IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

HAND SIGN RECOGNITION FOR SPECIALLY-ABLED USING CNN

[1]Akhilesh, [2]Ishant Bhatt, [3]Devesh Bisht, [4]Aditya Rawat, [5]Medhavi Pandey

[1]Student, [2]Student, [3]Student, [4]Student, [5]Guide
Department of Computer Science Engineering,
Delhi Technical Campus, Greater Noida, Uttar Pradesh, India

Abstract : One of the most important problems that our societies are facing is people who are specially-abled that are unable to communicate with normal people. More than 5% of the population around the world are deaf and dumb. The only means of communication for a person with special needs who is unable to talk or hear anything is sign language. People who are physically disabled can express their thoughts and emotions through signs using sign language. Some people can understand their thoughts and emotions but the other remaining are unable to understand their thoughts and emotions. So, we need a translator to understand what they speak and communicate with us. The main aim of this proposed work is to develop a cost-effective model to predict hand gestures using American sign language for communication between normal people and specially-abled people. In order to identify the alphabet and motions of an American sign language dataset, this research effort presents a revolutionary hand sign identification system. Using computer vision and neural networks, we can recognize the signs and output the proper alphabet.

***IndexTerms* - Specially-abled people, Sign language, CNN (Convolutional Neural Network), Open CV, AI- Artificial Intelligence, ANN (Artificial Neural Network)**

INTRODUCTION

The exchange of ideas and messages through speech, signaling, behavior, and pictures are all examples of communication. People who are deaf or dumb communicate with others by using distinct alphabets or words with their hands. Nonverbal signals are communicated through gestures, which are interpreted with the use of eyesight. Sign language is the name for this nonverbal communication used by the deaf and the hearing impaired. The hands and other body parts are employed as a way of communication in sign language recognition using computer vision and deep learning. According to the World Health Organization, 466 million people use sign language to communicate often because they have severe hearing loss.
Normal individuals have trouble communicating with the speech-impaired group because they use hand gestures and signs. Still, there will be problems when these people try to communicate with ordinary people who are unfamiliar with ASL. As a result, an automatic and real-time sign language interpreter is not only necessary but also in high demand for people with and without hearing loss alike.

With standard algorithms, it is challenging to recognise ASL expressions. Convolutional neural networks, on the other hand, enable us to recognise and translate sign language more effectively. Normal individuals have trouble communicating with the speech-impaired group because they use hand gestures and signs for communication which is not easily be understand by normal people.

METHODOLOGY

Artificial Intelligence is systems or machines that can imitate a human brain to perform the given tasks. The Artificial intelligence is bringing down the gap between machines and human beings. One of the fields is Computer vision. The goal of computer vision is to understand and view the world through machines as human beings do also the goal of the deep learning is to made a neural network that work exactly like a human brain

CNN is a subset of Machine Learning. CNN is a multi-layered neural network with an inbuilt architecture to recognise the images using deep neural network. A CNN Architecture consist of three main layers: Convolutional Layer, Pooling Layer and Fully Connected Layer. Convolutional Layer is the core of the CNN architecture. Majority if the Computation tasks happen in the convolutional layer. Pooling Layer helps in improving the efficiency and reduces complexity of a CNN. Fully Connected Layer is the layer in which Image classification happens on the features extracted in the previous layer.

A. Dataset Generation

For Data set generation we used the American Sign Languages. The gestures we aim to train are as given below:

