Dynamic Sign Language Interpreter Using Deep Learning

Siddhesh Garud[1], JayeshPatel[2], Mayur Wagh[3], Assistant Prof. Satish Kuchiwale[4]

Students, Dept. of Computer Engineering, Smt. Indira Gandhi College of Engineering, Navi Mumbai, Maharashtra, India.

Professor, Dept. of Computer Engineering, Smt. Indira Gandhi College of Engineering, Navi Mumbai, Maharashtra, India.

Abstract: Sign language recognition has emerged as one of the important areas of research in computer vision. As the community of speech and hearing impaired people has depended on sign language as a communication medium. New techniques have been developed from the past decade still now, to counter the problem of building a communication bridge between normal people and speech and hearing-impaired people. Sign gestures can be classified as static and dynamic. Static gesture recognition is easier than dynamic gesture recognition but both recognition systems are important for the human community This paper focuses on different techniques used for sign language recognition.

Keywords: **Sign Language Recognition, Computer Vision, ISL.**

I. INTRODUCTION

Sign language is the most natural ways of exchanging information among deaf and dumb people. It has been observed that deaf people are facing difficulty interacting with other people. The goal of sign language recognition is to provide an efficient and accurate system to convert sign language into text so that communication between deaf and normal people can be more efficient. Sign language consists of vocabulary of signs in exactly the same way as spoken language consists of a vocabulary of words. Indian sign language (ISL) is sign language used in India. ISL involves both static and dynamic gestures, single as well as double handed gestures, in addition to this the hands involved in gesturing may have complex motion. Some signs include facial expressions too. Because of these difficulties less research work has been done in ISL recognition system. A thorough literature survey covering almost all the aspects of the SLR is a primary step to build a ISL recognition system.

Any non-verbal communication like the motion of hands, facial expression and other body parts is a form of gesture. Gesture recognition enables devices to understand

human actions. Normal people do not want to learn sign language. That's why this community becomes isolated from others, they cannot express themselves. So if computer can be programmed in such a way that it can translate sign language to some speech or text format, the difference between the normal people and the deaf and dumb community can be minimized and then communication between Deaf and Dumb community will be easy. Now there is a requirement for such a type of system which can recognize and translate sign language into text or speech format. Almost all gestures have already assigned meaning and grammar is used to create a meaningful sentence from a set of recognized gestures. ISL gestures are mostly made up of two hands. So it is quite difficult to recognize. Continuous ISL or Continuous Sign language is a sequence of gestures that generate a meaningful sentence. Continuous ISL includes dynamic gesture recognition.

II. APPROACHES

There are basically two main approaches to solve sign language recognition problems.

- **Static gesture**

Static gesture recognition aims to identify particular kinds of posture which remain still for a period of time in the videos. Static gestures are single images which involve no time frame. We use the adaboost algorithm when training cascade classifiers, which can promote systems to be robust and realtime and can be used to identify static gestures..

- **Dynamic gesture**

In sign language recognition systems, gestures cannot be judged at the time they are detected, since the dynamic gestures should be taken into consideration. Namely, we do not judge gestures only by one image but just record the position of the hand because it may be one fragment in a series of image**s** which contains a dynamic gesture such as a wave**.** Dynamic gestures are single images which involves multiple time in frame

III. METHODOLOGY

Sign language recognition systems are being developed in order to provide an interface for the hearing impaired and mute persons. These automatic sign language recognition systems allow the non-signers to interpret the meaning of what the signer wants to convey and therefore facilitate the communication between them. Research in this direction comes under the category of Human Computer Interaction (HCI). Most of the previous work in Indian Sign Language has focused on static signs and images of signs with constant background and illumination. Some of these have used images in which only the hand is present so that segmentation is not hard, while others have used colored gloves that are needed to be worn by the users while signing in order to detect and segment the hand easily.

To Build Dynamic sign language interpreter, We are using three main components of this framework are:: preprocessing module, feature extraction module and **recognition modul**e.

- **Preprocessing:**

As these images are not taken in a controlled lighting environment, also images are taken with a digital camera, they have different sizes and different resolutions. So in image preprocessing is required. The first step of preprocessing a block is filtering. A moving average or median filter is used to remove the unwanted noise from the acquired image. Background subtraction forms the next major step in the pre-processing block. We therefore transform the images in HSV color space. It is a model which splits the colour of an image into 3 separate parts namely: Hue,Saturation and value. HSV is a powerful tool to improve stability of the images by setting apart brightness from the chromaticity The skin color Could be used to discriminate the hand region from the other moving object. The color of the skin is measured with the HSV model.

