International Journal of Control and Automation
Vol. 12, No. 6, (2019), pp. 17-25

Deep Neural Network and ABC Based Emotion Recognition

Himanshu Sharma, Rohit Agarwal

Department of Computer Engineering & Applications

GLA, University, Mathura, India-281406

Abstract

Facial expressions are used to show set of fundamental and universal emotions by humans. In the human emotion identification filed as well as in computer human interaction, automatic emotion recognition plays a major role. They are used in the applications like health care monitoring, personalized learning and surveillance etc. When features are mapped into high dimension, enhanced results are produced by Distance Metric Learning (DML) and in SVM detection.

However in which LBP features only extracted from the image and does not extracted the other features. And accurately detecting the emotion is still having some challenges. To avoid this issue in this work provide one new method to extract GLCM feature along with the LBP features. Additionally feature selection method also introduced in this work using ABC algorithm to improve the accuracy of the classifier. And noise removal using Gaussian Noise. In this Emotion detection will be done by using ensemble learning classifiers (EL-CF) method.

***Keywords**: Emotion detection, DML, LBP features, facial expressions and GLCM feature.*

1. Introduction

In the field of computer and human interaction, smart system exploration, emotion recognition is receiving a high attention [1]. Six fundamentals emotions are classified from facial expression recognition, depends on attributes of face. They are happiness, anger, surprise, fear, disgust and sadness. Stereo-scopic-perceptual conflict property is contained by every emotion as stated by Russel and Coren.

It is a very challenging task to design an effective automatic emotion recognition system. Smooth computer and human interaction can be done using a successful emotion recognition [2,3]. In the field of heterogeneous, various applications used human emotion recognition system [4,5,6]. Computer and human interaction, antisocial motives inspection, surveillance of patients are most commonly used applications. When features are mapped into high dimension, enhanced results are produced by Distance Metric Learning (DML) and in SVM detection.

And accurately detecting the emotion is still having some challenges [7]. To avoid this issue in this work provide one new method to extract GLCM feature along with the LBP features. Additionally feature selection method also introduced in this work using ABC algorithm to improve the accuracy of the classifier. And noise removal using Gaussian Noise. In this Emotion detection will be done by using EL-CF method[8].

Pao, [9] extracted image's facial features and face by using Harris corner key-points and Viola-Jones cascade object detectors to form a hybrid facial expression recognition and feature extraction technique. A multi-class predictor(MCP) is trained by using support vector machines (SVM), histogram-of-oriented-gradients (HOG) feature extraction, linear discriminant analysis, and principal component analysis in order to classify facial expressions of humans.

Dagar et al [10]used a live streaming to design an automated framework(AF). Neural network and Gabor feature extraction are used to process it. Principal emotion based extraction and detection of emotion facial attributes are challenging. Component analysis is utilized to overcome the requirements of facial expression recognition method which is

ISSN: 2005-4297 IJCA

having face with multiple variations in colour, expression and orientation. Based on emotion, clustering is performed. Already learned pattern classifiers is used to find the facial expression by feature vector which are processed.

2. Proposed Methodology

From the expressions of the face, Emotion recognition and classification is the most challenging and critical issue. Proposed model in order to detect emotion Gaussian Noise based noise removal and GLCM feature along with the LBP features and ABC algorithm based feature selection is used in this system. And noise removal using Gaussian Noise. In this Emotion detection will be done by using deep learning method. Figure 1 shows the proposed model's overall architecture.

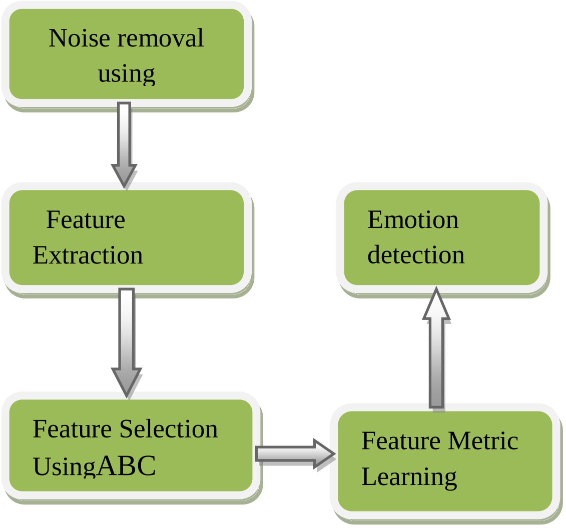


Figure: 1.Proposed System

2.1. Noise Removal Using Gaussian Noise

Assume, current pixel to be processed as X (i,j); sliding or filtering window of size (2L+1) x (2L+1) is given by S i, j and which is centered at X(i,j). The elements of this window are S i,j = {Xi-u,j-v , -L ≤ u,v ≤ L}[11,12].

