Human Facial Expression Recognition System Using Artificial Neural Network Classification of Gabor Feature Based Facial Expression Information

Bayezid Islam, Firoz Mahmud, Arafat Hossain, Md. Sumon Mia, Pushpen Bikash Goala
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
Emails: (bayezid.shouvik, fmahmud.ruet, arafat.ruet.bd, sumon.ruetbd, pushpen.bikash)@gmail.com

Abstract—**Facial expressions contribute highly in conveying the feelings of a person. An emotion recognition system through facial expression recognition is proposed in this paper. Preprocessed input images are segmented into four facial expression regions by following the proposed highly effective image segmentation method. 2D Gabor filter with different frequencies and orientations are used to extract features from the segmented parts. Reduction of the dimension of the extracted features is done using downsampling and Principal Component Analysis (PCA). Classification of the features is done using the artificial neural network (multilayer perceptrons with backpropagation). To evaluate the performance of the proposed method three widely used facial expression datasets (JAFFE, CK+, RaFD) are used. Performance on these datasets by the proposed method is compared with the performance on these datasets by other methods to indicate that state-of-the-art performance is achieved by the proposed method.**

Keywords—Image Segmentation, Gabor Filter, PCA, Artificial Neural Network (ANN), Multilayer Perceptrons (MLP), Backpropagation, Facial Expression Recognition (FER), Emotion Recognition

I. INTRODUCTION

Facial expressions are powerful tools of nonverbal communication which play a vital role in interpersonal communication. Human emotion, sentiment, mental condition are highly dependent on facial expressions. In his classic paper [1] A. Mehrabian analytically showed that what a speaker wants to say is 55% dependent on facial expression. So if the facial expression can be recognized properly then there is a very high possibility of knowing the feelings, sentiment, mental condition and thus emotion of a person.

Recognized facial expressions can be used in different sectors of our life to give us more comfort, better life experience with increased security. The emotion of a person can be a measure to define whether a person is fit to do a certain task or not. Social networks can incorporate the facial expression recognition (FER) mechanism and suggest user post status depending on it. Robots can be enriched with FER features to get a better experience from them. As we are heading towards automation day by day, the ultimate goal is to recognize emotions with machines as a human can do spontaneously. Basically, facial expressions can be categorized into seven classes [2] such as neutral, happy, sad, fear, anger, disgust, surprise.

As it has a lot of applications in varied fields, FER has been an active research topic in different fields such as computer vision, image processing, human-computer interaction etc. for the last few decades. Many efforts have been made and still many are being made to make systems as accurate as possible to recognize human facial expressions at the least possible time. The usual steps of FER systems are image preprocessing, feature extraction from the preprocessed image and classification of the extracted features. In the image preprocessing step color conversion, face detection, image resizing, image rotation and many other system dependent tasks are done. Popular feature extraction techniques include Linear Discriminant analysis (LDA) [3], Principal Component Analysis (PCA) [3-7], Gabor wavelets [8], moments, Speeded-Up Robust Features (SURF), Histogram of Oriented Gradients (HOG) features, Local Binary Patterns (LBP) and many others.

Among the popular classification algorithms, Support Vector Machines (SVM) [9], multiclass SVM [10], nearest neighbor classification [11], Extreme Learning Machine (ELM) [12], AdaBoost classification [13], random forests [14], decision trees [15], Artificial Neural Network (ANN) [16-18] are few to be mentioned. But all of these classifiers are not suited for all scenarios. Choosing appropriate regularization term and kernel function, to avoid overfitting, is a tough task while using SVM where the number of samples is far less than the number of features per sample. Overlapping classes are problems with SVM. Training time can be another issue in special cases with SVM. The appropriate value of K, proper distance metric are challenges while using nearest neighbor classifier. It is not memory efficient and also known as a lazy learner. Finding an optimal number of nodes in the hidden layer is a crucial task with ELM. At cases, it can get stuck at local minima. Systems with ELM and decision trees might be overfitted. Random forest is tough to visualize. Outliers and noisy data can be challenges while using AdaBoost.

To solve the above-mentioned problems with different classification methods ANN is used as the classification algorithm in the proposed method. ANN can imitate the learning process of human and as the target is to recognize expressions as a human can do, so it can be used for FER problems. Besides, ANN has the ability to model complex problems, so is suited for image processing problems like FER. An FER system is proposed where images are preprocessed first in three different steps, preprocessed images are segmented into four facial expression regions according to a newly proposed method, features are extracted from the expression regions using 2D Gabor filter, dimension reduction of the features are done and finally the features are used for classifying expressions by ANN.