- **Feature extraction**

In feature extraction phase, various hand shape features like circularity, extent, convex deficiency and hand orientation are considered Classification machine learning algorithms like SVM are used for supervised learning, which involves labeling the dataset before feeding it into the algorithm for training.The Support Vector Machine (SVM) is a machine learning algorithm which is used for binary classification. SVM uses supervised learning techniques and was proposed by Vapnik and his colleagues in 1979. The main concept of SVM is to find a hyperplane that distinguishes the positive and negative class with the objective to maximize the width of the hyperplane between the classes. The orientation of the hyperplane and the distance from the origin of the hyperplane are the parameters which need to be considered.

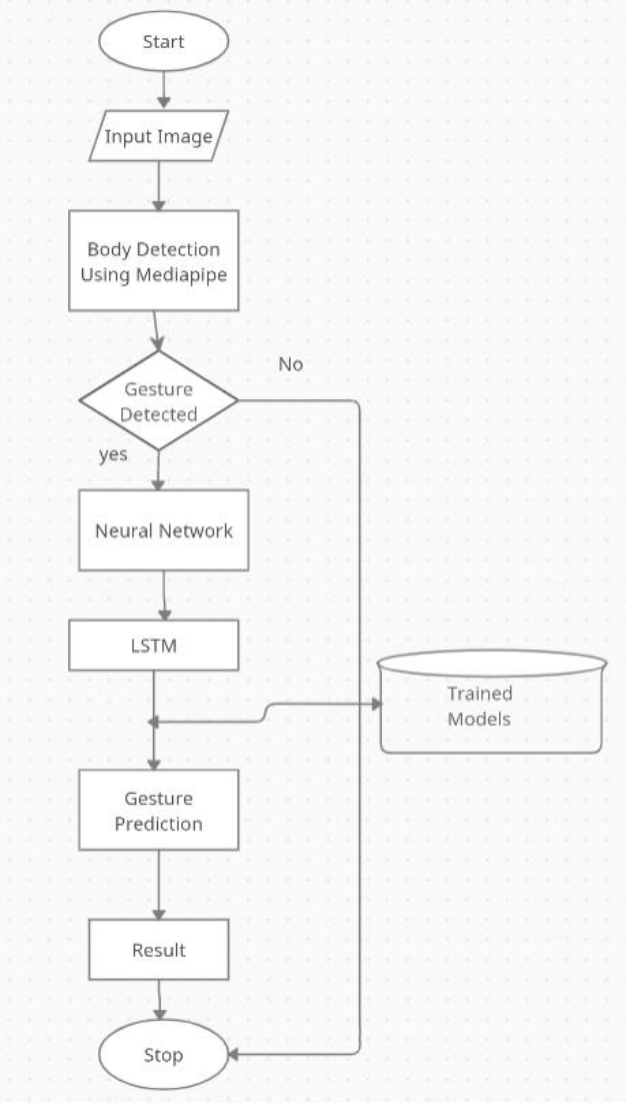
- **Recognition module**

In the recognition module we are using **Convolutional Neural Network** to extract features from the input image. It preserves the spatial relationship between pixels by learning image features using small squares of input data CNN's are very effective in reducing the number of parameters without losing on the quality of models. Images have high dimensionality (as each pixel is considered as a feature) which suits the above-described abilities of CNNs. All the layers of a CNN have multiple convolutional filters working and scanning the complete feature matrix and carry out the dimensionality reduction.

To make Static sign language recognition to Dynamic we are using 3D CNN, it's simple as **the 3D equivalent**: it takes as input a 3D volume or a sequence of 2D frames (e.g. slices in a CT scan), 3D convolutions applies a **3 dimensional filter to the dataset** and the filter moves 3-direction (x, y, z) to calculate the low level feature representations. Their output shape is a 3 dimensional volume space such as cube or cuboid. They are helpful in event detection in videos, 3D medical images etc.

After Recognising the gesture the output is shown on screen.

IV. System flowchart



V. DATASET

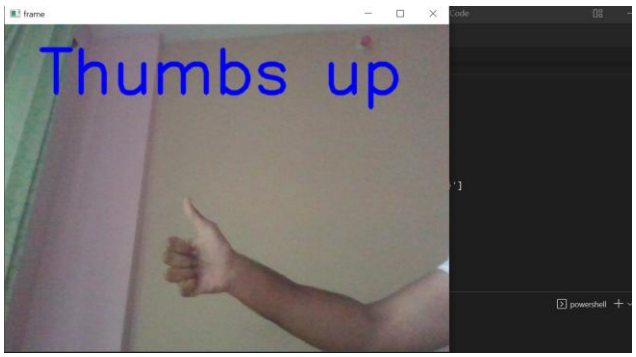
Dataset consists of 20 Indian Sign Language Hand Gestures in complex backgrounds. The entire dataset is inspired by the Indian Sign Language Research and Training Centre (ISLRTC). Twenty sign language gestures were selected from the ISL Dictionary. Selected signs were for the words- “Adult”, “Answer”, “Ask”, “Calm”, “Close”, “Compare”, “Complain”, “Confuse”, “Decide”, “Develop”, “Discuss”, “Doubt”, “Edit”, “Finish”, “Group”, “Important”, “Like”, and “Near”.

