A REVIEW ON SIGN LANGUAGE RECOGNITION USING DENSENET-A DEEP LEARNING APPROACH

Sneha Makode[1], Dr. G. D. Dalvi[2]

[1]*PG Scholar, Amravati, India*

[2]*Professor, P. R. Pote(Patil) Education & Welfare Trust, Engg, Amravati, India*

Abstract: *Sign language is used by deaf and hard hearing people to exchange communication between their own community and with normal peoples. Sign language works on sign gesture and its acquisition and continues till text/speech generation is performed in recognition by computer. Static and dynamic are the two types of Sign. Both recognition systems are important to the society, but static gesture recognition is simple than dynamic gesture. The sign language recognition and processing steps are reviewed in detail in methodology. The data acquisition, data preprocessing and transformation, feature extraction, classification and results obtained will be examined. Some future directions for research in this area also suggested. In this project we will used densenet with transfer learning.*

Keywords : **Sign, Acquisition, preprocessing, feature extraction.**

1. INTRODUCTION

Sign language is the hand gesture based visual language used by hearing-impaired people to communicate with other people. The sign language consists of the coordinated movement of different parts of our body [7]. These are hands, face and the body. Body posture can play important roles likes sign, single hand sign or combination of both hand movement, face expression. Sign Language Recognition (SLR) is the method of translation of the systematic and coordinated movements of one's body into lingual or textual phrases. Sign Language Recognition (SLR) is an active area of research due to its important role in Human Computer Interaction (HCI) or Human Robot Interaction (HRI). Sensor based, Vision based and Depth based approaches are SLR based systems. In these approaches, hand movements are captured by an external sensor connected to signer's body, a camera, a depth sensor, respectively. Speech and hearing impaired people are humans at the deepest psychological level. Many of these people are not even exposed to sign languages and it is observed that it gives a great relief on a psychological level, when they find out about signing to connect themselves with others by expressing their love or emotions. About 5% population in world are suffering from hearing loss. There are some certified sign language interpreters for a deaf population of around million of peoples. Sign language recognition (SLR) is an area where automation can provide tangible benefits and improve the quality of life for a significant number of people who rely on sign language to communicate on a daily basis [6]. Recently, basic machine learning approaches have been largely replaced with deeper architectures that employ several layers and pass information in vector format between layers, gradually refining the estimation until positive recognition is achieved. Such algorithms are usually described as ``deep learning'' systems or deep neural networks, and they operate on principles similar to the machine learning strategies described above, although with far greater complexity. recurrent neural networks (RNNs) that include at least one recurrent layer, and convolutional neural networks (CNNs) that include at least one convolutional layer based on the structure of the network these two architectures are commonly used. Depending on the number and type of layers, these networks can exhibit different properties while the training phase decisively impacts the performance of the algorithm. The general rule is that larger and more specific datasets allow for more robust network training, and therefore the quality of the training set is an important factor. Additional fine-tuning of a model can usually be achieved by changing some of the relevant hyper-parameters that define the training procedure.

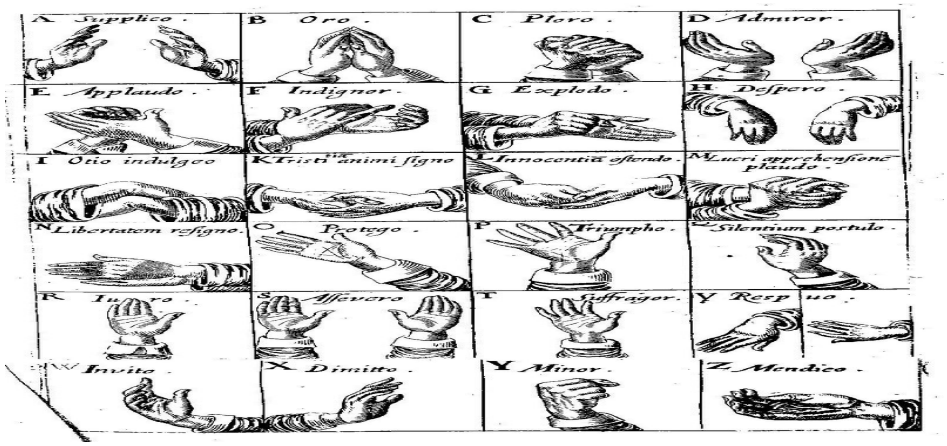

Fig 1: Chirogram from Chirologia, 1644, signed language, Wikipedia.

