International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 992-996 www.ijaem.net ISSN: 2395-5252

Hand Gesture Recognition Using Artificial Neural Network in Matlab

Shristi Priya
Department of Electronics and CommunicationEngineering
Galgotias College of Engineering and Technology, Greater Noida

Submitted: 15-10-2021 Revised: 26-10-2021 Accepted: 28-10-2021

ABSTRACT—In these recent few years, various hand gesture recognition systems received great attention because of its importantapplicationsuch ashuman computer interaction i.e automatic interpretation of gestures based on computer vision. In this paper we have proposed a general approach for hand gesture recognition using eight target images as static hand gesture and input image as dynamic gesture then we have used geometrical method for feature extraction. The extracted result from the geometric method is thus optimized usingnntool (Neural Network Tool) in Matlab and finally converting the data into the required command interpretation of various different hand gesture based on computer vision.

Keywords-- Neural Networks, Gesture Recognition System, Artificial Neural Network,nntool.

I. INTRODUCTION

Gestures are among the primary and expressive form of human communication. These gesture include various meaningful body motions which involves various physical movements of body parts which are intended of conveying various meaningful information or interaction with the environment using fingers, hands,head, face, arms or body[1]. Various Gesture recognition has wide-rangeof applications such as developing an aids for the hearing impaired, virtual environment control, designing techniques for forensic discovery; sign language translation etc. Here we are talking especially about different hand gesture recognition systems.

The main aim of building hand gesture recognition system is to create a natural interaction between human and computer where the recognized different hand gestures can be used for controlling any device or conveying any meaningful information using sign language [2]. Hand gesture are mainly divided into two different approached i.eglove based analysis and vision based analysis. In glove based analysis there is sensor attached to glove which acts as transducer and in vision based several feature extraction techniques are used to extract the features of the different gesture images. There are some different machine learning approaches used in hand gesture recognition, such as ANN (Artificial Neural Network) [3], HMM (Hidden Markov Model) [6], recursive induction [7] and SVM (support vector machine).In this paper we are using ANN.

An ANN isconfigured for a certain applications, such as language processing, making decisions and dataclassification, through a learning process. Advantages of Artificial Neural Networks include efficiency i.e when any neuron is not responding will still produce the output, input is stored in its own networks instead of a database.

A. Related Work

Some researches that handle hand gesture recognition problem using different neural networks systems have been done with detailed showing their advantages and disadvantages. Many researches have been done their work on hand gesture recognition using Artificial Neural Network techniques [3][8][9].

This paper is organized as follows. Section 2 briefly introduces Artificial NeuralNetworks (ANNs), Section 3describes about a proposed model for hand gesture recognition and Section 4 discusses aboutthe input data and result and Section 5 about the conclusion.

II. INTRODUCTION TO ARTIFICIAL NEURAL NETWORK

A. Artificial Neural Network

Our human brain is not capable of solving various complex task and extracting various details from any compound structures. So to overcome this lack of ability Artificial Neural Network came up with a mathematical model. Now-a-days, Artificial Neural Network is a very popular and one of the most usedapproaches for the hand gesture

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 992-996 www.ijaem.net ISSN: 2395-5252

recognition. In Artificial Neural Networksystems, there areinterconnectedthat can compute values from inputs by feeding information through the network. Thismodel of artificial neural network is inspired by central nervous systems of brain which is also capable of machine learning and pattern recognition. An Artificial Neural Network is an information processing paradigm that works by the way, as the biological nervous system (brain), process information [7]. The key element of this paradigm isits novel structure of the information processing system. The neurons in Artificial Neural Network are highly interconnected which are working in unison such that it can solve any specific problem. According to Haykin [25], and Marcus [26], an Artificial Neural Network (ANN) can be defined as a hugely parallel distributed processor, consisting of simple processing units, which has a natural tendency for storing experimental knowledge and making it available for use.

