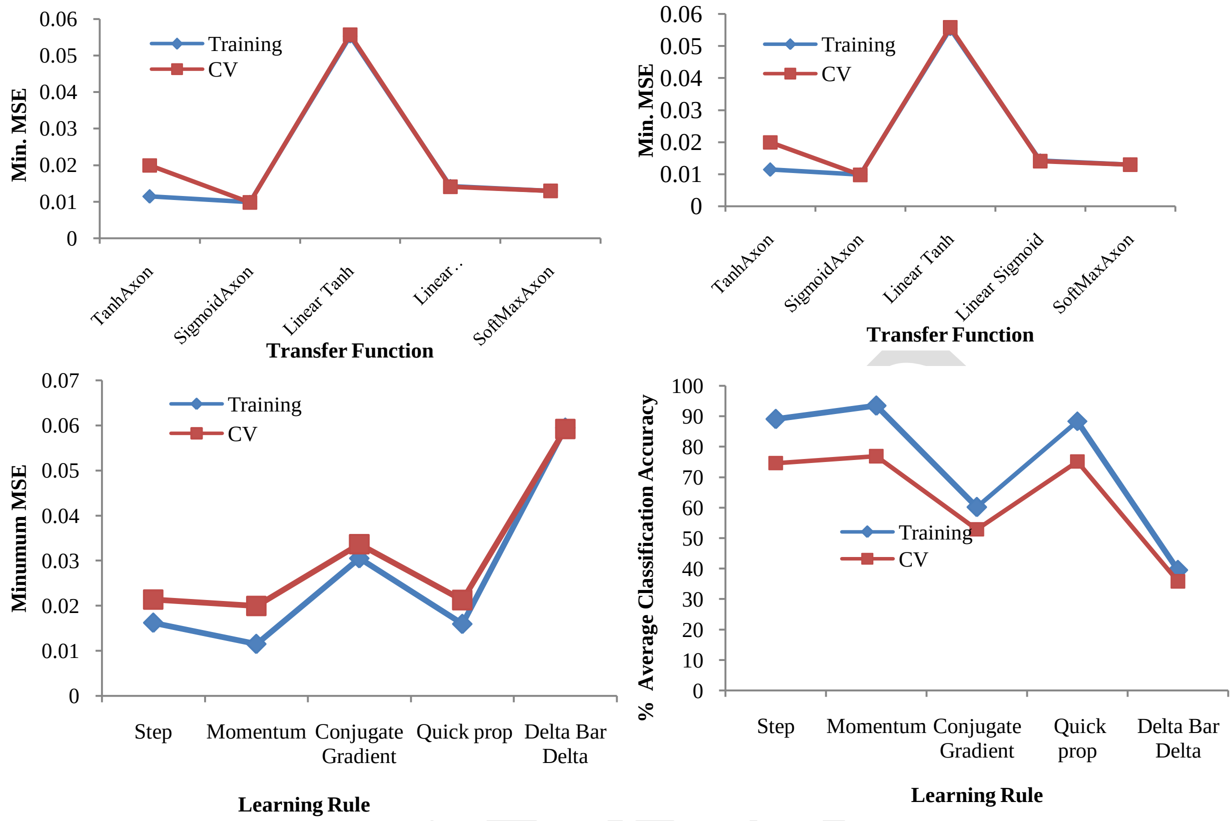

Fig. 6. Variation of Minimum MSE and Percentage average classification accuracy with different transfer functions and learning rules

C. Support Vector Machine

We have varied epoch & number of runs by fixing the step size at 0.1. It is observed that from epoch 37 onwards, there is very small change is MSE as shown in Fig 7. Maximum percentage of classification accuracy is obtained at step setting 0.1 as shown in Fig.8.

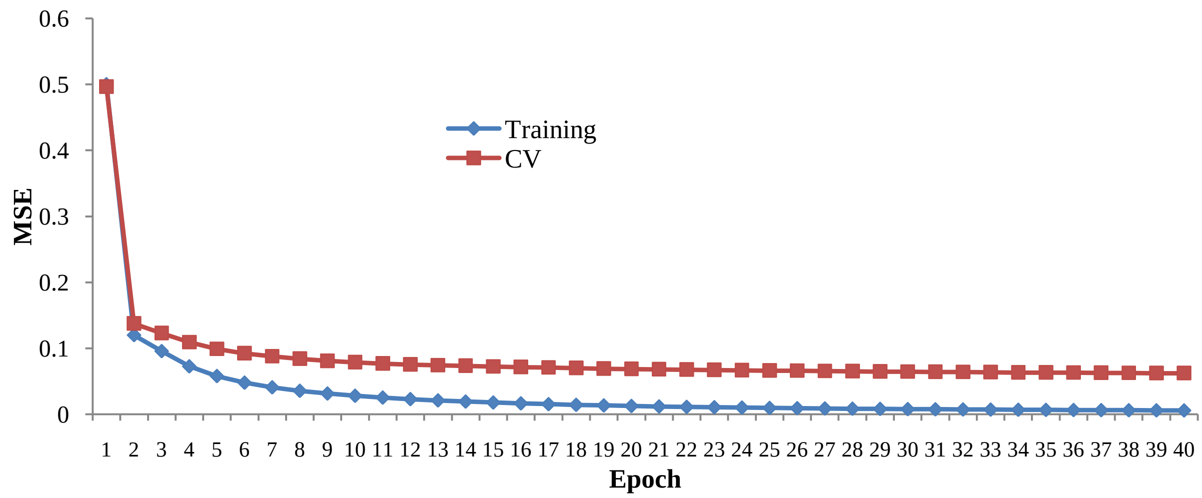

Fig. 7. Epoch Vs Minimum MSE

After experimentation it is observed that the best result obtained as 95.87 % on training and 75.70 % on CV data set with optimal parameter setting as below

