Exposition into Indian Sign Language Recognition using Deep Convolutional Neural Network

Aishwarya T[1] Dr. Ravi Kumar V[2]
[1]M. Tech Student [2]Professor and Head
[1,2]Vidyavardhaka College of Engineering, Mysuru, India

Abstract— A Sign Language is the most innate and articulate way to communicate for the hard to hear. Hand gesture is a method used in sign language for non-verbal communication. It is commonly used by deaf and dumb people who have hearing or speech problems, to communicate among themselves or with normal people. Pattern recognition and Gesture recognition are the emerging fields of research. Indian Sign language is the commonly used sign language among deaf people in India. Various parts of India has different signs but grammar is same throughout the country. Various sign language systems has been developed by many makers around the world and various methods are being used to recognize these signs. Hence, in this paper, we compare various methodologies that aid us in building models that help us in recognizing the Indian Sign Language.

Keywords: Sign Language, Hand Gestures, Indian Sign Language (ISL), Sign Language Recognition, Convolutional Neural Network, Artificial Neural Network, Eigen value, Eigen Vector

I. INTRODUCTION

Sign language recognition is helpful in communication between signing people and non-signing people. Various research projects are in evolution on different sign language recognition systems worldwide. There are country wide variations available but the research is limited to a particular country. Sign language is a visual-gestural language used by deaf and hard-hearing people. Three dimensional spaces and the hand movements are used to convey meanings. Own vocabulary and syntax are present which is entirely different from written and spoken languages. A gesture may be explained as a movement, usually of hand or face that can demonstrate an idea, sentiment or emotion, e.g., rising of eyebrows, shrugging of shoulders is some of the gestures we use in our everyday life. Sign language is a more organized and defined way of communication in which every word or alphabet is assigned to a particular gesture. With the rapid breakthroughs in technology, the use of computers in our daily life has increased manifolds.

There are two main approaches deployed in the sign language recognition, that is Glove or Device based and Vision based. In the glove based method the user has to wear a device which carries a huge numbers of cables so as to connect the device to a computer. Such devices are expensive and reduce the primitiveness of the sign language communication. While on the contrary, the Vision based method requires only a camera and directly works with image gestures. It is a two step process: sign capturing and sign analysis. Vision based methods provide a natural environment to the user and reduces the disadvantages and complications as in the glove based method [1].

Around 5% of world community in all parts of the world is using sign language as a medium of communication. Regionally different languages have been evolved as ASL (American Sign Language) in America, GSL (German Sign Language) in Germany, BSL (British Sign Language) in the UK or ISL (Indian Sign Language) in India [2]. Therefore, many researchers came forward with various techniques and models that could be used to recognize the sign languages efficiently.

II. INDIAN SIGN LANGUAGE

Indian Sign Language was developed so that the deaf people can interact with the normal people without any complexity or perplexity [3].

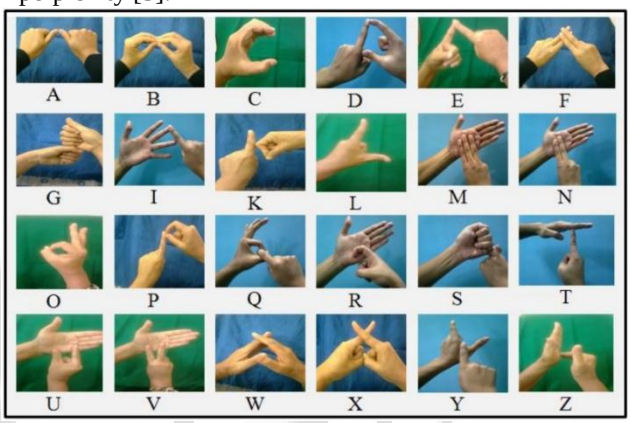

Fig. 1: Various alphabets of Indian Sign Language.

Figure 01 depicts the various signs for the corresponding English alphabet.

Gesture Recognitions and Sign Language recognition has been a well-researched topic for the ASL, but not so for ISL. Few research works have been carried out in Indian Sign Language using image processing or vision techniques. Most of the previous works found either analyzed what features could be better for analysis or reported results for a subset of the alphabets. No standard datasets for Indian Sign Language is present. Using two hands leads to occlusion of features. Variance in sign language with locality and usage of different symbols for the same alphabet by the same person is also a challenge.

III. ISL RECOGNITION USING MULTICLASS SUPPORT VECTOR MACHINE

An Indian sign language recognition method is based on the consolidation of invariant moment and shape descriptor features. Here, they look at only manual signs which comprises of hand gesture of isolated signs.

