International Conference on Communication and Signal Processing, April 4-6, 2019, India

Signet: A Deep Learning based Indian Sign Language Recognition System

Sruthi C. J and Lijiya A *Member, IEEE*

Abstract—**Sign language is the primary mode of communication between hearing and vocally impaired population. The government of India has enacted the Rights of Persons with Disabilities Act 2016 (RPwD Act 2016). This act recognizes Indian Sign Language (ISL) as an important communication medium for communicating with hearing impaired people. This also insists the need for sign language interpreters in all Government organizations and public sector undertakings in order to abide RPwD Act 2016. This can avoid their isolation from the rest of society to a great extent. In this work, we propose a signer independent deep learning based methodology for building an Indian Sign Language (ISL) static alphabet recognition system. Here, we review various existing methods in sign language recognition and implement a Convolutional Neural Network (CNN) architecture for ISL static alphabet recognition from the binary silhouette of signer hand region. We also discuss in detail, the dataset used along with the training phase and testing phase of CNN. The proposed method was successfully implemented with an accuracy of 98.64% which is better than most of the currently existing methods.**

Index Terms—**Convolutional Neural Network, Deep Learning, Gesture recognition, Indian Sign Language, Signer Independent.**

I. Introduction

PEOPLE with disabilities or differently able are often isolated from accessing proper health, education and other social interactions. Assistive technology can change the life of a differently able person to a great extent in all means. Deafness and vocal impairment are one of the major disabilities faced by human beings from centuries. This problem hinders a person to communicate in verbal languages to the outside world leading to isolation from the rest of the major verbally communicating society. They use sign language to communicate with people but it is limited within them or near relatives. To communicate in society they require a manual translator. A sign language converts natural language into hand gestures along with facial expression and eyebrow movements. Sign language is not universal, it changes from region to region[1-8].

In India, Indian Sign Language (ISL) is used by hearing and vocally impaired community. It is estimated that above two million people in India have this disability and among them, one million adults and half million children use ISL [1]. RPwD Act 2016 increased the demand of automated ISL recognition system as it insists all Govt. organizations and public sector undertaking to create a barrier-free environment by appointing interpreters. A manual sign language translator is not a practical solution always and also it hinders the privacy of the person. A solution to this problem is an automated sign language translator which can translate sign language into natural language and outputs in text and speech and also the reverse, speech to signs. An automated bi-directional sign language translator with good accuracy will be an effective assistive technology for hearing and vocally impaired society as it will equip them to lead an independent and very social life for rest of their age[9-18].

An automated sign language translator problem contains numerous challenges, dependencies, and sub-problems inside. The main challenges include hand region extraction, background, and elimination, distance and viewing angle between signer and camera, inter-class and intra-class variation between ISL signs, occlusions, and depth information. A complete recognition system must be able to identify alphabets, numerals, static and dynamic words, contexts, emotions, co-articulation phase, facial expressions, eyebrow movement, body posture, and numerous other situations. The paper discuss the related works followed by proposed method, experiment and results, limitations of the system and finally conclusion and future work in section II, III, IV, V and VI respectively.

II. Related Works

Hand gesture recognition problems are addressed in various different ways by researchers. These can be broadly divided into glove-based or sensor-based methods and appearance-based or vision-based methods. Glove or sensor based approaches provide good accuracy as these specific devices collect data or features directly from signer but, it will have an overhead of carrying and signing with the external device. Vision-based methods, on the other hand, do the recognition task from images or videos based on the features calculated using various image or video processing techniques. This can be again categorized into 2-D Vision based techniques and 3-D Vision based techniques[19-22].

A. Glove or Sensor based Approaches

In most of the sensor based approaches, there will be an input or receptor unit, a processing unit, and output unit. Input unit includes sensors for measuring hand orientation, movement, finger bending, abduction between fingers, etc. and a processing unit is a microcontroller unit which uses these sensor data for processing and recognition. Commonly used sensors in glove based methods include flexion sensors, proximity sensors, accelerometers, abduction sensors, and inertial

Sruthi C. J and Lijiya A. are with the Department of Computer science and Engineering, National Institute of Technology, Calicut, India (e-mail: sruthicj.edu@gmail.com, lijiya@nitc.ac.in).

