International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

ENHANCING THE PERFORMANCE OF SIGN LANGUAGE RECOGNITION USING CNN AND RNN CLASSIFIER

Sai Sampath Gunupuru1 , Akshidha2 , Pragna3 , Uma Sankara rao V 4, Aditya 5

1,2,3,4,5Department of Computer Science and Engineering, GST, GITAM University, Visakhapatnam, AP, India.

Abstract:

Sign language gave hope to deaf and dumb people to communicate with others. Although there is a significant development in sign language, Sign recognition systems are created to increase efficiency in sign recognition. Our sign language recognition technology was created to aid in communication for individuals suffering from hearing and visibility issues. We have created it using CNN (Convolutional neural network) and RNN (Recurrent neural network). After pre-processing the images, the images are used in the dataset containing the 25 English alphabets. We built CNN and trained it with the dataset. After training the model we have predicted the accuracy. To increase the accuracy we have enhanced the CNN model with RNN. The best accuracy obtained is 88.8% for CNN with RNN.

Keywords- Convolutional neural network, Recurrent neural network

Introduction:

In general, machine learning is utilized when working with picture recognition because numerous cutting-edge image processing methods, including deep neural networks, rely on machine learning models.

We face several challenges when we do so like choosing a suitable dataset many projects tend to choose a predefined dataset like MINST [5], and the size of the data also affects the accuracy of the model to get a good accuracy we need to include a good enough dataset [2]. The classifier we use also affects our accuracy so we need to choose an efficient

classifier [1][2][6][7].

This method seeks to consider these issues and recognize Sign Language Symbols as an alternative form of communication to spoken communication. They communicate with each other by signing. Yet it takes a lot of work to learn and interpret sign language, and not everyone can grasp what the gestures in sign language mean. Unfortunately, there is no trustworthy, portable technology available to learn sign language, thus it takes time. A translator who is also competent in sign language is required for hearing-impaired or speech-impaired people who use sign language to properly communicate with others. In order to help people with hearing loss or speech impairments get over these obstacles, this technique helps them learn and translate their sign language.

Contributions:

To overcome the above-mentioned problems, we propose to use our dataset obtained from the real-time feed of size (25 x 50) and also use CNN classifier and RNN classifier to train our model some other challenges faced to include:

1. Lighting conditions: As it has a significant impact on the accuracy of the system, the lighting conditions in the area where it is utilized must be taken into consideration. We acquired all of our data in a stable environment

2. Distance from the lens: Make sure the hand gesture is executed at the proper distance from the lens we have employed an ROI selection dimensions

3. Background: The camera must at least capture the wrist and cannot focus on objects in the background. the use of ROI abates the problem of background images by creating a region of interest

4. Recognizing the characteristics of the symbols can be a system to improve accuracy to focus on the feature better we convert our RGB images to greyscale.

In conclusion, we propose a system that collects data in a stable environment by creating an ROI for the images and effectively pre-processing them, using the CNN classifier to train them and further enhancing the results with the RNN classifier.

Literature review:

Mehreen hurroo et al[1]. presented a CNN and computer vision system for sign language recognition. With this system, the user must utilize a web camera to capture the hand motion photographs, and the system uses the HSV colour algorithm to forecast and show the name of the captured image. Only 10 American sign gesture alphabets can, however, be accurately recognised by the sign language recognition system employing the HSV colour algorithm and CNN.

Sunitha Nandhini et al[2]. proposed a technique using Convolutional neural networks that converts the gestures from sign language to a spoken language which is understood by listening.

Lionel Pigou et al[3]. presented a system using Microsoft Kinect, and convolutional networks, and GPU acceleration was used to suggest a recognition system. It accurately detects 20 Italian hand motions.

Rachana patil et al[4]. presented a system in which the user must be able to use a web camera to record images of hand motions. The model identifies the images after being taught using CNN.

Heramba Limaye et al[5]. presented a system based on CNN and computer vision to offer a graphic user interface and a solution for each user.

Dr.Thamaraiselvi et al[6]. proposed a sign recognition system using CNN classifiers where the signs are conveyed for relevant information. However, written communication is difficult while moving.

Jie Huang et al[7]. proposed a 3D CNN which extracts various spatial features from raw videos and validates the proposed model on a real dataset collected with Microsoft Kinect.

Vanita jain et al[8]. suggested a Support vector machine (SVM) and CNN-based system for recognizing signs. Compared to SVM, the double-layer CNN has greater accuracy.

