International Journal of Interdisciplinary Innovative Research &Development (IJIIRD)
ISSN: 2456-236X
Vol. 05 Special Issue 01 | 2020

Emotion Recognition using ANN classifiers in Machine Learning

Ajay K Talele[1], Dr S Tamil[2] , Dr Bharati Chaurasiya[3]

[1]*Research Scholar EC Department RKDF IST, SRK University Bhopal, India,*
[2]*Associate professor, EC Department, RKDF IST, SRK University Bhopal, India*
[3] *Head and Associate professor ,EC Department, RKDF IST, SRK University Bhopal, India,*

ABSTRACT

An emotion is a trigger of learning success, so the learning should be adapting to the students' emotions. The most of popular approach is the acquisition of facial-based features. Therefore, we present facial emotion recognition based on the Viola-Jones Algorithm in the learning environment. Basically, the Viola-Jones algorithm is a face detection algorithm. However, we use facial-based features to detect face and recognize emotion, thus we applied rectangular feature and DCT algorithm which are the main concept of the Viola-Jones Algorithm in those both of process. In this study, we compare accuracy, precision, recall, and time-consuming of the EARTHA algorithm and our previous methods [1] using 50 UM's learning images in student emotion recognition. The accuracy, precision, recall, and time-consuming of same algorithm reach 0.74, 0.73, 0.76 and 15 seconds per frame, whereas our previous methods [1] reach 0.46, 0.48, 0.52, and 42 seconds per frame. In emotional recognition, we can conclude that the proposed algorithm is superior to our previous work.

Keywords: *EARTHA (Emotion Affect Recognition Through Hybrid Approach), DCT,ANN classifier .*

1. INTRODUCTION

In the past few years, automatic facial expressions recognition has been extensively researched.[1]. It's basic concept is to detect human emotions based on the facial expressions. As per Charles Darwin theory A brief overview of emotions and expressions has been presented by, according to that these emotions and expressions species-specific rather than culture-specific [1][2]. Whereas in 1970's, studies performed by Paul Ekman indicated that there exist seven basic prototypical facial expressions as shown in Figure 1 which are Fear (FE), Disgust (DI), Angry (AN), Happy (HA), Neutral (NE), Sad (SA) and Surprise (SU) [3].

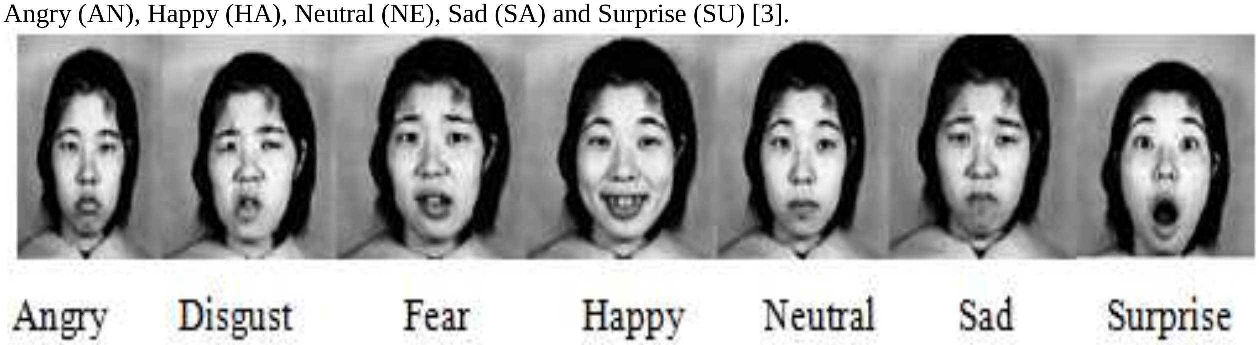

Figure 1 Emotional expressions

Emotions can be presented in two ways i.e. orally or though action [4]. The oral mode includes talked words in order to transmit message from each other efficiently while through action communication Facial Expressions or Gestures are used information between the groups of different humans [5]. One of the important way of expressing through action emotions is Facial expressions which generally provides a different approach to present features , goals and stream of emotions [6].