2. LITERATURE SURVEY

1 Sign Language Recognition Approaches

Sign language recognition systems focus on hands movements, face expression and the body posture, while hand gesture recognition systems rely only on the movements of hands. Hand gesture recognition systems can be categorized into three main branches: Sensor based, Vision based and Depth based approaches.

2 Sensor Based Approaches

In this approach, hand movements are measured by an external device connected to signer's body. External device, which has comprehensive sensors on it, can easily sense palm of hand and fingers. This decreases the level of user friendliness of the system. The second drawback is that these devices are expensive to manufacture. The last drawback is that sensors require high quality calibration to sense data correctly [1] [2] [3] [4].

3 Vision Based Approaches

In this approach, hand movements are captured by camera. This method is more user friendly and convenient in the sense that it provides a natural interaction method compared to sensor based methods, since signer does not need any cumbersome equipment to wear. Only a camera is required to capture images. In some of the vision based SLR systems, signer wears colored gloves in order to simplify the hand segmentation process. By the help of the colored gloves, segmentation errors resulting from other skin-colored objects are prevented. Also, segmentation process has lower computational cost than non-glove based systems. [2] [3] [1]

4 Depth Based Approaches

In this approach, depth images are used for hand gesture recognition. Cameras provide two-dimensional information on the captured space. To acquire data on the third axis, a second specialized equipment is necessary. There are two ways to acquire depth images. The first and the most frequently used one is by using depth measuring cameras such as: Microsoft Kinect, ASUS Xtion, Mesa SwissRanger. The other option is extracting depth information from stereo video cameras .There are several advantages of depth based approaches. Other skin colored objects, lighting conditions and variations, complex background don't cause any major problems in hand segmentation and tracking due to the usage of depth information [4]. Without these challenges, hand segmentation and tracking can be done easily and accurately

[1]. However, Depth cameras are much more expensive than basic cameras used in visual based approaches.

The proposed prototype TSL recognition system can recognize dynamic gestures in realtime; and the experience of developing posture and gesture models can be applied to other sign languages, and may help to interpret a large set of virtual reality commands[5]. In the centre of Winspect is the working-cum-dataglove, and its application in two use cases: data-collection and document-browsing, both of which can be completely carried out without the necessity to take off the gloves. The solution presented in this paper is a wearable one, because the worker is wearing computer hardware and input-devices on his body[6]. A database of images is first clustered using a k-mediod clustering algorithm with a distance metric based upon shape context. From this, a tree structure of boosted cascades is constructed. The head of the tree provides a general hand detector while the individual branches of the tree classify a valid shape as belong to one of the predetermined clusters exemplified by an indicative hand shape[7]. In summary, a systematic approach to building an extremely robust hand appearance detector, providing an important step towards easily deployable and reliable vision-based hand gesture interfaces is made[8]. There are no sufficient number of work on Indian language character recognition although there are 12 major scripts in India. A review of OCR work done on Indian language scripts[9]. It begins by describing typical imaging devices and the imaging process. Document analysis from a single camera-captured image as well as multiple frames and highlight some sample applications under development and feasible ideas for future development[10].they can capture images of thick books, historical manuscripts too fragile to touch, and text in scenes, making them much more versatile than desktop scanners[11]. The system has been used in a number of demonstration applications ranging from desktop applications to the control of a mobile robot[12]. The system was implemented and tested using a data set of 300 samples of hand sign images; 15 images for each sign and their Experiments revealed that system was able to recognize selected ASL signs with an accuracy of 92.3% The proposed system was able to reach a recognition rate of about 98.5% for training data and 80% for testing data[13]. Utilising HTK capabilities such as tree clustering procedure and character N-grams might eliminate the unseen tri-model problem and error resulting from using basic units in the recognition network[14]. The word level recognition is done on the basis of a string edit distance. For a number of characters, their handwritten images sometimes get segmented into two pseudo characters after removal of the Matra. Thus, the set of pseudocharacters consists of such pseudocharacters as well as true characters[15]. On experimentation with a dataset of 4900 samples the overall recognition rate observed is 92.80% and is compared with other recent methods for Handwritten Devnagari Character Recognition and it has been observed that this approach has better success rate than other methods[16].

