Deep Learning for Sign Language Recognition

Bhagyashri Pagar
Department of Information Technology, Sinhgad
College of Engineering Pune, Maharashtra, India

Rutuja Shelar
Department of Information Technology, Sinhgad
College of Engineering Pune, Maharashtra, India

Soumya Sheelavant
Department of Information Technology, Sinhgad
College of Engineering
Pune, Maharashtra, India

Avinash A. Utikar
Prof., Department of Information Technology,
Sinhgad College of Engineering
Pune, Maharashtra, India

Abstract:- A sign language is a way of communicating by using the hands instead of spoken words. Sign language is used by deaf and dump people to communicate with other individuals. People who are speech-impaired and also some people who have autism spectrum disorder face problem while communicating with normal as they can converse using only sign language. So, it becomes difficult for other individuals to understand this sign language. Each country usually has its own native sign language. The Indian Sign Language recognition application proposed here aims at solving the communication problem between people. System will capture the different hand gestures through camera of mobile phone. The signs used in sign language are identified by the features extracted from the hand gestures. Then, processing of the image takes place by using convolution neural network algorithm. After all the processing we finally get the output as a text which can be easily understood by all people. We are keeping purpose of developing system that will make communication between Deaf and Dumb person and normal person easy and convenient.

Keywords:- Sign Language Recognition Application, Convolution Neural Network, Deep Learning.

I. INTRODUCTION

The history of sign language in Western societies dates back to the 17th century, as a visual language although references to the forms of communication using sign language date back to the 5th century BC Greece. Sign language is made up of a system of general gestures, to represent the letters of the alphabet and emotions. Many sign languages are native languages, distinct from the structure of spoken languages used near them, and are mainly used by deaf people to speak. Hearing is the important sense among 5 human senses. Deafness hinders a person from understanding spoken languages Unlike spoken languages, where grammar is expressed using punctuation-based symbols, feature, attitude and syntax sign languages use gesture, punctuation, and body and facial expressions to form grammar. Sign language is for community that helps to interact with each other. They are used for conveying meaning through visual modality.

Due to lack of ease in sign languages, non-sign language interpreter speaks or pays less attention to communicate with dump and deaf people. Most of people know nothing about sign language except for people engaged in special education and most people will not take initiative to spend time and learn sign language. The Indian sign language (ISL) is one of the popular languages among Sign languages. It uses gestures for communication. Gestures are various movements used in the process of communication either hand or body.

Sign languages uses gestures which usually make use of visually transmitted patterns. ISL presents various hand movements using both right and left hands. This ISL gestures can be classified into two categories static and dynamic. In static gesture there is no movement of hand. Most of gestures of alphabets in ISL are static signs. Dynamic gestures involve hand movements during performing gestures. The gesture of hand recognition includes hand detection and recognition of that hand sign during processing. People find it very difficult to understand sign language hence it become important to design vision-based sing language translator. Many researchers try to combine recent models for deaf community to communicate more easily using AI in real world scenario which includes machine learning models and some dataset collection. There are two major methods for sign language translation. 1)Vision based 2) Glove based which includes sensors and gloves for implementation. Glove-based hand gesture recognition system has introduced as a beginning, due to wearing heavy devices and cable connections, it lacks naturalness. This research proposes a vision-based sign language device capable of translating ISL to text. It includes basic models like object detection, feature extraction and classification. The training model built for classifying hand gestures using ML. The hand gestures are captured dynamically using camera.

A Convolutional Neural Network (CNN) is a Deep Learning algorithm which is applied to Image and Video recognition, Image Analysis and Classification etc. We use Convolutional Neural Network to increase the accuracy of sign classification. Therefore, this system balances the communication between deaf people and normal people without any requirement of intermediate translator by converting gestures into text, which achieves our objective and reduces the barrier in communication. The aim of this paper is to recognize signs used in Indian Sign Language. It

is well studied subject in American Sign Language, but it has received little attention in Indian Sign Language. We want to solve this issue, but instead of using high-end technologies like gloves, we want to recognize gestures from photographs and then use computer vision and ML techniques to extract specific features for classification.