The working concept of Artificial Neural Networks (ANNs) is similar to our human nervous system. Artificial neural network consists of one input layer, one output layer and one hidden layer, making a total of three, with each layer fully connected. The training algorithm for this Artificial Neural Networks is classified in the three classes: gradient descent, quasi-Newton (Levenberg–Marquardt, LM) and genetic algorithm (GA). Two versions of gradient descent back-propagation algorithms are incremental back-propagation (IBP) and batch back-propagation (BBP)

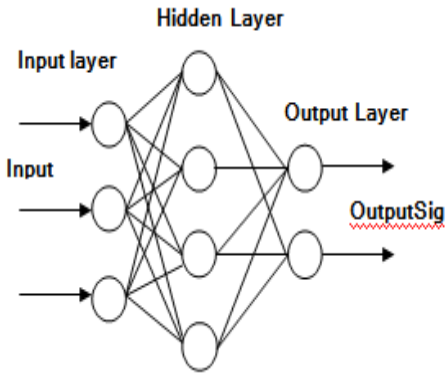

Fig. 1. Artificial Neural Network [3]

There are many various learning rules (algorithms) but the foremost often used is that the Delta-rule or back-propagation (BP) rule. A neural network is trained to map a group of input data by iterative adjustment of the weights. Information from inputs is fed forward through the network which is there to optimize the weights between different neurons. Optimization of the weights is formed by backward propagation of the error during training or learning phase. The ANN reads the input and output values within the training data set and changes the worth of the weighted links to scale back the differences between the anticipated and target (observed) values. Until this network reached the specified level of accuracy, the minimization of error in prediction is done. A complete round of forward–backward passes and weight adjustments using all input–output pairs within the data set is named an epoch or iteration. If a network is left to coach for too long, however, it'll be over trained and can lose the power to generalize.

III. A PROPOSED MODEL FOR HAND GESTURE RECOGNITION.

In this paper the system proposed five different stages: Image acquisition, skin detection, image pre-processing, feature extraction and artificial neural network, as shown in the figure.

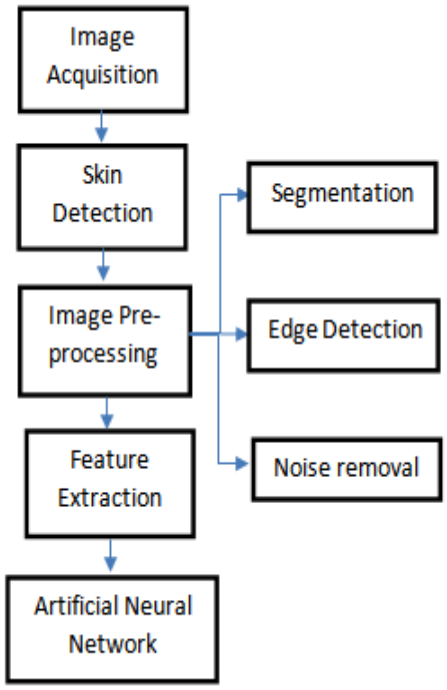

Fig. 2. Block Diagram of Proposed Model for Hand Gesture Recognition

The initial step of any computer vision system is the process of digital image acquisition.Digital image acquisition is a process in which there is creation of digitally encoded representation of visual characteristics of an image. From any image, various mathematical operations are applied to the data in order to create an enhanced

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 992-996 www.ijaem.net ISSN: 2395-5252

version of image i.e to perform some of the interpretation and recognition tasks.

After image acquisition we have done skin detection of captured image using thresholding method then we move to pre-processing of the image. Preprocessing is the very a vital part of the any gesture recognition process, as the captured image is made apt for the further processing i.e., for feature extraction. Pre-processing include mainly three different stages:segmentation[31], edge detection[30] and noise removal [40]. In pre-processing, first we do segmentation which involves detecting and extracting the hand region from the background. Here we have applied Ostu algorithm [41] for segmentation. The output of this step is a binary image in which skin pixels are having value '1' and non-skin pixels are having value '0'. If we take close proximity to the segmented image after applying the Otsu algorithm [41] on the original gray scale image we will see that the segmentation is not perfectly done. Background noise is given with 1s and hand gesture is given with 0s. As these errors can lead to a problem in contour detection of hand gesture, so we need to remove these errors. [42]

As we are using gesture based analysis so we will use feature extraction technique to extract the information. The features are one of the useful information that can be extracted from the segmented hand object by which the machine can understand the meaning of that posture. The numerical representation ofthese features can be procured from the vision perspective of the segmented hand object which manifests the feature extraction phase. [31]

Finally we will use Artificial Neural Network to optimize the result in Matlab. This we have done by nntool.

IV. INPUT DATA AND RESULT

These are the target samples, which are static images of 8 different orientations, as shown in Table I which are required to find out the orientations of input samples, which are dynamic images.

Figure 3: Gesture of 8 different orientations of target images used in our system. Top row [(1) th1=0 degree (Target 1), (2) th2=45 degree (Target 2), (3) th3=90 degree (Target 3),(4) th4=135 degree (Target 4), Bottom row [(5) th5=180 degree (Target 5),(6) th6=225 degree (Target 6), (7) th7=270 degree (Target 7), (8) th8=315 degree (Target 8)].

