Recognition of Indian Signed Language using Images of Hand Gestures

Shanthi M. B[1], NitishHegde[2],Omkar Patra[3], Meera Krishna P[4].
Department of Computer Science and Engineering, CMR Institute of Technology, Bengaluru[1,2,3,4]

Abstract:
Sign language predominantly uses hand gestures for nonverbal communication modalities. People with speech or hearing disabilities use them to communicate with others. Many companies have developed several different sign language system, but they are not flexible to adopt to various requirements or they are expensive for the user. In this paper we propose a prototype system to automatically recognise sign language in order to help people with speech or hearing disabilities toeffectivelycommunicate with others.
Gesture recognition and more generally pattern recognition are two fields of study that are still in their early stages of development. We humans tend to use hand gestures for nonverbalcommunication in our day to day life. The Hand Gesture Recognition System offers auser friendly, natural and innovative way of communicating with computers. The system we propose is based on a real-time application for hand gesture recognition that candetect features based on shapes like figure status like folded or raised fingers, folded fingers, centre of mass, orientation and centroid, basedon on the shape of a human hand, which has a thumb and four fingers.
Key words: hand gesture, non-verbal,pattern recognition, gesture recognition,sign language,

I. INTRODUCTION

Ease of communication for people with speech or hearing disabilities depend a lot on sign languages, because of whichautomatic recognition of sign language is studies for years.Gesture recognition is a rapidly expanding area of image processing and artificial intelligence. Gesture recognition is a method of identifying and using the movements or postures of human body parts to operate computers and other electronic devices. Every action in sign language has a specific meaning, due to which different basic actions can be combined to describe a complex meaning.Sign language is a non-verbal way of interaction that helps deaf and verbally challenged people to communicate more successfully with each other and with the general public. According to World Health Organization over 5% of the world population are suffering with loss of hearing including children.One tenth of world population will suffer from hearing loss by 2050 [10]. It has particular principles and language for efficiently expressing oneself. There are two basic ways to sign language recognition: vision-based and sensor-based. Vision based system particularity has attracted lot of researches they avoid the requirement for sophisticated setup like helmets,hand gloves and such other accessories that the sensor-based approaches extensively use. Image comprehension and sign recognition are linked. It is typically approached in2 phases, first being the detection, followed by sign recognition. It employs some image pre-processing

processes for elimination after receiving the input sequence of images collected via web camera.

II. RELATED WORK

There are two approaches for recognising hand movements in general. Recognition of gestures using eyes and sensors. Many experiments have been conducted on sensor-based systems such as gloves, wires, helmets, and so on, but because continuous wear is not feasible, image-based techniques are receiving greater attention. HMM (Hidden Markov Model), ANN (Artificial neural network), Eigen value based, and perceptual colour-based techniques have all been used to recognise gestures. For categorization and particle filtering, methods such as Support Vector Machines (SVM), KNN, and CNN have been proposed. In their studies, Aditya V.et al.[1] and T. L. Baldi et al. [2] employed data glove to extract hand gesture data. They discovered improved performance, but they ran into the issue of having to wear a hand glove all of the time to collect data.

Sakshi Sharma et al. [3] discussed an innovative and robust model for hand gesture detection called G-CNN. For feature learning, they employed CNN, and for classification, they used the SoftMax classifier. P. K Sharma and colleagues [4] created a user-programmable glove that registers hand and finger gestures in three dimensions. This Acela Glove is utilised in video games for sports training or physical rehabilitation [5]. Dong Xu Li et al. published a paper in which they worked on a large-scale Word-Level American Sign Language (WLASL) video collection with over 2000 words performed by over 100 signers. Their findings reveal that pose-based and appearance-based models perform similarly on 2,000 words/glosses, with top-10 accuracy of up to 62.63 percent.In paper [7] Saad M. et al. have used type-2 fuzzy HMM that uses shape information for recognition of the gesture. They could achieve 95% accuracy in recognizing the hand gestures. In [8], Phat Nguyen Hu et al. have discussed about the algorithm for recognizing gestures usinghistogram of oriented gradient (HOG) and support vector machines (SVM).Their algorithm has given accuracy of 99% with execution time of 70 ms for each frame, which is accepted for industrial scale applications.Most of the research works have given focus on collecting data based on video data or data from glove equipped with sensors. There is an ample scope for refinements in the current approaches to increase the accuracy in hand gesture recognition. In this paper we have given focus on implementing an image-based hand gesture recognition technique for sign language to analyse the accuracy level of hand gesture recognition.