978-1-5386-7595-3/19/$31.00 ©2019 IEEE

measurement unit (IMU) [2]. Table I shows some of the glove-based approaches.

B. Vision based Approaches

Most of the research in sign language recognition focuses on vision-based methods rather than glove based methods. This is because of the fact that these techniques reduce the dependencies of users on sensor devices along with cutting the extra cost for special sensors. 2-D Vision based technique works with the camera which is now available in mostly all mobile phones and laptops. 2-D system works well when combined with proper image or video processing techniques but it lacks depth information. Availability of 3-D or depth cameras in affordable prices had improved 3-D vision based research to a great extent. Computational and cost overhead of three-dimensional sign language recognition can be compromised with its high recognition rate. Addition of the third dimension actually solves many challenges in two-dimensional sign language recognition by providing extra information about the depth. Most of the research in sign language recognition focuses on vision-based methods rather than glove based and depth based methods. Many vision-based methods are proposed by various researchers and some of them are listed in Table II.

III. Proposed Method

This work addresses Indian sign language static alphabet recognition problem with a vision based approach. A Convolutional Neural Network (CNN) which is a deep learning technique is used to create a model named signet, which can recognize signs, based on supervised learning on data. The whole process can be divided into CNN training and model testing. The diagrammatic representation of Signet architecture is shown in Fig. 1.

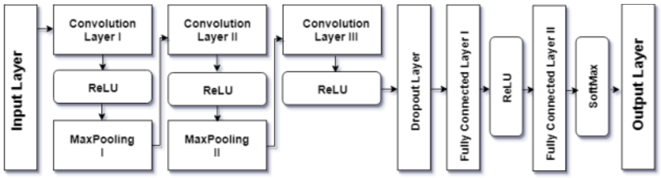


Fig. 1. Diagrammatic Representation of Proposed Signet Architecture

First and foremost step in all vision based SLR system is to preprocess the images in data set to obtain a noise-free hand region extraction. In this work, we use a dataset containing binary hand region silhouette of the signer images. These images were extracted and saved after the preprocessing stage in [23]. This preprocessing includes Viola-Jones face detection algorithm [24], [25] to detect the face of the signer and then eliminate it by replacing it with black pixels. This image is then processed with a skin color segmentation algorithm [26] followed by the largest connected component algorithm for hand region segmentation. These images are used in this work to train and test the signet architecture. Images along with their class labels are given to the developed CNN architecture to learn the classification model. The learned classification model can be tested and then saved for recognizing ISL static alphabets[27].

A. CNN Architecture

The proposed convolutional neural network architecture has six hidden layers along with one input layer, one dropout layer and one output layer with a total of nine layers. Table III shows the design of the proposed CNN architecture.

Each of the training images fed into this architecture is resized to a size of (200,100) and input layer consist of 20000 neurons. Next layer is a convolution layer with 32 different filters of size (3,3). A stride of size one is used in convolution layers without padding. Each of the 32 filter kernels, when convoluted with input from previous layer, gives 32 different feature maps as in equation 1.

$$F_{i,j,k} = b_k + \sum_{u=1}^{f_x} \sum_{v=1}^{f_y} \sum_{k'=1}^{f_n} x_{i',j'} . w_{u,v,k} \tag{1}$$

- $F_{i,j,k}$ is the output stored after convolution operation in row i, column j in feature map k of layer L.
- $f_x, f_y$are height and width of filter kernels in layer L-1.
- f_n is the number of filters in layer L-1.
- $x_{i',j'}$ is the pixel value with index i', j' in layer L-1 on which convolution is done.
- $w_{u,v,k}$ is the filter pixel value in k^{th} filter with index values as u,v.
- b_k is the bias term for k^{th} feature map.

