Hand Motion Analysis using CNN

Harsh Raj, Aditya Duggal, Aditya Kumar Shetty M, Sreekanth Uppara, Srividya M S

Abstract-Hand motion detection and gesture recognition research has attracted large interest due to its wide range of applications in the field of Human computer interaction such as sign language recognition, 3D printing, virtual reality. There have been several approaches to create a robust algorithm to ease human computer interaction and perform in unfavourable environments.The real time recognition and learning of the model are big challenges. In this work, we use Convolutional Neural Network architecture to detect and classify hand motions, the region of interest of the image is passed through the neural network for the hand motion analysis and detection.Our system has achieved testing accuracy of 98%.

Keywords: Convolutional Neural Network, Human Computer Interaction, Robust, Testing accuracy.

I. INTRODUCTION

Human computer interaction has expanded covering most parts of the forms of modern information technology design available.Among the several human computer interaction approaches, Hand motion detection and analysis is also one of them. It can be used as a way to ease the interaction between human and computer.In the existing system, hand motion detection and analysis are affected by the working conditions.Hence in this work we try to overcome the cluttered environment and produce a satisfactory result and also try to ease the Human computer interaction. For the hand motion tracking several techniques available like Haar cascade, sensor based, edge detection, etc. This work makes use of Neural networks instead of other mentioned techniques to obtain faster and accurate results. Hand motion detection works properly when the factors are favorable to the camera. There are several factors which affect the system to run smoothly such as environmental background, light and rotation, translation and scale. Therefore an effective real time hand motion detection system is developed which works overcoming the challenges faced in the earlier systems.

Revised Manuscript Received in April, 2020.

* Correspondence Author

Harsh Raj*, Dept. of Computer Science and Engineering, R.V. College of Engineering, Bengaluru, India. Email: rajharsh209@gmail.com

Aditya Duggal., Dept. of Computer Science and Engineering, R.V. College of Engineering, Bengaluru, India. Email: aditya.duggal7777@gmail.com

Aditya K Shetty M., Dept. of Computer Science and Engineering, R.V. College of Engineering, Bengaluru, India. Email: adishetty1995@gmail.com

Sreekanth Uppara, Dept. of Computer Science and Engineering, R.V. College of Engineering, Bengaluru, India. Email:upparasreekanth12@gmail.com.

Srividya M. S., Dept. of Computer Science and Engineering, R.V. College of Engineering, Bengaluru, India. Email: srividyams@rvce.edu.in

II. LITERATURE SURVEY

Jing-Hao Sun et al. [1] In this paper, the authors have done research on hand gesture recognition using Deep learning. At first, hand segmentation is done using a skin color based model, then the image classification is done using a type of classifier called Adaboost which is based on Haar Classifier. All the activities such as segmentation, classification and hand gesture tracking is done in real time using an algorithm called Camshift algorithm.Then the hand gesture detection is also recognised in real time by Convolutional Neural Network.This system has accomplished a testing accuracy of 98.3%. Raj Patel et al. [2] In this paper, the author presented a HGR system developed for Sign Language using CNN and deep learning for the specially abled people.The architecture of the system has to be static to make the system ambulatory and adaptable. That one may accomplish immense exactitude on live scenarios; we employ a few image processing techniques with the application of neural networks both which reinforces a pertinent background subtraction and frame segmentation. Here the author 's approach focuses on maneuverability, costless and straightforward stationing in a low computational surroundings.This system has accomplished a testing accuracy of 96%.M. K. Ahuja et.al. [3] This paper is proposed based on a vision-based system of recognition which consists mainly of three steps. Hand segmentation is performed using the skin color model which is the first step. By applying the Otsu threshold the hand portion of the input image is segregated. A template based matching for gesture recognition is built in the final stage, PCA.M. I. N. P. Munasinghe et al. [4] This paper contains the plan and improvement of a framework which can recognize motions through a web camera progressively utilizing movement history pictures (MHI) and feedforward neural systems. As a matter of first importance the foundation from secured outlines is disengaged utilizing Gaussian blend put foundation/closer view division calculation that one may hold onto moving zones in the edge and hereafter middle sifting is authorized to dispose of impromptu clamor from the edge. Later the parallel thresholding nearby Otsu's binarization has been implemented and it will pinpoint incredible edge esteem and these refined casings are merged and total movement history picture is created utilizing a developed calculation relied on the anatomical liking measure. Anatomical proclivity between the cumulated picture and the underlying casing is additionally decided and applied in this calculation. Unequivocally feed-forward neural system with a theoretical slope based analyzer used to isolate the dynamic signals continuously.Guillaume Plouffe and Ana-Maria Cretu [5]

