International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Design and Development of Indian Sign Language Character Recognition System

[1]Ifham Khwaja, [2]Naman Sanklecha, [3]Kaushik Rathod, [4]Prof. Pragya Sinha

[1]Student, Computer Engineering, Vishwakarma University, Kondhawa, Pune, Maharashtra, India.

[2]Student, Computer Engineering, Vishwakarma University, Kondhawa, Pune, Maharashtra, India.

[3]Student, Computer Engineering, Vishwakarma University, Kondhawa, Pune, Maharashtra, India.

[4]Professor, Computer Engineering, Vishwakarma University, Kondhawa, Pune, Maharashtra, India.

Abstract - The purpose of this study is to look into the challenges involved in categorizing Indian Sign Language (ISL) characters. While a lot of research has been done in the related field of American Sign Language (ASL), not as much has been done with ISL. Lack of standard datasets, obscured traits, and variance in language with geography are the key barriers that have hindered much ISL research. Our study aims to progress this field by collecting a dataset from a deaf school and applying various feature extraction techniques to extract useful information, which is then input into a range of supervised learning algorithms. Our current results for each approach include four fold cross-validation. What sets our work apart from earlier research is that the validation set in our four-fold cross-validation contains photographs of people who are not the same people as those in the training set. Hand gestures and signs are used by those with speech impairments to communicate. Understanding what they're trying to say is challenging for the average person. Though extremely uncommon, there are many systems that convert data to Hindi. Therefore, it is imperative to implement a system that enables the general public to understand and interpret all signals, gestures, and communications. It will close the communication gap that exists between normal people and those who have speech difficulty. The two primary research approaches centered on human-computer interaction are sign language recognition and learning. Multiple sensors are required in order for data flow to be understood in sign language. This research paper focuses on the development of a Hindi-language training tool that can detect images and interpret what someone else is trying to say to persons who have speech impairments.

Keywords: Indian Sign Language (ISL), American Sign Language (ASL), Feature Extraction, Supervised learning, Sign Language, etc.

I. INTRODUCTION

The goal of this project is to recognize Indian Sign Language alphabets by their associated movements. For American Sign Language (ASL), gesture and sign language recognition have been extensively studied; however, there aren't many published research studies on Indian Sign Language (ISL). However, rather of employing expensive gear like kinects or gloves, our goal is to overcome this issue with cutting-edge machine learning and computer vision methods.

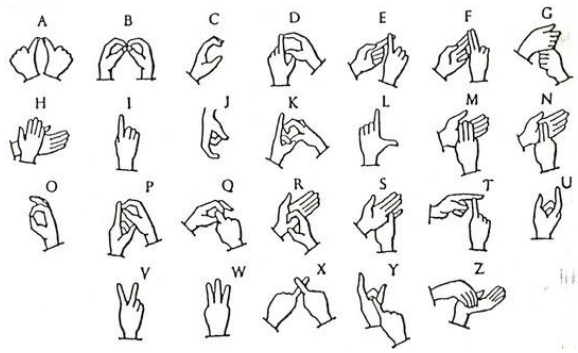


Figure 1: Indian Sign Language

Hearing impairment or deafness is prevalent throughout the nation and its territories. Over 1.5 billion people worldwide—nearly 20% of the global population—have hearing loss today. There are 430 million people that are hard of hearing. It is projected that more than 700 million individuals worldwide may suffer from deafness by the year 2050. The fact that many deaf persons have never learned sign language—which is essential for communicating with a sizable portion of the global population—affects them as well. [WHO] Individuals who suffer from hearing loss or speech impairments encounter numerous challenges while attempting to communicate or engage with others. The primary means of communication is through gestures, or physical motions, particularly hand With varying finger forms, gestures are the most common way to communicate. Look at Figure 1 for a variety of hand gestures.

According to WHO estimates, there are over 63 million individuals with significant hearing impairment in India. In an effort to assist these people, recent technical advancements have produced a variety of abilities that may be modified to enhance communication. The majority of people in India speak Hindi. A system of sign language must be established in order to transmit information in Hindi. This may be accomplished with Python QT, OpenCV, and CNN convolutional neural networks. It will be very beneficial for those who cannot read or speak Hindi thanks to our technology.