3. PROPOSED METHODOLOGY

Sign Language Recognition

In general, sign language recognition is composed of three main steps: Segmentation, Feature Extraction, and Classification as shown in Figure 2. Firstly, hand segmentation techniques must be used on acquired data. In the second step, features that are describing the performed sign must be extracted from output data of the first step by using feature extraction techniques. In the classification step, developed model and algorithm classify the performed sign according to extracted

features. In the following subsections, previous studies in the literature are given in details for each of the mentioned steps.

Fig:2 basic block diagram

The basic flow structure of any sign language reognition would looks like this 3 steps process. First, the input sign/data is acquired, via the phone camera or from some sensor. The next step is to extract the features from the input data. Finally, the sign is classified using some appropriate algorithm that is compatible with the extracted features and expected output. For each method we examine, we take a close look in how that method approaches the problems of feature extraction and recognition/classification.

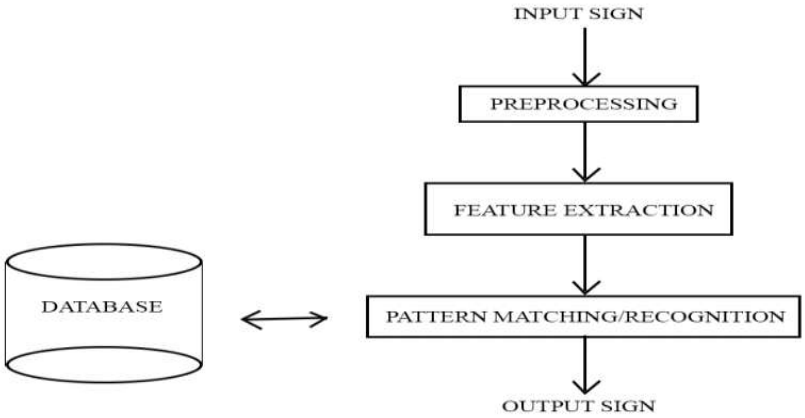

Fig: 3 Basic System Architecture

DenseNet :

DenseNet has been widely used for classification tasks due to the advantages it introduces such as alleviating the vanishing gradient – a common problem encountered with deep networks. Classification tasks can be done by DenseNet due to the advantages it introduces such as alleviating the vanishing gradient – a common problem encountered with deep networks.

Support Vector Machine (SVM) :

A support vector machine (SVM) is a supervised machine learning model that uses classification algorithms for two-group classification problems. After giving an SVM model sets of labeled training data for each category, they're able to categorize new given problem.

Dynamic gesture :

Dynamic gesture recognizing is mainly focusing on recording the position of hands in images and identifies the meaning of actions by analyzing them. After identifying hand automatically, the system itself can get the position of the hand in images with frames. We can used in static gesture recognizing to judge if any hands is in images or not. Comparatively, we do not judge any gesture only by one image but just record the position of hand because it may be one fragment in a series of images which contents a dynamic gesture such as a wave.