During this research some challenges are listed as follows, To properly detect faces using algorithm when given a Frontal face image of Size 256×256 pixels. To select Low Frequency DCT features, that can effectively and efficiently selected for detecting emotions. To model a ANN based solution which can be optimized in such a way to achieve the better prediction. To perform a comparative analysis of proposed and existing approach

The paper is arranged in the following manner first covers the used methodology, performance and then result and discussions .

2. METHODOLOGY

First we obtain a datasets of facial expression including all the basic 7 emotions. Then first step is preprocessing after pre- processing, we identify and extract potentially useful features using DCT for feature extraction and reduction of dimension of the images then ANN classifier is used for classification purpose which is shown in Figure 2

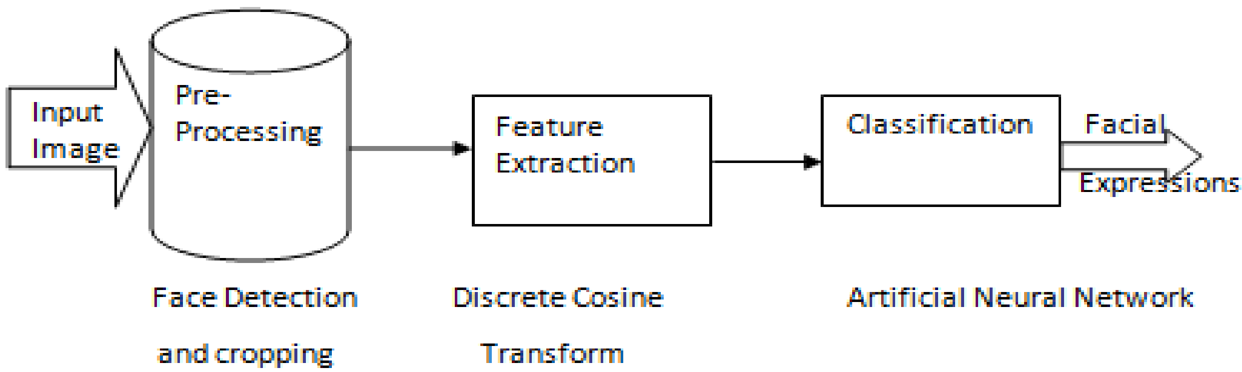

Fig 2: Pictorial representation of proposed methodology

In our research we are using Japanese Female Facial Expression (JAFFE) Database. It consists of ten different images, all consisting of only Japanese women. The database consists of total 213 gray scale images of size 256×256 pixels with female subject posing nearly 4 to 5 examples of the 7 basic facial expressions such as Angry, Fear, Disgust, Sad, Surprise, Happy and Neutral. Some sample images are shown in Figure 3.

In order to accurately and precisely extract the facial region a very efficient algorithm is used in our proposed methodology named as, EARTHA algorithm. The next step after the extraction process is to crop and resize to 240×260 pixels as shown in Figure 4. This data is further processed with a frequency domain technique known as Discrete Cosine transform (DCT) for feature extraction.

Fig 3: Sample images from JAFEE database

Figure 4: Face detection using voila Jones detector and cropping using detection.

The DCT has been widely applied to solve numerous problems among the digital image processing community. Discrete Cosine Transform with its one of a kind capacities can be utilized to extract distinguishing features from the given picture and furthermore can be utilized as an instrument for dimensionality reduction without losing a great part of the helpful data as illustrated in Figure.

DCT generally transforms the whole image into a set of different frequency co-efficient of cosines. The DCT frequency coefficient are categorized in three parts as, low, middle and high. Each category carry information and detail of the entire image. The most important feature required for facial recognition is intensity of image which is present in the low frequency part of the DCT. Moreover, in frequency the domain the most critical information (high magnitude) is carried by the low frequencies. Because of this feature high frequency part can neglected without affecting the original image. The two-dimensional DCT of an image results in matrix carrying low frequency component at top left and high at the bottom right corners.