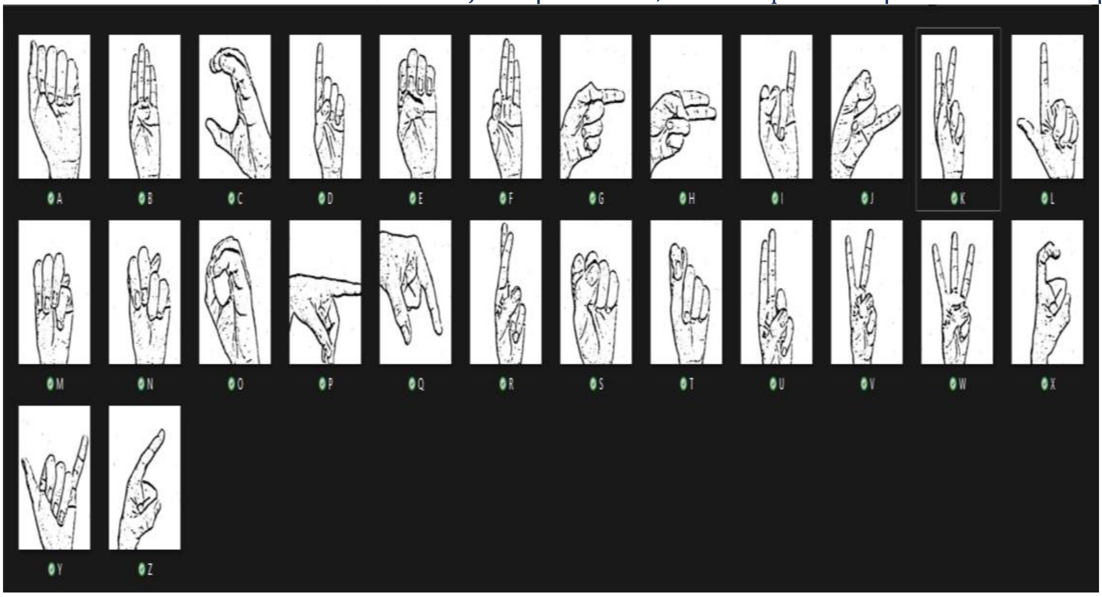

Fig 1: Hand Signs

We looked for pre-made datasets for the project but were unable in finding any that were in the form of raw photos and met our specifications. The datasets in the form of RGB values were the only ones we could locate. We made the decision to compile our own data set. Our dataset was created using the Open Computer Vision (OpenCV) framework. Using the computer's webcam, we take a picture of every frame and stored it in the training and testing folder.
From the whole image we extracted the hand signs that we denoted as a region of Interest (ROI). From the ROI extracted we converted each of the images to grayscale images as the CNN model can take much time for the RGB images.
After converting the coloured images to grayscale, we applied gaussian blur to images to extract the features from the such as border of the hand sign, lines of the palm or hand.
The data set images after applying gaussian blur:

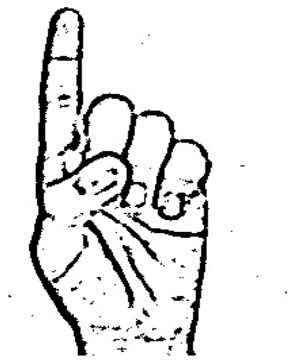

Fig 2: Applying Gaussian Blur

B. Computing Platform Specification

The experiment is performed using an Acer laptop Acer aspire 7 with processor Generation Intel(R) Core (TM) i5-9300H CPU @ 2.40GHz 2.40 GHz, memory type DDR4 8GB, Storage capacity 512 GB SSD, Windows 11 Operating System.
The Software application used for coding the application is implemented using PyCharm.

C. Feeding Images/ Data set to CNN

1. 1st Convolution Layer :The input picture has resolution of 128x128 pixels. It is first processed in the first convolutional layer using 32 filter weights (3x3 pixels each). This will result in a 126X126 pixel image, one for each Filter-weights.
2. 1st Pooling Layer : The pictures are downsampled using max pooling of 2x2 i.e we keep the highest value in the 2x2 square of array. Therefore, our picture is downsampled to 63x63 pixels.
3. 2nd Convolution Layer : Now, these 63 x 63 from the output of the first pooling layer is served as an input to the second convolutional layer.It is processed in the second convolutional layer using 32 filter weights (3x3 pixels each).This will result in a 60 x 60 pixel image.
4. 2nd Pooling Layer : The resulting images are downsampled again using max pool of 2x2 and is reduced to 30 x 30 resolution of images.
5. 1st Densely Connected Layer : Now these images are used as an input to a fully connected layer with 128 neurons and the output from the second convolutional layer is reshaped to an array of 30x30x32 =28800 values. The input to this layer is an array of 28800 values. The output of these layer is fed to the 2nd Densely Connected Layer.We are using a dropout layer of value 0.5 to avoid overfitting.
6. 2nd Densely Connected Layer : Now the output from the 1st Densely Connected Layer are used as an input to a fully connected layer with 96 neurons.

7. Final layer: The output of the 2nd Densely Connected Layer serves as an input for the final layer which will have the number of neurons as the number of classes we are classifying (alphabets + blank symbol).

D. In each of the hidden layers, we employed the SoftMax Function and the ReLu (Rectified Linear Unit) activation function. The option for the epoch is set to 10 epochs. After processing the CNN, the experiment's validation accuracy was 97.08%.

E. Model and Weights saving

After completion of the training process, we saved the weights of the model in a .h5 file that contains trained model weights and model architecture. We can save our CNN labels in .Json file.

F. Plot the accuracy graph

After the model is trained, we can use the matplotlib library of python to plot the accuracy and loss graph to see the performance of our trained model.

LITERATURE STUDY

For the need of the Literature Study there was the need for the previous related work towards Hand sign recognition. The following were the research papers used for study the Hand sign Recognition project:

1) In the study by Aggarwal and Prasad, the authors explore the use of deep learning techniques, specifically convolutional neural networks (CNNs), for hand gesture recognition. The authors used a dataset of dynamic hand gestures and compared the performance of the CNN-based approach to traditional hand gesture recognition methods such as template matching and hidden Markov models (HMMs). The results showed that the CNN-based approach outperformed the traditional methods in terms of accuracy and real-time performance. The authors also discussed the potential applications of the CNN-based approach, including human-computer interaction and sign language translation.

