An Efficient Indian Sign Language Recognition System using Sift Descriptor

Jasmine Kaur, C. Rama Krishna

Abstract: Communication is one of the basic requirements for living in the world. Deaf and dumb people convey through Sign Language but normal people have difficulty to understand their language. In order to provide a source of medium between normal and differently abled people, Sign Language Recognition System (SLR) is a solution . American Sign Language (ASL) has attracted many researchers' attention but Indian Sign Language Recognition (ISLR) is significantly different from ASL due to different phonetic, grammar and hand movement. Designing a system for Indian Sign Language Recognition becomes a difficult task. ISLR system uses Indian Sign Language (ISL) dataset for recognition but suffers from problem of scaling, object orientation and lack of optimal feature set. In this paper to address these issues, Scale-Invariant Feature Transform (SIFT) as a descriptor is used. It extracts the features that train the Feed Forward Back Propagation Neural Network (FFBPNN) and optimize it using Artificial Bee Colony (ABC) according to the fitness function. The dataset has been collected for alphabet from the video by extracting frames and for numbers it has been created manually from deaf and dumb students of NGO "Sarthak". It has been shown through simulation results that there has been significant improvement in accurately identifying alphabets and numbers with an average accuracy of 99.43%.

Index Terms: ISLR, ABC, FFBPNN, SIFT, SLR, ISLR, ASL, ISL.

I. INTRODUCTION

Sign Language Recognition is mainly used to communicate between people who are hard of hearing [1] but individuals who have disability or condition (augmented and alternative communication) also use sign language as communication medium. There are many native sign languages in India [1]. Each language has its own grammar and lexicon and are very different from one another. Normal people find it difficult to understand sign language and use interpreter as translation source.SLR is used to recognize the Sign Language. It has its advantages as it helps in making the system intelligent and easy to be taught and thus help everyone communicate freely. Sign language contains series of hand movements and facial features that express words [3]. Hand movements are largely classified into two types named as static and dynamic [4]. Static Motion is a definite posture

* Correspondence Author
Jasmine Kaur*, Dept. of Computer Science & Engineering, NITTTR, Chandigarh, India
C. Rama Krishna , Dept. of Computer Science & Engineering, NITTTR, Chandigarh, India

assigned with appropriate meaning. In Static Hand Gesture Recognition Technique, the hand posture along with finger, thumb and palm configuration is detected [5].

As syntax, speech and hand movement varies in each region so to convey various signs India has its own sign language. Therefore, it is necessary to design a system in India that helps other people to converse with dumb or deaf people [6]. The work has already been performed on Sign Language in European countries [7]. The SLR system has been already presented by number of researchers, but the accuracy rate is not as per need due to the lack of feature optimization concept. Since, the uniqueness of feature is less for different types of dataset hence we have tried to use ABC as a feature optimization technique in this work to improve the accuracy of the system [8].

In this work, an automated Indian Sign Recognition System has been optimized for differently abled people in India which classifies all 26 English Alphabets and numbers 0-9. The ISLR system processing has gone through two stages specifically training and testing which is followed by the classification mechanism. Fusion of ANN and SIFT is used for feature extraction and multiclass classification respectively. The main contributions of this paper are:

1) An efficient ISLR system is proposed and an average accuracy of 99.43% is reached.
2) The dataset used is created from scratch and has been made public at https://github.com/jasminek247/SLR/tree/master for future research.

II. RELATED WORK

The research in ISL's recognition has facilitated researchers to improve their recognition methods over the past decade. Four steps have been presented to recognize the alphabets of ISL [9]. In [10] authors have proposed a statistical approach in order to categorize and promote ISL's dynamic movements in real time. Histogram orientation has been utilized for classification. The techniques such as Euclidean distance and K-nearest have been used that provides classification with higher accuracy. In [11], Nandy A et al. have presented an automatic signed expression recognition scheme that works on the basis of modeling gestures along with subunits.

The characterization to articulate sign language gestures in different parts of the body as three-dimensional movements has been presented [12].In [13] Rajam et al. proposed recognizition of South Indian Sign Language gestures by considering 32 number of gestures with an obtained accuracy of 98.12% .