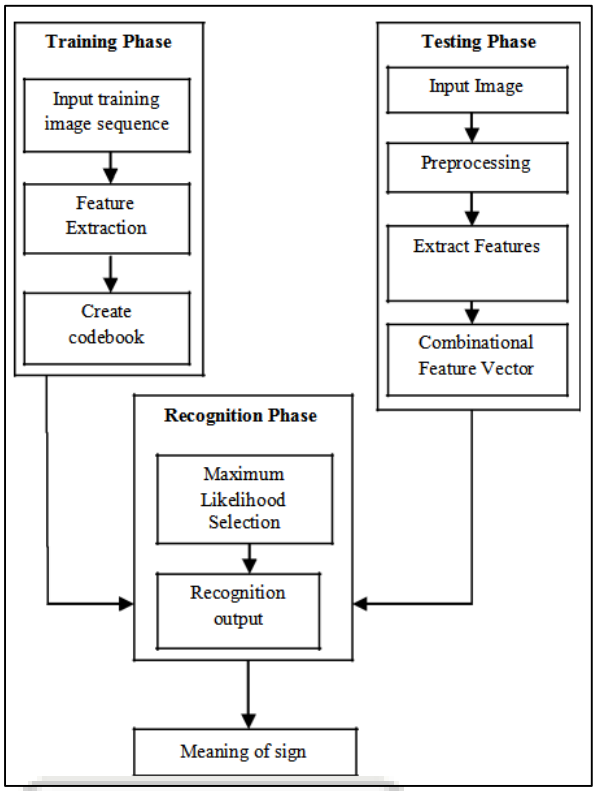

Fig. 2: Block diagram of hand gesture recognition system

Preprocessing is applied to images before extracting features from hand images. Preprocessing consists of two steps, segmentation and filtering. All the functions are applied on a gray scale image. The segmentation of an input image of a hand gesture is performed using Global thresholding algorithm. Hu invariant moment is deployed for scale and position invariant pattern identification. The pre-eminence of using Hu invariant moment is that it can be used for disjoint shapes. In particular, Hu invariant moment set consists of seven values computed by normalizing central moments through order three [4].

A multi-class support vector machine has been employed to correctly and appropriately classify the hand gestures among multiple classes. In the proposed approach, a binary classifier is converted into multiclass classifier. The ijth binary classifier uses the pattern of class i as positive and the patterns of class j as negative examples. For the final outcome, minimum distance of the generated vector has been calculated to the binary pattern representing each class [1].

IV. ISL RECOGNITION USING ORIENTATION HISTOGRAM WITH PRINCIPAL COMPONENT ANALYSIS

This paper focuses on continuous ISL gesture recognition system. The Dataset consists of a collection of signs where single hand or both the hands have been used for performing continuous ISL gestures. The Database has been created to have ten sentences. Every sentence is a combination of static and dynamic gestures. Extracting start frame and end frame of each gesture is the main problem in continuous sign language gesture recognition system because it consists of a group or collection of meaningful gestures and also a vague gestures that have no meaning. Gradient based key frame extraction method is used to deal with this problem. Here, major change in the gradient shows end of the one gesture and start of another gesture. Key frame helps to break each sentence into sequence of words and also obliging for extracting frames of meaningful gestures [5]. Orientation histogram, DWT and PCA is used for extracting features of those frames which comprises of meaningful gestures. General diagram of proposed framework is shown in figure 03.

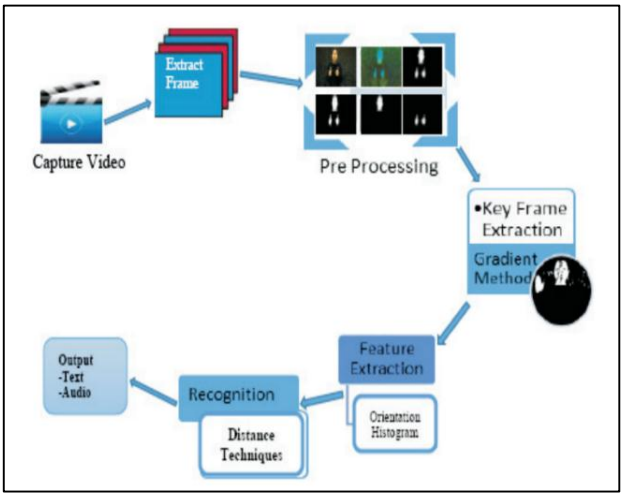

Fig. 3: Proposed general framework

Orientation histogram is used as a feature extraction technique for extracting most appropriate features of each gesture. It provides convenience to even light condition changes and scene illumination changes. The edges of the images of the scene will be still the same. All the continuous ISL gestures have been obtained in normal lighting mode where pixel intensities can be suggested to change the scene lighting. Another upper hand which is employed on orientation histogram is translation invariant property. It determines that the same feature vectors will be produced by the same frames at different position of gestures. It is achieved to measure the local orientation histogram for all the frames of the dynamic gestures.