The classifier we used is Feed-forward back propagation network or multi-layer perceptron neural networks. The number of input node are 128

Whereas there are 21 hidden nodes in the only one hidden layer and the output nodes are seven.

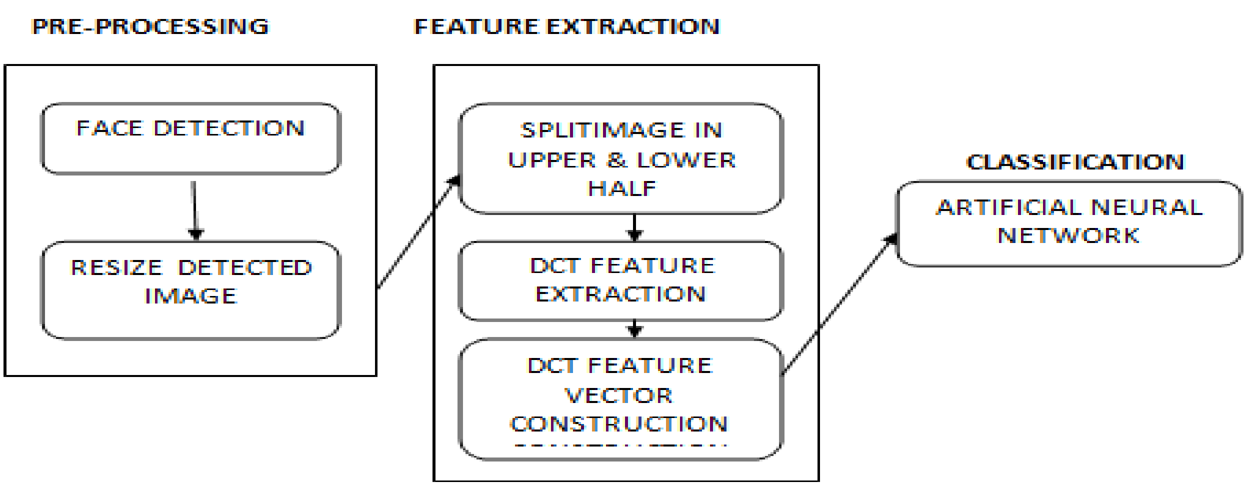

***Figure 5.* Proposed methodology**

Again, we use three division strategies as in [6] and the results obtained during these strategies are given next. The stepwise representation of the paper is a follows, First we acquired a dataset of facial expression including all the basic 7 emotions. Then after preprocessing, we identify and extract potentially useful features that show better recognition rates. After extracting features we use DCT for feature extraction & reduction of dimensionality then ANN classifier is used for classification purpose.

3. PERFORMANCE EVOLUTION

The human ERS model is analyzed and evaluated on the basis of different performance parameters like accuracy, false positive rate, false negative rate, sensitivity, receiver operating characteristics etc. Performance analysis of implemented methods and proposed method is presented in this chapter. Performance of proposed method is also compared with other methods surveyed through literature.

Performance Parameters Results obtained through experimentation are tabulated in the form of confusion matrix. Confusion matrix provides the values of TP-true positive, TN-true negative, FP-false positive and FN-false negative. The performance parameters like accuracy, FAR, FRR are obtained mathematically using equations .

True Positive Rate - TPR:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \text{------}(1)$$

False Positive rate - FAR/FPR:

$$\mathrm{FAR\backslash FPR} = \mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP+FN}} - - - (2)$$

False Negative Rate - FRR/FPR

$$\mathrm{FRP\backslash FNR} = \mathrm{FNR} = \frac{\mathrm{FN}}{\mathrm{FN+TP}} \text{-----}(3)$$

4 . RESULTS AND DISCUSSION

In this section we discuss our result which are made by our methodology. Our results in comparison with the existing work and found that our proposed methodology (EARTHA) of combining upper and lower facial features sets generate a very good recognition rate.