II. LITERATURE SURVEY

All people communicate with each other through the spoken language but, deaf and dumb people face problems while communicating with other people. They use sign language to communicate with normal people. As there are many people who are not familiar with sign language and unable to interpret the meaning of various signs used by deaf and dumb people. Also, unlike spoken languages sign languages users are limited and sign languages have their own vocabulary and grammar, different than the corresponding spoken language. Deaf and Dumb people dependent on sign language interpreters to communicate with others. However, for all daily life work finding qualified interpreters is kind of difficult task and also it is not affordable for those people. Hence, there is need of some kind of interface which help to ease the communication between deaf and dumb people and normal people. So, various researchers used different approaches to tackle this problem. There are mainly two approaches for sign language recognition. One is vision based where different signs are converted into text or speech using techniques like image processing and computer vision techniques. In this we compare different features of capture image with database of signs used in sign language. Light intensity and image resolution affects the accuracy of vision-based system and dynamic gesture recognition is complex in this type of approach. In second approach we used sensors which can be placed directly on the hand and then we use sensor values to classify gestures. This approach provides better accuracy but, wearing heavy devices and cable connection lacks naturalness of sign language recognition. Also, those sensors are very expensive and difficult to use in day-to-day life.

Template matching algorithm which converts sign language into text output. In this they used canny edge detection algorithm which removes the data which is not useful and gives sign alphabet as output. So, we can extend this implementation to mobile phones [1]. Deep CNN architecture to discover and categorize signs used in sign language from hand gesture images was used in another model. However, the real time detection and classification of hand gestures has not been implemented in this paper. Therefore, we should be directed towards development of intelligent methodologies for real time sign language detection. In this regard, region-based CNN can be implemented for more convenient detection of sign languages [2]. Convolutional Neural Network (CNN) to classify Myanmar Sign Language image datasets, discusses skin colour enhancement method for skin detection and Viola Jones algorithm for face detection was used in other paper. But Viola Jones algorithm can only detect frontal image. So if we detect with unconstrained face, the system will be difficult to detect the face and will not be able to work [3].

Hand gesture recognition framework was proposed for Bangla sign language to minimize the communication gap between deaf and dumb people. Hand detection and gesture recognition from the detected hand is targeted output in this study. To gain better accuracy, they have created a dataset of 3219 images. The development of this dataset will reduce the limitation of a prospective study in this field. Hence, they worked with this dataset and emphasizing implementing a more realistic real-time sign language recognition technique [10]. A vision-based American sign language translator that incorporates three concatenated modules- SSD, CNN and FCN was proposed. The translator is trained using cross validation technique. Several experiments were performed but the results of the experiment with the best setting were given in that paper. Experiments also showed that the sign labelling algorithm was capable of automatically detecting and differentiating the different signs within a short period of time with very high accuracy [11].

In all these different methods they have developed model for converting sign language into text or they have classify signs according to sign language used in their country. Although, these models provide better accuracy but, every time using laptop for sign language recognition is tough in day-to-day life. Therefore, we need some application which can be implemented on the handheld device like mobile phones. Nowadays, everyone has their own mobile phone that they can easily carry everywhere. So, we have proposed an application which we can easily install in our phones and through that we can easily recognize different signs used in sign language which will ultimately help to bridge communication gap between deaf and dumb people with other individuals.

III. PROPOSED SYSTEM

According to the reference of available studies there are different models for sign language recognition Although, these models provide better accuracy but, every time using laptop for sign language recognition is tough in day-to-day life. Therefore, we need some application which can be implemented on the handheld device like mobile phones. Nowadays, everyone has their own mobile phone that they can easily carry everywhere. So, we have proposed an application which we can easily install in our phones and through that we can easily recognize different signs used in sign language which will ultimately help to bridge communication gap between deaf and dumb people with other individuals.