II. PROPOSED METHOD

The steps of the proposed method are illustrated in Fig. 1. At first, a color or grayscale image is taken as input of the system. The image is converted to a grayscale one if it is a color image. Face region from the image is detected using Viola-Jones face detection method [19]. The detected face region is then resized into a fixed size for using it in the image segmentation step. At the segmentation step, the image is segmented into four expression regions (right eye, left eye, nose, mouth) by applying the proposed segmentation method which is highly effective. From the segmented parts, features are extracted by using 2D Gabor filter with different frequencies and orientations. The dimension of the feature vector is reduced at first by downsampling and then by PCA. Finally, at the classification step, the features of some images are used to train the system and feature of rest of the images are used to test the system for classifying the expressions. Details of each step mentioned in this section can be found in the following sections.

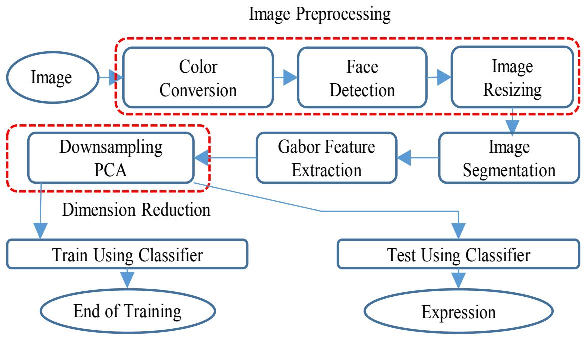

Fig. 1. Proposed Method of facial expression recognition.

III. IMAGE PREPROCESSING

The image preprocessing step is shown in Fig. 2 on a sample image from RaFD dataset [20].

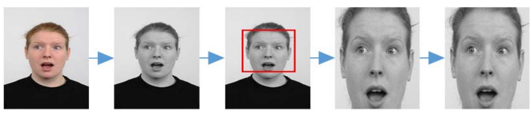

Fig. 2. Steps of image preprocessing.

At first, an input image is checked whether it is a color image or a gray level image. If the image is color then it is converted into the gray level image but if the input image is gray level then this step is omitted. Then face region is detected from the gray level image if it has any face in it and for doing this task the popular Viola-Jones face detection technique [19] is used. Later the detected face region is converted to a fixed size of 150 × 150 pixels for using it in the image segmentation step. These three steps are cumulatively named as image preprocessing.

IV. IMAGE SEGMENTATION

One of the most contributory and unique parts of this paper is this image segmentation step. In this step a preprocessed image is segmented into four facial expression regions such as right eye, left eye, nose and mouth. There are different available methods to do this task. Viola-Jones object detection method [19], Active Appearance Models (AAM) [21] are among the popular ones. But each have some problems to be handled. For example when the eyes are almost closed then Viola-Jones method fails to detect the eye region. So to avoid the problems and segment the four parts properly a manual segmentation method is proposed. According to this method, at first four specific coordinate points are defined for four regions. Then from the four coordinate points four regions with defined width and height are selected to segment the facial image into desired four parts. For understanding the process let us consider the segmentation of nose. To segment nose from image at first the coordinate point 54.33,81.84 is selected from the preprocessed image of size 150 × 150. Then from this point a region of width 45.43 and height 38 is selected and segmented and used as the nose expression region of an image. Same procedure is followed with different values for all the four parts of an image. The coordinate points and respective width, height for these four regions are given in TABLE I. These values are defined by analyzing many facial images and position of these four parts in these images.

TABLE I. VALUES FOR SEGMENTATION

Facial Part	Coordinate (x,y)	Width (w)	Height (h)
Right Eye	23.55,46.58	46.01	30.67
Left Eye	88.29,46.58	44.00	29.15
Nose	54.33,81.84	45.43	38.00
Mouth	50.24,114.0	57.00	34.18

Almost 4000 facial images of people from different race, ethnic groups and countries were analyzed to define these coordinate values. What really happens when these values are applied to an image of 150 × 150, is shown in Fig. 3.

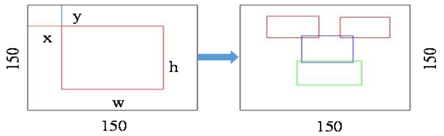

Fig. 3. Proposed image segmentation method on a block of size 150 × 150.

For a better understanding of the proposed image segmentation method, the method is applied to a sample image from RaFD dataset [20] and the step by step segmentation is illustrated in Fig. 4.