Retrieval Number F8124088619/2019©BEIESP
DOI: 10.35940/ijeat.F8124.088619
Journal Website: www.ijeat.org

1456

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

In [14], the work has been performed by Deora et al. to recognize ISL gesture by considering both hand and overlapping hand. In [15] [16], authors have used Support Vector machine to classify SLR gestures.

In [17] Jayashree et al. has presented an ASL system by using the concept of Euclidean distance to distinguish 26 hand gestures that are static from the composite background. Also, in [18] authors proposed an algorithm that helps to distinguish the static and dynamic gesture. In [19], the hand segmentation scheme has been utilized by utilizing the concept of lab color space and hence provided valley as well as peaks as key points. Randive et al. [20] designed a new system that has been used to recognize the 36 different gestures using SIFT as a feature extractor.

In [21], Camgöz et al. proposed an end to end hand shape identification system. The problem of the sequence to sequence learning has been resolved which is referred to as Sub-Unets. Sub-Unet allows the developer to emphasize expert proficiency in the system regarding appropriate representations. In this research, Convolutional Neural Network (CNN) is used to extract the features of the input image. At last, Connectionist Temporal Classification (CTC) has been used to train the network with the dissimilar length video. In [22], an automatic recognition of the static gesture of sign language that has been integrated from the kinect sensors. The CNN structure has been used from the kinect depth images. Using CNN, the system has been trained to classify 24 numbers of alphabets along with the numbers from 0-9 by utilizing 33000 numbers of images. The system has obtained accuracy rate of 94.67%. The accuracy is low due to the occlusion problem of alphabet A, M and N.

In [23], authors presented multiple feature extraction methods for SLR. The movement information has been converted into a single image by using the concept of temporal estimation and accumulated differences. K nearest neighbour (KNN) and Bayesian have been used to extract the features with a classification rate upto70% approximately. The rate is low since no optimization algorithm has been used which helps to select the best features like valley points of fingers or skin color of hand.

A new way to make sign language quicker in application used by mobile platforms is introduced [24]. Selfie is considered as capturing video signal language by restricting its figuring power for smartphones. 'Pre-filtering', retouching, and extracting the characteristic on video frames produces a space for the sign language characteristic. The performance rate up to 90% has been achieved when ANN is used. In [25], authors proposed a neural network to classify the ISL (Indian sign language). It has been examined that the if data has been combined from two sensors the accuracy rate is more than that of the data examined from the single sensor. The detection accuracy rate when both hands have been included is 91%.

III. PROPOSED OPTIMIZED ISLR SYSTEM

After extensive literature survey it has been found that the problem of accuracy, optimization and recognition of all letters and numbers persists. There is also a problem of scaling and orientation which makes image distorted if it increases, decreases or is rotated. Being scale and orientation invariant we have used SIFT as feature descriptor in order to extract the feature sets according to different hand gestures. In case of feature selection, FFBPNN is used as classifier for training of ISLR system with Artificial Bee Colony as an optimizer. Fig. 1 depicts working of proposed ISLR system.

A. Upload Hand Gesture Images
i. Data Collection

There are several sign languages presented all over the world like ASL, Japanese Sign language, Australian Sign language and many more [1]. Similarly, ISLR was also designed for hearing impaired people in India. Following are the facts which make ISL different from other sign languages.
- Most of the alphabets in ISL are represented by both hands.
- The movement of both the hands is not static.
- Static as well as dynamic hand gesture along with facial expression is also included in ISL.

