# Chinese Sign Language Alpha-Numeric Character **Classification using Neural Network**

## R. B. Mapari<sup>1</sup>, G. U. Kharat<sup>2</sup>

<sup>1</sup>Anuradha Engineering College,, Chikhli-443201, MS, India

<sup>2</sup>Sharadchandra Pawar College of Engineering, Otur-412409, MS, India

Abstract: The Chinese Sign Language (CSL) alpha-numeric character classification/recognition without using any aid (embedded sensor, color glove) is really difficult task. This paper describes a novel method to classify static sign by obtaining feature set based on DCT (Discrete Cosine Transform) and Regional properties of hand image. We have collected dataset (alpha numeric character) from 60 people including students of age 20-22 years and few elders aged between 25-38 who have performed 30 signs resulting in total dataset of 1800 signs. Feature set of size 1860×74 is later trained and tested using different classifiers like Multilayer Perceptron (MLP), Generalized Feed Forward Neural Network (GFFNN), Support Vector Machine (SVM). Out of this 90% dataset is used for training and 10% dataset is used for testing/Cross validation. We have got maximum classification accuracy as 89.84 % on CV dataset using GFF Neural Network. NWW.ijs

Keywords: CSL,GFFNN, MLP, SVM, DCT

#### 1. Introduction

Millions of Deaf-Mute people in the world are using sign language as a primary means of communication. A Sign language is a non-verbal languages where information is conveyed using gestures and vision. Information can be conveyed using manual sign or Non Manual signs where apart from hands other body parts also included while performing sign. Sign Language recognition Systems are mainly categorized in two classes as instrumented/Data Glove based and vision (Camera) based. It is observed that hardware (Instrumented glove/Data Glove) based systems can recognize sign more correctly than vision as it has direct information of positioning of fingers and hand movement in coordinate format. Object identification is not the issue in instrumented based system as sensors are directly mounted on elbow, hand, fingers etc.

In comparison to this, vision based system need to first identify the object from an image based on color space selection may be based on skin color or color glove used in segmentation process. Skin color based segmentation is mainly done with plain background or with cloths of dark color where complete hand is covered and only palm, fingers are uncovered. However due to advancement in technology new devices like Leap Motion Sensor & Kinect, researchers have no barrier of background as expected depth can be programmed and 3D information with RGB color information solves most of the problems in traditional methods of sign language recognition.

#### 2. 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.

#### 2.1 Vision Based System

In Vision based system the hand is segmented using color space like RGB, YCbCr, HSV and used skin color as base. In 2007 [1] Yikai Fang et al. have proposed a robust realtime hand gesture recognition method. Hand tracking is achieved using optical flow and color cue. Hand detection uses extended Adaboost method which adopts a new type of feature four box. Using HSV color space hand is segmented with the help of single Gaussian model. Using the features of binary image and Fourier transform gestures are recognize.

In 2012 [2], Serban Oprisescu et al. proposed static hand gesture recognition using depth and intensity information provided by a time-of-flight (ToF) camera. The combined depth and intensity information facilitates the segmentation process, even in the presence of a cluttered background. Hand is segmented using region growing algorithm using distance property. Gesture classification is based on a decision tree using structural descriptions of partitioned contour segments. In 2014 [3], Jingzhong Wang, Meng Li. have recognized 30 finger gestures recognition of Chinese phonetic alphabet using contour features. After images preprocessing, edge features and contour characteristics are used as for matching.

#### 2.2 Instrumented Glove based System

Despite lots of research work carried out using traditional vision-based hand gesture recognition methods [4]-[6] they are still far away from real-life applications. Optical sensing based system is mainly fail due to poor lightening conditions and cluttered backgrounds. So these methods are usually unable to detect and track the hands robustly, which degrades the performance of hand gesture recognition.

Using instrumented glove, In 2002[7], Chunli Wang system two CyberGloves and a Pohelmus 3-D tracker with three receivers positioned on the wrist of CyberGlove and the waist

#### International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

are used as input device to recognize continuous Chinese sign language recognition(CSL). The raw gesture data include hand postures, positions and orientations.