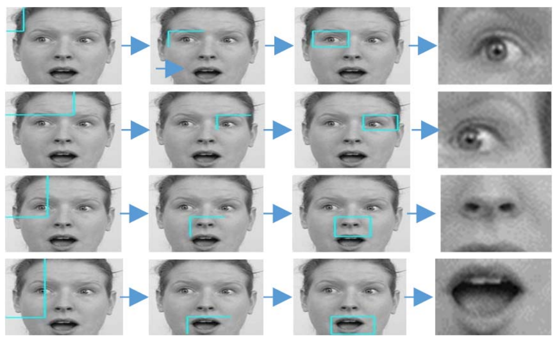

Fig. 4. Proposed image segmentation method (step by step) on a sample image.

The figure shows all the steps to segment all the four parts using the proposed method.

V. FEATURE EXTRACTION

Feature extraction is a crucial step because classifiers make a decision based on the features made available to them.

The proposed method uses 2D Gabor filter with different orientations and frequencies to extract useful features from the segmented parts. Some advantages [22] with Gabor filters are their invariance to rotation, scale, translation. They are also robust against photometric disturbances such as illumination changes, noise [22]. These reasons made us choose 2D Gabor filter as the feature extraction technique. A 2D Gabor filter is a Gaussian kernel function modulated by a complex sinusoidal plane wave [22], defined as:

$$G(x,y) = \frac{f^2}{\pi\gamma\eta}\exp(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2})\exp(j2\pi f x' + \phi)$$
$$x' = x\cos\theta + y\sin\theta \qquad (1)$$
$$y' = -x\sin\theta + y\cos\theta$$

Where θ is the orientation of the normal to the parallel stripes of a Gabor function, f is the frequency of the sinusoid, σ is the standard deviation of the Gaussian envelope, ϕ is the phase offset and γ is the spatial aspect ratio which specifies the ellipticity of the support of the Gabor function [22].

During the implementation, a Gabor filter bank with 5 scales and 8 orientations was generated using equation (1) with $\gamma = \eta = \sqrt{2}$, $f = 0.25, 0.1768, 0.125, 0.0884, 0.0625$, $\theta = 0$, $\frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}, \frac{\pi}{2}, \frac{5\pi}{8}, \frac{3\pi}{4}, \frac{7\pi}{8}$. The generated Gabor filter bank was then used to extract features from each of the segmented parts. As there was 40 filters in the filter bank so each segmented part was segmented 40 times. So if the size of any segmented part is $p \times q$ then after filtering its size gets $40 \times p \times q$. For our specific FER problem the size of the feature vector was 276960, after segmenting each four parts of an image 40 times. So it was the dimension of the extracted feature's feature vector per sample image. The dimension is so huge that it would create a great computational overhead for the classifier and be an obstacle for creating a real time system. So some kind of dimension reduction technique is needed to reduce the dimension of the feature vector to make the system computationally efficient.

VI. Dimension Reduction

Usually, adjacent pixels of an image are highly correlated. To reduce the dimension of feature vector the redundant information can be reduced by downsampling [22]. During the implementation, downsampling by a factor of 2 column wise and by a factor of 2 row-wise is used. After downsampling, the number of features per sample image was 69240 ($= 276960 \div (2 \times 2)$) which means dimension is reduced 75%. But still, a feature vector with 69240 features is huge.

To further reduce the dimension of the feature vector, PCA [23] is used which is famous for its ability to reduce dimensions by analyzing principal components. The steps of PCA are as follows:

- Normalization of the data
- Calculation of covariance matrix
- Calculation of eigenvectors, eigenvalues of the covariance matrix
- Choosing principal components and translating the data in terms of the components

During the implementation, PCA was used aiming at retaining 99% variance ratio which resulted in up to 99.72% dimension reduction. For example, while using images from JAFFE dataset [24] after applying PCA the dimension of feature vector was 191 which means dimension was reduced by 99.72% ($1 - (191 \div 69240)$). At this stage, a point worth mentioning is that although the dimension of each segmented region is different, they are same for all the images. The dimension of feature vector per sample image was also of the same dimension for all sample images. So when features are used in classification step, all sample images had the same number of features and as a result, there was no problem in comparing them for classification purpose.