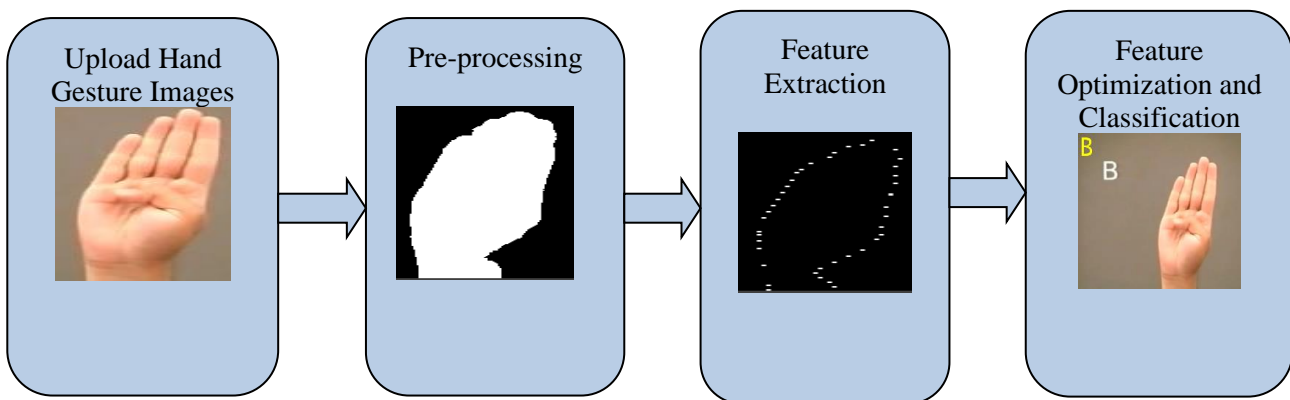In this work we need to upload the hand gesture images for

Fig. 1 Working Mechanism of Proposed ISLR System

for ISLR. This makes it a challenge as resources are not available to download ISLR alphabets and numeric image dataset. In Indian Sign Language dataset, video is only available for SLR system. But we have worked on static images.

This helps in reducing frame extraction time caused in video and images have been extracted from video itself leading to creation of fixed dataset. It also removes complexity caused due to movement of hand in video. Similarly, the ISL dataset for numbers has been created manually from differently abled students of NGO "Sarthak". The PNG images have been used with dimension of 640×480 pixels. All images are uncompressed and noise free which combines background and foreground part. The size of each image varies from 60 KB to 70 KB. There are total of 26 alphabets and 10 numbers (0-9) used in this research. Each test data comprises of 5 images. In total we have 130 subjects. For the ratio 60% of images from the prepared dataset are required for training and the remaining are employed for validating the system during testing process.

ii. Data Acquisition

This is the initial step and helps to create the database by using high-quality leaf images, in which the images are uploaded for testing. Some sample of loaded test images are shown in Fig. 2. Here the image captured is passed through digitizer that converts the analog image into a digital image. Mainly two steps are performed in digitization one is sampling and second is quantization that converts color image into a gray image [15].
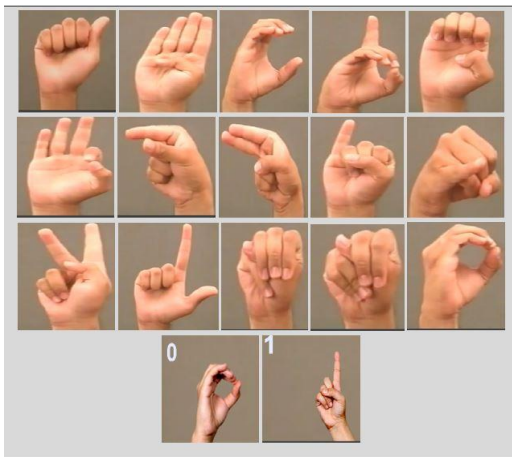

Fig. 2 Uploaded Samples for Alphabets and Number

A. Image Pre-processing

For pre-processing of image the hand area from the surrounding environment is extracted. The pre-processing performs the following steps(i) Skin Color segmentation (ii) Morphological Filtering. In any image of gesture of hand the region of hand is red in color and red intensity of those regions is high. Therefore, the selection of skin color-based segmentation is best approach to differentiate between foreground and background of image. Also, morphological operations are used to make a better RoI of hand gesture using opening and filling technique. The pre-processed images are depicted in Fig. 3.
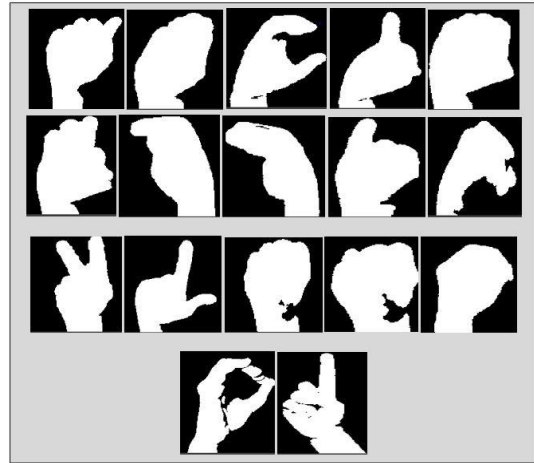