Retrieval Number: F3409039620/2020©BEIESP
DOI:10.35940/ijsce.F3409.059620
Journal Website: www.ijsce.org

26

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication
© Copyright: All rights reserved.

This paper discusses the evolution of a natural gesture user interface that uses the data collected by a kinetic sensor to recognize hand gestures in real time. Using an algorithm to locate the first pixel on hand region within that area, the scanning time is improved. For recognition of the entire hand area a directional search algorithm is used. When checked, a mean recognition rate of 92.4 percent was detected over 55 gestures. Two potential applications of this work are explored and evaluated: one for the perception of sign digits and gestures for a comfortable interaction between humans and machines and the other for the natural control of a software interface.Nabeel Vandayar, et al. [6] The paper has proposed a system to perceive the two i.e, static and dynamic hand motions in distinction to the users in real time by applying Convolutional Neural Networks. This system is adequate to accurately analyze 19 static hand gestures and recognise 6 hand motions with certainty of better than 95%. Nevertheless the structure composition is of very low cost, it has a fluctuating latency due to incompetent assimilating with the operating system as a shortcoming. It is also cited that this shortcoming can be overcomed by adopting an optimised user interface which would enhance the system performance and the user experience regarding this project.

III.   METHODOLOGY

A.   Image Pre-Processing

The First phase of the Hand Motion detection System is Image pre-processing. The Images from the real time hand motion capture by the camera must be processed using openCV python libraries before sending to the Neural Network model so that further latency, inaccuracy and difficulty in detection of pre-defined Hand gestures can be diminished. In this paper, Background subtraction, binary thresholding, Grayscaling along with Gaussian filtering are used for converting images received from the camera into Binary Images fit to be fed to the Deep Convolutional Neural Network. The Background Subtraction helps in faster recovery. The desired region of interest is calculated after the image cropping and hand region is segmented and the segmented region is drawn using contours.

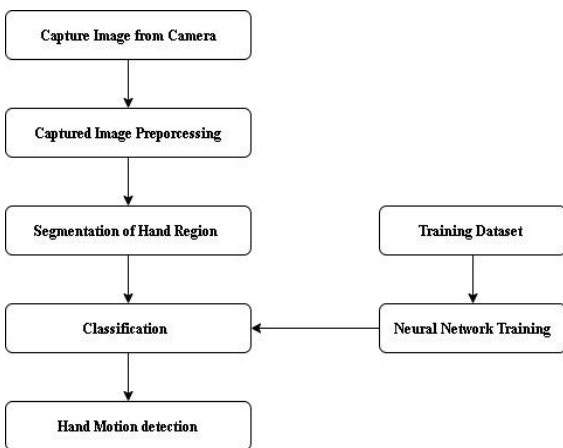### B.   Feeding processed images to Deep CNN

Fig. 1. Methodology Summary

Step 1: Random Weights Initialized

Step 2: The pre-processed images of size 215 x 240 pixels are fed to the neural network through forward propagation phase. There are 7 Convolutional layers and 2 fully connected layers. The features are extracted for Hand motion detection using Convolutional layer and fully connected layer is used to classify the hand gestures for hand motion detection. The convolutional operator uses Conv2D filter for extraction of different features. The Activation function used is Rectified Linear function and max pooling layer is used for dimensionality reduction of features and selecting best features.

Step 3: Output layer error is calculated using categorical cross entropy and Adam optimizer is used and the learning rate is set to 0.001.

Step 4: DNN is used to update all the weights and reduce loss.