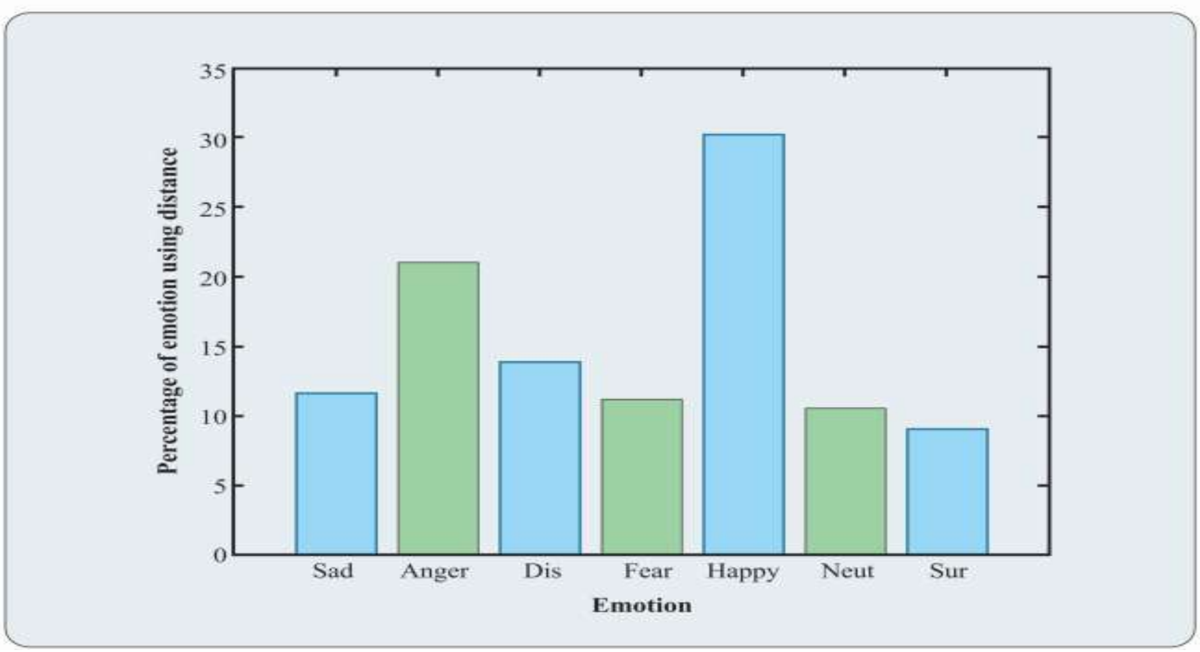

Figure 6 Sample emotion profile.