Convolution layer is accompanied by a ReLU activation function given by

$$y = max(x, 0) \tag{2}$$

Next is a pooling layer which reduces the dimensionality of the data fed into it. This layer uses a max operator for this over a receptive field of (5,5) and with a stride of 5 pixels.

$$y = max_i(\{x_i\}) \tag{3}$$

where $\{x_i\}$ denote set of all x_i which is an element of (5,5) receptive field. Third, fourth and fifth hidden layers are convolution pooling and convolution layers respectively, with same activation function, stride and filter size and shape combination. Next layer is a fully connected neural network layer, but before passing, data from convolution layer to fully connected layer a 20% dropout is applied to avoid overfitting. Remaining 80% data are then flattened and are given to the next hidden layer or first fully connected layer of our CNN architecture. This layer consists of 128 neurons which are fully connected with the next layer having 24 neurons. ReLU activation function is used for activation in first fully connected layer neurons and softmax activation function for the next layer. Softmax activation function can be formulated as in equation 4.

$$y_j = \frac{e^{x^T w_j}}{\sum_{k=1}^{K} e^{x^T w_k}} \tag{4}$$

Softmax activation function calculates probability values for all 24 classes and the class with the highest probability is assigned value one and all others are given value zero to form one-hot-vector. This is then given to the output layer having the same number of neurons symbolizing all 24 static alphabets.

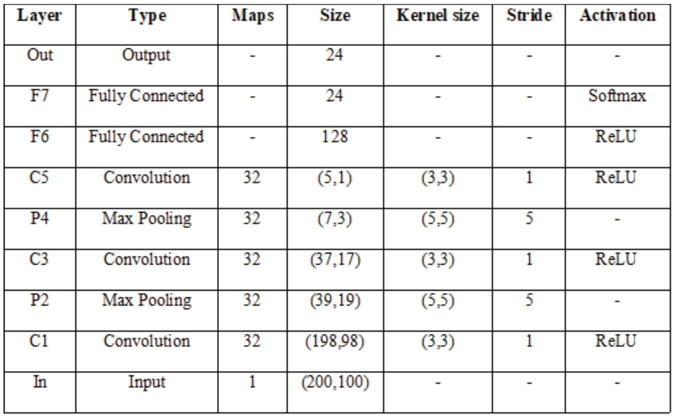
TABLE I
GLOVE BASED APPROACHES

Reff.	Sign Language	Sensors	Gestures recognized	Detection Accuracy
[3]	American	6-IMUs and accelerometer	Digits: 0-9 and Alphabets: A-Z	92%
[4]	Indian	3-axis accelerometer and Flex sensors	8-common words	99%
[5]	Australian	Multiple flex and contact sensors	120 different static gestures	-
[6]	Common gestures	Multiple IMU sensors	5-static and 8-dynamic	99.4%
[7]	American	Flex sensors and gyroscope	3 basic gestures	86.67%
[8]	-	Bend and hall effect sensors with accelerometer	0-9 symbols	96%
[9]	American	3-axis accelerometer and Flex sensors	Finger spelling alphabets	94%

TABLE II
VISION BASED APPROACHES

Reff.	3-D/2-D	Sign Language	Techniques Used	Gestures recognized	Accuracy
[10]	3-D	Indian	3-D position trajectories and adaptive matching	20 actions with 10 subjects	98%
[11]	3-D	Indian	3-D motionlets and adaptive kernal matching	500 signs	98.9%
[12]	2-D	Indian	Discrete wavelet transforn and HMM	10 Types of sentences	91%
[13]	2-D	American	Skin-color segmentation, zernike moment, finger tip detection and SVM	24 static alphabets and 4 dynamic signs	static: 93% and dynamic: 100%
[14]	2-D	American	Finger tip and palm position, PCA, optical flow, CRF	9 Alphabets	96%
[15]	2-D	Indian	Scale Invariant Feature Transform, Key point extraction	26 Alphabets	Time factor was given importance than recognition accuracy
[16]	2-D	Indian	Skin color segmentation,distance transform, fourier descriptor, Artificial Neural Network	26 Alphabets and 0-9 Digits	91.11%
[17]	2-D	Indian	Direct pixel value, Hierarchial centroid, k-Nearest Neighbour, neural Network Pattern Recognition Tool	Digits 0-9	97.10%
[18]	2-D	Indian	Zernike moments, Sequential Minimal Optimization	5 Alphabets	94.4%
[19]	2-D	American	Zernike moment, SVM	24 Alphabets	96%
[20]	2-D	Indian	Histogram of edge frequency, SVM	26 Alphabets	98%
[21]	2-D	Indian	Eigen value weighted euclidean distance	24 Static alphabets	95%
[22]	2-D	Indian	B-spline approximation and SVM	29 signs (A-Z and 0-5)	90%