III. PROPOSED WORK

Gesture acquisition, segmentation and tracking, extraction of features, identification of gestures are the main processes in creating a sign language recognition system. The acquisition of gestural data is the first stage in a gesture recognition system. Images can be captured using a simple digital cameraor an external webcamor abuild in webcam in a computer or a laptop. A pre-existing database or one created by the researchers themselves can be used. Fig.1 shows the commonly used hand gestures in Indian Signed Languages.

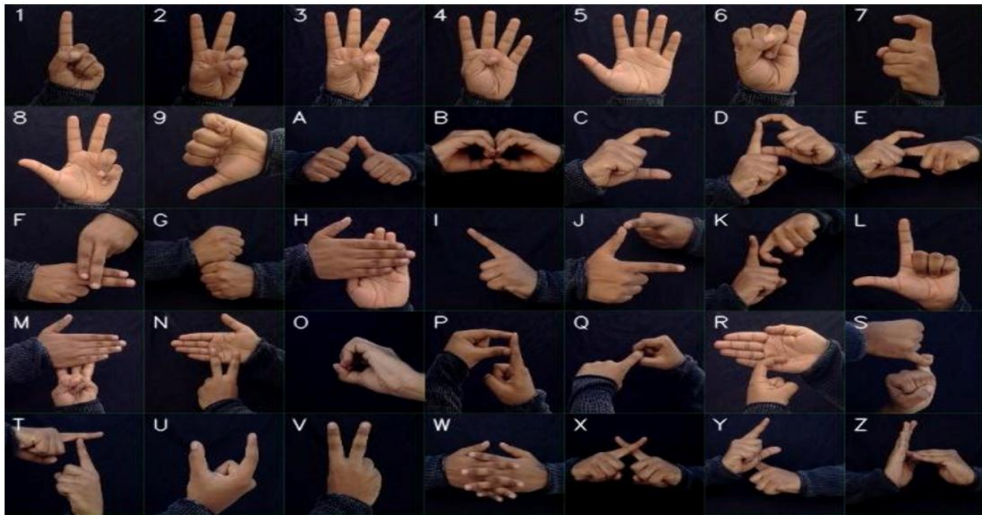

Fig1. Hand gestures used in Indian Signed Languages (Source: Internet images)

Hand segmentation is required to track the movement of the hand. ROI and background are separated through segmentation. Hand tracking is a method for determining the position of a hand.Following the successfully finishing the trackingand its segmentation of hand, the next stage is extraction of features, which is used to extract significant features. The entire recognition process may be broken down into the following two stages: training and testing. The classifier is trained using the training database at the initial stage. The development of a database, pre-processing, feature extraction, and classifier training are all important phases in the training stage. Gesture acquisition, pre-processing, feature extraction, and classification are all important processes in the testing phase. The block diagram [2][5] of a sign language recognition system is shown in Figure 2.

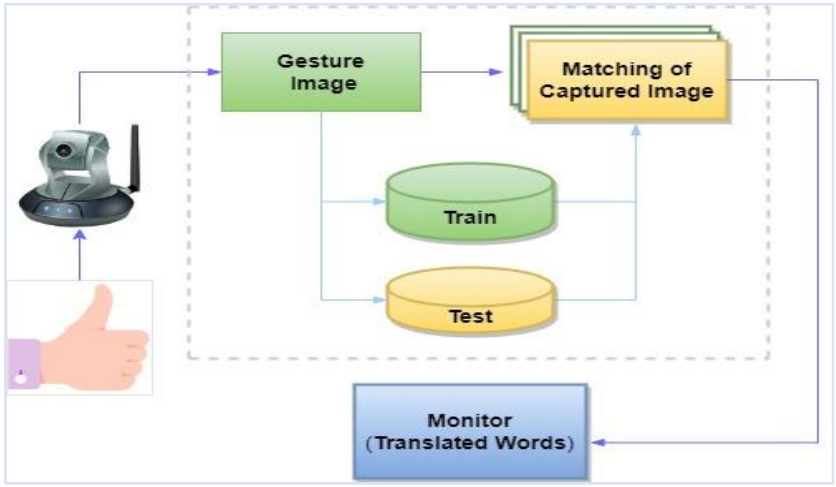

Fig.2 Block diagram of the ISL recognition system

IV. IMPLEMENTATION AND RESULTS

The proposed work has been implemented using Convolution Neural Network (CNN) classifier as it performs well in image classification problems asthey can reach a much tighter fit to the data. Convolution layer is forming the elementary block which takes image input and generates feature map after convolving the filter of particular size.Kernel convolution has been computed as per the Eq.1.