The structure of the proposed CNN architecture is given below:

LAYERS	DETAILS
First Convolutional Layer	32 filters of size 2x2, ReLU
First Max Pooling Layer	Pooling size 2x2
Second Convolutional Layer	64 filters of size 2x2, ReLU
Second Max Pooling Layer	Pooling size 2x2
Third Convolutional Layer	128 filters of size 2x2, ReLU
Third Max Pooling Layer	Pooling Size 2x2
Fourth Convolutional Layer	256 filters of size 2x2, ReLU
Fourth Max Pooling Layer	Pooling size 2x2
Fifth Convolutional Layer	256 filters of size 2x2, ReLU
Fifth Max Pooling Layer	Pooling size 2x2
Sixth Convolutional Layer	128 filters of size 2x2, ReLU
Sixth Max Pooling Layer	Pooling size 2x2
Seventh Convolutional Layer	64 filters of size 2x2, ReLU
First Fully Connected Layer	1000 nodes, ReLU
Dropout Layer	Exclude 75% of Neurons Randomly
Output Layer	9 Nodes for 9 classes, Softmax

Fig. 2. CNN Architecture used.

To improve the learning process and quality of output Hyper parameter optimization is used.

Table 1. Hyper Parameter values

Hyper-Parameter	Value
Number of Epochs	10
Dropout Ratio	0.75
Convolutional Kernel	2x2
Max Pooling Kernel	2x2
Loss Function	Categorical Cross Entropy
Learning Rate	0.001

IV. EXPERIMENTAL RESULTS

Tensorboard 1.14 is used for the visualization and debugging of the training model. Tensorflow is the machine learning framework created by Google for designing, building and training deep learning models is also used in training of the Deep CNN model. TFLearn is a modular and transparent deep learning library built on top of Tensorflow is used to facilitate and speed up experimentation. On the x-axis, the number of training steps are shown. The dataset used contains 1144 Testing dataset and 1144 validation dataset. The dataset comprises binary images of hand-gestures. While testing the trained model, almost all hand gestures were recognized and respective navigation was given to the car simulation successfully. The shadows and light source orientation might affect the hand motion detection by the model.

Following are the hardware specifications used.

Table 2. Hardware Specifications

Processor	Intel i5
RAM	4 GB
System Type	64-bit Windows 10
GPU	1GB Nvidia-920

Following are the Graphical analysis of the training of the Deep Convolutional Neural Network model.
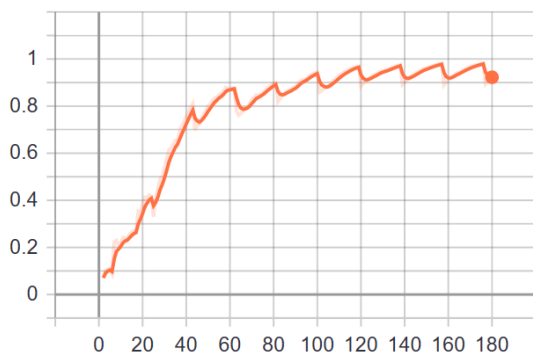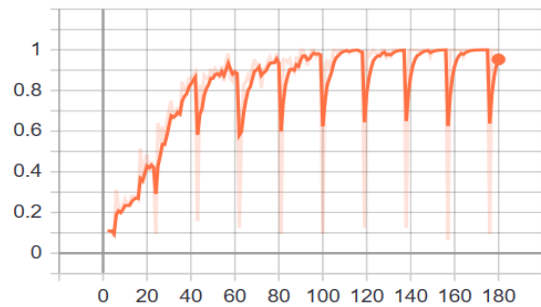

