iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42626

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Sign Language Recognition System

Prof. Pritesh Patil[1], Ruchir Bhagwat[2], Pratham Padale[3], Yash Shah[4], Hrutik Surwade[5]
[1, 2, 3, 4, 5] *Department of Information Technology, AISSMS Institute of Information Technology, Pune, Maharashtra, India*

Abstract: ***A large number of deaf and mute people are present around the world and communicating with them is a bit difficult at times; because not everyone can understand Sign language(a system of communication using visual gestures and signs). In addition, there is a lack of official sign language interpreters. In India, the official number of approved sign language interpreters is only 250[1] . This makes communication with deaf and mute people very difficult. The majority of deaf and dumb teaching methods involve accommodating them to people who do not have disabilities - while discouraging the use of sign language. There is a need to encourage the use of sign language. People communicate with each other in sign language by using hand and finger gestures. The language serves its purpose by bridging the gap between the deaf-mute and speaking communities. With recent technological developments, sign language identification is a hard subject in the field of computer vision that has room for further progress. In this project, we propose an optimal recognition engine whose main objective is to translate static American Sign Language alphabets, numbers, and words into human and machine understandable English script and the other way around. Using Neural Networks, we offer a machine learning-based technique for identifying American Sign Language.***

Keywords: deep learning; convolutional neural network; recognition; comparison; sign language;

I. INTRODUCTION

The majority of the world's deaf community uses American Sign Language(ASL), a visual gesture language. The use of a sign language recognition system is also quite important and necessary in a country like India. ASL is a globally recognised standard for sign language, but only a small number of people understand it, limiting the user's ability to converse in real-life scenarios. We present a system for ASL recognition in this project to address this issue and thus make communication between the speaking and non-speaking communities much easier and simpler. A variety of pre-processing stages are used in the proposed system to turn a gesture image into boundary highlighted images, which are then fed into a machine learning algorithm for identification. Existing solutions necessitate the use of external devices to capture finger movements, such as motion sensing gloves or the Microsoft Kinect. As the result, the feasibility and accessibility are reduced. A variety of pre-processing stages are used in the proposed system to turn a gesture image into boundary highlighted images, which are then fed into a machine learning algorithm for identification. The systems rely heavily on hardware components in their operation, which raises the system cost and makes it difficult to use for ordinary people.

II. LITERATURE SURVEY

Sr. No.	Title	Methodology	Conclusion	Year
1.	Bangla Sign Digit s : A Data set for Real Time Hand Gesture Are cognition	In this present paper, the detailed system for the recognition of Bangla sign digits has been elaborated. The proposed system detects hand area from a real-time video, extracts hand features after preprocessing, and trains a deep CNN model to recognize	A real-time hand gesture recognition system has been proposed to lessen the communication troubles of hearing-impaired people. The proposed methodology incorporates four convolution layers together with four pooling layers including four fully connected layers and a dropout layer with a 40% dropout.	2020