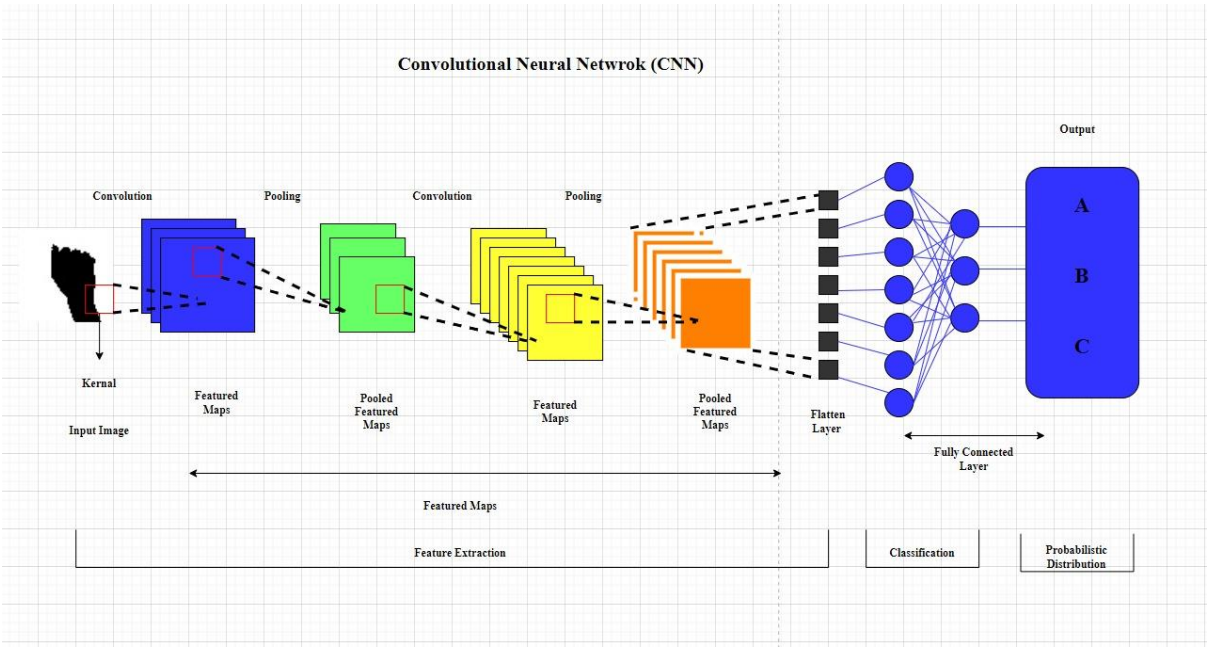
Abdul Mannan et al[9]. presented a deep CNN-based method for recognizing signs. Through the use of data augmentation techniques, deep CNN's performance is enhanced.

Ayush Kumar et al[10]. CNN classifier is used in a proposed sign language recognition system, which passes a hand image via a filter before the CNN classifier determines the type of hand gesture it represents.

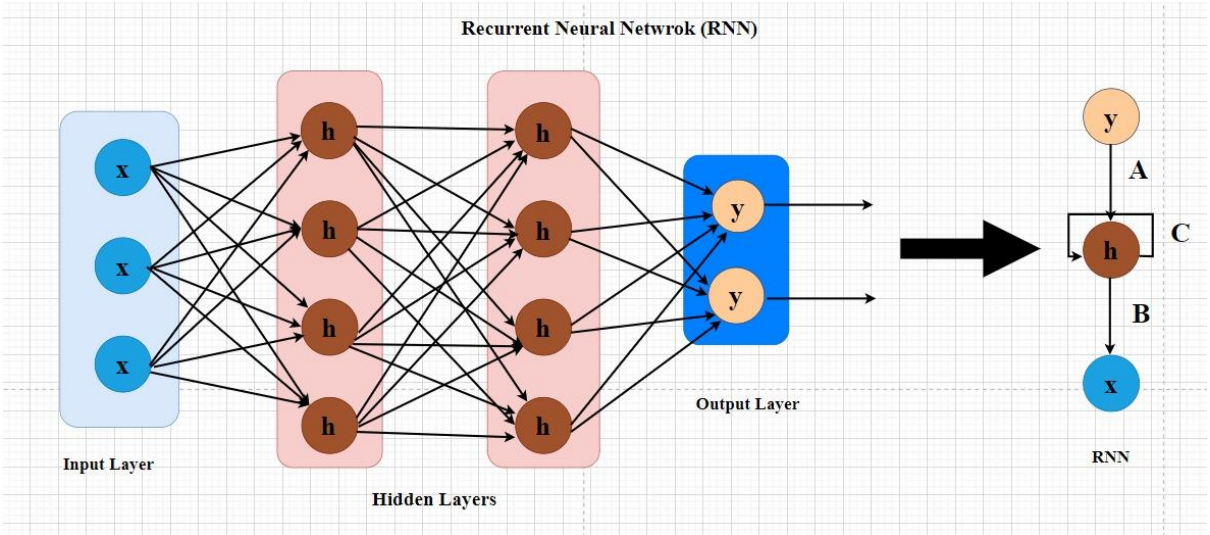
Proposed method:

We have built the CNN model. The construction of CNN includes five steps. Initially, the data is converted into a sequential model. A fully connected neural network is created by adding a convolutional layer, a max pooling layer, and a flattening layer. A group of layers makes up the sequential layer. It delivers the output to the convolutional layer after receiving the image as input. The key elements of the feature map are extracted by the Max pooling layer. The feature map's two-dimensional arrays are flattened into a single, long continuous linear vector. The images are classified by dense layers and produce the output. The CNN model is further enhanced by adding RNN. The RNN receives the output picture from CNN as input. RNN with long short-term memory(LSTM) is used as it can handle sequential data. The CNN with RNN model is trained with the dataset. And the accuracy is predicted. The accuracy obtained for CNN with RNN is high.

CNN :



RNN :

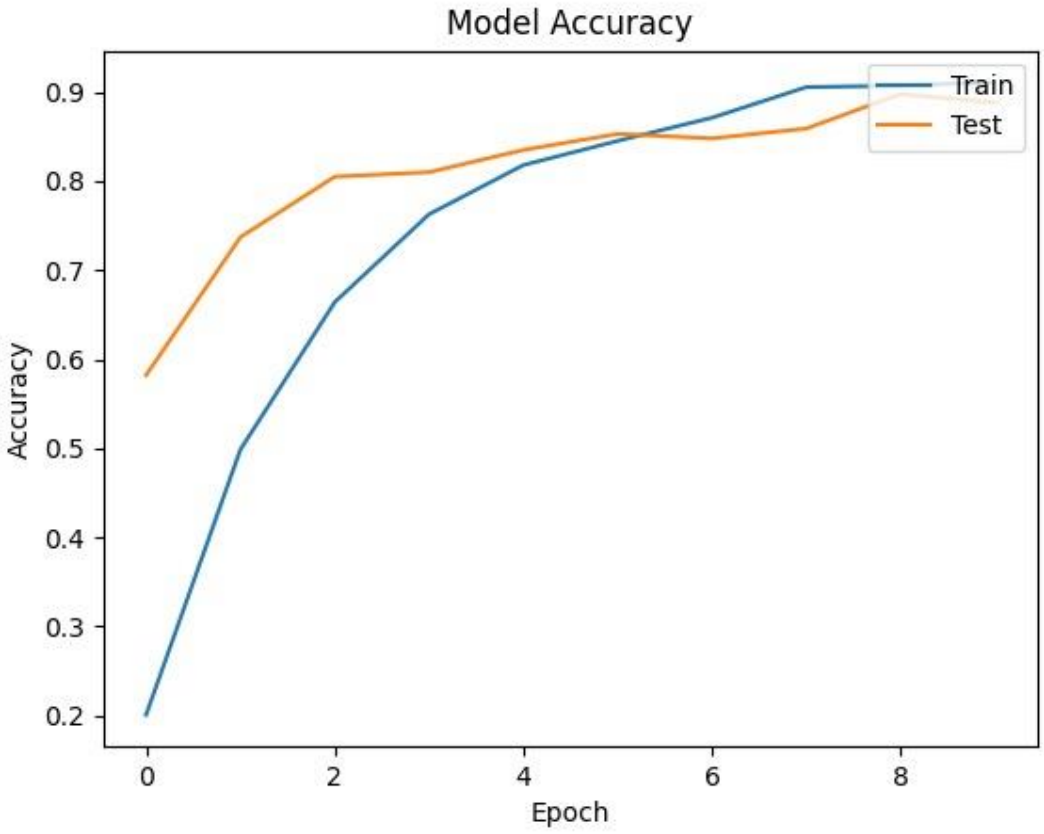


Results:

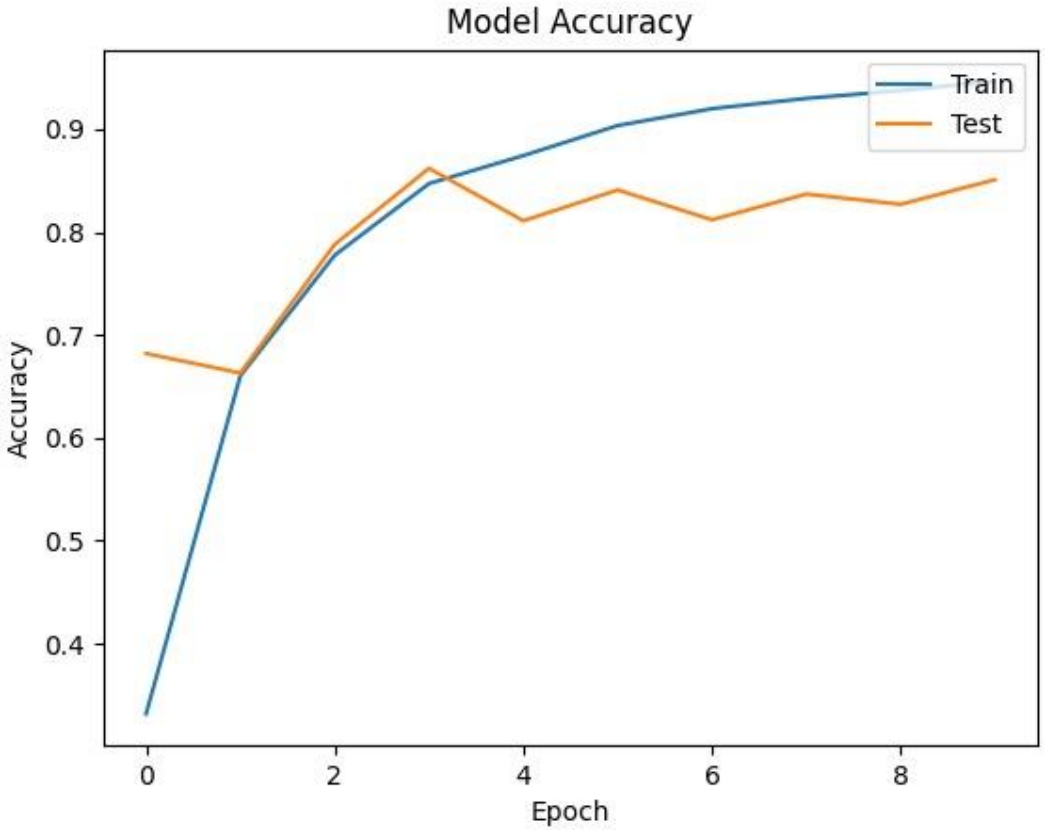
This study uses accuracy to evaluate the performance of our dataset of hand movements associated with American sign language. We have developed CNN and RNN classifiers for recognizing sign language. The sign language dataset contains 25 classes. There are 140 images for each alphabet, with 100 images used for testing and 40 images for training. The classification results of the proposed Cnn and Rnn classifiers are shown in figure-1. The accuracy of Cnn at 10 epochs is 85.1%. The accuracy of CNN with RNN at 10 epochs is 88.8%. The accuracy is high for Rnn with Cnn.

Figure-1:

Rnn with Cnn

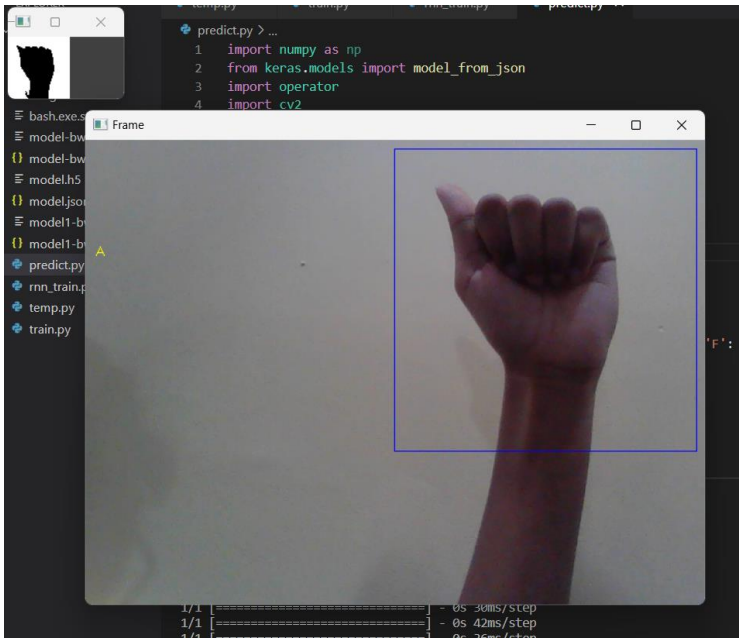


Cnn:

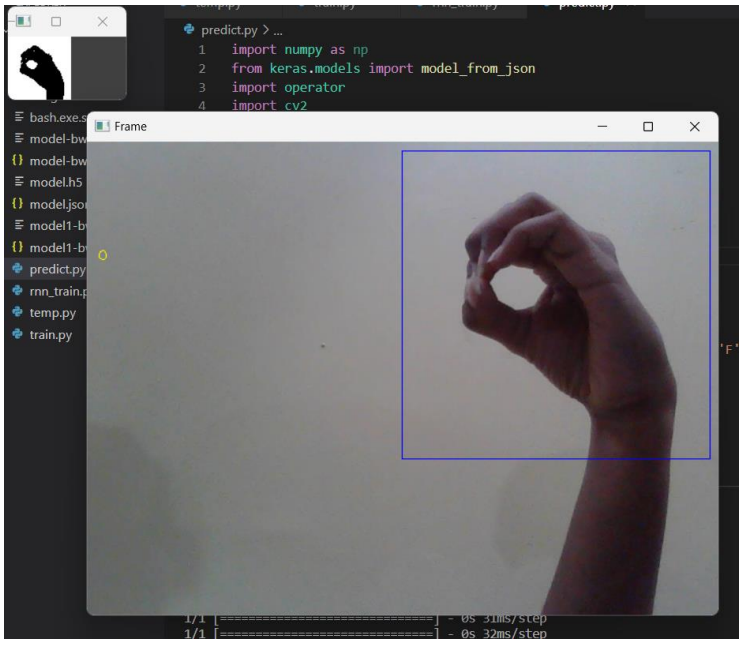

Model Accuracy
Train
Test
Accuracy
0.9
0.8
0.7
0.6
0.5
0.4
0
2
4
6
8
Epoch

Output:

For alphabet A



For alphabet O



Conclusion:

The identification of hand gestures in everyday situations has been a challenge for many researchers. Due to the various efficiency, robustness, and accuracy requirements, it is difficult. In this study, we suggested a reliable ASL recognition method that uses 25 sign language alphabets. The suggested method uses CNN and RNN to recognize the hand gestures of sign language. The proposed CNN with RNN can recognize sign language with an accuracy of 88.8%.

References:

[1] Dhavale, Prathamesh. “Sign Language Recognition Using Machine Learning Approach.” International Journal for Research in Applied Science and Engineering Technology 7.4 (2019): 853–861.

[2] Raval, Jinalee Jayeshkumar and Ruchi Gajjar. “Real-time Sign Language Recognition using Computer Vision.” 2021 3rd International Conference on Signal Processing and Communication (ICPSC) (2021): 542-546.

[3]https://www.sciencedirect.com/science/article/pii/S2405959519 303455 batch size

[4] Hasan, M., et al. "Comparative analysis of SVM, ANN and CNN for classifying vegetation species using hyperspectral thermal infrared data." The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 42 (2019): 1861-1868.

[5] M. Bilgin and K. Mutludoğan, "American Sign Language Character Recognition with Capsule Networks," 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2019, pp. 1-6, doi: 10.1109/ISMSIT.2019.8932829.

[6] Kumar, S.sravan and Vedant L. Iyangar. “SIGN LANGUAGE RECOGNITION USING MACHINE LEARNING.” (2018).

[7] J. Ekbote and M. Joshi, "Indian sign language recognition using ANN and SVM classifiers," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017, pp. 1-5, doi: 10.1109/ICIIECS.2017.8276111.

[8] Deep Residual Learning for Image Recognition" by Kaiming He et al. (2016):

[9]https://www.ttested.com/ditch-mnist/

[10]https://analyticsindiamag.com/is-more-data-always-better-for-building-analytics-models/#:~:text=We%20need%20big%20data%20mostly,datasets%20to%20get%20high%20accuracybig