Fig. 3 Segmented Image After Pre-Processing

It also extracts unique features leading to more classification accuracy. To obtain them the selected image goes through different processes including finding key points. Here key points are found based on the maxima/minima of Difference of Gaussians (DoG) [9] which is calculated using Gaussian Filter as given in Equation 1 where P(u,v,σ)

$$Q(u, v, \sigma) = P(u, v, L_i, \sigma) - P(u, v, L_j, \sigma) \quad (1)$$

is the convolution of the hand ROI part R(u,v) with the Gaussian blur G (u,v,L_σ,) at scale L_σ. Equation 1 is used to filter the key-points which are affected by scaling of image.

$$P(u, v, \sigma) = G(u, v, L_\sigma,) \times R(u, v) \quad (2)$$

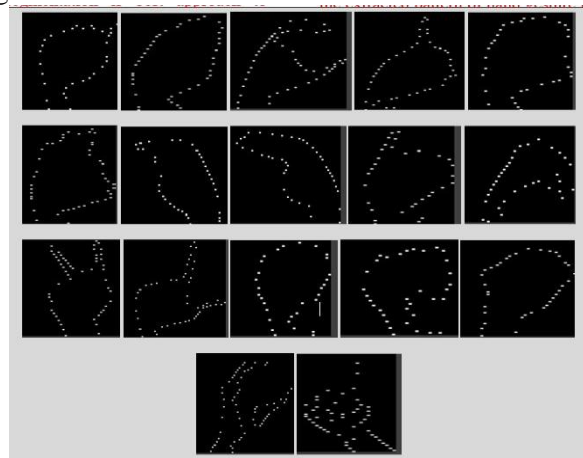The extracted images using feature extraction are shown in Fig. 4.

Fig. 4 Feature Extracted Images

B. Feature Optimization and Classification

In the optimization phase Artificial Bee Colony is used to optimize the extracted features using SIFT based on the fitness function which is called feature selection method. The uniqueness of SIFT feature is more and the chances of availability of common feature still remains. The algorithm is used to select the better feature set using the fitness function of ABC optimization technique. Equation 3 defines the fitness function of ABC optimization which indicates that if any feature in feature set fulfills this condition then it is considered as useful feature otherwise replaced by threshold value of feature set.

Retrieval Number F8124088619/2019©BEIESP
DOI: 10.35940/ijeat.F8124.088619
Journal Website: www.ijeat.org

1458

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

For the fitness function, we use the concept that uses mean value of feature set as threshold because the similarity of feature is more for lower feature value and higher feature value always differs from others. Using this concept, we create an optimal feature set for each category like A, B, C…Z and 0, 1, 2...9. For classification optimized feature sets act as inputs of Feed Forward Back Propagation Network to train the proposed ISLR system. FFBPNN works well with multi-class and has the capability to adjust weight automatically with high tendency to discard irrelevant feature. In this research, optimized feature set is utilized as input of the network and diverse types of hand gestures are used as target of network. On the basis of training data and

target, neural network trains the system. The combination of ABC and FBPNN results in optimizing Feed Forward with different validation parameters and cross-validating it with Back Propagation mechanism. Fig. 5 represents classified images.
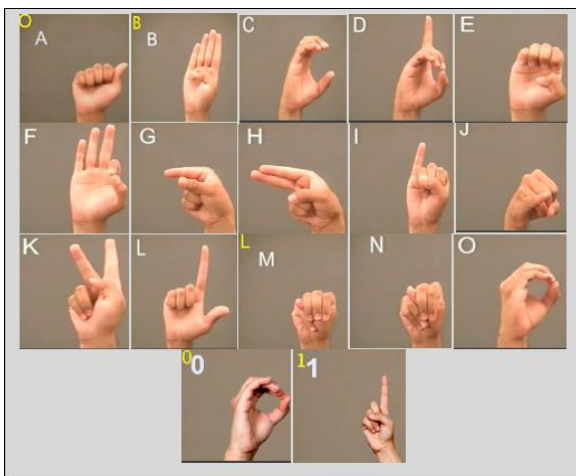

