

SIGN LANGUAGE RECOGNITION USING MACHINE LEARNING

Sneha Prabhu¹, Sriraksha Shetty², Sushmitha P Suvarna³, Vindya Sanil⁴, Dr. Jagadisha N⁵

^{1,2,3,4}Student, Dept. of Information Science and Engineering, CEC, Karnataka

⁵HOD & Associate Professor, Dept. of Information Science and Engineering, CEC, Karnataka

Abstract - Communication must happen in both ways. Communicating with the people having hearing disability is a difficult task so we developed the real time sign language recognition system using machine learning. This system contains the hand gestures performed by people which is captured using the web camera. Out of all images the region of interest is identified and the system will predict and show the name of the sign captured. The captured signal will undergo some of the processing steps inclusive of data acquisition, image processing, characteristic extraction and classification. We will use the Convolutional Neural Network (CNN) to train and build our model. Hence it is very useful for people with hearing and speech disability.

Key Words: CNN, Sign Language, Machine Learning.

1. INTRODUCTION

Sign language is a language for the deaf and dumb which uses simultaneous orientation and movement of hand shapes instead of acoustically conveyed sound patterns. Deaf and Dumb human beings depend on sign language translators for communications. However, finding experienced and qualified translators for their day to day affairs throughout life period is a very difficult task and also overpriced. Sign language is the basic means of communication for those with hearing and vocal disabilities. Those disadvantaged have difficulty in their day to day lives. In other scenario large group of people who had no problem in communicating but are unable to speak because of injury like accidents or disease are not able to speak. Sign language translates to text or speech, would give them everyday communication. In regional and rural areas, it is very effortful for the deaf and dumb people to communicate to normal people so to bridge the gap between them a communication tool needs to be provided. The purpose is to increase the ability of deaf and dumb people to acquire services wherever they want without taking the help of any sign language experts. This systems can reduce the urge for professional translating services and can actually makes the life better. This can also be used for training sign language and for people who are dumb and deaf due to sickness or injury because of accidents. We aim to develop a system that would ease this difficulty in communication. Sign language consists of making shapes or movements with your hands with respect to the head or other body parts along with certain facial cues. A recognition system would thus have to identify specifically the head and hand orientation or movements, facial expression and even body pose. We propose the design for a basic yet extensible system that is

able to recognize Sign Language. American Sign Language was chosen since it is utilized by a majority of those disabled. We believe that it will benefit for deaf and dumb people by providing them a flexible translator. Sign Language Recognition System (SLR) utilize computer vision and machine learning method convolutional neural network (CNN).



Fig 1: Sign Language for Alphabets A-Z

2. RELATED WORK

In[1],Citra Suradi et al., This paper implements a design of sign language recognition using E-CNN. In this study SIBI alphabetical images were used datasets. The total datasets used is 38656 data which are then divided into three parts of data, namely training data, validation data and testing data. Handkey point was used for filter and cropping process before the use of dataset. Since Handkey point will detect the hand joints from the image and it will crop and filter image data so that the dataset used focuses on hand object. In this study, there were 5 CNN (Convolution Neural Network) model combined with 3 Ensemble model were used. Also observed that combining several algorithms can increase the accuracy hence ensemble method was used along with CNN model.

In[2],Lionel Pigou et al., This paper sign language recognition system using CNN. For building the automated recognition system two main steps are necessary. The first step is extracting the features from the frame sequence. The second step is actions will be classified. In this they have used CNN for extraction of frame sequence from the image and ANN is used for the classification. Datasets had been used from the Chalearn Looking at human beings. Data units consists of 20 Italian gestures achieved from the 27 users with variant in surrounding, garb, mild.. There was 6600 gestures in development set used, out of which 4600 for training set and 2000 for validation set. The methodology used are dataset

collection, pre-processing, CNN, proposed architecture, generalization and training, temporal segmentation and results. In proposed architecture there were two features one for extracting hand gestures and other one for upper body features extraction.

