WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Review on Conversion of Sign Language to Text

Veda Chiluka*, Vaishnavi Mate*, Sakshi Jagtap*, Om Darvitkar* and Prof. K. S. Jetha*

*Student,Department of Information Technology,ABMSP's,Anantrao Pawar College of Engineering & Research, Pune, India
**HOD,Department of Information Technology,ABMSP's,Anantrao Pawar College of Engineering & Research, Pune, India

ABSTRACT

Sign languages are communication systems used by the deaf community in which different types of hand gestures are used. In India, there are between 10 and 15 million deaf people. There is no universal sign language in India for the deaf community to communicate. A deaf person who communicates solely through gestural sign language will find it difficult to communicate with someone who does not understand sign language. This results in a variety of issues. Despite the fact that many sign languages exist, the general public is unaware of them. As a result, communicating with deaf people becomes more challenging. There is a requirement for a translator to translate what they want to say. Our work attempts to improve deaf-to-deaf communication. The proposed concept (sign language conversion) represents a significant step forward in assisting the deaf community.

Keywords—Convolutional Neural Network (CNN), Image Processing, Gesture Recognition, Machine Learning.

1.INTRODUCTION

Dumb and deaf people can communicate with one other and the rest of the world via sign languages. Around the world, there are approximately 135 different sign languages, including American Sign Language (ASL), British Sign Language (BSL), and Australian Sign Language (Auslan), among others. According to the World Health Organization, 466 million individuals worldwide have impaired hearing loss (more than 5% of the global population), with 34 million of them being teens (WHO). According to studies, by 2050, these numbers will have surpassed 900 million. Furthermore, the majority of cases of profound hearing loss, which afflict millions of individuals, occur in low- and middle-income nations. Hearing impaired people are shut out of internet conferences, office meetings, and school. They normally communicate via basic text chat, which isn't the most efficient approach. Deaf people need to be able to connect naturally with their healthcare network, coworkers, and peers, regardless of whether the other person knows sign language, as telehealth adoption grows. Digital Image Processing is used in vision-based Sign Language Translation. It is a framework for perceiving and interpreting uninterrupted gesture-based communication in relation to English content. A camera is utilized as input in vision-based gesture recognition. Before being processed, videos are broken down into frames. As a result, vision-based methods are favoured over gesture-based methods since anyone with a smartphone can convert sign language to text or speech, and it is reasonably inexpensive.

II.METHODOLOGY

Image Processing

Image processing is a technique for performing operations on a picture in order to improve it or extract relevant information from it. It's a sort of signal processing in which the input is an image and the output is either that image or its characteristics/features. Image processing is one of the most quickly evolving technology today. It is also a critical research field in engineering and computer science. The three phases that make up image processing are as follows:

- Using picture acquisition tools to import the image
- Examining and modifying the image
- Output that may include a changed image or a report based on image analysis.

Analogue and digital image processing are the two types of image processing methods employed. Hard copies, such as prints and photographs, can benefit from analog image processing. When employing these visual tools, image analysts employ a variety of interpretive fundamentals. Digital image processing techniques allow for computer-assisted alteration of digital images. Preprocessing, augmentation, and presentation, as well as information extraction, are the three general processes that all sorts of data must go through when using digital techniques.

Image processing in digital format: The alteration of photographs using digital computers is known as digital image processing. In the previous few decades, its popularity has skyrocketed. Its uses vary from medicine to entertainment, with geological processing 2 and remote sensing thrown in for

good measure. Digital image processing is significantly used in multimedia systems, which are one of the pillars of the current information society. The manipulation of those finite precision numbers is what digital image processing is all about. Digital image processing can be classified into four categories: image enhancement, picture restoration, image analysis, and image compression. An image is enhanced by using heuristic techniques to modify it such that a human viewer can extract relevant information from it. 22 The term "digital image processing" refers to the use of a computer to manipulate images. The process of subjecting a numerical representation of an item to a series of operations in order to produce a desired outcome is known as digital image processing. The translation of a physical image into a digital image and the extraction of data are both part of digital image processing. by utilizing multiple algorithms to extract meaningful information from a digital image

Pattern Recognition: On the basis of image processing, it is important to use pattern recognition technology to separate objects from images, then use statistical decision theory technologies to identify and classify these things. When an image contains many items, pattern recognition is divided into three phases, as shown in Fig.5.2.1. The image segmentation and object separation are part of the first phase. Different objects are recognized and separated from the backdrop during this phase. The feature extraction phase is the next step. Objects are measured in this phase. During feature extraction, a group of features is merged to form a feature vector, and the measuring feature is used to quantitatively estimate some significant features of objects. The classification phase is the third step. The result of this step is just a choice about which category each object belongs to. As a result, photos are the input for pattern recognition, and object kinds and structural analysis of images are the output. The structural analysis is a description of images that allows you to accurately interpret and judge the key information included within them.