Fig. 5 Classified Images

Algorithm: ABC for ISLR Optimization

Input: Extracted SIFT key points
Call the ABC with their function such as
Employ Bee (E_{bee})
Onlooker Bee (O_{bee})
Scout Bee (S_{bee})
Define the fitness function,

$$FitnessFunction, fit(x) = \begin{cases} True \ (E_{bee}), \ E_{bee} > O_{bee} \\ False \ (O_{bee}), \ E_{bee} \leq O_{bee} \end{cases}$$

(3)

A blank array is designed to save an optimized keypoints, $ABC_{data}= []$
Initialize Cycle = 0
Key Feature = keypoints of image which act as food source
Determine rows (R) and columns (C) of key features = Size (E_{bee})
For i = 1 → R
For j =1 → C
E_{bee} =Required Key Feature (i, j) // *The current feature from the feature set is consider as employ bee*
O_{bee} = Average (key Feature) // *Onlooker is marked as threshold value which helps in the feature selection*
Fit_bee = $fit(E_{bee}, O_{bee})$// *using the equation 6 optimal feature is selected*
ABC $_{data}$(i, j) = Desired_bee
End

End
Return; ABC_{data}as an optimized hand gesture features
End

Algorithm: FFBPN for ISLR Classification

Input: Trained ANN using Optimized key points of hand gesture (T), Target (G) and initialize Neurons (N)
The structure of ANN is initialized using parameters;
– Number of Epochs (E)
– Set MSE as a desired Performance parameters
– Defined Levenberg-Marquardt as a training approach
–Random selection of trained data
Initialize for loop to all trained data
Group (1st)= If Category belongs to 1st group of T such as letter (a)
Group (2nd)= If Category belongs to 2nd group of T such as letter (b)
Group (3rd)= If Category belongs to 3rd group of T such as letter (c)
.
.
.
.
Group (nth)= If Category belongs to nth group of T such as letter (z)

End
Initializing the ANN using T and G
Net = Newff (T, G, N)
ANN set the training parameters automatically and trains the system
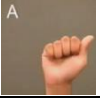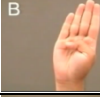Net = Train (Net, T, G)
Return; provided trained ANN structure
End

IV. RESULTS AND DISCUSSION

Proposed ISLR system has been simulated in MATLAB version 2019. The performance of the system is measured in terms of error, precision, recall, F-measure and accuracy as shown in Table 1. Accuracy is the major parameter on which research is focused on as it leads to efficiency. From the Table 1average accuracy determined for the proposed ISLR system is about 99.43%. The detection accuracy in the existing work by Rao et al. [24] is about 90%. The increase in accuracy is due to the additional use of feature optimization with feature classification algorithm. The comparative graph for the detection accuracy is shown in Fig. 6 (a). The average increase in the detection accuracy of the proposed work is 10.48% as shown in Fig. 6 (b). Also, error observed during simulation error rate is 0.56% which results due to uniqueness of features. The more unique are the feature sets less is the chance of irrelevant features during the classification process which helps in getting high accuracy with minimum error rate. Similarly, by using the concept of feature selection technique, most of irrelevant features are replaced with the threshold feature value of similar type of data, so the rate of true samples is more in classification process resulting in precision rate of 0.9867 and recall rate of 0.9803.As both rates are high, this results in F-measure being 0.9835 which is the harmonic mean of precision and recall.