II. LITERATURE SURVEY

Sign language Recognition gained a lot of researchers' attention in the past years. Many Researchers had done research in this area for minimizing the communication gap between deaf people and normal people. Various methods have been tried so that the system can work efficiently. Sign language recognition can be done either by using the glove-based method or the vision-based method, the former one is having better accuracy but has limited movement and setup is quite sensitive. In the vision-based method, there are two steps, first one is feature extraction that uses image processing for capturing static frames from video capture and the second is the recognition part that performs learning by identifying patterns in training data. Many researchers have used various techniques to develop a system that can translate static sign language into its corresponding word.

The paper by Karlo et al. [1] aimed to develop a Convolutional Neural Network-based model that is trained on static sign language for American Sign Language alphabets (A to Z), numbers (1 to 10) and some basic words with an average testing accuracy of 93.67%.

The paper published by Mehreen Hurroo and Mohammad Elham Walizad [2] provides a method that requires low computational power. They have used American Sign Language(ASL) alphabets to create a dataset and a Static recognition vision-based method to capture sign gestures. A dataset of 2000 ASL images of 10 static sign alphabets A, B, C, D, H, K , N ,O, T , Y was made to train the Convolutional Neural Network in which 1600 are training images and 400 are testing images. The images are shot using a web camera on which the HSV color algorithm is applied to eliminate the background. Then Segmentation is performed and a binary image of size 64*64 is obtained where the white area of the image shows the sign gesture and other parts are shown in black. Later Feature extraction using CNN extracted the binary pixels from each frame and CNN is applied for training and classification. They have used 2D CNN model with a tensor flow library in their proposed system. The results used in this paper gave an accuracy of over 90.0% and alphabetical classification using CNN accuracy of 98%. They also proposed in their research to normalize and rescale the images to 64 pixels to extract features (binary pixels) and make the system more robust.

In the study done by Dolas [3], they used K- Nearnest Neighbor's algorithm for classification on a training data set of images of 5 gestures. A series of preprocessing events were used to remove noise, filling unwanted holes and smoothening the images for better recognition. They have used API for text to speech conversion so that final output of classified result can be given in speech form also. Hand movement in gestures was the challenged faced by them.

Some of these researches have been known to be successful and such systems are developed using different methodologies but still, they are unable to commercialize because of their cost. Suharjito et al. [4] published a paper that helps researchers to research that method and develop a system that can be used commercially. This paper provides a comparison between various techniques and methods that has been used by researchers. So that the best method can be chosen. It gives information about the data acquisition and classification approaches which are the mandatory steps in developing a sign language recognition system. It states that for data acquisition the researchers mainly used two methods named Camera and Microsoft Kinect. All other methods are; 3-axis accelerometer and flex sensors, a 5-Dimensional Tracker, 2 Cyber gloves, and a 3SPACEposition tracker. These methods are equipped with sensors attached to the gloves which provide more accurate data and finger movement data. But these all are very costly so it is difficult to be used commercially. There are also various classification methods like hidden Markov Model, Convolution Neural Network, Self-Organizing Map, and SimpSVM(Simple Support Vector Machine) There are some methods that use a combination of SOFM(Self Organizing Feature Map), SRN(Simple Recurrent Network), and HMM. Other researchers use the wavelet family method, Eigen Value, and Euclidean distance to classify the sign language.

It can be seen that sign language recognition has been a well-researched topic for American Sign Language (ASL), but in the case of Indian Sign language(ISL), the number is very low. There is an extensive difference between American Sign language(ASL) and Indian Sign Language(ISL) because ASL uses a single hand whereas ISL uses both hands for communicating. One of the researches that use ISL is done by Madhuri Sharma et al. [5]. This paper talks about the idea of the project which is to design a system that can interpret the Indian sign language in the domain of alphabets accurately. In their work they have described four for recognition of sign laungauge. The first method used was the hand tracking system in which the CAM-SHIFT method was employed. The second was the hand segmentation in which HSV color model and neural network. The third one was used for feature extraction in which Generic Fourier Descriptor (GFD) method was used. The fourth was gesture recognition in which a Genetic algorithm was used. The data set was divided into two groups, one used for training and the other for testing. The training set consists of 70% of the aggregate data and the remaining 30% are used as testing. We also perform experiments on the same (30% or 70%) dataset which is training as well as testing for the KNN classifier. The results of these experiments have a 100% accuracy rate. The gesture recognition system is capable of recognizing only alphabetical signs with 97.10% accuracy.