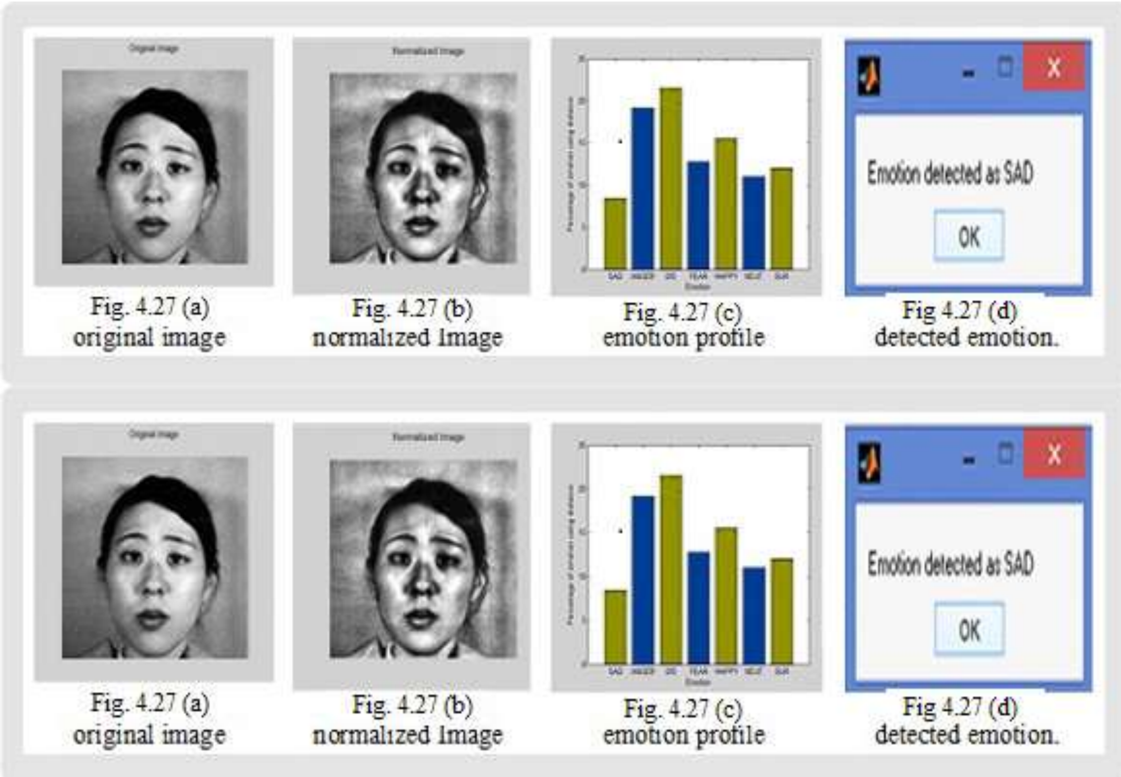

Figure 7 screen shots of the proposed system in MATLAB simulator

Classification is used for the cross validation, person independent , leave on out We present our results obtained from ANN classifier and compare the results with the existing work in the tabular form. We also construct confusion matrices and graphs for each strategy. Again, we use three division strategies as in [7] and the results obtained during these strategies are given next table 1.

Table 1 Result obtain from confusion matrix.

PC → AC ↓	SA	AN	DI	FE	HA	SU
SA	29	0	0	1	0	0
AN	0	27	2	1	0	0
DI	0	0	28	0	2	0
FE	0	1	0	31	0	0
HA	1	0	1	0	29	0
SU	0	0	0	0	1	29
Total	30	0	31	33	32	29
Accuracy	96.66	90.00	93.33	96.87	93.54	96.66

in cross validation method, the whole procedure is repeated for 10 times and the average recognition factor 97.14% is achieved. Whereas in second leave one out cross validation method, the whole process only one image is kept for testing while the others are used for preparation. Total 210 images are used out of them 209 for preparation and one is used for testing. In the third technique, the same strategy of segmentation is used as used in Cross Validation Method, be that as it may, this time rather than arbitrary division like our first strategy, we thoroughly avoid a man's looks one by one in each part to test the generalization capacity over various subjects.