Table I: Computed performance parameters

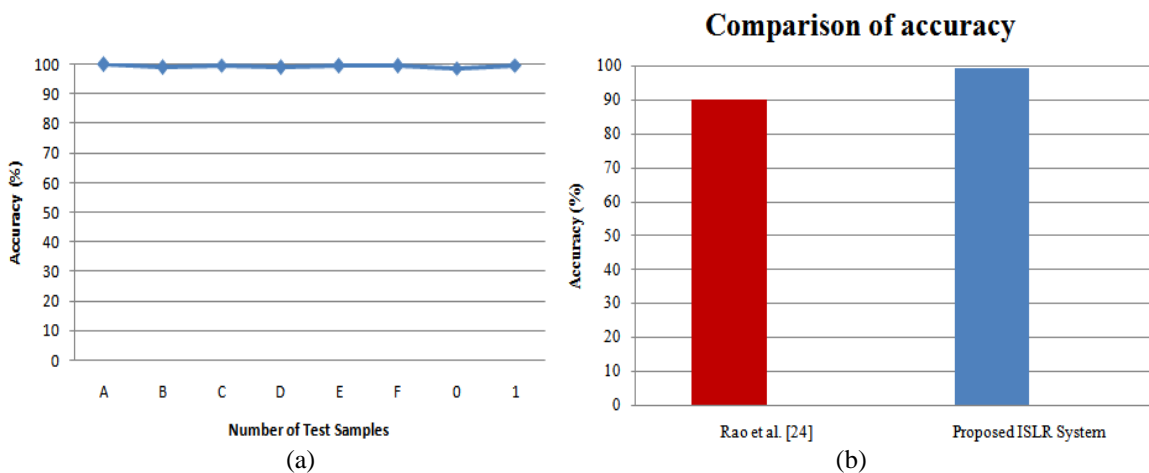| Test Samples	Error (%)	Precision	Recall	F-measure	Accuracy (%)
A	0.0080	0.98642	0.98134	0.98388	99.991
B	0.96143	0.98718	0.9797	0.9834	99.03
C	0.0921	0.9868	0.98119	0.9839	99.90
D	0.9808	0.9868	0.9785	0.98269	99.01
E	0.5135	0.9867	0.9793	0.9830	99.48
F	0.5403	0.9868	0.9794	0.98309	99.45
0	1.25594	0.98667	0.98154	0.9841	98.74
1	0.1566	0.98636	0.98178	0.98408	99.84

(a)   (b)

Fig. 6:(a) Accuracy Obtained for Proposed ISLR System (b) Comparison of Accuracy of Proposed ISLR System with Rao et al. [24]

V. CONCLUSION

The barrier of communication is always present between normal and differently abled people. The aim of this research is to provide a medium by contribution in the ISLR research. An ISLR system is proposed for all English alphabets (A,B…Z) and numbers (0-9). Here by utilizing ABC and FBPNN with SIFT as a descriptor, average accuracy of 99.43%is reached. This is also compared with Rao et al [24] and an increase in 10.48% is detected. Furthermore, the created ISL dataset has also been made publicly online for further research. In future, research can be expanded in dynamic images using classification approach like SVM which uses neural network and optimization technique such as ABC with Greedy Algorithm.