4. SUMMARY

Traditional machine learning and deep learning algorithms are the major approaches used for SLR in the literature. When we survey the dissimilarities between traditional machine learning and deep learning algorithms, we decided to use traditional machine learning algorithms in our system for having high accuracy and low computational cost. However, deep learning based SLR system will be studied because deep learning algorithms have gained popularity and attention in recent works for SLR. The first SLR system, which is based on traditional machine learning, consists of three main step: hand segmentation, feature extraction and classification. System starts with hand segmentation algorithm. In this algorithm, firstly region of interest is found by using HSV thresholding. Then, post processing operations will be used to correct errors caused by clustering algorithm and lighting conditions. For feature extraction step, the most frequently used appearance based feature descriptors across multiple criteria which are computational cost, scale and rotation invariance, and accuracy will be taken into consideration. Using feature descriptor for describing the hand shape because it is scale and illumination invariant, has low computational cost, and has high recognition accuracy for hand shape detection in the literature is studied. In the classification step, multi-class SVM can be used to classify feature vectors by considering dataset dependency, computation time, number of parameter tuning needed for optimization, and accuracy.

REFERENCES

[1] J. S. Sonkusare, N. B. Chopade, R. Sor, and S. L. Tade, "A review on hand gesture recognition system," in 2015 International Conference on Computing Communication Control and Automation, pp. 790–794, Feb 2015.

[2] A. Er-Rady, R. Faizi, R. O. H. Thami, and H. Housni, "Automatic sign language recognition: A survey," in 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), pp. 1–7, May 2017.

[3] Omkar Govalkar, Pratik Gaikar, Pramod Gavali, "Sign Language Recognition Techniques: A survey", International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056".

[4] J. Suarez and R. R. Murphy, "Hand gesture recognition with depth images: A review," in 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication, pp. 411–417, Sept 2012.

[5] Rung-Huei Liang Ming Ouhyoung, "A Sign Language Recognition System Using Hidden Markov Model and Context Sensitive Search", Proc. of ACM VRST'96 (ACM International Symposium on Virtual Reality and Software Technology), pp. 59-66, July, Hong Kong, 1996.

[6] Michael Boronowsky, Ansgar Marco Schmidtin, "Winspect: A Case Study for Wearable Computing-Supported Inspection Tasks", 2001 https://www.researchgate.net/publication/221240920

[7] A Boosted Classifier Tree for Hand Shape Detection, Eng-Jon Ong and Richard Bowden Centre for Vision, Speech and Signal Processing University of Surrey, Guildford, GU2 7XH. 2004

[8] Mathias K¨olsch and Matthew Turk, " Robust Hand Detection" , Department of Computer Science, University of California, Santa Barbara, CA 93106, 2004.

[9] U. Pal , B.B. Chaudhuri, "Indian script character recognition: a survey", Elsevier Pattern Recognition 37 (2004) 1887 – 1899

[10] Jian Liang, David Doermann , Huiping Li, " Camera-based analysis of text and documents: a survey", International journal on Document Analysis and Recognition", Springer-Verlag 2005.

[11] Jian Liang, David Doermann, Huiping Li,"Camera-based analysis of text and documents: a survey" , IJDAR International Journal on Document Analysis and Recognition 7: 84–104, (2005)

[12] Holger Kenn Friedrich Van Megen and Robert Sugar , A glove-based gesture interface for wearable computing applications, " TZI University at Bremen, funded by European Commission through IST Project" IEEE Xplore Applied Wearable Computing (IFAWC), 2007, 4th International Forum.

[13] Qutaishat Munib, Moussa Habeeb, Bayan Takruri, Hiba Abed Al-Malik "American sign language (ASL) recognition based on Hough transform and neural networks", Elsevier Journal, Expert Systems with Applications 32 (2007) 24–37

[14] M.S. Khorsheed , "Offline recognition of omnifont Arabic text using the HMM ToolKit (HTK)", Elsevier Journal, Pattern Recognition Letters 28 (2007) 1563–1571

[15] Bikash Shaw, Swapan Kr. Parui, Malayappan Shridhar, "Offline Handwritten DevanagariWord Recognition: A Segmentation Based Approach" , 978-1-4244-2175-6/08/$25.00 ©2008 IEEE

[16] Sandhya Arora, Debotosh Bhattacharjee, Mita Nasipuri,"Combining Multiple Feature Extraction Techniques for Handwritten Devnagari Character Recognition", 2008 IEEE Region 10 Colloquium and the Third ICIIS, Kharagpur, INDIA December 8-10.