1. X represents the noisy image.
2. Immerkaer's fast method is used to compute, standard deviation (SD) of noise.
3. From noisy image, select, a 2-D filtering window $S_{ij}$with size 3x3 and center pixel is assumed as X(i,j). Absolute value of difference AD = | S_{ij} – X(i,j)| is computed by subtracting every pixel from centre pixel.
4. In a one dimensional array, the pixel is stored as DA(x), if the absolute difference AD < (SF*SD)).
5. Mean of DA(x) is calculated, if number of elements in DA(x) is at least (2*W) -1 and replaced it at the center pixel X(i,j) of window.
6. Else increase the window size and repeat the same process.
7. untilprocessing the entire image, repeat step 3 to 6.

2.2. LBP and GLCM Feature Extraction

In variation of light, a useful algorithm called Local binary pattern (LBP) feature extraction algorithm can be used. The process of LBP is as follows, Over an image, traverse a window which has specific neighbourhood values. Assignment of centre pixel is done. Based on the adjacent pixel of the centre pixel, threshold value is chosen.

In anti-clock wise or clockwise direction, based on local neighbourhood, Compute the LBP matrix. Mathematically compute the textural or structure and statistical model. The change in gray level and simplicity of computation are the major features of LBP algorithm. In real time applications they are used because of this.

In an image, there will be G gray levels. The number columns and rows of GLCM matrix equals this gray levels G. [13].In experiments the GLDM feature descriptor was computed from five values of the angle ө and displacement distance d=1 therefore, the implementation contains 20 features. The texture features are computed as below: Energy: Energy indicates the uniformity observed in the mammographic image. Generally, energy is computed from the value of the mean squared signal. It is computed as below

Contrast*:* The separation among brightest and darkest area defines the contrast.

Contrast = $\sum_{ij=0}^{n-1} P_{ij}(i-j)^2$ (1)

Correlation*:* Correlation is computed and if it lies in the range between -1 to +1, then it is termed as correlation coefficient.

Correlation = $\sum_{ij=0}^{n-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2}$ (2)

Homogeneity*:*It is defined as the quality or state of being homogeneous

Homogeneity = $\sum_{ij=0}^{n-1} \frac{P_{ij}}{1+(i-j)^2}$ (3)

Entropy*:*In a random variable, uncertainty is measured by entropy.

Entropy = $\sum_{ij=0}^{N-1} -\ln(P_{ij})P_{ij}$ (4)

Energy:

In GLCM, squared elements sum is given by this. It is also termed as angular second moment and uniformity.

Energy = $\sum_{ij=0}^{N-1} (P_{ij})^2$ (5)

Shape*:* Object's geometric properties are referred by shape. It also refers external boundaries when compared to material composition, texture and colour properties.

Colour*:* Light's components are given by colour. When object reflects it, it will get separated. The coordinates of it are used for its identification.

Intensity*:* Purity of colour or strength of colour is given by this.

Texture*:* Surface's visual characteristics are given by this. There may be smooth or rough surface.

2.3. Feature Selection Using ABC

The honey bee's foraging behaviour is inspired in ABC. There are three class of bees in the colony in ABC algorithm. They are, employed, scout and onlooker bees. Twice the amount of food sources equals the population. The optimization problem's possible solutions are represented by food source count. The solution's quality is represented by food source's nectar amount.

Food sources are exploited by employed bees and the information about nectar quantity in that food source is given to onlooker bees. The count of employed bees equals the count of onlooker bees. Based on this from employed bees, exploitation of food sources are decided by onlooker bees. The employed bees of abandoned food sources are made as scout bees to select new food sources [14].

Every feature is allocated with every employed bee. From objective function f I, feature's fitness function is evaluated by employed bee as,

$$fit_i = 1/(1+f_i) \quad (6)$$

The employed bees will give information to onlooker bees. Using equation (7), selection probability of feature is computed.

$$P_i = \frac{fit_i}{\sum_{i=1}^{m} fit_i} \quad (7)$$

The new solution v_i , for feature's pointed by employed bees and predictive accuracy is computed by onlooker. The employed bee will point to feature subset, if new solution v_i is greater than x_i . The newly selected feature will be neglected. Using equation (8), compute the new solution vi.

$$V_i = X_i + \varphi_i(x_i - x_j) (8)$$

Where, predictive accuracy of feature allocated to employed bee is represented by xi and predictive accuracy of feature onlooker has selected is represented as xj. Uniformly distributed real random number is given as ϕi and it lies in range [0, 1]. A new subset of features is formed by this way every time and it is produced by exploiting onlooker.