The Convolutional Neural Network (CNN)

A convolutional neural network is a feed-forward ANN in which the structure of the human visual cortex influences the association pattern between perceptrons. Perceptrons are arranged in recurrent blocks throughout space and time in CNNs. These recurrent clusters of perceptrons in photographs can be thought of as two-dimensional convolutional kernels that are continually applied over each area of the image. They are viewed as onedimensional filters installed across windows for speech. The weights of these replicated clusters are shared during training, i.e. the mean of weight gradients learned from different image parts is calculated. CNNs are a type of neural network that is particularly useful in the field of computer vision. They were inspired by the actual perception of vision that takes place in our brain's visual cortex. They employ a filter/kernel to scan over the whole pixel values of the image and do computations by assigning appropriate weights to enable feature detection. Convolution layer, max pooling layer, flatten layer, dense layer, dropout layer, and a fully connected neural network layer are all included in the CNN. Together, these layers form a highly strong tool for identifying features in an image. The lower layers detect low-level features, which are subsequently replaced with higher-level features that are more sophisticated.

Dataset

To find the bounding boxes of diverse objects, we employed Gaussian historical past subtraction, which is a technique that uses a mixture of K Gaussian set distributions to version each history pixel (k varies from 3 to 5). The colorations that have a longer history are those that have a higher static. We're working on those swaying pixels.

Create a square-shaped bounding field. After gathering all of the gesture and heritage data, a Convolutional NN model was created to separate the gesture symptoms and indicators from the historical context utilizing those photographs. These function maps show how the CNN can recognize common unexposed structures in some of the gesture indicators in the training set, allowing it to differentiate between a gesture and the past.

III REAL-TIME CONVERSION OF SIGN LANGUAGE TO TEXT AND SPEECH

This paper compares different techniques and chooses the most optimal approach for creating a vision-based application for sign language to text/speech conversion for deaf/dumb people. The proposed system could efficiently recognize the alphabets from images using a customized SVM model. This project is aimed at societal contribution.

IV SIGN LANGUAGE TO TEXT AND SPEECH TRANSLATION IN REAL TIME USING CONVOLUTIONAL NEURAL NETWORK

The project is a straightforward instance of how CNN may be used to handle computer vision problems with exceptional precision. By creating the relevant dataset and training the CNN, a finger spelling sign language translator with an accuracy of 95sign languages is obtained. The approach is able to solve a portion of the Sign Language translation problem since sign languages 10 are spoken in context rather than as finger spelling languages. The major goal has been accomplished, namely, the requirement for an interpreter has been removed. When we're working on the project, there are a few details to keep in mind. The thresh must be kept under control to avoid distorted grayscales in the frames. If this happens, we'll need to either reset the histogram or hunt for a location with adequate illumination. We could also employ gloves to solve the issue of the signee's fluctuating skin tone. We were able to obtain precise prediction in this project after we began testing using a glove. Another challenge that people may have is their understanding of ASL gestures. The use of poor gesture postures will not result in accurate prediction. This project can be improved in the future in a few ways. It could be constructed as a web or mobile application enabling users to access the project more easily. Also, the existing project only works for ASL, but with enough dataset and training, it could be extended to work for additional native sign languages. This project implements a finger

spelling translator, however sign languages are also spoken in context, with each gesture representing an object or a verb, therefore identifying this type of contextual signing would necessitate a higher level of processing and natural language processing (NLP). This project's scope does not include this.

V SIGN LANGUAGE RECOGNITION SYSTEM USING CONVOLUTIONAL NEURAL NETWORK AND COMPUTER VISION

Artificial intelligence, machine learning, and computer vision have all seen significant advancements. They've done it. greatly influenced how we interpret the world around us and improved how we implement their techniques in our daily lives. Many studies on sign gesture recognition have been undertaken utilizing various algorithms such as ANN, LSTM, and 3D CNN. However, the majority of them necessitate additional computing power. On the other hand, in order to extract features (binary pixels) and make the system more resilient, our research article requires little computational power and achieves a stunning accuracy of over 90 to 64 pixels. We used CNN to classify ten alphabetical American sign motions and were able to attain a 98 percent accuracy rate.