III. PROPOSED SYSTEM

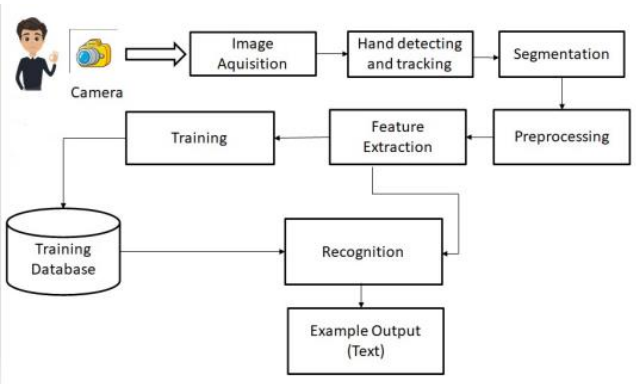


Figure 2. Proposed Methodology.

Image Acquisition:
It is the process of taking an image out of a source—usually one that is hardware-based—in order to process it. The hardware-based component of our concept is WebCamera. It is the initial stage in the workflow sequence since a picture is necessary for all processing to be done. The resulting image has not undergone any kind of processing.

Segmentation:
Segmentation is the process of taking items or signs out of the background of an image that has been taken. The segmentation process makes use of edge detection, context subtraction, and skin-color detection. Recognizing gestures requires the detection and segmentation of hand motion and location.

Features Extraction:
Preprocessed images are used to extract predefined features, such as form, contour, geometrical feature (position, angle, distance, etc.), color feature, histogram, and others, which are then utilized for sign identification or classification. A step in the dimensionality reduction process that separates and arranges a sizable amount of raw data is feature extraction. Lowered to more manageable, smaller classrooms Processing would be easier as a result. The most significant aspect is the sheer quantity of variables present in these enormous data sets. This data requires a significant amount of processing power to process. So, by choosing and combining variables into functions, function extraction helps to extract the optimal feature from enormous data sets. These features accurately and uniquely describe the actual data collection process, and they are also very user-friendly.

Preprocessing:
Preprocessing techniques such as erosion, dilation, and Gaussian smoothing are applied to each image frame in order to remove noise. Converting a color image to grayscale results in a smaller image. Grayscale picture conversion is a popular way to minimize the quantity of data that needs to be processed. Preprocessing goes through the following phases:
a) *Morphological Transform (Morphological Transform):*
To produce an output image with a comparable size, morphological processes employ a structural feature on the input image. To find the value of each pixel in the output image, it compares the matching pixel in the input image with its neighbors. Morphological alterations come in two flavors: erosion and dilation.
Dilation: The output pixel's value is the maximum value of all the nearby pixels. If every pixel in a binary image has the value 1, then that pixel is set to 1. Morphological dilatation closes tiny gaps and makes artifacts more visible.
Erosion: The value of the o/p pixel is the lowest of all the nearby pixels. In a binary image, a pixel is set to 0 if every one of its neighbors is also 0. Morphological erosion removes minor artefacts, leaving behind larger objects.

b) Blurring:
Blurring occurs when an image is passed through a low-pass filter. In computer vision, the term "low-pass filter" describes a technique for removing noise from an image while preserving the integrity of the remaining portion. Before moving on to more complex tasks, including edge detection, a blur is a basic process that needs to be finished.

c) Thresholding:
Thresholding is a type of image segmentation where an image's pixels are altered to facilitate picture interpretation. The act of turning a grayscale or color image into a binary—basically, black and white—is known as thresholding. Thresholding is most frequently used to choose regions of interest in an image while disregarding the parts we don't care about.