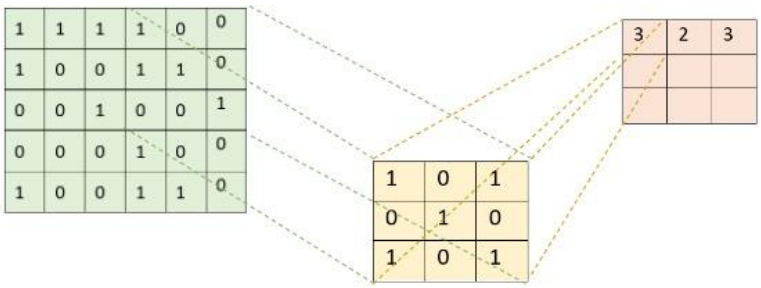

Fig.3 Kernel Convolution

$$G[m,n] = (k * h)[m * n] \quad (1)$$

$$ie, \quad G[m,n] = \sum_i \sum_j h[i,j]k\,[m-i][n-j] \quad (2)$$

Kernel convolution process normally reduces the image size continuously. To avoid this problem, To maintain the information contained in the image, CNN adds a border as padding by adding zero bits to the filter. The padding factor P is decided based on the size of the filter as given in the Eq.3.

$$p = \frac{f-1}{2} \quad (3)$$

Where, P is the padding factor and f is the size of the filter. The dimension of the output matrix is calculated by considering the padding and stride by using the Eq.4.

$$m_{out} = \left\lfloor \frac{m_{in} + 2p - f}{s} + 1 \right\rfloor \quad (4)$$

Convolution over volume is carried out by using multiple filters at the same layer and stacking the results one on top of the other to get a combined output. The dimension of the resulting tensor is computed by using Eq.5.

$$[m, m, m_c] * [f, f, m_c] = \left[\left\lfloor \frac{\lfloor m + 2p - f \rfloor}{s} + 1 \right\rfloor, \left\lfloor \frac{\lfloor m + 2p - f \rfloor}{s} + 1 \right\rfloor, m_f \right] \quad (5)$$

Where, m is the image size, f is the filter size, m_c is the number of channels in the image with p padding factor using stride s and using m_f number of filters. An activation function is used to generate output activation map from the feature map generated by the convolution layer. It decides whether a neuron is to be activated or not based on the value of weighed sum and added bias.

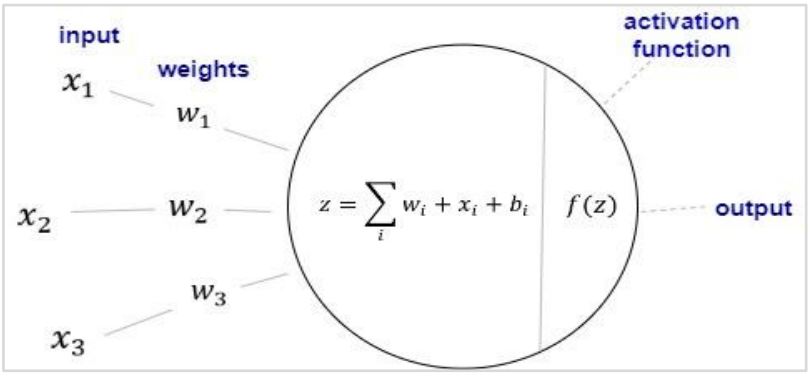

Fig. 4 Activation function in CNN

Where,x_1, x_2, x_3 are the vectorized inputs, w_1, w_2, w_3 are the vectorized weights assigned to the neurons and b is the vectorized bias assigned to the neurons in the hidden layer and z is

the vectorized output of the layer. We have used RELU (Rectified Linear Unit) activation function for the implementation of the proposed project. The experimental observations are given in the following section.