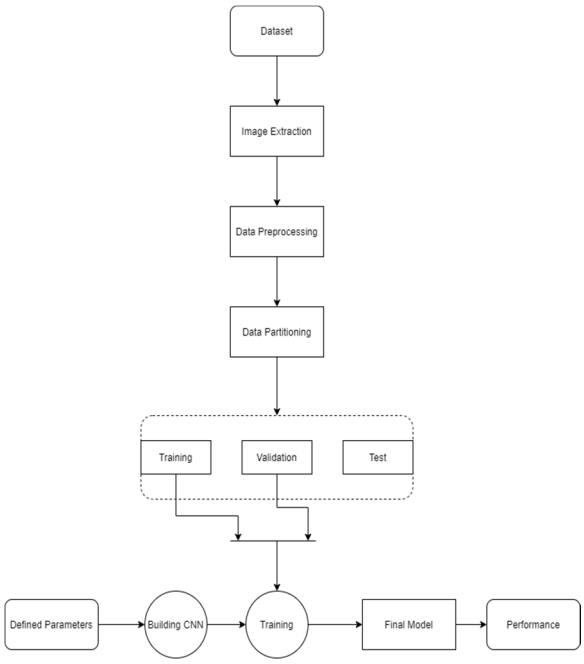
©IJRASET: All Rights are Reserved

2.	Study of Convolutional Neural Network in Recognizing Static American Sign Language	This paper is motivated towards applying Convolutional Neural Network.	a CNN architecture to recognize 24 letters in American Sign Language. The experimental results show the proposed method is effective in predicting static alphabetical gestures, which in return can serve as a beginning step to bridge the communication gap between the deaf-mute person and the community.	2019
3.	A Static Hand Gesture Based Sign Language Recognition System using Convolutional Neural Network	In this base paper, they used and described hand segmentation and detection phase then Database and last CNN algorithm to accomplish desired output	This paper presents a system of static hand gesture movement which can be used to recognize sign languages. They have used ISL for case study. This model has used the Back Projection Histogram algorithm for setting the histogram of the image.	2021
4.	Digit Recognition in Sign Language Based on Convolutional Neural Network and Support Vector Machine	A pre-trained Inception-v3 deep CNN architecture is used for feature extraction from sign digit images. The Inceptionv3 deep CNN architecture was pre-trained on an image dataset with more than 1 million images to achieve state-of-the-art accuracy	This paper presented CNN-SVM model for digit recognition in sign language which achieved recognition accuracy 98.20% in ASL dataset and 98.30% in SLD dataset	2020
5.	Sign Language Detection from Hand Gesture Images using Deep Multi-layered Convolution Neural Network	The images in the database are non-uniform in size and have a constant background. Therefore, in this paper, the images have been rescaled and busy background has been added to make the proposed system more robust	In this paper, a deep CNN architecture consisting of 5 layers has been proposed to detect and classify sign languages from hand gesture images. The proposed methodology uses both static (0 - 9 and A - Z) and dynamic (alone, afraid, anger etc.) gestures in the training, validation and blind testing phase to make the system more robust.	2021
6.	Hand Gesture Recognition for Bangla Sign Language Using Deep Convolution Neural Network	This research objective is to promote the learning process of deaf peoples of Bangladesh. To overcome the restrictions of prior research Using Deep Convolutional Neural Network	Hand detection and gesture recognition from the detected hand is our targeted output in this study. Using HSV and YCbCr color space, the input image was segmented. This segmented hand image gives a better input to our system. The output image has subsequently converted to a normalized image	2020

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

III. SYSTEM DESIGN AND WORKFLOW

A dataset is created by capturing an ample amount of images through a web-cam. When the system starts up, it begins collecting gesture images from the previously generated dataset. The images are then pre-processed and cropped so that only the hand gesture is visible. This is accomplished by drawing contours and determining the contour with the greatest area. The contour region is then photographed and resized to 128 x 128 pixels. These images are then divided into three categories: Training, Validation, and Testing. After this, the model starts building by adding different layers to the image for the accuracy purpose to all the input images. Then, different layers are added to the Model. After the addition of the layers, the training of CNN models began. With the help of CNN, the model builds a network and creates a prediction model which will be used in the prediction of the result. After this, Validation and testing of the model is performed to check whether the model is overfitting or under fitting. Further steps are taken to increase the accuracy of the model later on.



System Architecture

IV. METHOD

A. Data Collection

We are using the OpenCV library to capture images through web-cams. The main objective of using Opencv library is to apply various filters provided by the library. Filters like grayscale conversion and gaussian blur were applied on the target images. Grayscale conversion simply converts the image into black and white image. The input image is in color which is converted into a grayscale image. Gaussian blur removes some of the noise before further processing the image. We are using the 'adaptive-threshold' function to highlight the image borders. The above functions are discussed in detail in further subsections. We are using a Region of interest of 300 x 300 pixels for capturing the hand signs. Below is a code Snippet of how the gestures in Region of interest are captured. We have resized all the images to 128 x 128 resolution to minimize the training time of the model. We have created two folders in our dataset, namely Train and Test, both consisting of individual folders having 600 images of each symbol.

B. Image Pre-Processing

We are using two types of images to train the dataset namely binary images and canny edge images. Binary image consists of pixels of only two colours, usually black and white. Canny edge image shows all the edges present in the image which increase the amount of data to be processed but in turn helps in getting useful structural information from an image. For less symbol classes (10 symbols), using binary images to train the dataset yields good results. But for all 26 English alphabets, canny edge data was found to be more useful than binary data images. For Creating a dataset, we implemented a program and captured the live feed through a web-cam.

C. Implementation

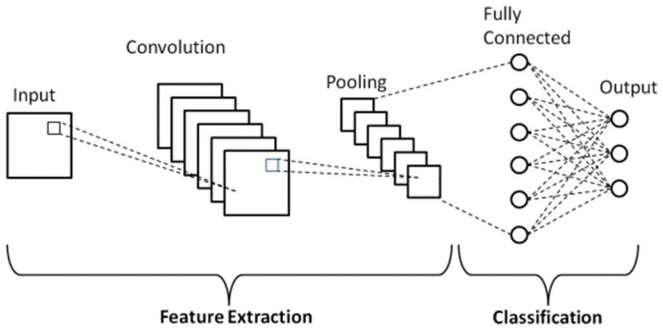
These steps are followed while preprocessing dataset:

1) Capture a RGB image
2) Convert the RGB image into grayscale and apply gaussian blur to it.
3) Applying desired threshold to convert grayscale into binary image. Applying a threshold means selecting some specific value for pixels in image and converting the pixels in image such that: if (pixel) is less than (threshold) : change the color of the pixel to white else : change the color of the pixel to black.
4) Cropping and resizing the (Region of Interest)ROI to desired size. In our case, the size of the image is 128 x 128 pixels.
5) Converting the grayscale image into canny edge using adaptive-threshold filter. Adaptive-threshold method calculates threshold value for smaller regions and helps in finding more features in the image. As a result, different threshold values for different places when it comes to lighting change.
6) Cropping and resizing the ROI to desired size. In our case, the size of the image is 128 x 128 pixels.
7) Saving both images into Local disk for training and testing purposes.

V. ALGORITHM

A. CNN Model

We are using the Convolutional Neural Network(CNN) to train the model. CNN is a deep learning system that takes photos as input and assigns weights or biases to various components of the image, allowing it to distinguish between them.



B. CNN (Convolutional Neural Network):

CNN stands for "Convolutional Neural Network" and is largely utilized in image processing. The strength of the CNN is determined by the number of hidden layers used between the input and output layers. A set of features is extracted by each layer. A series of filters are applied to the input to create feature maps. After running through the entire input, each filter multiplies its weights by the values. The result is sent to a Rectified Linear Unit (ReLU), sigmoid, or the activation function. The set of weights is evaluated using a loss function. The feature maps created by the filters highlight different features of the input.

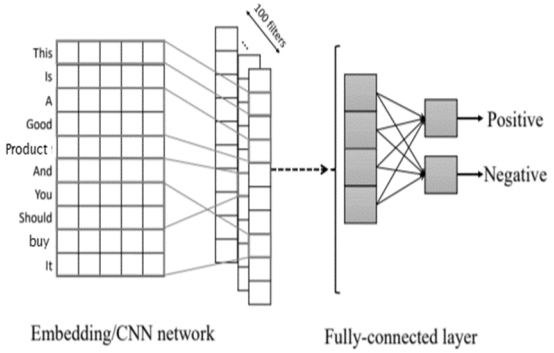


Fig 4: CNN Model

Even though CNN is commonly utilized in image and video processing, recent approaches to NLP utilize CNN. A pre-processing step in NLP converts the text input to a matrix representation. In the matrix structure, sentence characters serve as rows while alphabet letters serve as columns. In NLP, a filter is slid across the matrix's words. As a result, the words are detected using the sliding window technique.

1) CNNs are automatic feature extractors.
2) We can use max pooling to reduce size of image without compromising with image quality.
3) CNN provides usage of convolutions, which act as feature emphasizers.