VII. Classification

The importance of classification step cannot be denied. Despite performing all the previous steps properly, a system might not perform desirably just due to selection and application of improper classification method. As mentioned previously for the classification task artificial neural network (ANN) is used. To be more precise multilayer perceptrons (MLP) are used with backpropagation (for training). For weight optimization limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) method is used. It is a member of quasi-Newton methods family. Only a finite number of vectors are stored and iterative method is used to approximate objective function by the L-BFGS method. Considering weight vector at time t as $w^{(t)}$, tolerance as ε, $d^{(t)}$ as presented input at time t, $g^{(t)}$ as activation function output, then the algorithm for multilayer perceptrons with L-BFGS optimization can be summarized [25-26] as follows:

1. Initialize all weights $w^{(0)}$with small random values, set constant ε, set $t = t + 1$.
2. Present the network all training set and calculate $E(w^{(t)})$ as the error function for the current weight configuration.
3. If t=0
   then $d^{(t)} = -\nabla E^{(t)}$
   else $d^{(t)} = -\nabla E^{(t-1)} + Ap + Bv$
   where $p = w^{(t+1)} - w^{(t)}$, $v = g^{(t+1)} - g^{(t)}$
4. Calculate $w^{(t+1)} = w^{(t)} - \alpha d^{(t)}$,
   where $\alpha^{(t)} = -\frac{d^{(t)T} g^{(t)}}{d^{(t)T} H d^{(t)}}$
5. Calculate A and B for the next iteration as
   $A = -\left(1 + \frac{v^T v}{p^T v}\right)\frac{p^T g^{(t+1)}}{p^T v} + \frac{v^T g^{(t+1)}}{p^T v}$, $B = \frac{p^T g^{(t+1)}}{p^T v}$
6. If $E(w^{(t+1)}) > \varepsilon$ then $t = t + 1$ and goto 2, else stop.

During the implementation, the logistic sigmoid function was used as the activation function, as the stopping criteria of the algorithm a tolerance limit of 1e-8 was used, two hidden layers were used and each of them contained 50 nodes. The number of nodes in each of the layers were selected by trial and error. The number of hidden layers was chosen according to [27].

VIII. Result

For carrying out the implementation process, a computer with 4GB of RAM, Core i5 processor and 64-bit system was used. For evaluating the effectiveness of the proposed method, three widely used publicly available facial expression datasets JAFFE [24], CK+ [28], RaFD [20] were used. All 213 facial expression images from JAFFE dataset, 1219 images of 22 people from CK+ dataset and 1407 front facing facial expression images of 67 people were used. All seven facial expressions were considered but the contemptuous expression images were not considered from RaFD dataset for comparing the performance properly with other datasets and methods. The Correct Recognition Rate (CRR) or accuracy of the system was measured using K fold cross-validation to avoid biased results. Splitting into K folds were done randomly. TABLE II represents the accuracy achieved on different datasets using different folds.

TABLE II. ACCURACY OF DIFFERENT DATASETS

Dataset Used	No. of Folds	Proposed Method		Traditional Method	
		Avg.	*F.L.(no D.R.)*	*Avg.*	*F.L.(no D.R.)*
JAFFE	5	94.82		94.67	
JAFFE	10	94.81		94.98	
CK+	5	99.59	266800	99.63	900000
CK+	10	99.51		99.80	
RaFD	5	98.72		99.04	
RaFD	10	99.15		99.21	

The average (avg.) accuracy columns represent the average accuracy achieved using K folds. For example, when using 2 folds there are 2 splits namely 1, 2. Using 1 as training set, 2 as testing set and 2 as training set, 1 as testing set are the possible combinations. The average accuracy of these two combinations is represented in the average. The efficiency of the proposed image segmentation method is evident from TABLE II. F.L. (feature length without any dimension reduction) refers to the initial length of the extracted features without any dimension reduction. Traditional method refers to feature extraction from the full facial image without any dimension reduction. The size of the feature vector differs highly but the CRR differs slightly. So by considering CRR very few, the F.L. can be reduced a lot. Yes, dimension reduction techniques can be applied to reduce the dimension of feature vector but reducing 266800 features to a few hundred and reducing 900000 features to a few hundred is not computationally equivalent. The latter requires way more time and computations which is a hindrance in developing a real-time system. Fig. 5 is a comparison of achieved average CRR on different datasets using a different number of splits in the dataset.

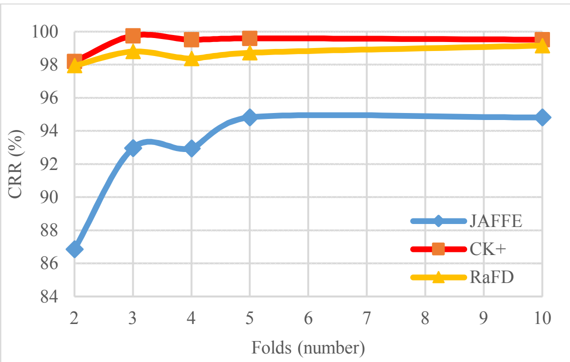

Fig. 5. Experimental results on JAFFE, CK+ and RaFD dataset.