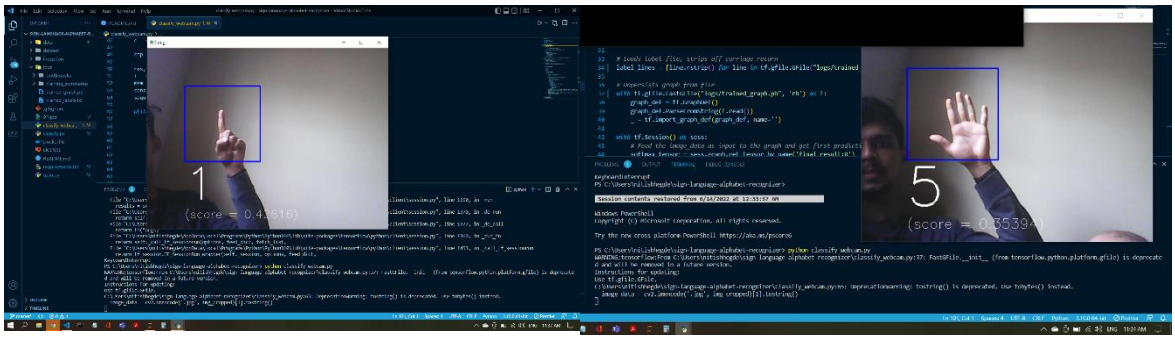

Fig.5 Recognition of Digits based on captured hand gestures

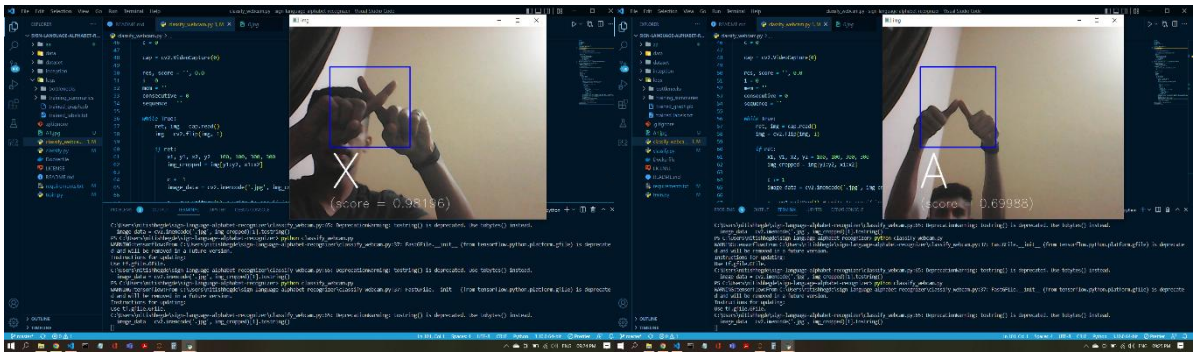

Fig.6 Recognition of Characters using hand gestures

V. CONCLUSION

In this paper, we have proposed an approach to detect alpha-numeric signs of Indian sign language. With the help of inception v3 architecture and CNN techniques we have achieved accurate results in recognizing every ISL signs for multiple static gestures. This work could be extended in future to analyse dynamic gestures under varying conditions and also on recognizing different phrases of Indian sign language.

REFERENCS

1. V. Adithya, P. R. Vinod and U. Gopalakrishnan, "Artificial neural network-based method for Indian sign language recognition," Information & Communication Technologies (ICT), 2013 IEEE Conference on, JeJu Island, 2013, pp. 1080- 1085.
2. T. L. Baldi, S. Scheggi, L. Meli, M. Mohammadi and D. Prattichizzo, "GESTO: A Glove for Enhanced Sensing and Touching Based on Inertial and Magnetic Sensors for Hand Tracking and Cutaneous Feedback," in IEEE Transactions on Human-Machine Systems, vol. 47, no. 6, (Dec. 2017) pp. 1066-1076.
3. SakshiSharma, SukhwinderSingh, ‘Vision-based hand gesture recognition using deep learning for the interpretation of sign language’, Expert Systems with Applicationsdoi.org/10.1016/j.eswa.2021.115657, Volume 182, 15 November 2021, 115657.
4. P. K. Sharma and S. Sharma, “Evolution of Hand Gesture Recognition: A Review,” International Journal of Engineering and Computer Science, vol. 4, no. 1, p. 9963, Jan. 2015.