Now we will capture random dynamic image as input image as shown in fig. by the process of image acquisition.

Figure 4: Image acquition for random input gesture

Here we are doing skin detection then segmentation to get more and more information related to the images. The process of dividing the image segmentation into multiple parts is known as image segmentation.This method is generally used to identify objects or various other information indigital images. The methods used here for the segmentation purpose is the ostualgorithm.

The images shown below are the output images generated during the implementation of the segmentation algorithm.

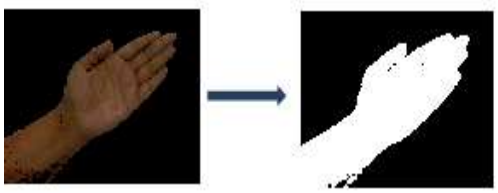

Figure 5:Image after segmentation and skin detection

Noise removal is done by various methods like linear filtering, median filtering and adaptive filtering. Here we have used median filter for noise removal.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 992-996 www.ijaem.net ISSN: 2395-5252

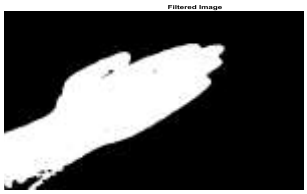

Figure 6: Filtered image

The filtered image undergoes feature extractions. The filtered image converted into binary image and then centroid of the image is calculated further using the region props function. Then the pixel value of the centroid is obtained as shown in the figure.

The pixel value is converted to polar values for all the target and the input values.

In this we apply the algorithm and then compare the result with the target images such that to obtain the orientation of the input samples, where x1, x2 and so on are defined as follows.

x1=th9-th1,x2=th9-th2, x3=th9-th3 and so on till x8=th9-th8. And again further among the various target images these values will be compared likex9=th2-th3, x10=th2-th4, x11=th2-th5, x12=th3-th4, x13=th3-th5, x14=th4-th5 and so on.

The input is then compared with all the these targets and then the result is observed

x=th9>th1; x1=1 1

x=th9>th2; x2=1 1

x=th9>th3; x3=0 0 and so on

It is thus calculated that the image is having least inclination of all three target images.

To further decide these images are among themselves, to know which one is smallest and according to that the image will have orientation in the direction of least th value like

x=th3>th2; x9=1 1

x=th4>th3; x14=1 1so on.

Thus we conclude th2>th9>th3 thus the image is in orientation of target 2 that is towards right with 45 degree i.e Target 2.

These polar values are given to the nntool for the optimization process and thus needed for obtaining the orientation of the input image with respect to the target images.Then we verify the result and create a number of any random input images in nntool using matlab. The error calculated of input image with different target network will classify in which orientation the input image exit. These error will depends on the number of epoch. Here epoch value selected is 40. A feed-forward neural network is having three hidden layers each containing 100 neurons. In order to train the feed-forward neural network, motion history images for different hand orientation gesture were created and saved. Then these images were loaded and fed to the neural network. After the training has been completed, the trained neural network is in classification phase of any random input image.Then for different orientation we can assign different commands.

V. CONCLUSION

In this paper, we have discussed how using camera can be used for detecting hand gestures and we have also used artificial neural network algorithm for optimizing results. We are using camera as a detecting device as well as input device for Reality System. In hand gesture recognition, artificial neural networks is generally used. The input for all the selected methods was digitized image and went under segmentation. Then some preprocessing was made on the input image like edge detection, background removal, noise removal, and binarisation. Then feature extraction must be made, different methods can be made, geometric features or non-geometric features, geometric features that use angles and orientations, palm center. Non geometric features like color, silhouette and textures, but they are not adequate in recognition. Here we have used geometric feature extraction using various geometric algorithms. Then the geometric feature of input sample images were compared with the eight target images of different orientations as given in figure 3. As the input target image was towards and the result comes out was correct, hence experiment was successful. Finally the result was optimized using nntool in which we have import the value of three target images and one input images. Then it form three different networks, now on training the network we see that our result was more optimized with right orientated image which was correct output.

REFERENCES

[1]. J. Canny, “A computational approach to edge detection,” IEEE Trans. Pattern Anal. Machine.Intel1, vol. 8, no. 6, pp. 679–698, Nov. 1986.

[2]. Kouichi Murakami and Hitomi Taguchi, Gesture recognition using recurrent neural networks,In CHI’91 Conference Proceedings,pages 237-242.1991.