In[3], Adithya V et al., In this paper a method for automation recognition of fingerspelling in Indian sign language was proposed. Digital Image processing strategies and artificial neural community for recognizing special symptoms were used to enforce the gadget. This paper presents a method for both English alphabets as well as numbers. This technique had included 26 letters of English alphabets and zero-9 numbers. This approach had four foremost steps namely Image Acquisition and pre-processing, hand gestures, characteristic extraction, classification. Images have been captured with black heritage. Data units were created for education and checking out the device. Next extracting the hand signs and symptoms from the captured photograph using skin colour-based totally segmentation and produce the binary image with signs or skin pixels with white colour and background become set to black colour. The form function was derived from the space remodel the binary picture. Euclidean, City block and chessboard used for locating the distance transform of an photograph. ANN turned into used for classification. The feature vector acquired from inside the feature extraction become used as enter for the classifier to apprehend the Sign.

In[4], Muthu Mariappan H et al., This system is implemented for recognising the gestures of Indian Sign Language (ISL). System uses non-invasive vision-based recognition method. Using skin segmentation feature of OpenCV the region of interest (ROI) are identified and tracked. By applying the Fuzzy C-means Clustering (FCM) machine learning algorithm the training and prediction of hand gestures are performed. The data samples of 80 words and 50 sentences of daily usage terms of Indian sign language were collected and 10 volunteers videos were recorded therefore there was $80 \times 10 = 800$ total samples of words and $50 \times 10 = 500$ total samples of sentences. The raw videos were taken in dynamic background and given to the system as the input. During pre-processing noise eliminated and blurred image was obtained by performing convolution operation with low-pass box filter. BGR image is converted into HSV color space and morphological operations are performed on HSV image and gives the binary image. Area of all contour was calculated and using them three were selected and the orientation between the contours is also kept track. Those three contour are face, right and left hand which contributes gestures. For each frame in video required features are extracted from these regions as vector feature. During training, extracted features are given to the c-mean algorithm and during testing, it matches the text file with existing cluster.

In[5], Soma Shrenika et al., This system implemented to overcome the barriers faced by the impaired hearing people using sign languages. ASL (American sign language) dataset was used to identify the sign made by gesture. The dataset

was consist of numbers zero - nine and alphabets a - z. There were 70 samples for each 36 symbols. This system uses camera as there source which captures various gestures of hand. The captured image changed into in RGB format and by using the common approach the RGB image turned into converted into grayscale photo. Then pr-processing of the image takes region wherein edges are detected the usage of Canny facet detection set of rules. Canny area detection algorithm had 5 extraordinary stages specifically, Image smoothing, Gradient Magnitude, Non-maximum Suppression, Double threshold, Hysteresis. Finally template matching algorithm turned into used to become aware of the sign and show the textual content. Sum of Absolute Difference (SAD) method was used in comparison of template image with the images present in the data set.

In[6], Divya Deroa et al., In this paper simultaneously combining hand shapes, orientation and movement of hands, fingers or frame, and facial features to fluidly specific a speaker's thought. Here data is acquired through a 3 Mega pixel camera and person wears red and blue coloured gloves. system is made only to recognise alphabets and numbers Therefore total number of images in the database is 510 which include 25 alphabets and 9 numbers. The steps included in this experiment are (1) segmentation, (2) fingertip finding algorithm and (3) PCA. Recognition is done using PCA for live data frames where every 20th frame is given as input to be recognized. The system can recognise 94% of the signs made correctly. Fingertip algorithm is used to make database on the basis of number of finger tips. PCA compares the input image with the new database formed with the help of finger tips.

In[7], G. Anantha Rao et al., This paper proposes the recognition of Indian sign language using CNN. The captured method used in this paper was Selfie mode continuous sign language video. Datasets were created with 5 different subjects performing 200 signs in 5 different viewing angles under various circumstances. Out of which 3 different samples size was used for performing CNN training where as remaining 2 samples were used for testing the trained CNN. This model was constructed with input layer where data's are given as input, in feature extraction stage 4 convolution neural network of different window size, five rectified linear unit for non-linearities, 2 stochastic pooling layer to avoid a substantial information loss, in classification stage one dense and one fully connected layers. In classification Softmax regression was adopted.