TABLE III
PROPOSED CNN ARCHITECTURE



Layer	Type	Maps	Size	Kernel size	Stride	Activation
Out	Output	-	24	-	-	-
F7	Fully Connected	-	24	-	-	Softmax
F6	Fully Connected	-	128	-	-	ReLU
C5	Convolution	32	(5,1)	(3,3)	1	ReLU
P4	Max Pooling	32	(7,3)	(5,5)	5	-
C3	Convolution	32	(37,17)	(3,3)	1	ReLU
P2	Max Pooling	32	(39,19)	(5,5)	5	-
C1	Convolution	32	(198,98)	(3,3)	1	ReLU
In	Input	1	(200,100)	-	-	-

IV. EXPERIMENT AND RESULTS

Implementation of the proposed system was done using python 3.2. We also used keras and tensorflow for implementing the CNN part. Details regarding dataset, training & testing and results are discussed in following subsections.

A. Dataset

Indian sign language lacks a standard signer image or video dataset to study upon. Even though we have standard signs in ISL it may vary slightly from region to region. In this research, we used binary hand region silhouette of the signer images from [23]. Original dataset before preprocessing consisted of images collected from Rahmaniya HSS Special School Calicut. Seven different signers were used to create a total of 2500 images. Binary hand region silhouette dataset also contained 2500 images initially. These extracted images were then augmented with numerous random translations and scaling operations to produce an augmented dataset of 5157 images.

B. Training & Testing

In the training phase, 80% of images from each alphabet set was selected randomly and was given to the CNN model with their corresponding labels. We fixed the learning rate as 0.01 initially and used a parameter updating stochastic gradient descent method for optimization which iteratively approximates the CNN weight vector over the training set with the equation,

$$w := w - \eta . \sum_{i=1}^{N} \delta F_i(w) \tag{5}$$

- w is the weight vector
- η is the learning rate
- $F_i(w)$ is the objective function

A collection of 4125 randomly selected images was used for training and obtained a training accuracy of 99.93%. This was validated and tested with 1032 images which come around 20% of our dataset. A remarkable validation accuracy of 98.64% was obtained. The system was also tested with images from outside the dataset.

C. Results and Discussions

The proposed method produced a remarkable result compared to current state of art methods. Training accuracy of 99.93% and validation accuracy of 98.64% was achieved.

Confusion matrix obtained for the method is given in Fig. 2. Figure shows sparse entries in confusion matrix else where other than its dense diagonal which highlight very less miss classifications. Dense diagonal signifies the correctly classified predictions. Precision, recall, and f1-score obtained after testing is displayed in Table IV. Precision actually shows the ability of the classifier to predict only the real positive samples as positive and recall shows the amount of positive samples that are correctly classified as positive.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{6}$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \tag{7}$$