An Artificial Neural Networks (ANNs) are the neural networks which mimic the computational structure and working of biological neurons. ANN is a computing system made up of number of simple, highly interconnected processing elements, which process information by their dynamic state response to external input. Neurons are interconnected forming the layers, when there are more than 3 layers of neurons it is called a Deep Neural Network. This neural network is trained over a dataset to form a model that detects Indian Sign Language (ISL) hand gestures. ISL is a language of deaf community in India. ISL contains many

static gestures, in our main dataset we have focused on dataset of common ISL phrases.
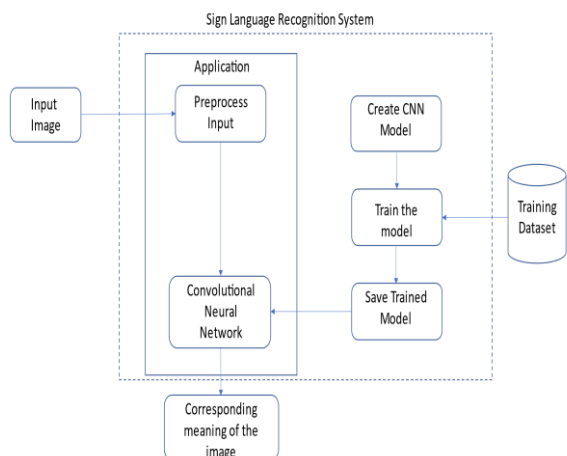

Fig 1. System Architecture

System Architecture is shown in figure 1. The entire system works in stages.

We have collected real-world samples for recognition system. The list of observed ISL were gathered via Kaggle website. Kaggle is a website which provides various datasets of various types. Our dataset contains both high and low resolution images. The dataset is divided into training and testing dataset. The dataset is trained using CNN algorithm. It is important for the model to extract only valuable features while ignoring the other objects or unimportant features in the background, such as person's clothes, skin color and shadows.

➤ *Convolution Neural Network:*
Convolution Neural Network (CNN) is a neural network which is used in image processing. The model recognizes what is in an image or what an image is. CNN has two layers that are Convolution layer and Pooling layer.

A. Convolution
Convolution layer maps or creates a feature map from a dataset. Feature map uses pixels, a filter of stride (n*n) is predefined, which moves across the images and extracts the pixels which contains a value or feature.

B. Pooling
Pooling layers down samples the output of previous layers by sliding the filter of stride size and calculates the maximum or average of the input.

The model is trained using CNN algorithm, this model is saved in TensorFlow Lite (tflite) API. tflite model can be accessed using Flutter, Flutter is a plugin for accessing TensorFlow Lite API in Android or iOS. After importing the trained model in mobile application, the main function of detection uses camera. When a user wants to detect a sign, he/she will start the application. Detection can be done in two ways firstly, using the camera we capture the image and second way, we can use a pre-existing photograph from

gallery. Once the image is provided or captured the image is sent for pre-processing, the trained model detects the sign in the image and gives corresponding output in the text format.
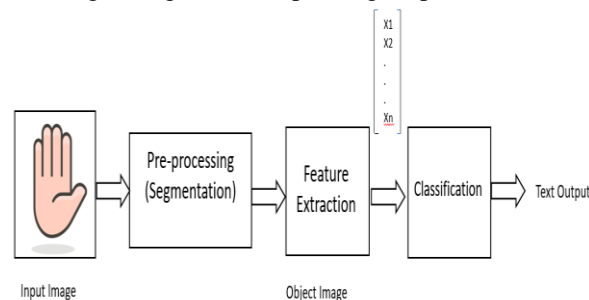

Fig 2. Phases of sign Recognition

In the phases of sign language recognition first, we capture the input image then, we pre-process that input image and perform segmentation to extract important feature from the image. After feature extraction we classify that image according to the signs used in sign language and show the output as the text which can easily interpret the meaning of particular sign that we have used as the input.

➤ *Pre-Processing:*
Captured image is prepossessed, pre-processing is a step which do not increase image information, its aim is to improve the image data and supress undesired distortions or enhances some image features. We will be using the image segmentation technique, segmentation partitions an image into multiple parts based on pixels. Important features are grouped and data is sent to feature extraction.

➤ *Feature Extraction:*
The pre-processed data is taken as input. The important features are extracted by transforming raw data into numerical features that can be processed. The extracted features are in numerical form called as feature vector.