In[8], Surejya Suresh et al., In this paper the proposed architecture is analysed using two models Stochastic Gradient Descent optimizer and Adam optimizer. The algorithm used is Convolution Neural network. CNN contains several layers such as pooling, dropout and fully connected layer. All the important features are extracted from the bottom layer and abstract features are computed from high level layers of CNN. The proposed model architecture has the input layer which receives the images of size 64×64 pixels. Later two convolution operations are applied to the images to

generate feature maps. The feature maps are flattened which results in feature vector. Python is the programming language used. Some of the libraries used are keras, scikit-learn, matplotlib and numpy.

In[9],Kusumika Krori Dutta et al.,In this paper Indian Sign Language is communicated using hand gestures made both Single hand and Double hands . In this, only main words are taken into consideration grammar usage is not considered. Articles such as 'a', 'an', 'the' are omitted and also doesn't include tense form .This paper deals with classification of Indian sign language using machine learning. The system is trained with double handed sign language by using a Principal Component Analysis(PCA) and Artificial neural network (ANN) algorithm in MATLAB. The hand gestures are acquired from different sized and complexion hands and made an image database of having ease image of size 768 by 1024. Once the images were preprocessed the size of the image was reduced to 256 by 256. Sign language patterns are trained separately using two machine learning techniques. A. K Nearest Neighbours Algorithm B. Back Propogation Algorithm. This experiment was conducted over 220 images of double handed Indian sign language alphabets and 800 images of single handed Indian sign language alphabets. English letters and numbers using KNN and Back propagation techniques where PCA is used for dimension reductionality .For K-NN techniques with K=1, achieved 100% recognition rate whereas using back propagation technique achieved 94-96.

In[10],Lihong Zheng et al., In this paper Auslan is the primary language used by the Australian Deaf Community, with an estimated 30,000 users in Australia Sign language translation to text or speech would give them everyday communication. The fashionable framework of imaginative and prescient based SLR includes pre-processing, sign gesture illustration, function extraction and type. The widely used 2D features include SIFT, HOG, HOF, STIP, and kernel descriptors. For the purpose of dealing with occlusion 3D/4D space of time (HON4D) and Random Occupancy Pattern(ROP) features were proposed. Sign gesture recognition system were divided into two types based on image acquisition: touched based or untouched/vision based There are about seven main large data sets of sign language. They are American Sign Language Lexicon Video Dataset MSR Gesture3D , Auslan data set, LTI-Gesture Database, RWTH German Fingerspelling Database, DEVISIGN Chinese Sign Language dataset, and Indian Sign Language dataset.

3. PROPOSED SYSTEM

The system uses the web camera for capturing the hand gestures of the hearing and speech impaired people. The raw images taken is given as an input to the system. The image frames are resized to maintain the equality among all the images. The proposed system includes several steps namely, Image Acquisition, Image Pre-Processing, Segmentation ,Feature Extraction, Recognition and Text Output.

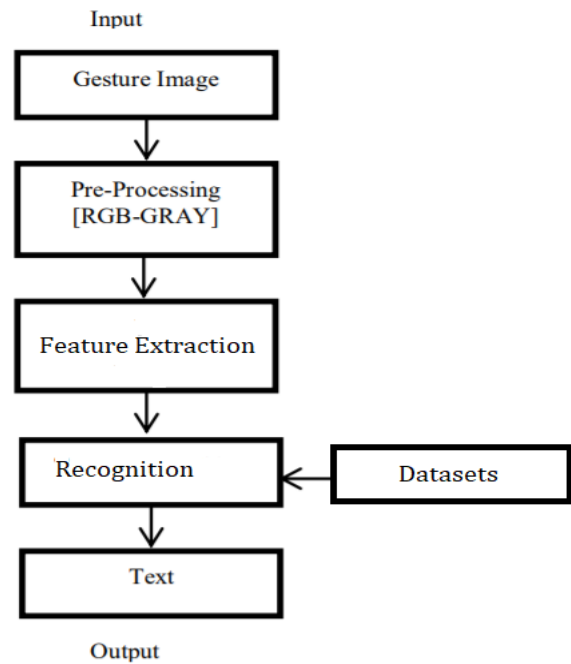


Fig 2: Schematic Diagram of Sign Language Recognition

3.1 Image Acquisition:

This is the first step in the workflow as no processing can be done without getting an image. Image Acquisition is the process of extracting an image from a source, here the source is the hardware-based source that is the web camera. This image is not processed in any way so it is the raw image.