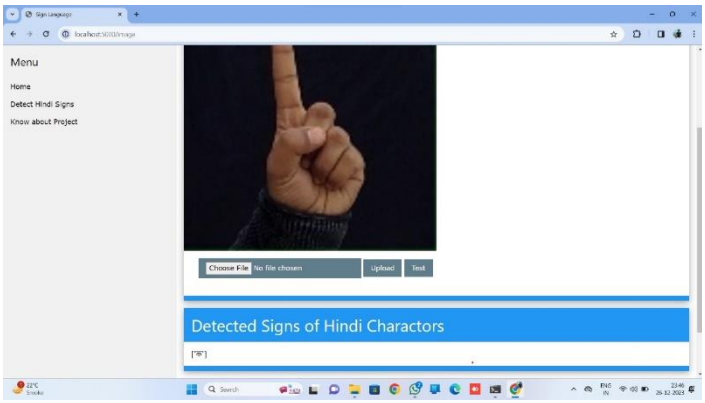
d) Recognition:
In this instance, classifiers will be used. The techniques or algorithms used to interpret the signals are called classifiers. The Hidden Markov Model (HMM), K-Nearest Neighbor classifiers, Support Vector Machine (SVM), Artificial Neural Network (ANN), and Principle Component Analysis (PCA) are a few popular classifiers that recognize or comprehend sign language. CNN will be used as the classifier in this research, nevertheless. CNNs are utilized for picture recognition and classification due to their high precision. Utilizing a hierarchical paradigm, the CNN creates a network akin to a funnel before producing a fully-connected layer that processes the output and connects all of the neurons.

e) Text output:
Recognizing and translating into text different body postures and movements, as well as comprehending human behavior.

IV. RESULTS AND DISCUSSIONS

It can be inferred from the results that the validation test of the Hindi Sign Language Recognition System was considered legitimate that the system is a valuable tool for translating sign language into Hindi.
We have developed a model that can recognize gestures in Indian sign language and provide textual and auditory output. Three possibilities were created in the sign predictor. such as life prediction, image loading, and image classification. If the user wishes to create a forecast based on photographs, we offer real-time gesture detection using life prediction. To facilitate this process, we have included the opportunity for the user to upload images. We obtained 98% accuracy for both the alphabet and numbers and 96% accuracy for words when we trained the model. When we provide the system with input, we need new light for the system to function at its best. 90% of the data were utilized for training, and 10% were used for testing. More data might be used for testing and training, which would improve the accuracy and output of the system.

Figure 3.



V. CONCLUSION

An approach for identifying a group of formed signs and converting them into text or voice with the proper context is the Hindi Sign Language Recognition System. The development of efficient human-machine interactions demonstrates the importance of gesture recognition. For this research, we tried to create a model using a convolutional neural network.

REFERENCES

[1] Tolentino, Lean Karlo & Serfa Juan, Ronnie & Thio-ac, August & Pamahoy, Maria & Forteza, Joni & Garcia, Xavier. (2019). Static Sign Language Recognition Using Deep Learning. International Journal of Machine Learning and Computing. 9. 821-827. 10.18178/ijmlc.2019.9.6.879.

[2] Mehreen Hurroo , Mohammad Elham, 2020, Sign Language Recognition System using Convolutional Neural Network and Computer Vision, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 09, Issue 12 (December 2020).

[3] Vijay More, Sanket Sangamnerkar, Vaibhav Thakare, Dnyaneshwari Mane, & Rahul Dolas. (2021). SIGN LANGUAGE RECOGNITION USING IMAGE PROCESSING. JournalNX - A Multidisciplinary Peer Reviewed Journal, 85–87.

[4] Suharjito, Suharjito & Anderson, Ricky & Wiryana, Fanny & Ariesta, Meita & Kusuma Negara, I Gede Putra. (2017). Sign Language Recognition Application Systems for Deaf-Mute People: A Review Based on Input-Process-Output. Procedia Computer Science. 116. 441-448. 10.1016/j.procs.2017.10.028.

[5] Sahoo, Ashok. (2014). Indian sign language recognition using neural networks and kNN classifiers. Journal of Engineering and Applied Sciences. 9. 1255-1259.