REFERENCES

1. Baus, Cristina, Eva Gutiérrez, and Manuel Carreiras, "The role of syllables in sign language production," *Frontiers in Psychology*, vol. 5, pp. 1254, 2015.
2. Mohandes, S. Mohamed, Aliyu, and M. Deriche, "Arabic sign language recognition using the leap motion controller," *IEEE 23rd International Symposium on Industrial Electronics*, pp. 960-965, 2014.
3. A. R. Vega, A. Vasquez, W. Amador, and A. Rojas, "Deep learning for the recognition of facial expression in the Colombian sign language," *Annals of Physical and Rehabilitation Medicine*, vol. 61, 2018.
4. Karayılan, Tülay, and ÖzkanKılıç, "Signlanguage recognition," *Computer Science and Engineering (UBMK),* pp. 1122-1126, 2017.
5. Ahmed, Washef, KunalChanda, and Soma Mitra, "Vision-based hand gesture recognition using dynamic time warping for Indian sign language," *Information Science (ICIS), International Conference*, pp. 120-125, 2016.
6. D. Karaboga, and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm," *Journal of Global Optimization*, vol. 39, no. 3, pp. 459-471, 2007.
7. D. Karaboga, and B. Basturk, "On the performance of artificial bee colony (ABC) algorithm," *Applied Soft Computing*, vol. 8, no.1, pp, 687-697, 2008.
8. X. He, and S. Xu, "Artificial neural networks," *Process Neural Networks: Theory and Applications*, pp. 20-42, 2010.
9. S. Agatonovic-Kustrin, and R. Beresford, "Basic concepts of artificial neural network (ANN) systeming and its application in pharmaceutical research," *Journal of Pharmaceutical and Biomedical Analysis*, vol.22, no.5, pp.717-727, 2000.
10. AS. Ghotkar, R. Khatal, S. Khupase, S. Asati, M. Hadap (2012) "Hand gesture recognition for Indian sign language," *ICCCI International Conference on Computer Communication and Informatics*, pp 1–4,2012.
11. A. Nandy, JS. Prasad, P. Chakraborty, GC. Nandi, S. Mondal ,"Classification of Indian sign language in real time," *Information Technology and Intelligent Transportation Systems,* vol 2,pp 52–57, 2010.
12. P. V. V. Kishore, "Motionlets Matching with Adaptive Kernels for 3D Indian Sign Language Recognition," *IEEE Sensors Journal,* 2018.
13. PS. Rajam, G. Balakrishnan, "Real time Indian sign language recognition system to aid deaf-dumb people*," IEEE 13th International Conference on Communication Technology (ICCT)*, pp 737–742, 2011.
14. D. Deora, N. Bajaj, "Indian sign language recognition," *in IEEE 1st International Conference on Emerging Technology Trends in Electronics, Communication and Networking (ET2ECN)*" pp. 1–5, 2012.
15. SC. Agrawal, AS. Jalal, C Bhatnagar, "Redundancy removal for isolated gesture in Indian sign language and recognition using multi-class support vector machine," *International Journal of Computational Vision and Robotics,* vol 4, no. 2, pp. 23–38, 2014.
16. N. H. Dardas, & Georganas "Real-time hand gesture detection and recognition using bag-of features and support vector machine" *IEEE Transactions on Instrumentation and Measurement*, vol 60, no 11, pp. 3592-3607,2011.
17. R. Jayashree Pansare, H. Sharavan Gawande, Maya Ingle, "Real-Time Static Hand Gesture Recognition for American Sign Language (ASL) in Complex Background", *Journal of Signal and Information Processing,* vol.3, pp 364-367, 2012.
18. Kumar, Anup, Karun Thankachan, "Sign language recognition," *Recent Advances in Information Technology (RAIT), 3rd International Conference,* 2016.
19. Rajesh Mapari, Dr.Govind Kharat, "Hand Gesture Recognition using Neural Network", *International Journal of Computer Science and Network*, vol.1, no. 6, December 2012.
20. A. A. Randive, H. B. Mali, & S. D. Lokhande, "Hand gesture segmentation," *International Journal of Computer Technology and Electronics Engineering*, vol 2,no. 3, 2012.
21. N. C. Camgöz, S. Hadfield, O. Koller, and R. Bowden, "SubUNets: End-to-End Hand Shape and Continuous Sign Language Recognition," *International Conference on Computer Vision (ICCV),* vol. 1, pp. 22-27, 2017.
22. M. V. Beena, M. A. Namboodiri, and P. G. Dean, Automatic Sign Language Finger Spelling Using Convolution Neural Network: Analysis," *International Journal of Pure and Applied Mathematics*, vol. 117, no. 20, pp. 9-15,2017.
23. T. Shanableh, K. Assaleh and M. Al-Rousan, "Spatio-Temporal Feature-Extraction Techniques for Isolated Gesture Recognition in Arabic Sign Language," *in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),* vol. 37, no. 3, pp. 641-650, June 2007.
24. G. Rao, Ananth, and P. V. V. Kishore, "Selfie video based continuous Indian sign language recognition system," *Ain Shams Engineering Journal,* vol 9, no. 4, pp -1929-1939 2017.
25. P. Kumar, H. Gauba, P. P. Roy, and D. P. Dogra, "Coupled HMM-based multi-sensor data fusion for sign language recognition," *Pattern Recognition Letters*, vol. 86, pp. 1-8, 2017.

AUTHORS PROFILE

Name: Jasmine Kaur, Designation : ME Research Scholar Qualifications: ME from NITTTR Chandigarh Contact Detail : Dept. of Computer Science & Engineering, NITTTR Chandigarh -160019 Specialization: Digital Image Processing, Machine Learning, Data Analysis
Email : jasminekaur741@gmail.com

Name : C. Rama Krishna, Designation : Professor Qualifications: Ph.D. from IIT Kharagpur Contact Detail : Dept. of Computer Science & Engineering, NITTTR Chandigarh -160019 Specialization: Wireless Communications & Networks, Computer Networks, Distributed Computing, Cyber Security and Cryptography
Email : rkc_97@yahoo.com