3.2 Segmentation:

The method of separating objects or signs from the context of a captured image is known as segmentation. This process includes Con-text subtracting, skin-color detection, and edge detection. The location of the hand must be detected and separated in order to recognize the hand gestures.

3.3 Features Extraction:

Predefined features such as form, contour, geometrical feature like position, angle, distance, etc, colour feature, histogram, and others are extracted from the pre-processed images. Later it is used for sign language recognition. Feature extraction is a step that divides and organizes a large collection of raw data. It reduces to smaller and easier-to-manage size so that processing would be simpler. There is massive amount of data sets that have a lot of variable features in it. To process these variable features, a large amount of computation is needed. By Reducing the size of the data these features are simple to use.

3.4 Pre-Processing:

Pre-processing is the process of eliminating high-intensity noises from the image frames. The first step is smoothing or blurring. Then the image is converted to gray scale from RGB. The main aim of converting the image into grey scale is to reduce the size and amount of processing.

The different phases of pre-processing are:

3.4.1 Morphological Transform

It is a simple operation based on the shape of the image which are performed on binary images. There are forms of morphological transforms specifically dilation and erosion.

3.4.2 Blurring

Blurring is the process of eliminating the noise in an image and keeping required feature intact.

3.4.3 Thresholding

The process of converting grey scale image into binary image is called Thresholding. Thresholding is used to pick the area of interest ignoring the area that is not required.

3.4.4 Recognition

Classifiers are used in this section. Classifiers are the algorithm used to interpret the alerts. The classifier used in this challenge is Convolution Neural Network. Convolution neural network is a deep learning algorithm which takes the photograph as input, images the critical aspects from the photo and differentiates from one another. CNN algorithm offers better precision compared to other algorithms in photo type and reputation.

3.4.5 Text output

The hand gestures are recognized and the system gives an output in the human understandable form that is in text.

4. CONCLUSION

The Sign Language Recognition (SLR) system is a method for recognizing a collection of formed signs and translating them into text with the appropriate context. The significance of gesture recognition can be seen in the development of effective human-machine interactions. We attempted to build a model using a Convolutional Neural Network in this project.

5. REFERENCES

[1] Citra Suradi, Anik Nur Handayani, Rosa Andrie Asmara, Aji Prasetya Wibawa, Lilis Nur Hayati, Huzain Azis (2021). "Design Of Sign Language Recognition Using E-CNN".

[2] Lionel Pigou; Sander Dieleman; Pieter-Jan Kindermans; Benjamin Schrauwen (2014). "Sign Language Recognition Using Convolution Neural Network".

[3] Adithya V, Vinod P.R, Usha Gopalakrishnan (2013). "Artificial Neural Network Based Method For Indian Sign Language Recognition".

[4] Muthu Mariappan H, Dr. Gomathi V (2019). "Real Time Recognition of Indian Sign Language".

[5] Soma Shrenika, Myneni Madhu Bala (2020). "Sign Language Recognition Using Template Matching Technique".

[6] Divya Deroa, Nikesh Bajaj (2012). "Indian Sign Language recognition".

[7] G. Anantha Rao, K Syamala, P V V Kishore, A S C S Sastry (2018). "Deep Convolution Neural Network for Sign Language Recognition".

[8] Surejya Suresh, Mithun Haridas T P, Supriya M H (2019). "Sign Language Recognition using Deep Neural Network".

[9] Kusumika Krori Dutta, Sunny Arokia Swamy Bellary (2017). "Machine Learning Techniques for Indian Sign Language Recognition".

[10] Lihong Zheng, Bin Liang, Ailian Jiang (2017). "Recent Advances of Deep Learning for Sign Language Recognition".

[11] Rachana Patil, Vivek Patil, Abhishek Bahuguna, Mr. Gaurav Datkhile (2021). "Indian Sign Language Recognition Using Convolution Neural Network".