In 2011,[8] Yun Li et al. have worked on Chinese Sign Language(CSL) recognition system to interpret sign components from ACC and sEMG data only. Three basic components hand shape, orientation and movement have been analyzed to identify gesture. Similar type of work carrier in 2012 [9], Deen Ma et al. have proposed Hidden Conditional Random Field (HCRF) for Sign Language Recognition (SLR) based on surface electromyography (sEMG) and acceleration (ACC) signals. In the proposed method, after the periods of data acquisition, data segmentation, feature extraction, and preliminary recognition on the decision-tree level, HCRF was utilized in the bottom layer to classify an observation sequence into a specific class. 4 sEMG & one 3-D accelerometer placed on wrist to acquire data for words. These data glove based systems are sometimes inconvenient to use, expensive and may hamper the natural articulation of hand gesture. As a result, it has gained less popularity. .

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 [10], Zhou Ren et al. have used advanced sensors like Kinect to recognize signs from 1 to 10. The hand is detected using distance threshold. Using one black color belt wear on wrist, hand shape is extracted. Later hand shape is represented as a time-series curve. Using Template matching and Finger-Earth Mover's Distance (FEMD) experiments carried out. A.S.Elons et al. [11] have captured hands and fingers movements in 3D digital format using Leap motion. The sensor throws 3D digital information in each frame of movement. These temporal and spatial features are fed into a Multi-layer perceptron Neural Network (MLP).

# **Experimental Setup**

#### 3.1 Data Collection

We have kept Black background using black cloth and Signers have wear 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 camera of 1.3M pixels (Interpolated 12M pixels still image resolution). 60 different signers of different age categories performed 30 signs as per the chart mentioned in Fig. 1



Figure 1: Sample Signs of CSL

## 3.2 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  $256 \times 256$  size RGB image.

Later on image is converted to gray scale image. The gray scale image is divided in to  $32\times32$  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, ConvexArea, 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

#### 3.3 Classifiers

Following three Neural Network classifiers are tested.

igns of CSL Volume 5 Issue 5, May 2016

<u>www.ijsr.net</u> Licensed Under Creative Commons Attribution CC BY

#### 3.3.1 Generalized Feed Forward Neural Network

Many trials have been performed to get optimal parameters for minimum MSE and maximum percentage of Average Classification Accuracy.



Figure 3: Processing Element (PE) Vs Minimum MSE

Feature vectors are divided into two part as 90 % for training (TR) and 10% for Cross validation. 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 minimum MSE is given by processing element (PE) number 22.

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.

Following parameter setting gives maximum Percentage classification accuracy of 98.57 % on training and 89.84 % on CV dataset.

Input Layer : Input Processing Element - 74 Exemplars - 1620 Hidden Layer :

Processing Elements - 22 Learning Rule - Momentum

22 Transfer Function - Tanh tum Momentum - 0.7

Step Size - 0.1 Output Layer :

Output PE's - 30 Transfer Function - Tanh Learning Rule - Momentum Momentum - 0.7 Step Size - 0.1



**Figure 4:** a) Minimum MSE Vs Transfer b) Percentage of Average classification Accuracy Vs Transfer Functioin c) Minimum MSE Vs Learning Rule d) Percentage of Average classification Accuracy Vs Learning Rule

#### International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

			1	[abl	e 1:	Coi	nfusi	ion 1	Mat	rix f	or C	Cross	s Va	alid	atio	ı (C	V) d	ata s	set u	sing	GFI	F Ne	eura	l Ne	etwo	rk				
O/p	01	04	05	07	04	OB	00	OD	OF	OF	06	ОН	01	01	OK	OT.	OM	ON	00		00	OR	05	от	OU	ov	ow	ov	ov	07
O1	8	04	0.5	0/	1	0	0	0	0	0	00	0	0	0	0	0	0	0	0	0	00	0	0.0	0	0	0	0	3	0	02
04	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
05	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
07	0	1	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
OA	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
OB	0	0	0	0	0	8	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
OC	0	0	0	0	0	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
OD	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	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OF	0	0	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
OG	0	0	0	0	0	0	0	0	0	0	8	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	8	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	6	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
OJ	0	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
OK	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
OL	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
OM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	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	1	6	0	0	0	0	0	0	0	0	0	0	0	0
00	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
OP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	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	2	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	9	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	5	0	0	0	0	0	0	0
ОТ	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	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	3	0	0	0	0	0
OV	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	6	0	0	0	0
OW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	4	0	0	0
OX	1	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	2	0	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	5	0
OZ	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	0	0	5