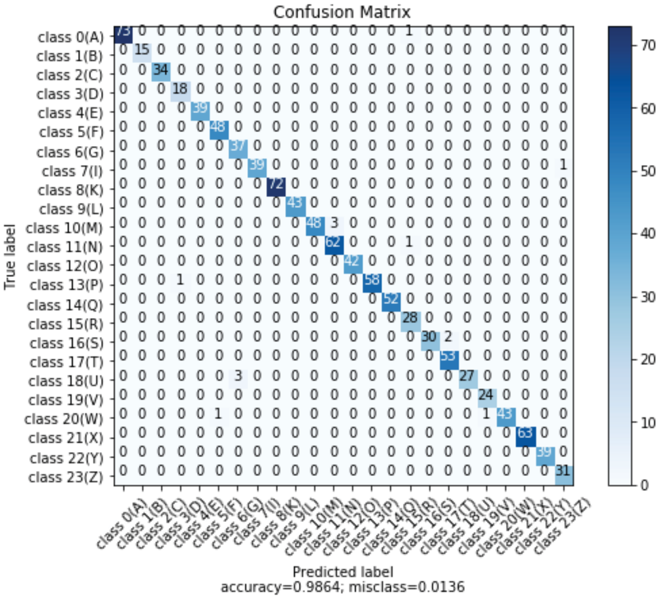
Here we can find that both precision and recall have high values and F1-score which is the harmonic mean of both precision and recall also shows values very close to 1.0 signifying good classification result.

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{8}$$

The proposed method is compared over existing methods and it was found that it outperforms other methods over the size of datasets and accuracy. Comparison with other methods is shown in Table V.

V. Limitations

In this work we used binary silhouette of signer hand region as input to the CNN. This can be extended to a method which can process videos or images taken directly from mobile camera or any external camera.



	class 0(A)	class 1(B)	class 2(C)	class 3(D)	class 4(E)	class 5(F)	class 6(G)	class 7(I)	class 8(K)	class 9(L)	class 10(M)	class 11(N)	class 12(O)	class 13(P)	class 14(Q)	class 15(R)	class 16(S)	class 17(T)	class 18(U)	class 19(V)	class 20(W)	class 21(X)	class 22(Y)	class 23(Z)
class 0(A)	73	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
class 1(B)	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
class 2(C)	0	0	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
class 3(D)	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
class 4(E)	0	0	0	0	39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
class 5(F)	0	0	0	0	0	48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
class 6(G)	0	0	0	0	0	0	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
class 7(I)	0	0	0	0	0	0	0	39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
class 8(K)	0	0	0	0	0	0	0	0	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
class 9(L)	0	0	0	0	0	0	0	0	0	43	0	0	0	0	0	0	0	0	0	0	0	0	0	0
class 10(M)	0	0	0	0	0	0	0	0	0	0	48	3	0	0	0	0	0	0	0	0	0	0	0	0
class 11(N)	0	0	0	0	0	0	0	0	0	0	0	62	0	0	0	1	0	0	0	0	0	0	0	0
class 12(O)	0	0	0	0	0	0	0	0	0	0	0	0	42	0	0	0	0	0	0	0	0	0	0	0
class 13(P)	0	0	0	1	0	0	0	0	0	0	0	0	0	58	0	0	0	0	0	0	0	0	0	0
class 14(Q)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	52	0	0	0	0	0	0	0	0	0
class 15(R)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28	0	0	0	0	0	0	0	0
class 16(S)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30	2	0	0	0	0	0	0
class 17(T)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53	0	0	0	0	0	0
class 18(U)	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	27	0	0	0	0	0
class 19(V)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24	0	0	0	0
class 20(W)	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	43	0	0	0
class 21(X)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	63	0	0
class 22(Y)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	39	0
class 23(Z)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31

Fig. 2. Confusion matrix

TABLE IV
PRECISION, RECALL AND F1-SCORE

Class	precision	recall	f1-score
class (A)	1.00	0.99	0.99
class (B)	1.00	1.00	1.00
class (C)	1.00	1.00	1.00
class (D)	0.95	1.00	0.97
class (E)	1.00	1.00	1.00
class (F)	0.98	1.00	0.99
class (G)	0.93	1.00	0.96
class (I)	1.00	0.97	0.99
class (K)	1.00	1.00	1.00
class (L)	1.00	1.00	1.00
class (M)	1.00	0.94	0.97
class (N)	0.95	0.98	0.97
class (O)	1.00	1.00	1.00
class (P)	1.00	0.98	0.99
class (Q)	1.00	1.00	1.00
class (R)	0.93	1.00	0.97
class (S)	1.00	0.94	0.97
class (T)	0.96	1.00	0.98
class (U)	1.00	0.90	0.95
class (V)	0.96	1.00	0.98
class (W)	1.00	0.96	0.98
class (X)	1.00	1.00	1.00
class (Y)	1.00	1.00	1.00
class (Z)	0.97	1.00	0.98