All the samples of gestures have been taken manually from 3 different users in different light conditions.

All the video clips are captured in different environmental conditions, such as lightning variation and background scenes (e.g., simple and complex). This makes the dataset more suitable for the purpose of the study. Each gesture consists of more than 500 image samples. Data used for testing the trained network model was isolated at the time of training. Frames extracted from the video samples were found to be more than 150000.

VI. RESULTS

1) Thumbs up



2) Thumbs down

3) Hello

4) I Love You

VII. CONCLUSION

We aimed to make a deep neural network that would model and recognize the Hand Gestures of standard Indian Sign Language. The base 3D CNN architecture is used for analyzing the modeling exercise for these dynamic gestures. Every person does not commonly know hand gestures in society. Moreover, each country having its own set of symbols is a challenge. Standardization at a global level is not there. We are using Indian Sign Language in this Project. Sign language is one of the tool of communication for physically impaired, deaf and dumb people. From the above consideration it is clear that the vision based sign gesture recognition has made remarkable progress in the field of body gesture recognition.

VIII. REFERENCES

[1] Deepika Tewari, Sanjay Kumar Srivastava; " A Visual Recognition of Static Hand Gestures in Indian Sign Language based on Kohonen Self-Organizing Map Algorithm" ; International Journal of Engineering and Advanced Technology , ISSN: 2249 – 8958, Volume-2, Issue-2, December 2012.

[2] Kenny Morrison, Stephen J. McKenna; "An Experimental Comparison of Trajectory-Based and History-Based Representation for Gesture Recognition"; In Proceedings of the International Gesture Workshop,2004.

[3] P. Kakumanu, S. Makrogiannis, N. Bourbakis; “Asurvey of skin-color modeling and detection methods”; Elsevier, The journal of the pattern recognition society, 40 ,1106 – 1122, 2007.

[4] Y. Wang, B. Yuan;”A novel approach for human face detection from color images under complex background”, Pattern Recognition 34(10) ,1983– 1992,2001.

[5]Ketki. P.Kshirsagar, Dharmpal Doye; “ Object Based key Frame Selection for Hand Gesture recognition”; 3176 International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 10, October - 2013 IJERT ISSN: 2278-0181 IJERTV2IS100891 www.ijert.org International Conference on Advances in Recent Technologies in Communication and Computing, 2010.

[6] R. Rojas ; “Neural Networks”; Springer-Verlag, Berlin, 1996.

[7] Siddharth S. Rautaray , Anupam Agrawal; “Vision based hand gesture recognition for human computer interaction: a survey”; Springer Science+Business Media Dordrecht, November 2012.

[8] Z. Cao, G. Hidalgo, T. Simon, S. -E. Wei and Y. Sheikh, "OpenPose: Realtime Multi-Person 2DPoseEstimationUsing Part Affinity Fields," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 1, pp. 172-186, 1 Jan. 2021, doi: 10.1109/TPAMI.2019.2929257.

[9]H. Elleuch, A. Wali, A. Samet and A. M.Alimi,"Astatic hand gesture recognition system for real time mobile device monitoring," 2015 15th International Conference on Intelligent Systems Design and Applications (ISDA), Marrakech, 2015, pp. 195-200, doi: 10.1109/ISDA.2015.7489224.

[10]Ghanem,Sakher&Conly, Christopher&Athitsos,Vassilis. (2017). A Survey on Sign Language Recognition Using Smartphones. 171-176. 10.1145/3056540.3056549.

[12] R. Y.Hakkun, A. Baharuddin, et al. Sign language learning based on android for deaf and speech impaired people. In Electronics Symposium (IES), 2015 International, pages 114–117. IEEE, 2015. [6] Zong-yuan Zhao, Wan-lin Gao, Miao-miao Zhu, Lina Yu, A Vision Based Method to Distinguish andRecognize Static and Dynamic Gesture, Procedia Engineering, Volume 29, 2012,

[13] O. Aran, I. Ari, L. Akarun, B. Sankur , A. Benoit , A. Caplier, P. Campr, A. H. Carrillo, and F. Xavier Fanard, “SignTutor: An Interactive System for Sign Language Tutoring,” IEEE feature article, pp. 81-93, 2009.

[14] T. Dasgupta, S. Shukla, S. Kumar, S. Diwakar, and A. Basu, “A Multilingual Multimedia Indian Sign Language Dictionary Tool,” The 6‟Th Workshop on Asian Language Resources, pp. 57-64, 2008.