Tagging of Data: 90% for Training & 10% Cross validation
No. of Epoch - 37 No. of Runs - 1 step size - 0.1 I/p processing Elements - 74
Exemplars - 1782 Output Processing Elements - 33 Kernel Algorithm: Adatron

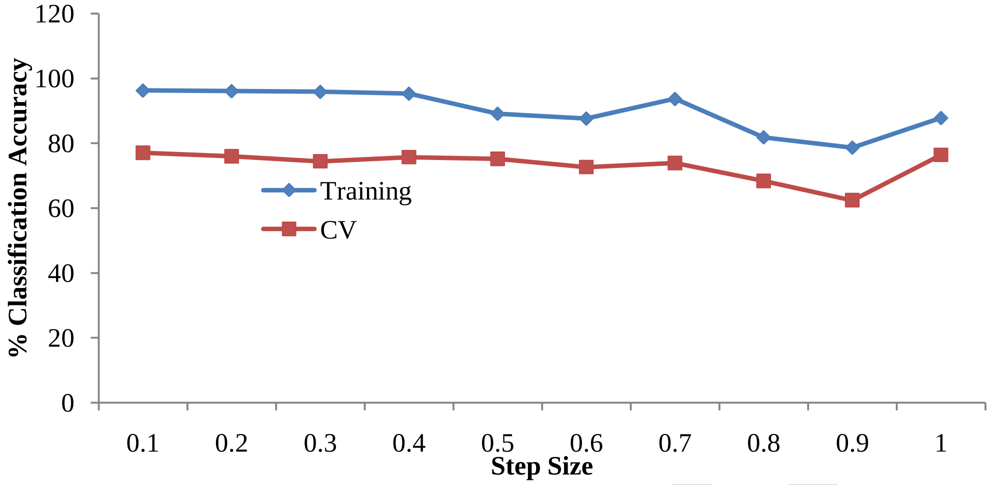

Fig. 8. Percentage average classification accuracy Vs Step size

V. RESULT

We have obtained maximum Average classification accuracy as 86.27 % on CV dataset using MLP Neural Network as per the Table I. The approach presented in this paper is able to classify all alphabets A to Z and numbers 0 to 9 of ISL. However number posture 0, 1, 2 are not included in recognition system due to similarity with alphabet posture O, I, V respectively. The approach has one constraint that background of signer must be fixed black and only static signs are classified.

Table 3.Performance measure of different Neural Network classifiers

Sr. No.	Neural Network Classifier	Percentage of Average Classification Accuracy		Elapsed Time (Sec.)	Database Signs
		Training	CV		
1	MLP	94.63	86.27	214	A to Z 1 to 9
2	GFF	93.52	76.96	194	
3	SVM	99.87	75.70	127	

VI. CONCLUSION

Finally we came to conclude that although MLP Neural Network classifier gives good percentage classification accuracy as compared to other two classifiers but from Table 2 it can be observe that few signs (K, S, T, Y) have very poor classification accuracy. It can be observed from Table 3 that the time required for classification is also maximum in comparison to other classifiers discussed.

REFERENCES

[1] Dixit, Karishma, and Anand Singh Jalal, "Automatic Indian sign language recognition system." IEEE 3rd International Conference on Advance Computing (IACC), pp. 883-887, 2013.
[2] Adithya V, Vinod P. R., Usha Gopalakrishnan, " Artificial Neural Network Based Method for Indian Sign Language Recognition ",Proceedings of IEEE Conference on Information and Communication Technologies, pp. 1080-1085, 2013.
[3] Rekha, J., J. Bhattacharya, and S. Majumder, "Shape, texture and local movement hand gesture features for indian sign language recognition." IEEE 3rd International Conference on Trendz in Information Sciences and Computing (TISC), pp. 30-35, 2011
[4] Saxena, Ankur, Deepak Kumar Jain, and Achintya Singhal. "Sign Language Recognition Using Principal Component Analysis." Fourth IEEE International Conference on Communication Systems and Network Technologies (CSNT), pp. 810-813, 2014.
[5] Bhuyan, M. K., Debashis Ghosh, and P. K. Bora. "A framework for hand gesture recognition with applications to sign language." Annual IEEE In India Conference, pp. 1-6, 2006.
[6] Bhuyan, Manas Kamal, Mithun Kumar Kar, and Debanga Raj Neog. "Hand pose identification from monocular image for sign language recognition." IEEE International Conference on Signal and Image Processing Applications (ICSIPA), pp. 378-383, 2011.

[7] P.V.V.Kishore, M.V.D.Prasad, Ch.Raghava Prasad, R.Rahul,"4-Camera Model for Sign Language Recognition Using Elliptical Fourier Descriptors and ANN", International Conference on Signal Processing And Communication Engineering Systems (SPACES), pp.34-38, 2015.
[8] Ren, Zhou, Junsong Yuan, Jingjing Meng, and Zhengyou Zhang. "Robust part-based hand gesture recognition using kinect sensor." IEEE Transactions on Multimedia, pp. 1110-1120, 2013.
[9] A.S.Elons, Menna Ahmed, Hwaidaa Shedid and M.F.Tolba "Arabic Sign Language Recognition Using Leap Motion Sensor" 9th International Conference on Computer Engineering & Systems (ICCES), pp. 368-673, 2014.
[10] Marin, Giulio, Fabio Dominio, and Pietro Zanuttigh. "Hand gesture recognition with leap motion and kinect devices." IEEE International Conference on Image Processing (ICIP), pp. 1565-1569, 2014.