1. Positions of food source are initialized
2. Food sources are evaluated
3. For employed bees, new food sources are produced
4. Greedy selection is applied
5. Fitness and probability values are computed
6. For onlookers , new food sources are predicted
7. Greedy selection is applied
8. Abandoned food sources are computed and new food sources are find by allocating employed bees of abandoned food source as scout bees.
9. The found best food source is memorized
10. For a pre-determined number of iterations, steps 3-9 are repeated

2.4. Distance Metric Learning

The features with same class are brought together by DML method. This is the major objective of this. This makes the features from different class with high distance between them. Emotion's all samples are used to compute the mean template of it.

2.5. Ensemble learning method

A decision making system made up of an analogy in daily life demonstrates the multiple classifier system's usage. Before making a final decision in an important issue, experts need suggestions from other experts. The individual decisions from all experts are combined to make a final decision.

Various errors are produced by individual classifiers or diverse features. So this can be avoided by combining them to form an ensemble classifier which reduces the error by an averaging function of result. Single models are not able to produce good prediction and classification performances in multiclass problems. But it can be produced by ensemble methods [15].

Extreme learning Machine

Singlehidden layer feed forward neural network (SLFNN) corresponds to Extreme Learning Machine (ELM). In this, without training, hidden neuron biases and input weights are selected randomly. General linear system's Moore-Penrose inverse and norm least-square solution are used to compute the output weights analytically. This reduces the training time required. For neurons of hidden layer and output layer's activation functions are chosen sine, sigmoid and Gaussian activation functions.

From image features, input is extracted in training process. Target values corresponds to components with high frequency which are extracted from high resolution original image. The model is learned by ELM. Interpolated image can be mapped by this. On high frequency components, they are imposed. Using images with low resolution are used to predict components of high frequency by learning model, which is done after training [15].

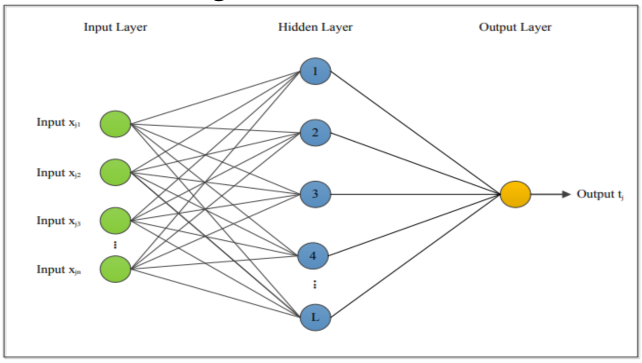
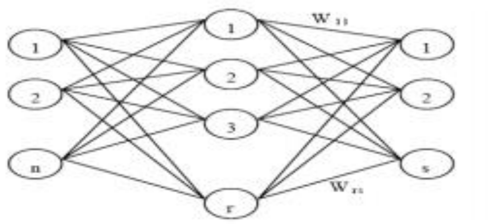


Figure 2. Overview of ELM

RBF Natural Network

As shown by Figure 6, there are one input, hidden and output layer in RBFNN. In hidden layer, radial basis functions are used. Radial basis function has, centre parameters, they are adjusted for training purpose. The strength of connection between output and hidden layer is computed by this function [16].



Input Layer RBF Units Output Layer

Figure: 3. the construction of RBFNN algorithm

$$R_i(X) = R_i\left(\frac{\|X - C_i\|}{\sigma_i}\right), i = 1,2, \ldots \ldots r \qquad (9)$$

Where, n-dimensional input feature vector is represented as x, n dimensional vector is represented as c_i and it is the center of RBF unit, width of RBF unit is given by σ_i and number of RBF units is represented as r. Gaussian function is chosen as a RBF's typical activation function and it has mean vector c_i and variance vector Ri (x) as,

$$R_i(X) = exp\left(-\frac{\|X-C_i\|^{\ 2}}{\sigma_i}\right) \qquad (10)$$

The Gaussian function's covariance matrix has diagonal entries and they are represented as σ_i 2. Linear behaviour is shown by the output unit. For a given input x, J^{th} output unit's response is given by,

$$y_j(X) = b_j + \sum_{i=1}^{r} \quad R_i(x)w_2(i,j) \qquad (11)$$

Where, Connection weight of i-th RBF unit to j^{th} output node is represented as W2 (i, j),bias of j^{th} output is given by b(j). The complexity of the network is reduced by eliminating this bias.

