Sign Language Detection Using Deep Learning

G. Sreenivasulu and Srilatha Kadari

February 24, 2023

Sign Language Detection Using Deep learning

Mr. G. Sreenivasulu

Associate Professor

Department of computer science& engineering

J. B. Institute of Engineering & Technology

g.sreenivasulu@jbiet.edu.in

Srilatha Kadari

Department of computer science& engineering

J. B. Institute of Engineering & Technology

kadarisrilatha31@gmail.com

Abstract

The presented paper deals with automatic Deep learning model based Sign Language Detection. Detecting the Sign Language Detection we design a real-time human computer interaction system based on hand gesture using an auto camera, whereby, based on the captured image, the neural network recognizes whether the driver is awake or tired. The convolutional neural network (CNN) technology has been used as a component of a neural network, where each frame is evaluated separately and the average of the last 20 frames is evaluated, which corresponds to approximately one second in the training and test dataset. First, we analyze methods of image segmentation, and develop a model based on convolutional neural networks. Using an annotated dataset of more than 2000 image slices we train and test the segmentation network to extract the driver emotional status from the images.

Keywords: - HCI, CNN, Hand Guessers, Sign Language Detection.

1. INTRODUCTION

In daily life, the hand gesture is a natural form of communication that is typically exclusively employed by those who have some difficulty hearing or speaking. Yet, there are numerous application situations for a gesture-based human-computer interaction system. For instance, control robots in situations where using speech or standard input devices is difficult, such as underwater environments, mouse- and keyboard-free games, or simply provide translation for persons who speak different gesture languages. Gesture recognition is the basis for using such a system. Since many years ago, gesture recognition has gained popularity. To accomplish gesture recognition today, essentially two methods are used. One is based on wearable electromagnetic devices for work, including specialised gloves. The alternative makes use of computer vision.

Speech impaired people use hand signs and gestures to communicate. Normal people face difficulty in understanding their language. Hence there is a need of a system which recognizes the different signs, gestures and conveys the information to the normal people. It bridges the gap between physically challenged people and normal people. The model proposed in this paper performs convolutions to learn about the spatial features and the temporal features. The deep architecture created pulls information of different kind from the adjacent layers and then separately runs the convolution and subsampling. Next the for the final feature representation the information from every channel is gathered and combined. Finally by using the multilayer perceptron classifier the feature representations are classified. The effectiveness of the method is shown by the results obtained when CNN are trained on the same dataset.

2. RELATED WORK

Image detection will perform when it needs to differentiate between healthy and unhealthy leaves. Here the CNN will identify the variance between plant images to determine the abnormality which is resided in the plants *Barbedo et al.* (2018). Due to curtailment in the plant, it will increase the risk of global warming in the world, so it requires to development of the modern convolutional system which helps to image detection and detection of plant disease. This approach *Ramcharan et al.* (2019) provided scientists several research applications with the required information.

The Machine Learning classifiers applied detection of disease with K-means, ANN, SVM, and DT, among others *Wallelign et al.* (2018). Here the machines cannot learn from images directly, so it should be converting into pixel format which is understood by the computers. Therefore these pixels will be feed to classifiers to determine the Sign Language. As well as, the deep learning classifier *Dhakal et al.* (2018) was also *used* to determine abnormalities in Sign with pixel-wise operations. These operations are used to gathering the information from leaves which is acquired from Sign and depends on disease affected area pattern it will decide which Language affected that Sign.

3. IMPLEMENTATION

Methodology and algorithm:

Convolutional Neural Networks:

The CNN is a special type of feed-forward artificial neural network in which the connectivity pattern between its neurons is inspired by the visual cortex. The advantage of CNN is the neuron in a layer will only be connected to a small region of the layer before it, instead of all of the neurons in a fully connected manner. CNN is a kind of deep neural network which were designed from biological-driven models. Regarding the CNN model, the researchers found that how humans perceived images in the brain is in different layers than how the CNN model was designed. Hence it has been proven that it was very efficient for all image processing pattern recognition kinds of applications. The diagram of basic convolutional neural network (CNN) architecture was shown in figure.2.

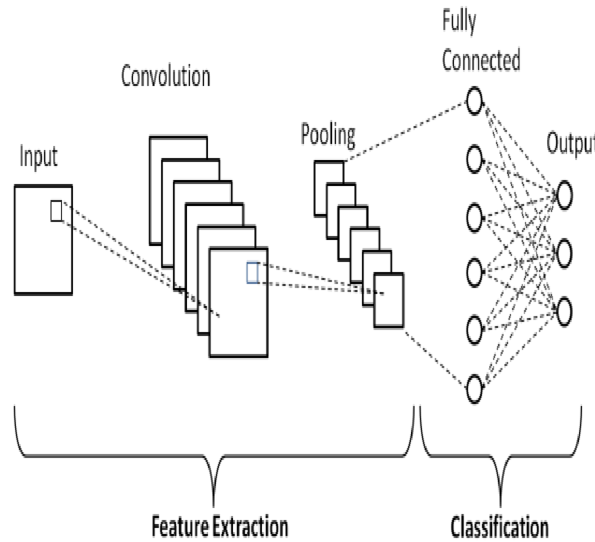

Figure.1 Basic CNN architecture

From figure.1, the input let us consider as image pixels like features are forwarded to Convolution layer then it will gather the value able features from image then after these features are forwarded to Pooling layer. The pooling layer can reduce the image size with the required features. Finally, a fully connected layer will be performing the image classification and it returns the classification result with a high percentage.