V. ISL RECOGNITION USING EIGEN VALUES, EIGEN VECTORS AND EUCLIDEAN DISTANCE

In this paper a novel approach for recognizing various alphabets of Indian Sign Language is proposed where continuous video sequences of the signs have been considered. The proposed system comprises of three stages: Preprocessing stage, Feature Extraction and Classification. Preprocessing stage includes skin filtering, histogram matching. Eigen values and Eigen Vectors were considered for feature extraction stage and finally Eigen value weighted Euclidean distance is used to recognize the sign. It deals with bare hands, thus allowing the user to interact with the system in natural way [3].

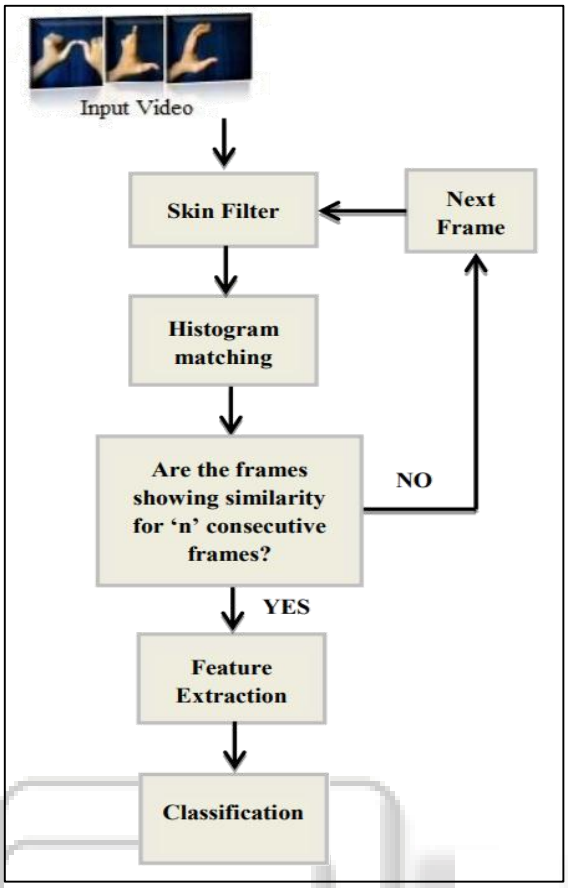

Fig. 4: Proposed system of the paper

The proposed system is shown in Figure 04 which comprises of 3 major stages-preprocessing stage which includes the skin filtering and histogram matching to find out the similarity between frames, Feature Extraction stage in which the Eigen values and Eigen vector are being considered as features and finally Eigen value weighted Euclidean distance based classification technique [6].

VI. ISL RECOGNITION USING NEURAL NETWROKS AND KNN CLASSIFIERS

After gathering and collecting the database of images, the images are pre-processed. The RGB images are converted to gray scale image by rgb2gray function available in Matlab environment. It converts the true color image RGB to the gray scale intensity image. The function converts RGB images to gray scale by eliminating the hue and saturation information while retaining the luminance. Derivative Sobel edge detector method was used because it computes gradient by using discrete difference between rows and columns of 3×3 neighbors.

Feature extraction is a kind of dimensionality reduction. Input images are too large for processing, so to process these images in time, reducing the dimension of the input image by feature extraction is done. Transforming input data into feature is called feature extraction. Feature extraction is chosen in such a way that image information must be retained. Feature extraction techniques used here are direct pixel value and hierarchical centroid [7].

K-nearest neighbor (kNN) classifier classifies objects on the basis of feature space. kNN employs supervised learning algorithm. Nearest neighbor algorithm is most popular classification technique proposed by Fix and Hodges. kNN classify method classifies each row of the data in sample into one of the groups in training using the nearest-neighbor method.

Nueral Network Pattern Recognition Tool (NPR Tool) leads through solving a pattern-recognition classification problem using a two-layer feed-forward patternnet network with sigmoid output neurons. Twenty five neurons were used on the hidden layer. The network has ten output neurons because there are ten target values associated with each vector. Pattern recognition networks are feed forward networks that can be trained to classify inputs according to target classes.