Feed forwarded neural network

Neural network is a nonlinear statistical data modelling tool which is used in practical terms. The data patterns and complex relation among input and output can be modelled using NN. In data mining, from dataset, information can be harvested by data warehousing using neural networks.

The data is cross-fertilized and manipulated actually in data warehouse. This makes the major difference between ordinary databases. Users can take a clear decision by this data warehouse. Feed forward Neural Network is a simplest of NN [17,18]. It is shown in figure 6. It has three layers namely, input, hidden and output.

3. Result and Discussion

The above mentioned techniques are modelled as a real-time graphical user interface. It requires a laptop with i3 third-generation processor including 2 GB RAM and hard disk space of 500GB. Test sample is created by taking a snap by using webcam of a laptop. The CK+ is used to train the model.

For test case, college student images are used. In real time, the performance of the system is tested. For the detection of emotion in extended Cohn–Kanade database, this method can be utilized. Results of the proposed EL-CF are measured with MCP, AF.

Table 1. Performance comparison

Classifiers	**Accuracy (%)**	**Precision (%)**	**Recall (%)**
EL-CF	75.00	76.92	90.91
AF	70.00	76.47	86.75
MCP	69.76	73.39	86.67

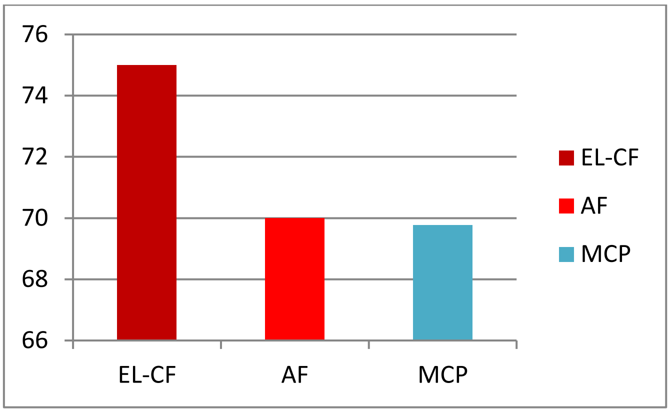


Figure: 4. Accuracy results comparison vs. classifiers

Figure. 4 illustrate performance comparison of Accuracy metrics with respect to three different classifiers such as EL-CF, AF, and MCP. From the results it concludes that the proposed EL-CF classifier produces higher Accuracy results of 75.00%, whereas other methods such as AF, and MCP. Classifier produces only 70.00% and 69.76% values respectively.

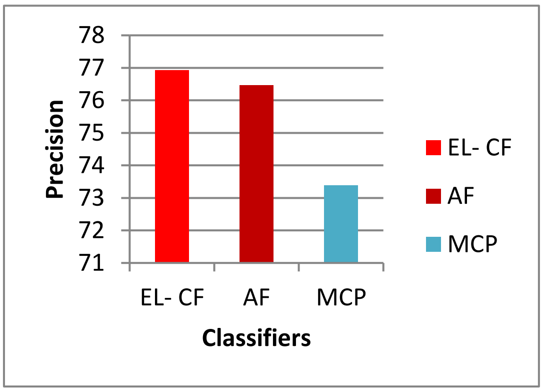


Figure: 5. Precision results comparison vs. classifiers

Figure. 5 illustrate performance comparison of precision metrics with respect to three different classifiers such as EL-CF, AF, and MCP. From the results it concludes that the proposed EL-CF classifier produces higher precision results of 76.92%, whereas other methods such as AF, and MCP. Classifier produces only 86.75% and 86.67% values respectively.

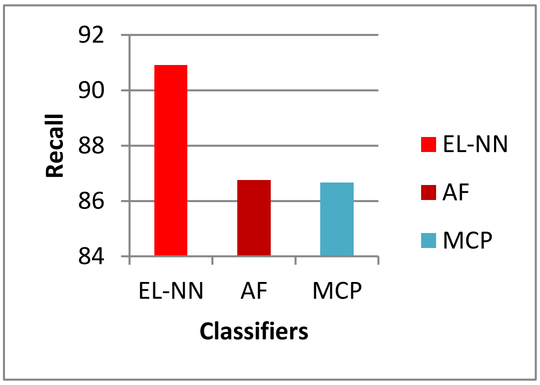


Figure: 6. Recall results comparison vs. classifiers

Figure 6 illustrates performance comparison results of precision metrics with respect to three different classifiers such as EL-CF, AF, and MCP. From the results it concludes that the proposed EL-CF classifier produces higher Recall results of 90.91%, whereas other methods such as AF, and MCP.Classifiers produce only 76.47% and 73.39% values respectively.