For image processing problems like FER, accuracy is not the only metric to analyze the performance of a system. Confusion matrices are also analyzed to find out which expression images are classified properly and which are not. TABLE III is a confusion matrix on JAFFE dataset with 3 folds and a CRR of 94.25%. Sad expression image had the lowest CRR among all expressions. TABLE IV is for CK+ dataset with 99.75% CRR and lowest CRR with fear expression images. TABLE V is for RaFD dataset with a CRR of 99.57% and having lowest CRR with happy expression images.

TABLE III. CONFUSION MATRIX OF CRR ON JAFFE DATASET WITH 3 FOLDS

	Ne	Ha	An	Su	Fe	Di	Sa	Acc.
Ne	100	0	0	0	0	0	0	100
Ha	0	100	0	0	0	0	0	100
An	12.50	0	87.50	0	0	0	0	87.50
Su	0	0	0	100	0	0	0	100
Fe	0	0	0	0	100	0	0	100
Di	0	0	11.11	0	0	88.89	0	88.89
Sa	8.33	0	0	0	0	8.33	83.33	83.33

TABLE IV. CONFUSION MATRIX OF CRR ON CK+ DATASET WITH 3 FOLDS

	Ne	Ha	An	Su	Fe	Di	Sa	Acc.
Ne	100	0	0	0	0	0	0	100
Ha	0	100	0	0	0	0	0	100
An	0	0	100	0	0	0	0	100
Su	0	0	0	100	0	0	0	100
Fe	0	1.79	0	0	98.21	0	0	98.21
Di	0	0	0	0	0	100	0	100
Sa	0	0	0	0	0	0	100	100

TABLE V. CONFUSION MATRIX OF CRR ON RAFD DATASET WITH 3 FOLDS

	Ne	Ha	An	Su	Fe	Di	Sa	Acc.
Ne	100	0	0	0	0	0	0	100
Ha	0	98.41	0	0	0	1.59	0	98.41
An	0	0	100	0	0	0	0	100
Su	0	0	0	100	0	0	0	100
Fe	0	0	1.43	0	98.57	0	0	98.57
Di	0	0	0	0	0	100	0	100
Sa	0	0	0	0	0	0	100	100

IX. STATE-OF-THE-ART

From the result section, it is apparent that the proposed method has a high CRR. But how this performance is compared to other methods can be known by comparing this performance with other method's performance on the same dataset.

TABLE VI. STATE-OF-THE-ART COMPARISON

Study	Technique	Dataset	Acc.
Proposed Method	Viola-Jones face detection, **facial region segmentation**, Gabor filter, PCA, ANN, MLP, backpropagation	JAFFE CK+ RaFD	**94.81** **99.51** **99.15**
2017 [16]	Viola-Jones face detection, Gabor filter, ANN	JAFFE	85.70
2017 [17]	Appearance based models, shape signature, grid, MLP	JAFFE CK+	92.00 97.90
2015 [18]	Gabor filter, Radial Bases Functions Networks (RBFN), ANN	JAFFE CK+	92.88 92.40
2015 [29]	SURF, gentle AdaBoost	RaFD	90.64
2017 [30]	Visual saliency,deep learning	RaFD	95.70

X. CONCLUSION

Result and state-of-the-art comparison section describe the effectiveness of the proposed method. Unique and highly effective facial region segmentation technique along with proper feature extractor and classifier makes the system capable of recognizing facial expressions quite accurately. Some facial expression images are so subtle that they are tough to recognize even for human let alone machines. A controversy of the proposed method might be its image segmentation method. Even at first, the proposed segmentation method might seem to be absurd. But the method was tested with around different type of 3000 standard front-facing images and the method performed pretty well. But if the facial structure of any person is unusual from any usual person's facial structure then the system might be unable to segment the expression regions properly because the method was developed aiming at handling standard images.

A noticeable flaw of the proposed system, for the time being, is its ability to handle only front-facing images for recognizing expressions. But the system was developed aiming at recognizing expressions of front-facing images so this flaw is not creating any problem as of now. But developing a more robust system capable of handling images rotated at any angle would be a future work. The ultimate purpose of the work is to recognize human emotions. But recognizing facial expression is not the only to do that. So developing a multimodal information fusion system combining facial expressions, vocal information, and textual messages to recognize human emotions would be a challenging future work.