THE PROBLEM:

The movements, body language, and facial expressions used in sign languages vary greatly from country to country. The grammar and structure of a sentence might also differ significantly. Learning and capturing gestures was a difficult task for us in our study because hand movement has to be accurate and on point. It's tough to duplicate some gestures. It was also difficult to keep our hands in the same posture as we created our dataset.

FUTURE WORK: We hope to include new alphabets in our datasets and develop the model so that it can recognize more alphabetical features while maintaining a high level of accuracy. We'd like to improve the system even more by include speech recognition so that blind individuals can benefit as well.

VI SIGN LANGUAGE CONVERTER RECOGNITION

The goal of this project is to use a Convolution Neural Network to recognize hand gestures. Identification of sign language, a lively tool of communication for physically impaired, deaf, and dumb people, is one of the most important applications of hand gesture recognition. This application will aid in the communication between deaf/dumb people and the general public.

We may deduce from the results that the Convolution Neural Network has a high level of accuracy in identifying sign language characters such as alphabets and numerals. This work can be expanded upon to create a real-time program that can detect sign language and recognize words and sentences rather than just characters or single words.The many algorithms and strategies for recognizing hand gestures are discussed in this study. A hand gesture recognition system is seen as a technique to make human-computer interaction more intuitive and effective. Virtual prototyping, sign language analysis, and medical training are among the applications. Physically disabled, deaf, and stupid persons use sign language as one of their communication tools. It is obvious from the preceding discussion that vision-based hand gesture recognition has made significant progress in the field of hand gesture recognition. The C, C++, and Java programming languages are used to implement the gesture recognition system. To make job easier, in particular MATLAB with the image processing toolbox is utilized when image processing activities are required.

VIII. SIGN LANGUAGE TEXT TO SPEECH CONVERTER USING IMAGE PROCESSING AND CNN

The technology will give a user interface that will allow deaf individuals to converse with ease using sign recognition. The technique can be used not only in a household setting, but also in a public setting. This strategy is particularly useful for deaf and dumb persons in social situations. We'll use the OpenCV toolbox to create a simple gesture recognizer that will be integrated into the Visionary framework. We'll price and down hand gestures regardless of which hand is used as a yes gesture

IX. CONCLUSION AND FUTURE SCOPE

This study evaluates multiple strategies and determines the best approach for developing a vision-based application for deaf/dumb persons to convert sign language to text/speech. Using a customized SVM model, the proposed system could efficiently recognize alphabets from photos.

Today's applications require a variety of image types as sources of data for explanation and analysis. Several characteristics must be retrieved in order to conduct various tasks. Degradation occurs when a picture is converted from one form to another, such as when digitizing, scanning, sharing, storing, and so on. As a result, the resulting image must go through an image enhancement process, which consists of a collection of approaches aimed at improving an image's visual presence. Image enhancement improves the interpretability or awareness of information in images for human listeners while also giving superior input for other automatic image processing systems. The image is then subjected to feature extraction using a variety of approaches in order to make it more computer-readable. A sign language recognition system is a useful tool for preparing expert knowledge, detecting edges, and combining erroneous data from

REFERENCES

[1] Kohsheen Tiku, Jayshree Maloo, Aishwarya Ramesh and Indra R."Real- time Conversion of Sign Language to Text and Speech".Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020).

[2] Dhanashree Bhandarkar, Amit Kale, Prajakta Mukhekar, Prof.Rohini Nere."Sign Language Converter Recognition".International Journal of Advanced Research in Computer and Communication Engineering Volume: 9, Issue: 05 | May 2020.

[3] Mangesh B, Mayur K, Rujali ."Sign Language Text to Speech Converter using Image Processing and CNN".International Research Journal of Engineering and Technology (IRJET) Volume: 07 Issue: 04 | Apr 2020.

[4] Ms. Kamal Preet Kour, Dr. (Mrs) Lini Mathew."LITERATURE SURVEY ON HAND GESTURE TECHNIQUES FOR SIGN LANGUAGE RECOGNITION".International Journal of Technical Research Science.ISSN No.: 2454- 2024.

[5] Lionel Pigou(B), Sander Dieleman, Pieter-Jan Kindermans, and Benjamin Schrauwen."Sign Language Recognition Using Convolutional Neural Networks".