4. Conclusion and Future Work

Recognizing emotion using facial expressions is a key element in human communication. An ensemble classifier is a predictor used in processing of images. It is trained by face images. In which noise removal using Gaussian Noise is used for eliminating the salt and pepper noise from the human image. In this work provide one new method to extract GLCM feature along with the LBP features. Additionally feature selection method also introduced in this work using ABC algorithm which enhances the accuracy of classifier. Metric learning technique is used to mapping the features and finally Emotion detection will be done by using EL-CF method. With respect to accuracy, precision and recall, better results are produced by proposed method as shown by experimentation. In future, techniques are compared and fine-tunedusing commondataset.

References:

[1] Yu, Z. and Zhang, C., 2015, November. Image based static facial expression recognition with multiple deep network learning. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction (pp. 435-442). ACM.

[2] Wood, A., Rychlowska, M., Korb, S. and Niedenthal, P., 2016. Fashioning the face: sensorimotor simulation contributes to facial expression recognition. Trends in cognitive sciences, 20(3), pp.227-240.

[3] Mollahosseini, A., Chan, D. and Mahoor, M.H., 2016, March. Going deeper in facial expression recognition using deep neural networks. In 2016 IEEE Winter conference on applications of computer vision (WACV) (pp. 1-10). IEEE.

[4] Calvo, M.G. and Nummenmaa, L., 2016. Perceptual and affective mechanisms in facial expression recognition: An integrative review. Cognition and Emotion, 30(6), pp.1081-1106.

[5] Zhang, K., Huang, Y., Du, Y. and Wang, L., 2017. Facial expression recognition based on deep evolutional spatial-temporal networks. IEEE Transactions on Image Processing, 26(9), pp.4193-4203.

[6] Barsoum, E., Zhang, C., Ferrer, C.C. and Zhang, Z., 2016, October. Training deep networks for facial expression recognition with crowd-sourced label distribution. In Proceedings of the 18th ACM International Conference on Multimodal Interaction (pp. 279-283). ACM.

[7] Barsoum, E., Zhang, C., Ferrer, C.C. and Zhang, Z., 2016, October. Training deep networks for facial expression recognition with crowd-sourced label distribution. In Proceedings of the 18th ACM International Conference on Multimodal Interaction (pp. 279-283). ACM.

[8] Mollahosseini, A., Hasani, B., Salvador, M.J., Abdollahi, H., Chan, D. and Mahoor, M.H., 2016. Facial expression recognition from world wild web. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 58-65).

[9] Pao, J., 2016. Emotion Detection Through Facial Feature Recognition. Tech. Report Project_Autumn_1617.

[10] Dagar, D., Hudait, A., Tripathy, H.K. and Das, M.N., 2016, May. Automatic emotion detection model from facial expression. In 2016 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT) (pp. 77-85). IEEE.

[11] Wang, X., Shen, S., Shi, G., Xu, Y. and Zhang, P., 2016. Iterative non-local means filter for salt and pepper noise removal. Journal of visual communication and image representation, 38, pp.440-450.

[12] Kaur, S., 2015. Noise types and various removal techniques. International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE), 4(2), pp.226-230.

[13] Zhang, L., Jing, J. and Zhang, H., 2015. Fabric defect classification based on LBP and GLCM. Journal of Fiber Bioengineering and Informatics, 8(1), pp.81-89.

[14] Babu, R.L. and Balasubramanie, P., 2019. Fuzzy Rule Selection Using Hybrid Artificial Bee Colony with 2-Opt Algorithm for MANET. Mobile Networks and Applications, pp.1-11.

[15] Tang, J., Deng, C. and Huang, G.B., 2015. Extreme learning machine for multilayer perceptron. IEEE transactions on neural networks and learning systems, 27(4), pp.809-821.

[16] Zhao, J., Wei, H., Zhang, C., Li, W., Guo, W. and Zhang, K., 2015. Natural gradient learning algorithms for RBF networks. Neural computation, 27(2), pp.481-505.

[17] Ehlers, R., 2017, October. Formal verification of piece-wise linear feed-forward neural networks. In International Symposium on Automated Technology for Verification and Analysis (pp. 269-286). Springer, Cham.

[18] Wan, Y. and Gao, Q., 2015, November. An ensemble sentiment classification system of twitter data for airline services analysis. In 2015 IEEE international conference on data mining workshop (ICDMW) (pp. 1318-1325). IEEE.