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

Performance         O1         O4         O5         O7         OA         OB         OC         OD         OE         OF         OG         OH         OI         OJ         OL         ON         OD         OP         OQ         OR         OS         OT         OU         OV         OV         OV         OZ           Percent Correct         89         80         100         80         100         100         100         100         100         86         100         71         100         75         86         100         100         80         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100         100																															
Percent Correct         89         80         100         100         80         100         100         100         100         86         100         71         100         75         86         100         75         100         83         100         75         100         40         100         100	Performance	01	<b>O</b> 4	05	07	OA	OB	OC	OD	OE	OF	OG	OH	OI	OJ	OK	OL	OM	ON	00	OP	OQ	OR	OS	OT	OU	ov	OW	OX	OY	OZ
Correct 89 80 100 100 100 100 100 100 100 100 100	Percent	80	80	100	100	80	100	80	100	100	100	100	100	86	100	71	100	75	86	100	75	100	100	83	100	75	75	100	40	100	100
	Correct	09	80	100	100	80	100	80	100	100	100	100	100	80	100	/1	100	15	80	100	15	100	100	85	100	15	15	100	40	100	100

#### 3.3.2 Multilayer Perceptron Neural Network

Like GFF Neural Network we have performed similar trials using Multilayer Perceptron (MLP) Neural Network. It is observed from graph as shown in Fig 5 that Minimum MSE (Mean Square Error) is for CV data is at 29 numbered PE.

Here also 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. 6. MLP with the following parameter setting gives maximum Percentage classification accuracy of 97.25 % on training and 85.58 % on CV dataset.

Tagging of Data: 90% for Training & 10% Cross validation Input Layer:

- Input Processing Element 74 Exemplars 1620 Hidden Layer:
  - Processing Elements 29 Transfer Function Tanh Learning Rule - Conjugate Gradient

Output Layer:

Output PE's:30 Transfer Function - Tanh Learning Rule:- Conjugate Gradient



International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391



Figure 6: a) Minimum MSE Vs Transfer Function b) Percentage of Average classification Accuracy Vs Transfer Function c) Minimum MSE Vs Learning Rule d) Percentage of Average classification Accuracy Vs Learning Rule

#### **3.3.3 Support Vector Machine**

We have varied epoch & number of runs by fixing the step size at 0.1. It is observed that from epoch 25 onwards, there

is very little change is MSE for CV data as shown in Fig 7. It can be observed from Fig. 8 that maximum Percentage classification accuracy is obtained at step size 0.1.



After experimentation we have observed that the best classification accuracy obtained was 99.84 % on training and 84.91% on CV data set with optimal parameter setting as below

Exemplars: 1620 No. of Epoch: 25 No. of Runs: 1 Input Processing Elements: 74 Output Processing Elements 30 Step Size: 0.2 Kernel Algorithm: Adatron

Tagging of Data: 90% for Training & 10% Cross validation

# Volume 5 Issue 5, May 2016



Fig. 8- a) Minimum MSE Vs Step size b) Percentage of classification Accuracy Vs Step size

#### Result

As the maximum classification accuracy is obtained using GFF Neural Network so we have shown details of confusion matrix and performance matrix of CV data only for GFF neural network. It can be observed from confusion matrix shown in Table 1 that percentage of correct classification of signs like K, M, P, U, V, X is not much satisfactory because of samples are misclassified. So Average classification accuracy of these signs is poor as shown in Table 2. We have obtained maximum Average classification accuracy as 89.84 % on Cross Validation data with the optimal parameter setting as explained earlier using GFF Neural network as shown in Table 3. We have not considered few signs like 2,3,6,8,9 because these signs have similar postures like V,W,Y,L,J respectively.

 Table 3: Performance measure of different Neural Network classifiers

Sr. No.	Neural Network Classifier	Percentage of Classification Training	f Average Accuracy CV	Elapsed Time (Sec.)	Database Signs
1	MLP	97.25	85.58	272	A to Z
2	GFF	98.57	89.84	214	1 to 9
3	SVM	99.84	84.91	67	

### Conclusion

In this paper, we have presented two techniques namely DCT and Regional Properties of Sign images for the accurate classification of signs. From the Table 3 it can be concluded that although GFF neural network is more precise in classification as compared to the other classifiers but, the computational time required for the classification is almost 3 times greater than SVM Neural Classifier.