TABLE V
COMPARISON WITH OTHER METHODS

Reff.	Methodology	Classes/Dataset	Accuracy	Remarks
[18]	Zernike moment, SMO	(5,720)	94.4	Only five alphabet
[20]	Histogram of edge frequency, SVM	(26,1560)	98	Use wrist band
[21]	Eigen value weighted euclidean distance	(24,240)	95	Small dataset
[22]	B-spline approximation and SVM	(29,290)	90	Small dataset,complex computation
[27]	YCbCr, wavelet, multi-class SVM	(23,230)	86.3	Small dataset, low accuracy
Proposed method	Skin color segmentation, CNN	(24,4125)	98.64	Large dataset and higher accuracy

VI. CONCLUSION AND FUTURE WORK

The paper presented a vision based deep learning architecture for signer independent Indian sign language recognition system. The system was successfully trained on all 24 ISL static alphabets with a training accuracy of 99.93% and with testing and validation accuracy of 98.64%. The recognition accuracy obtained is better than most of the current state of art methods.

Sign language recognition problem is a broad research area which includes recognition problems like finger spelling dynamic alphabets, dynamic words, co-articulation detection and elimination for sentence identification. Proposed architecture can be extended with additional modules and techniques to form a fully automated sign language recognition system in the future. Facial expression and context analysis are the other part to be included in sign language recognition. An automated ISL recognition system with speech translator which can process videos in real time to produce voice output of the same can become a most effective assistive technology in near future.

ACKNOWLEDGMENT

We would like to thank teachers and students in Rahmania HSS special school for their support with regard to data collection and information gathering. We also express our gratitude to Karuna special school and National Institute of Speech and Hearing (NISH) for their valuable suggestions and support.

REFERENCES

[1] Tavari, Neha V., A. V. Deorankar, and P. N. Chatur. ”Hand gesture recognition of indian sign language to aid physically impaired people.” International Journal of Engineering Research and Applications (2014): 60-66.

[2] Ahmed, Mohamed Aktham, et al. ”A review on systems-based sensory gloves for sign language recognition state of the art between 2007 and 2017.” Sensors 18.7 (2018).

[3] Abhishek, Kalpattu S., Lee Chun Fai Qubeley, and Derek Ho. ”Glove-based hand gesture recognition sign language translator using capacitive touch sensor.” Electron Devices and Solid-State Circuits (EDSSC), 2016 IEEE International Conference on. IEEE, 2016.

[4] Lokhande, Priyanka, Riya Prajapati, and Sandeep Pansare. ”Data gloves for sign language recognition system.” International Journal of Computer Applications (2015): 11-14.

[5] Ahmed, Syed Faiz, Syed Muhammad Baber Ali, and Sh Saqib Munawwar Qureshi. ”Electronic speaking glove for speechless patients, a tongue to a dumb.” Sustainable Utilization and Development in Engineering and Technology (STUDENT), 2010 IEEE Conference on. IEEE, 2010.

[6] Han, Rui, et al. ”A Data Glove-based KEM Dynamic Gesture Recognition Algorithm.” International Journal of Performability Engineering 14.11 (2018).

[7] Das, Abhinandan, et al. ”Smart glove for Sign Language communications.” Accessibility to Digital World (ICADW), 2016 International Conference on. IEEE, 2016.

[8] Chouhan, Tushar, et al. ”Smart glove with gesture recognition ability for the hearing and speech impaired.” 2014 IEEE Global Humanitarian Technology Conference-South Asia Satellite (GHTC-SAS),2014.

[9] Cabrera, Maria Eugenia, Juan Manuel Bogado, Leonardo Fermin, Raul Acuna, and Dimitar Ralev. ”Glove-based gesture recognition system.” In Adaptive Mobile Robotics, pp. 747-753. 2012.

[10] Kumar, D. Anil, et al. ”Indian sign language recognition using graph matching on 3D motion captured signs.” Multimedia Tools and Applications (2018): 1-29.