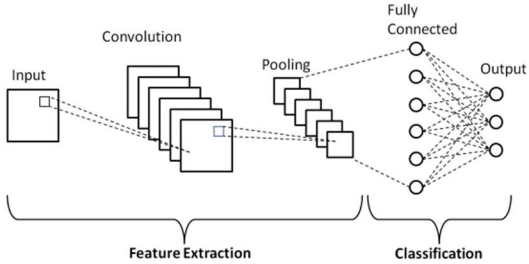


Fig 5: Basic Architecture of CNN

C. Training a CNN Model

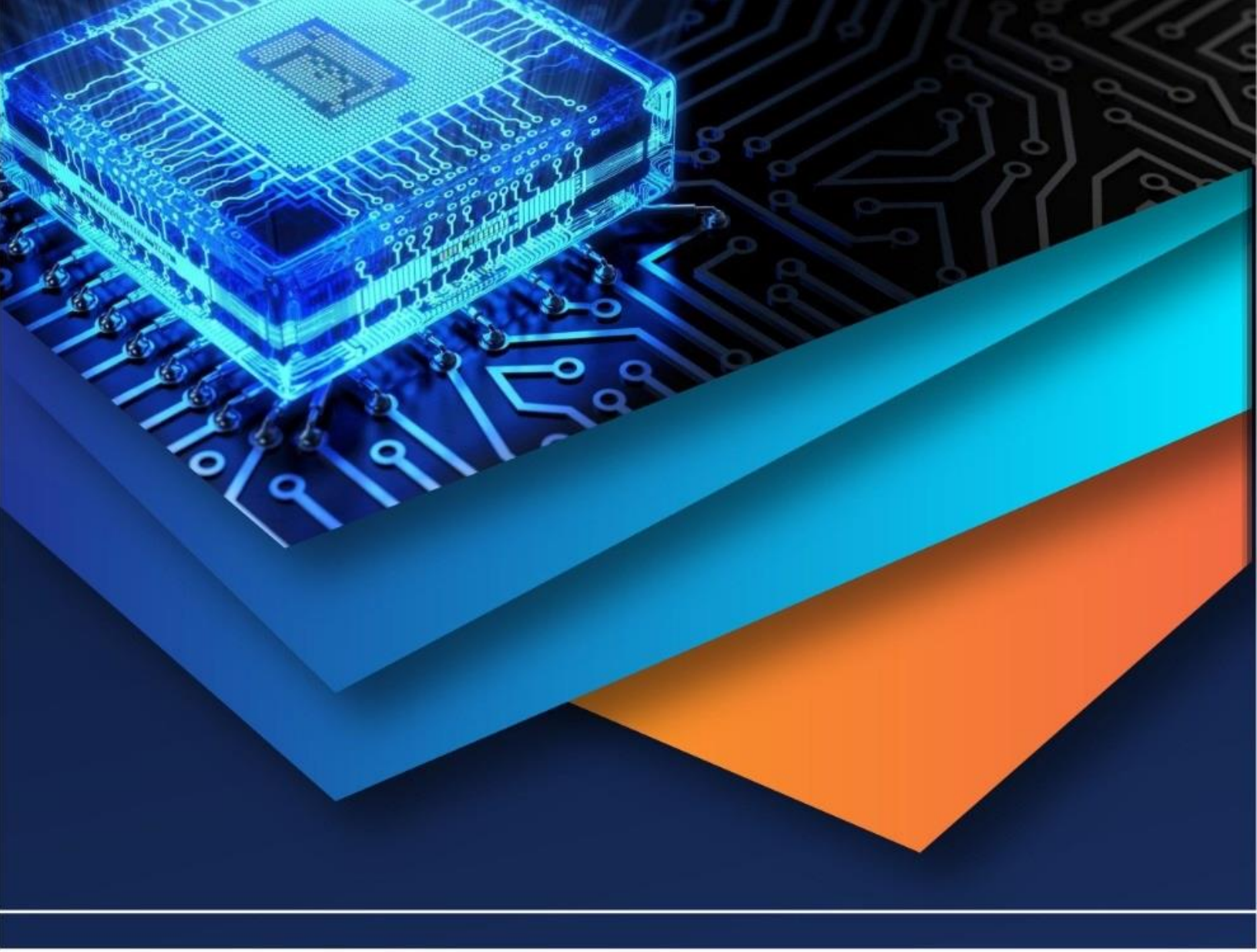
In our model, we have added one convolution layer with 32 liters of size (3,3) which will help in highlighting different features like gesture boundaries in Binary images and minute finger features like emerging thumb between finger and middle finger. The convolution layer produces a feature map that highlights the gesture features. We have added a max pool layer, which helps in reducing the size of input data to learning layers of model while ensuring that the details in image are not lost. This is done in order to reduce the computing time. Lastly, we have 3 fully connected layers with 128, 96 and 64 neurons each and an output layer. This layer learns from the input data and produces weights which are then used to classify signs.

VI. CONCLUSION

In this project, We developed a system that can correctly recognise ASL Alphabets and Numbers, which are mostly based on hand and finger movements. The model used in this project uses CNN for detection of 26 English ASL alphabets using different image enhancement techniques. This model also uses a Con-evolution layer to highlight the dominant features and max pooling layers to reduce computing cost. The current version of the model has one convolution, one max pooling and three fully connected layers. These layers may change while increasing the accuracy of the model. The current version of the model has accuracy of 98%.

REFERENCES

[1] Bangla Sign Digits : A Data set for Real Time Hand Gesture Are cognition 2020 11th International Conference on Electrical and Computer Engineering (ICECE) — 978-1-6654-2254-3/20[1]
[2] Study of Convolutional Neural Network in Recognizing Static American Sign Language This paper is motivated towards applying Convolutional Neural Network Proc. of the 2019 IEEE International Conference on Signal and Image Processing Applications (IEEE ICSIPA 2019), Malaysia, September 17-19, 2019 [2]
[3] A Static Hand Gesture Based Sign Language Recognition System using Convolutional Neural Network[3]
[4] Digit Recognition in Sign Language Based on Convolutional Neural Network and Support Vector Machine[4]
[5] Sign Language Detection from Hand Gesture Images using Deep Multi-layered Convolution Neural Network IEEE International Conference on Multimedia and Expo (ICME) (pp. 1-6). Turin: IEEE.[5]
[6] Hand Gesture Recognition for Bangla Sign Language Using Deep Convolution Neural Network 2017 IEEE Conference on (pp. 4724-4733). IEEE. Honolulu. [6]



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)