The CNN model will be categorized into four layers.
They are:

Convolution Layer:

The Convolution layer is the first layer of CNN and will be applied to the filters to extract the image features and perform the convolution operation (dot product), this layer will be used by several filters. Here it can treat the images as pixel values with matrix form.

ReLU Layer:

After completion of feature maps extraction, then it will be forwarded to rectified linear unit (ReLU) layer. This layer will be replacing all negative pixels with 0. By performing the element-wise operation it can return the rectified feature map.

Pooling Layer:

This layer will reduce the feature map dimensionality and the output of the ReLU layer feature map can be passed to the pooling layer to get pooled feature map.

Fully connected layer:

This layer will convert all two-dimensional arrays into one dimensional long continuous linear vector. This process is called flattening. This layer can be providing the image classification with the highest percentage of similarities.

Architecture diagram:

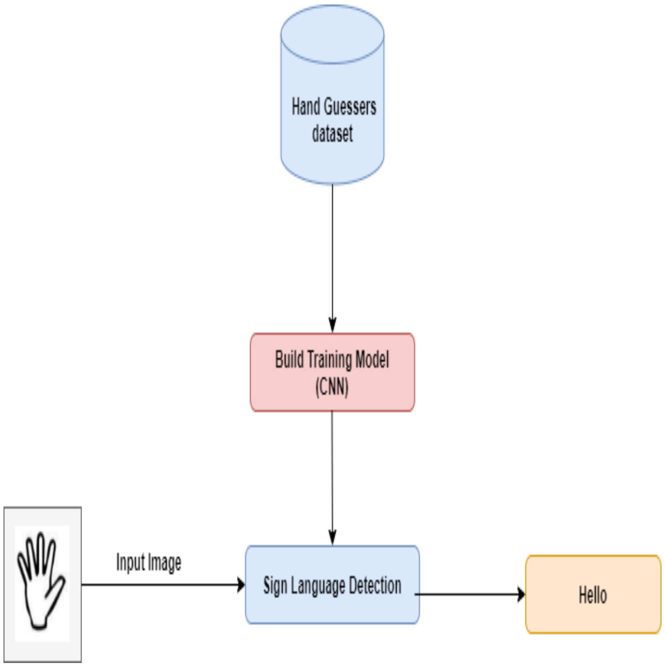

Figure.2 Architecture diagram

Create Hand Guessers:

In this system, we are using a customized dataset, where the dataset is created manually. Here the system webcam will be open and it will capture the given sign and store it in the local folder. For demo purposes, we have created a few dataset folders only based on various signs such as Hello, Yes, Thank you, etc. and we have created 50 images for each sign.

Training Deep learning model:

In this module, after creating the hand guessers dataset, the CNN architecture will be prepared. Here the hand guessers dataset will be read based on the image processing technique and extract the features. Later, these features are passed to the CNN model, then this model will train based on image features and generate the Sign language detection model file which is High dimensional file format(.h5)

Sign Language Detection:

In this module, this system will detect the various signs. Here the user will provide the signs through a webcam then this system will load the trained CNN model and give the sign test image then the model will return or predict the Sign name like Hello, thank you, etc. which is matched with dataset features.