[11] Kishore, P. V. V., et al. ”Motionlets Matching with Adaptive Kernels for 3D Indian Sign Language Recognition.” IEEE Sensors Journal (2018).

[12] Tripathi, Kumud, Neha Baranwal, and Gora Chand Nandi. ”Continuous dynamic Indian Sign Language gesture recognition with invariant backgrounds.” Advances in Computing, Communications and Informatics (ICACCI), 2015 International Conference on. IEEE, 2015.

[13] Kumar, Anup, Karun Thankachan, and Mevin M. Dominic. ”Sign language recognition.” Recent Advances in Information Technology (RAIT), 2016 3rd International Conference on. IEEE, 2016.

[14] Hussain, Imran, Anjan Kumar Talukdar, and Kandarpa Kumar Sarma. "Hand gesture recognition system with real-time palm tracking." India Conference (INDICON), 2014 Annual IEEE. IEEE, 2014.

[15] Patil, Sandeep Baburao, and G. R. Sinha. ”Distinctive feature extraction for Indian Sign Language (ISL) gesture using scale invariant feature Transform (SIFT).” Journal of The Institution of Engineers (India): Series B 98.1 (2017):19-26.

[16] Adithya, V., P. R. Vinod, and Usha Gopalakrishnan. ”Artificial neural network based method for Indian sign language recognition.” Information & Communication Technologies (ICT), 2013 IEEE Conference on. IEEE, 2013.

[17] Sharma, Madhuri, Ranjna Pal, and Ashok Kumar Sahoo. ”Indian Sign Language Recognition Using Neural Networks and KNN Classifiers.” ARPN Journal of Engineering and Applied Sciences 9.8 (2014).

[18] Sharma, Kalpana, Garima Joshi, and Maitreyee Dutta. ”Analysis of shape and orientation recognition capability of complex Zernike moments for signed gestures.” Signal Processing and Integrated Networks (SPIN), 2015 2nd International Conference on. IEEE, 2015.

[19] Otiniano-Rodrıguez, K. C., G. C ámara-Chávez, and D. Menotti. ”Hu and Zernike moments for sign language recognition.” Proceedings of international conference on image processing, computer vision, and pattern recognition. 2012.

[20] Lilha, Himanshu, and Devashish Shivmurthy. ”Evaluation of features for automated transcription of dual-handed sign language alphabets.” Image Information Processing (ICIIP), 2011 International Conference on. IEEE, 2011.

[21] Singha, Joyeeta, and Karen Das. ”Indian sign language recognition using eigen value weighted Euclidean distance based classification technique.” arXiv preprint arXiv:1303.0634 (2013).

[22] UC, Geetha M. Manjusha, Vishwa Vidyapeetham Amrita, and Amritapuri Amritapuri. ”A Vision Based Recognition of Indian Sign Language Alphabets and Numerals Using B-Spline Approximation.” International Journal on Computer Science and Engineering (IJCSE),Vol. 4,2012.

[23] P. K. Athira, ”Indian sign language recognition,” Phd thesis, Dept. CSE., NITC., Calicut, India, 2017, Accessed on: May, 2017.[online] Available: http://192.168.240.208:8080/xmlui/handle/123456789/35892

[24] Viola, Paul, and Michael J. Jones. ”Robust real-time face detection.” International journal of computer vision 57.2 (2004): 137-154.

[25] Yun, Liu, and Zhang Peng. ”An automatic hand gesture recognition system based on Viola-Jones method and SVMs.” 2009 Second International Workshop on Computer Science and Engineering. IEEE, 2009.

[26] Bhuyan, Manas Kamal, et al. ”A novel set of features for continuous hand gesture recognition.” Journal on Multimodal User Interfaces 8.4 (2014): 333-343.

[27] Rekha, J., J. Bhattacharya, and S. Majumder. ”Shape, texture and local movement hand gesture features for indian sign language recognition.” Trendz in Information Sciences and Computing (TISC), 2011 3rd International Conference on. IEEE, 2011.