5. P. Premaratne, Human Computer Interaction Using Hand Gestures, 2014th ed. Singapore: Springer, 2014.
6. Dongxu Li , Cristian Rodriguez Opazo, Xin Yu, Hongdong Li, Word-level Deep Sign Language Recognition from Video: A New Large-scale Dataset and Methods Comparison, https://doi.org/10.48550/arXiv.1910.11006, 2020.
7. Saad M. Darwish, Magda M. Madbouly, and Murad B. Khorsheed, Hand Gesture Recognition for Sign Language: A New Higher Order Fuzzy HMM Approach, IACSIT International Journal of Engineering and Technology, Vol. 8, No. 3, June 2016.
8. Phat Nguyen Huu , Tan Phung Ngoc , Hand Gesture Recognition Algorithm Using SVM and HOG Model for Control of Robotic System, Hindawi Journal of Robotics Volume 2021, Article ID 3986497, 13 pages https://doi.org/10.1155/2021/3986497.
9. Rady El Rwelli, Osama R. Shahin, Ahmed I. Taloba, Gesture based Arabic Sign Language Recognition for Impaired People based on Convolution Neural Network, International Journal of Advanced Computer Science and Applications, Vol. 12, No. 12, 2021.
10. who.int, "Deafness and hearing loss," 2019. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss. [Accessed 1 February 2020].
11. [10] Ashish S. Nikam; Aarti G. Ambekar, Sign language recognition using image-based hand gesture recognition techniques., 2016 Online International Conference on Green Engineering and Technologies (IC-GET).DOI: 10.1109/GET.2016.7916786.
12. Ashish Sharmaa, AnmolMittala, SavitojSingha, VasudevAwatramania, 'Hand Gesture Recognition using Image Processing and Feature Extraction Techniques', International Conference on Smart Sustainable Intelligent Computing and Applications under ICITETM2020. Procedia Computer Science 173 (2020) 181–190.
13. S. Mitra and T. Acharya. Gesture recognition: A survey. Tsmcc 37(3), pp. 311-324. 2007. Available: http://ieeexplore.ieee.org/document/4154947. DOI: 10.1109/TSMCC.2007.893280.
14. M. Al-Rajab, "Hand Gesture Recognition for Multimedia Applications.", ProQuest Dissertations Publishing, School of Computing, University of Leeds, 2008.
15. [14] P. Premaratne, Chapter2: Human Computer Interaction Using Hand Gestures, 2014th ed. Singapore: Springer, 2014.
16. [15] JayeshGangrade, Jyoti Bharti, Vision-based Hand Gesture Recognition for Indian Sign Language Using Convolution Neural Network, IETE Journal of Research,
17. [16] Ghotkar, Archana&Gajanan, K &Kharate, Gajanan. (2014). Study of vision-based hand gesture recognition using Indian sign language. International journal on smart sensing and Intelligent System.
18. K. S. Varun, I. Puneeth, and T. P. Jacob, "Hand Gesture Recognition and Implementation for Disables using CNN'S," 2019 International Conference on Communication and Signal Processing (ICCSP), 2019.

A. A. Alani, G. Cosma, A. Taherkhani and T. M. McGinnity, "Hand gesture recognition using an adapted convolutional neural network with data augmentation," 2018 4th International Conference on Information Management (ICIM), Oxford, pp. 5-12, 2018. doi: 10.1109/INFOMAN.2018.8392660.

19. Xing L. Human-robot interaction based on gesture and movement recognition. Signal Process-Image 2020; 81: 115686.

20. Neiva DH and Zanchettin C. Gesture recognition: a review focusing on sign language in a mobile context. Expert SystAppl 2018; 103: 159–183
21. Dinh DL, Ngoan PTK, Thang ND, A single depth silhouette-based hand gesture recognition for appliance interfaces in smart home environment. In: 6th International Conference on the Development of Biomedical Engineering in Vietnam (ed Vo Van T, et al.). Ho Chi Minh, Vietnam, June 2016. IFMBE Proceedings, vol 63, pp.369–373. Singapore: Springer. doi: 10.1007/978-981-10-4361-1_62.
22. Sagayam KM and Hemanth DJ. Hand posture and gesture recognition techniques for virtual reality applications: a survey. Virtual Real 2017; 21: 91–107.
23. Oyedotun OK and Khashman A. Deep learning in vision-based static hand gesture recognition. Neural Comput&Applic 2017; 28: 3941–3951.
24. Verma B and Choudhary A. Grassmann manifold based dynamic hand gesture recognition using depth data. Multimed Tools Appl 2020.
25. Dominio F, Donadeo M and Zanuttigh P. Combining multiple depth-based descriptors for hand gesture recognition. Pattern RecognLett 2014.