4. EXPERIMENTAL RESULT

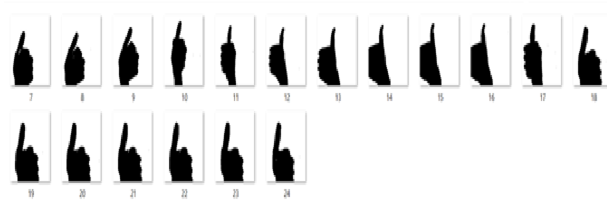

Figure.3 Module -1

Figure.4 Module-2

Figure.5 Module-3

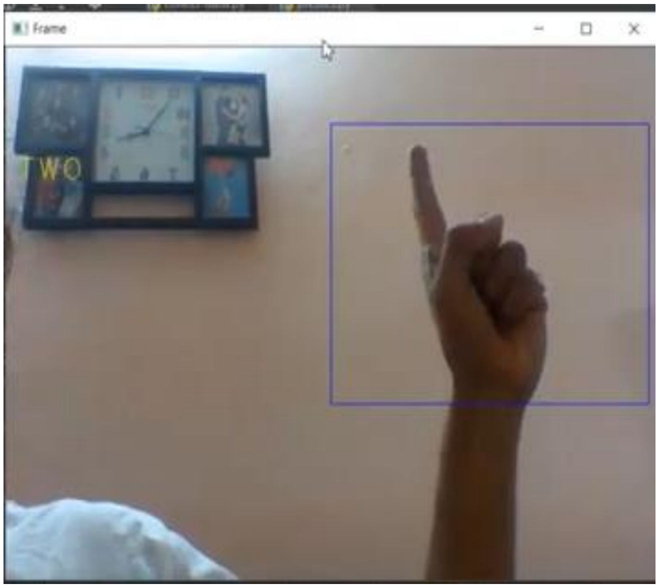

Figure.6 Result-1

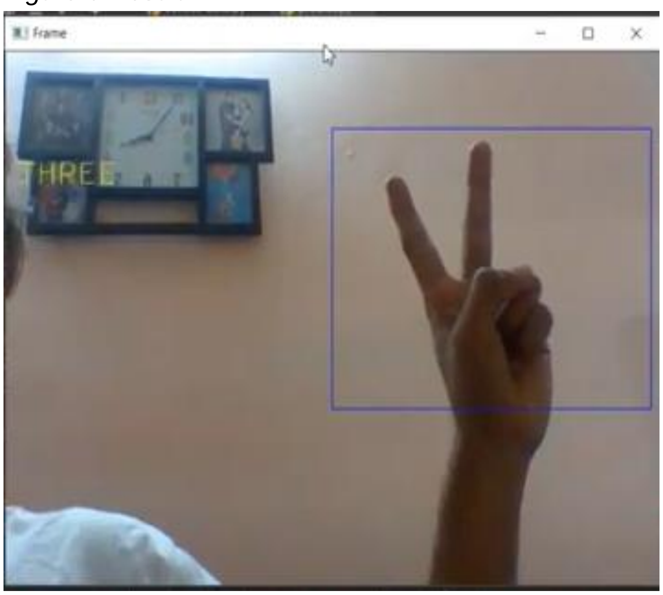

Figure.7 Result-2

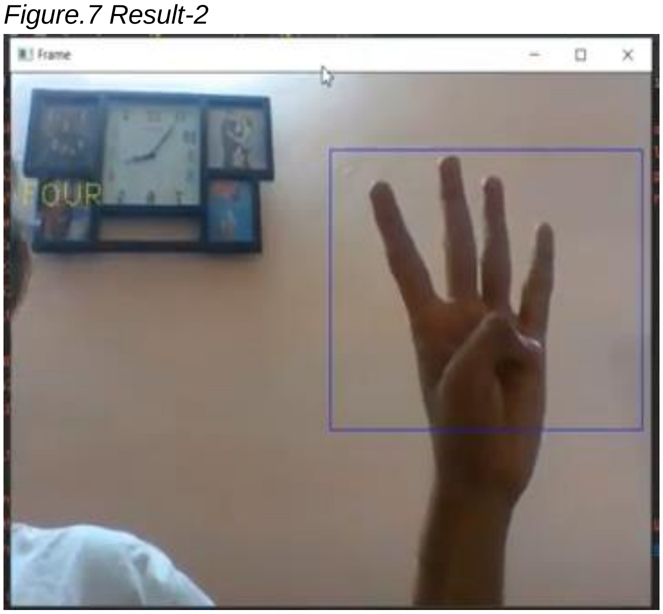

Figure.8 Result-3

5. CONCLUSIONS

This paper had successfully presented the most prominent techniques, applications, and challenges in hand gesture recognition. These include the gesture acquisition methods, the feature extraction process, the classification of hand gestures, the applications that were recently proposed in the field of sign language, robotics and other fields, the environmental surroundings challenges and the datasets challenges that face researchers in the hand gesture recognition process, and the future of hand gesture recognition.

6. REFERENCES

[1] W. T. Freeman and M. Roth, Orientation histograms for hand gesture recognition. International workshop on automatic face and gesture recognition. 1995, 12: 296-301.

[2] T. Starner and A. Pentland, Real-time american sign language recognition from video using hidden markov models. Motion-Based Recognition. Springer Netherlands, 1997: 227-243.

[3] L. Bretzner, I. Laptev and T. Lindeberg, Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering. Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on. IEEE, 2002: 423-428.

[4] N. H. Dardas and N. D. Georganas, Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques. IEEE Transactions on Instrumentation and Measurement, 2011, 60(11): 3592-3607.

[5] Y. Wu and T. S. HuangVision-based gesture recognition: A review. International Gesture Workshop. Springer Berlin Heidelberg, 1999: 103- 115.

[6] S. Mitra and T. Acharya, Gesture recognition: A survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2007, 37(3): 311-324.

[7] P. Y. Simard, D. Steinkraus and J. C. Platt, Best Practices for Convolutional Neural Networks Applied to Visual Document Analysi. ICDAR. 2003, 3: 958-962. [8] A. Krizhevsky, I. Sutskever and G. E. Hinton. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems. 2012: 1097-1105.

[9] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[10] J. Sattar and G. Dudek, Towards quantitative modeling of task confirmations in human-robot dialog. Robotics and Automation (ICRA), 2011 IEEE International Conference on. IEEE, 2011: 1957-1963.