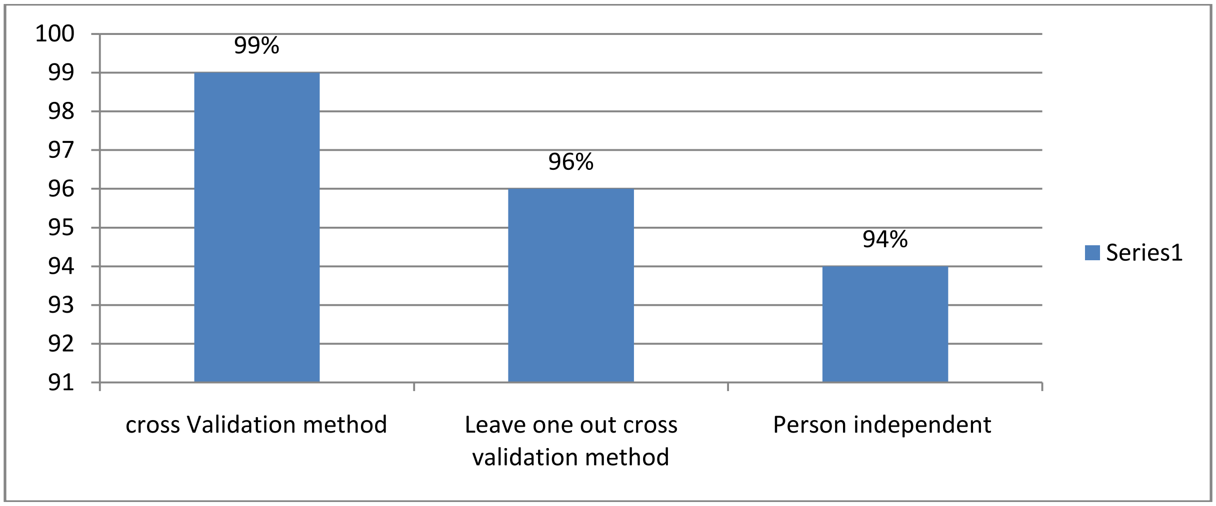

Figure 8 Comparison of different validation methods.

6. REFERENCES

[1]. T. M. Mahmoud, "A new fast skin color detection technique," In Proceedings of world academy of Science, Engineering and Technology, page 33, 2008.
[2]. C.Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, S. Lee, U. Neumann, and S. Narayanan, "Analysis of emotion recognition using facial expressions, speech and multimodal information," Proceedings of the 6th international conference on Multimodal interfaces, pages 205, New York, NY, USA, 2004.
[3]. D. Maltoni, D. Maio, A. K. Jain, S. Prabhakar, "Handbook of Fingerprint Recognition", 2nd Edition, Springer, New York, 2009.
[4].. M.Gargesha and P.Kuchi, "Facial expression recognition using artificial neural networks," J. Artificial Neural Computing System, pp. 1–6, 2002.
[5].. Mohammed Yeasin, Baptiste Bullot, and Rajeev Sharma, "Recognition of Facial Expressions and Measurement of Levels of Interest from Video," IEEE Transactions on Multimedia, Vol. 8, No. 3, June 2006.
[6]. W. Zhao, R. Chellappa, A. Rosenfeld, P. J. Phillips, "Face Recognition: A Literature Survey." ACM Computing Surveys, Vol. 35, No. 4, pp. 399–458, 2003.
[7].. Y.Azoz, L. Devi, M. Yeasin, and R. Sharma, "Tracking human amusing constraint fusion and multiple cue-based localization", Machine Vision Application., vol. 13, pp. 286–302, 2003.
[8].. M. Dahmane and J. Meunier, "Prototype-based modeling for facial expression analysis" IEEE Trans. Multimedia, vol. 16, no. 6, pp. 1574-1584, 2014.
[9]. A. Shaukat, M. Aziz and U. Akram, "Facial Expression Recognition Using Multiple Feature Sets," 2015 5th International Conference on IT Convergence and Security (ICITCS), Kuala Lumpur, 2015, pp. 1-5.
[10] G. Guo and C.R. Dyer, "Learning from Examples in the Small Sample Case: Face Expression Recognition," IEEE Trans. Systems, Man, and Cybernetics Part B, vol. 35, no. 3, pp. 477-488, 2005.
[11] J. Whitehill and C.W. Omlin, "Haar Features for FACS AU Recognition," Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (AFGR '06), pp. 217-222, 2006.
[12] K. Anderson and P.W. McOwan, "A Real-Time Automated System for Recognition of Human Facial Expressions," IEEE Trans. Systems, Man, and Cybernetics Part B, vol. 36, no. 1, pp. 96-105, 2006.
[13] M. Valstar, M. Pantic, and I. Patras, "Motion History for Facial Action Detection from Face Video," Proc. IEEE Int'l Conf. Systems, Man, and Cybernetics (SMC '04), vol. 1, pp. 635-640, 2004.
[14] M. Pantic and I. Patras, "Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments Form.
[15] Face Profile Image Sequences," IEEE Trans. Systems, Man, and Cybernetics Part B, vol. 36, no. 2, pp. 433-449, 2006.