2) In the second Research paper by Saha and Pauri. The use of machine learning techniques for hand gesture recognition, specifically using support vector machines (SVMs). The authors used a dataset of dynamic hand gestures and trained an SVM model to recognize the gestures. The results showed that the SVM-based approach achieved high accuracy in recognizing hand gestures and was able to perform in real-time. The authors also discussed the potential applications of the SVM-based approach, including human-computer interaction and sign language translation.

3) In the third research paper the authors provide an overview of hand gesture recognition techniques, including traditional approaches such as template matching and machine learning techniques such as neural networks. The authors discuss the advantages and limitations of each method and suggest future research directions. The authors also discuss the potential applications of hand gesture recognition, including human-computer interaction, virtual reality, and sign language translation.

IV. RESULTS AND DISCUSSION

The CNN model was able to recognise the hand pattern easily with an accuracy of 95.08%.

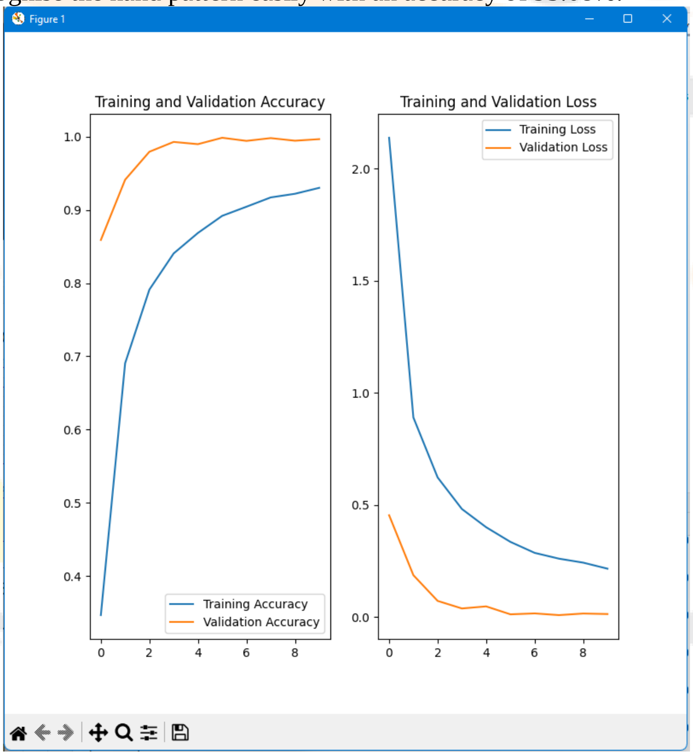

Fig 3: Training Accuracy and Loss results

CONCLUSION

Machine Learning, Artificial intelligence and computer vision-based hand sign and hand gestures recognition has many proven advantages over traditional devices. However, hand sign recognition for specially-abled community and its conversion to text or speech is difficult problem. This project is a small contribution towards achieving the results needed in the field of Artificial intelligence, Machine Learning, and deep learning.

The current process uses CNN and RNN. CNN for training the model and extracting the unique features from the data set of ASL given to the CNN model and RNN is used to save the model. TensorFlow and Keras are prominent level and lightweight machine learning APIs used in the experiment that have the predefine modules for training the models.

For future work one can focus on combining the hand sign recognition to online-meeting based applications so that the specially-abled community is easily able to communicate and can easily share their emotions to the world.

REFERENCES

[1] Pratibha Pandey, Vinay Jain, Hand Gesture Recognition for Sign Language Recognition, A Review, International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 3, March 2015.

[2] Mandeep Kaur Ahuja, Amardeep Singh, Hand Gesture Recognition Using PCA, International Journal of Computer Science Engineering and Technology (IJCSET), Volume 5, Issue 7, pp. 267-27, July 2015.

[3] Huang, J., Zhou, W., & Li, H, Sign Language Recognition using 3D convolutional neural networks, IEEE International Conference on Multimedia and Expo, 2015.

[4] Mehreen Hurroo , Mohammad Elham , Sign language Recognition System using Convolutional Neural Network and Computer Vision , Volume 09, Issue 12 (December 2020), IJERT.

[5] G Ananth Rao and PVV Kishore. Selfie video based continuous Indian sign language recognition system. Ain Shams Engineering Journal,2017.

[6] V. Vardhan1 and P. Prasad, “Hand Gesture Recognition Application for Physically Disabled People,” International Journal of Science and Research, vol. 3, no. 8, pp. 765–796, Aug. 2014.

[7] G. A. Rao, K. Syamala, P. V. V. Kishore, and A. S. C. S. Sastry, "Deep convolutional neural networks for sign language recognition," 2018 Conference on Signal Processing and Communication Engineering Systems (SPACES), Vijayawada, 2018, pp. 194-197. doi: 10.1109/SPACES.2018.8316344

[8] H. Chung, Y. Chung, and W. Tsai, "An Efficient Hand Gesture Recognition System Based on Deep CNN," 2019 IEEE International Conference on Industrial Technology (ICIT), Melbourne, Australia, 2019, pp. 853-858. doi: 10.1109/ICIT.2019.8755038