#### References

 Fang, Yikai, Kongqiao Wang, Jian Cheng, and Hanqing Lu. "A real-time hand gesture recognition method." In Multimedia and Expo, 2007 IEEE International Conference on, pp. 995-998. IEEE, 2007.

- [2] Oprisescu, Serban, Christoph Rasche, and Bochao Su. "Automatic static hand gesture recognition using tof cameras." In Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European, pp. 2748-2751. IEEE, 2012.
- [3] WANG, Jingzhong, and Meng LI. "A Method of Sign Language Gesture Recognition Based on Contour Feature." In Proceedings of the World Congress on Engineering and Computer Science, vol. 1. 2014.
- [4] Murthy, G. R. S., and R. S. Jadon. "A review of vision based hand gestures recognition." International Journal of Information Technology and Knowledge Management,2(2), 2009, 405-410.
- [5] Mitra, S. and Acharya, T., Gesture recognition: A survey. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 37(3), 2007,311-324.
- [6] Erol, Ali, George Bebis, Mircea Nicolescu, Richard D. Boyle, and Xander Twombly. Vision-based hand pose estimation: A review. Computer Vision and Image Understanding, 108(1), 2007, 52-73.
- [7] Wang, Chunli, Xilin Chen, and Wen Gao. "Expanding training set for chinese sign language recognition." In Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on, pp. 323-328. IEEE, 2006.
- [8] Li, Yun, Xiang Chen, Xu Zhang, Kongqiao Wang, and Jihai Yang. "Interpreting sign components from accelerometer and sEMG data for automatic sign language recognition." In Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, pp. 3358-3361. IEEE, 2011.
- [9] Ma, Deen, Xiang Chen, Yun Li, Juan Cheng, and Yuncong Ma. "Surface electromyography and acceleration based Sign Language Recognition using Hidden Conditional Random Fields." In Biomedical Engineering and Sciences (IECBES), 2012 IEEE EMBS Conference on, pp. 535-540. IEEE, 2012.
- [10] Ren, Zhou, Junsong Yuan, Jingjing Meng, and Zhengyou Zhang. "Robust part-based hand gesture recognition using kinect sensor." Multimedia, IEEE Transactions on 15, no. 5 (2013): 1110-1120.

Licensed Under Creative Commons Attribution CC BY

sr.ner

- [11] Elons, A. S., M. Ahmed, H. Shedid, and M. F. Tolba. "Arabic sign language recognition using leap motion sensor." In Computer Engineering & Systems (ICCES), 2014 9th International Conference on, pp. 368-373. IEEE, 2014.
- [12] Marin Giulio, Fabio Dominio, and Pietro Zanuttigh. "Hand gesture recognition with leap motion and kinect devices." In Image Processing (ICIP), 2014 IEEE International Conference on, pp. 1565-1569. IEEE, 2014.

#### **Author Profile**



**Prof. Rajesh B. Mapari** has received his B.E Degree from Sant Gadge Baba Amravati University Amravati, India, M.Tech.Degree from Rashtrasant Tukadoji Maharaj Nagpur University and currently doing Ph.D. Engineering and Technology faculty at Amravati

degree in Engineering and Technology faculty at Amravati University. He is currently Assistant Professor in Anuradha Engineering College, Chikhli, India. He has published many papers in National & International Journals and conferences. His interests include Image processing, Sign language reorganization/study and analysis, Human computer interaction etc.. He is currently Life member of Technical bodies like ISTE, IEI.



**Dr. G. U. Kharat** received his B.E. Degree in 1989 from Amravati University & M.E. Degree in 1999 from Shivaji University, India Ph.D. degree in 2009 in Engineering & Technology faculty at Amravati

University. He has teaching experience of more than 25 years & over 10 years of administrative experience. He has guided over 150 projects at U.G. level. He has over 25 papers in his credit at National & International Conference and Journals. He is Fellow member of Technical bodies like ISTE, IETE, IJERIA, IJMSEA. He was awarded as a best teacher in the year 1993. He is currently guiding 5 Ph.D. scholars. He is Professor/Principal at SharadChandra Pawar college of Engineering Otur, India. His current research interests include Artificial Neural Network, Signal processing and Human Computer Interaction (HCI), etc.

Online): 23,9