Fig 3. Accuracy Summary for Training Model

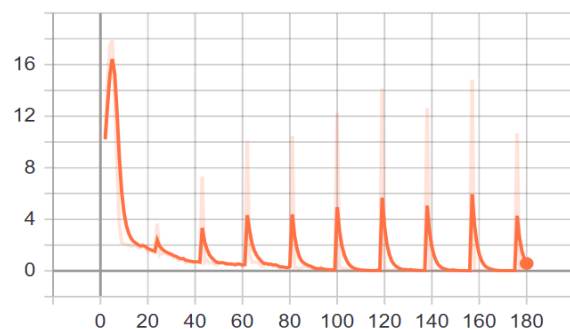**Fig 4. Randomized accuracy summary for Training Model**
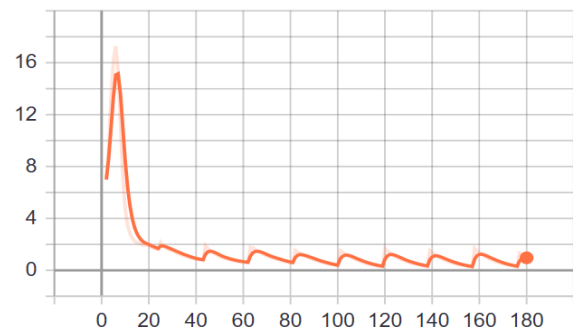
Fig 5. Loss Summary for Training Model

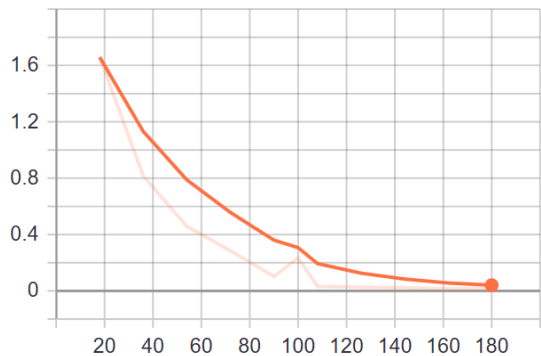**Fig 6. Randomized loss summary for Training Model**

Fig 7. Validation Loss summary for the training model

The Maximum accuracy reached by the model for

Adam optimizer is 98% while the maximum raw accuracy reached was 95%. It took about 1233 seconds to reach the highest accuracy at the 176th training step while training the Deep CNN model. The loss reached its minimum of 0.3 at the 157th training step and kept almost the same for the further training steps.The raw loss reached its minimum 0 at 116th training step. The Validation loss kept on decreasing until the last training step to the value of 0.04. Overall the training of the Deep CNN model was executed as expected.

V.   CONCLUSION

This work extracts the circumstances and issues in analysis of hand motions using an application of artificial neural networks. In this paper, a novel architecture is proposed to get the best accuracy and to provide a reliable and practical Hand motion detection system. We have implemented techniques such as hyper parameter optimizations to help improve the suitability of the system in real world scenarios. Although the system can recognize the numerous hand motions successfully, some extensions are still possible with respect to the various types of dynamic gestures which we can analyze for the input to our Software simulation model(car made by the application of computer graphics). Background subtraction has been applied to increase the model performance,the accuracy and to reduce the latency. It was assumed that the background should not be complex as the hand segmentation is sensitive to light intensity as well as the orientation of the camera. The recognition of dynamic gestures in complex backgrounds can be another extension to the work. This paper can be helpful for the researchers in further analysis of Hand motion detection using Convolutional Neural Networks.

ACKNOWLEDGMENT

REFERENCES

1. Raj Patel, Jash Dhakad, Kashish Desai, Tanay Gupta, Stevina Correia, 'Hand Gesture Recognition System using Convolutional Neural Networks ', 'IEEE/ICCCA 4th International Conference on Computing Communication and Automation,' 2018.
2. Jing-Hao Sun, Ting-Ting Ji, Shu-Bin Zhang, Jia-Kui Yang, Guang-Rong Ji, 'Research on the Hand Gesture Recognition Based on Deep Learning', 'IEEE/ISAPE 12th International Symposium on Antennas, Propagation and EM theory,' 2018.
3. M.I.N.P. Munasinghe 2018 3rd International Conference for Convergence in Technology (I2CT) The Gateway Hotel, XION Complex, Wakad Road, Pune, India. Apr 06-08, 2018.
4. Guillaume Plouffe and Ana-Maria Cretu,IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, 2019.
5. Nabeel Vandayar, Timothy James McBride, Kenneth John Nixon, 'Low Cost Hand Gesture Recognition System Design and Implementation,' 'IEEE/SAUPEC/RobMech/PRASA South African Universities Power Engineering Conference,' Bloemfontein, South Africa, 2019.
6. Mandeep Kaur Ahuja Dr. Amardeep Singh, 'Hand Gesture Recognition Using PCA,' 'IEEE/IJCSET International Journal of Computer Science Engineering and Technology,' 2015.
7. Felix Zhan, 'Hand Gesture Recognition with Convolutional Neural Networks,' 'IEEE/IRI 20th International Conference on Information Reuse and Integration for Data Science (IRI),' 2019.
8. Nico Zengeler 1, Thomas Kopinski 2, Uwe Handmann, 'Hand Gesture Recognition in Automotive Human–Machine Interaction Using Depth Cameras', MDPI, 2018.
9. Pei Xu, A 'Real-time hand gesture Recognition and Human-Computer Interaction System', 2017.
10. Opan Köpüklü, Ahmet Gunduz, Neslihan Kose,Gerhard Rigoll, 'Real time Hand gesture detection and classification using convolutional network (cs.CV)', 2019.