[3]. Kouichi Murakami and Hitomi Taguchi, Gesture recognition using recurrent neural networks,In CHI’91 Conference Proceedings,pages 237-242.1991

[4]. E. R. Dougherty, “An Introduction to Morphological Image Processing”, Bellingham, Washington: SPIE Optical Engineering Press, 1992.

[5]. X.D. Huang, "Phoneme classification using semi-continuous hidden Markov models,''IEEE Trans. on ASSP, Vol. 40, No. 5, pp. 1062-1067, 1992.
[6]. W. Zhang and F. Bergholm, "Multi-Scale Blur Estimation and Edge Type Classificationfor Scene Analysis", International Journal of Computer Vision, 24, 3, 219 – 250,1997
[7]. Joseph J. LaViola Jr., "A Survey of Hand Posture and Gesture Recognition Techniques andTechnology," Master Thesis, NSF Science and Technology Center for Computer Graphics and Scientific Visualization, USA, 1999.
[8]. KirstiGrobel and MarcellAssan, _Isolated sign language recognition using hidden Markovmodels," In Proceedings of the International Conference of System, Man and Cybernetics, pages 162-167.1999
[9]. C. L. Lisetti and D. J. Schiano, "Automatic classification of single facial images," Pragmatics Cogn., vol. 8, pp. 185–235, 2000.
[10]. Lalit Gupta and Suwei Ma "Gesture-Based Interaction and Communication: Automated Classification of Hand Gesture Contours", IEEE transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 31, no. 1, February 2001.
[11]. Marcus Vinicius Lamar, "Hand Gesture Recognition using T-CombNET A Neural NetworkModel dedicated to Temporal Information Processing," Doctoral Thesis, Institute of Technology, Japan, 2001.
[12]. Andrew Blais and David Mertz, "An Introduction to Neural Networks – Pattern Learning with Back Propagation Algorithm" Gnosis Software,July 2001
[13]. Plumb, A.P., Rowe, R.C., York, P., Brown, M.,Optimisation of the predictiveability of artificial neural network (ANN) models: a comparison of three ANN programs and four classes of training algorithm. Eur. J. Pharm. Sci. 25, 395–405,2005.
[14]. S. Mitra, and T. Acharya."Gesture Recognition: A Survey" IEEE Transactions on systems, Man and Cybernetics, Part C: Applications and reviews, vol. 37 (3), pp. 311-324,2007.
[15]. G. R. S. Murthy, R. S. Jadon. "A Review of Vision Based Hand Gestures Recognition," International Journal of Information Technology and Knowledge Management, vol. 2(2), pp. 405-410. 2009.
[16]. P. Garg, N. Aggarwal and S. Sofat. "Vision Based Hand Gesture Recognition," WorldAcademy of Science, Engineering and Technology vol. 49, pp. 972-977.2009.
[17]. M. M. Hasan, P. K. Mishra."HSV Brightness Factor Matching for GestureRecognition System", International Journal of Image Processing (IJIP), vol. 4(5), pp. 456-467.2010
[18]. Shigeo Abe, "Support Vector Machines for Pattern Classification, second edition", Kobe University, Graduate School of Engineering, 2nd edition, Springer-Verlag London Limited 2005, 2010
[19]. Zhi-hua Chen,Jung-Tae Kim, Jianning Liang, Jing Zhang, and Yu-Bo Yuan, "Real Time Hhand Gesture Recognition Using Finger Segmentation", The Scientific World Journal, Vol. 2014,pp. 1-9.
[20]. Zhi-hua Chen,Jung-Tae Kim, Jianning Liang, Jing Zhang, and Yu-Bo Yuan, "Real Time Hhand Gesture Recognition Using Finger Segmentation", The Scientific World Journal, Vol. 2014,pp. 1-9
[21]. Kingma, D. P. and J. Ba, "Adam: A Method for Stochastic Optimization," in 3rd International Conference for Learning Representations, San Diego, 2014.
[22]. NuwanMunasinghe, "Dynamic Hand Gesture Recognition Using Vision and Neural Networks", Conference Paper, IEEE 3rd International Conference for Convergence in Technology (I2CT), April 2018.

Author

ShristiPriya has received her B.E Degree in Electronics and Communications from BIT, Mesra in 2014, and M.E Degree in Wireless Communications from BIT, Mesra in 2016. Presently she is working as Assistant Professor in GCET, greater Noida, ECE Department. Her area of interest is wireless communication and networking, antenna design, soft-computing, cloud-computing.
Corresponding author. Email: shristipr7@gmail.com