VII. ISL RECOGNITION USING DEEP CONVOLUTIONAL NEURAL NETWORK

This paper proposes the recognition of Indian sign language gestures using a powerful artificial intelligence tool, convolutional neural networks (CNN). Selfie mode continuous sign language video is the capture method used, where a hearing-impaired person can use the SLR mobile application independently. CNN training is performed with 3 different sample sizes, each consisting of multiple sets of subjects and viewing angles. The remaining 2 samples are used for testing the trained CNN. Different CNN architectures are designed and tested with selfie sign language data to obtain better accuracy in recognition. The model is constructed with input layer, four convolutional layers, five rectified linear units (ReLu), two stochastic pooling layers, one dense and one SoftMax output layer. Figure 05 shows the proposed system architecture [8].

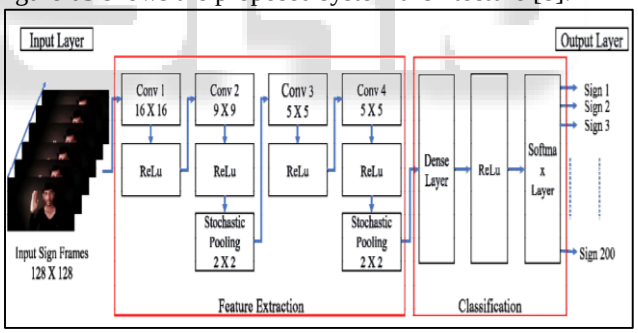

Fig. 5: Deep CNN architecture

The proposed CNN architecture uses four convolutional layers with different window sizes followed by an activation function and a rectified linear unit for non-linearities. The convolutional windows are of size 16*16, 9*9, 5*5 and 5*5 from layer 1 to 4 respectively. Three kinds of pooling strategies are tested via mean pooling, max pooling, stochastic pooling and found that stochastic pooling is suitable for our application. The feature representation is done by considering two layers of stochastic pooling. Only two layers of pooling is initiated to avoid a substantial information loss in feature representation. Classification stage is implemented with dense/fully connected layers followed by an activation functions. Softmax regression is adopted in classification. Using CNNs are advantageous when it comes to image recognition as they are curated to help in this process. Therefore, they are more compatible and easier to use and provide a good accuracy when compared to other machine learning algorithms.

VIII. CONCLUSION

Sign Language Recognition has emanated as one of the important domains of research in Computer Vision. The complexity faced by the researchers is that the instances of signs differ with both motion and appearance. A real-time sign language translator is an important milestone in helping and facilitating communication between the deaf community and the general public. There are numerous methods that have been designed that aids in recognising the sign languages. Different Machine Learning and Deep learning algorithms have been employed that provides a means to build a platform that helps in ISL recognition.We see after the analysis of these methods that CNNs have an upper hand when it comes to recognition.

REFERENCES

[1] Karishma Dixith, Anand Singh Jalal, "Automatic Indian Sign Language Recognition", 3rd IEEE International Advance Computing Conference (IACC), 2012.
[2] Madhuri Sharma, Ranjna Pal and Ashok Kumar Sahoo, "Indian Sign Language Recognition Using Neural Networks and KNN Classifiers", ARPN Journal of Engineering and Applied Sciences, VOL 9 - No. 8, August 2014.
[3] Joyeetha Singha, Karen Das, "Recognition Of Indian Sign Language In Live Video", International Journal of Computer Applications (0975 – 8887) Volume 70 - No.19, May 2013.
[4] M. K. Hu, "Visual Pattern Recognition by Moment Invariants," IRE Transactions on Information Theory, vol. 8, pp. 179-187.
[5] Kumud Tripathi, Neha Baranwal and G. C. Nandi, "Continuous Indian Sign Language Gesture Recognition and Sentence Formation", Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015).
[6] Singha, J. and Das, K. "Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique", International Journal of Advanced Computer Science and Applications, Vol. 4, No. 2, 2013, pp. 188-195.
[7] Madhuri Sharma, Ranjna Pal and Ashok Kumar Sahoo, "Indian Sign Language Recognition Uisng Neural Netwroks And KNN Classifiers", ARPN Journal Of Engineering And Applied Sciences, Volume 08, No - 9, August 2014.
[8] G. Anantha Rao, K. Syamala, P.V.V.Kishore1, A.S.C.S.Sastry1, "Deep Convolutional Neural Networks for Sign Language Recognition", IEEE, SPACES-2018.