AUTHORS PROFILE

Harsh Raj, is an Undergraduate scholar pursuing Computer science and Engineering from R.V. College of Engineering, Bengaluru, Karnataka, India. He is passionate about Machine Learning, Image Processing and Network Security. He has worked on java projects on Fast Food management system and Machine learning project on Credit card fraud detection. He is well versed with programming languages like C, C++, JAVA, PYTHON. He is aiming to do further studies and research in the field of Network security and Artificial Intelligence and work on application based projects in these fields. He is currently under the guidance of prof. Srividya M.S. Assistant Professor, CSE department, R.V. College of Engineering.

Aditya Duggal is an Undergraduate Scholar pursuing Computer Science & Engineering from R.V College of Engineering, Bengaluru, Karnataka, India.His areas of interests are Artificial Intelligence,Data Science and Machine Learning ,Cryptography and Cyber Security.He has worked on different projects like Car Showroom Database Management System,Loan Origination Database System(during internship for Tata Consultancy Services Company) and with respect to Machine learning he made a project for Twitter Spam Detection.His programming knowledge consists of several languages like C,C++,PYTHON.His passion is to become an Ethical Hacker and is currently pursuing CEH course for the same from EC Council .He is currently under guidance of Prof.Srividya M.S,Assistant Professor,CSE Department,R.V. College of Engineering.

Aditya Kumar Shetty M is an Undergraduate Scholar pursuing Computer Science & Engineering from R.V College of Engineering, Bangalore, Karnataka, India. His areas of interest are Machine Learning, Image Processing and Networking. He has worked on projects such as Hospital Database Management, Books Management System in library, Facial Image Classification based on age and gender. He is well versed with programming languages such as C, C++, JAVA, PYTHON. His passion is to work in the field of Image Processing in future and has been working on such projects. He is currently under guidance of Prof.Srividya M.S, Assistant Professor, CSE Department, R.V College of Engineering.

Retrieval Number: F3409039620/2020©BEIESP
DOI:10.35940/ijsce.F3409.059620
Journal Website: www.ijsce.org

29

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication
© Copyright: All rights reserved.

Sreekanth Uppara is an Undergraduate Scholar pursuing Computer Science & Engineering from R.V College of Engineering. His areas of interest are Machine learning and Networking. He has worked on various projects like Car comparison system, Student management system, Facial age and gender estimation and emotion detection and recognition from text in his time at college. He has strong knowledge in programming languages such as C, Java, Python and HTML. He plans to do more studies in the areas of machine learning and artificial intelligence. He is currently under guidance of Srividya M.S., Assistant professor, CSE department, R.V college of engineering.

Prof. Srividya M S is an Assistant Professor at R.V College of Engineering.She has over 8 years of experience in teaching and 8years of experience in industry.Main area of research interest is Image Processing, Video Processing and Neural Networks.She has guided 5 UG projects and has following publications in international journals and conferences: "Underwater Video Processing for Detecting and Tracking Moving Object" International Journal of Engineering And Computer Science ISSN:2319-7242 Volume 3 Issue 5, May 2014".

"A Survey on: Underwater Video Processing for Detecting and Tracking Moving Objects" International Journal of Soft Computing and Engineering ISSN:2231-2307, Volume-4, Issue-1, March 2014