➤ *Classification:*
The feature vector is then sent for classification. Image classification is the process of taking an input and outputting its class or probability that input is of a particular class. The classification is done according to the class. The classified object is our output. The output is displayed into text format.

IV. CONCLUSION

Deaf and dumb people communicate with other using sign language, however it is difficult for normal people to understand them. So, there is a need of a system which recognizes the different signs, gestures and conveys the information to the normal people. Hence, we are proposing the idea of sign language recognition application which uses convolution neural networks to recognize different signs used in sign language. This application will be develop for individuals who cannot use gloves, sensors and other equipment for communication. In this proposal we have designed an idea of application which will convert sign language into text. We want to bridge the communication gap between physically challenged people and normal people, but

instead of using high-end technologies like glove, we want to recognize gestures from photographs (which can be accessed from a webcam), and then use computer vision and machine learning techniques to extract specific features and classify them and produce output as a text.

REFERENCES

[1]. Soma Shrenika, Myneni Madhu Bala: Sign Language Recognition Using Template Matching TechniqueUniversity College London July 05,2020 from IEEE Xplore.

[2]. Rajarshi Bhadra, Subhajit Kar: Sign Language Detection from Hand Gesture Images using Deep Multi-layered Convolution Neural Network. 2021 IEEE Second International Conference on Control, Measurement and Instrumentation (CMI), India

[3]. Sai Myo Htet, Bawin Aye, Myo Min Hein: Myanmar Sign Language Classification using Deep Learning.California State University Fresno. Downloaded on June 21,2021 from IEEE Xplore

[4]. Sarika Jain, Deepika Pahuja: Recognition of Sign Language Symbols using Templates, 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) Amity University, Noida, India. June 4-5, 2020

[5]. Ebey Abraham, Akshatha Nayak, Ashna Iqbal: Real-Time Translation of Indian Sign Language using LSTM 2019 Global Conference for Advancement in Technology (GCAT) Bangalore, India. Oct 18-20, 2019

[6]. Ashok K Sahoo, Gouri Sankar Mishra and Kiran Kumar Ravulakollu: Sign Language Recognition: State of the Art. ARPN Journal of Engineering and Applied Sciences. Vol. 9, No. 2, February 2014

[7]. Galib Ibne Haidar, Hasin Ishraq Reefat: Glove Based American Sign Language Interpretation Using Convolutional Neural Network and Data Glass. 2020 IEEE Region 10 Symposium (TENSYMP), 5-7 June 2020, Dhaka, Bangladesh

[8]. Necati Cihan Camgoz, Oscar Koller, Simon Hadfield, Richard Bowden: Sign Language Transformers: Joint End-to-end Sign Language Recognition and Translation. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)

[9]. Dr.L.Priya, A.Sathya, S. Kanaga suba raja: Indian And English Language To Sign Language Translator- An Automated Portable Two Way Communicator For Bridging Normal And Deprived Ones.2020 2nd International Conference on Power, Energy, Control and Transmission Systems May 19,2021 from IEEE Xplore.

[10]. Dardina Tasmere, Boshir Ahmed: Hand Gesture Recognition for Bangla Sign Language Using Deep Convolution Neural Network. 2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI), 19-20 December, Dhaka

[11]. Rahib Abiyev, Rahib Abiyev, Murat Arslan: Reconstruction of Convolutional Neural Network for Sign Language Recognition. Proc. of the 2nd International Conference on Electrical, Communication and Computer Engineering (ICECCE) 12-13 June 2020, Istanbul, Turkey

[12]. Rachana Patil1, Vivek Patil1,, Abhishek Bahuguna1, Mr. Gaurav Datkhile: Indian Sign Language Recognition using Convolutional Neural Network. ITM Web of Conferences 40, 03004 (2021) ICACC-2021

[13]. Ying Xie, Kshitij Bantupalli: American Sign Language Recognition using Deep Learning and Computer Vision Department of Computer Science Kennesaw State University Kennesaw, USA IEEE International Conference on Big Data (Big Data) 2018