REFERENCES

[1] A. Mehrabian, "Communication without Words", *Psychology Today*, Vol. 2, No. 4, pp.53-56, 1968.

[2] G. Sandbach, S. Zafeiriou, M. Pantic and L. Yin, "Static and Dynamic 3D Facial Expression Recognition, A Comprehensive Survey", *Image Vision Compute*, Vol. 30, No. 10, pp. 683-697, 2012.

[3] F. Mahmud, M. T. Khatun, S. T. Zuhori, S. Afroge, M. Aktar and B. Pal, "Face recognition using Principle Component Analysis and Linear Discriminant Analysis", *2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, Dhaka, pp. 1-4, 2015.

[4] F. Mahmud, M. E. Haque, S. T. Zuhori and B. Pal, "Human face recognition using PCA based Genetic Algorithm", *2014 International Conference on Electrical Engineering and Information & Communication Technology*, Dhaka, pp. 1-5, 2014.

[5] F. Mahmud, S. Afroge, M. A. Mamun and A. Matin, "PCA and back-propagation neural network based face recognition system", *2015 18th International Conference on Computer and Information Technology (ICCIT),* Dhaka, pp. 582-587, 2015.

[6] A. Matin, F. Mahmud and M. T. B. Shawkat, "Recognition of an individual using the unique features of human face", *2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, Pune, India, pp. 57-60, 2016.

[7] A. Matin, F. Mahmud, T. Ahmed and M. S. Ejaz, "Weighted score level fusion of iris and face to identify an individual", *2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Cox's Bazar, pp. 1-4, 2017.

[8] A. Matin, F. Mahmud, S. T. Zuhori and B. Sen, "Human iris as a biometric for identity verification", *2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)*, Rajshahi, pp. 1-4, 2016.

[9] B. Gite, K. Nikhal and F. Palnak, "Evaluating facial expressions in real time," *2017 Intelligent Systems Conference (IntelliSys)*, London, pp. 849-855, 2017.

[10] Borui Zhang, Guangyuan Liu and Guoqiang Xie, "Facial expression recognition using LBP and LPQ based on Gabor wavelet transform," *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, Chengdu, pp. 365-369, 2016.

[11] A. S. Md. Sohail and P. Bhattacharya, "Classification of facial expressions using K-nearest neighbor classifier", *MIRAGE'07 Proceedings of the 3rd international conference on Computer vision/computer graphics collaboration techniques*, Rocquencourt, pp. 555-566 ,2007.

[12] Z, -T. Liu, G. –T. Sui, D. –Y. Li and G. –Z. Tan, "A Novel Facial Expression Recognition Based on Extreme Learning Machine", *34th Chinese Control Conference (COC)*, pp. 3852-3857, 2015.

[13] Y. Xing and W. Luo, "Facial expression recognition using local Gabor features and adaboost classifiers," *2016 International Conference on Progress in Informatics and Computing (PIC)*, Shanghai, pp. 228-232, 2016.

[14] Ju Jia, Yan Xu, Sida Zhang and Xianglong Xue, "The facial expression recognition method of random forest based on improved PCA extracting feature," *2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, Hong Kong, pp. 1-5, 2016.

[15] F. Z. Salmam, A. Madani and M. Kissi, "Facial Expression Recognition Using Decision Trees," *2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV)*, Beni Mellal, pp. 125-130, 2016.

[16] K. Verma and A. Khunteta, "Facial expression recognition using Gabor filter and multi-layer artificial neural network," *2017 International Conference on Information, Communication, Instrumentation and Control (ICICIC)*, Indore, pp. 1-5, 2017.

[17] A. Barman and P. Dutta, "Facial expression recognition using shape signature feature," *2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, Kolkata, pp. 174-179, 2017.

[18] R. Saabni, "Facial expression recognition using multi Radial Bases Function Networks and 2-D Gabor filters," *2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC)*, Sierre, pp. 225-230, 2015.

[19] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, pp. I-511-I-518, Vol.1, 2001.

[20] O. Langner, R. Dotsch, G. Bijlstra, D.H. Wigboldus, S.T. Hawk and A. van Knippenberg, "Presentation and validation of the Radboud Faces Database", *Cognition and Emotion*, Vol. 24, pp. 1377-1388, 2010.

[21] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active Appearance Models, " *IEEE Trans on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 6, pp. 681-685, 2001.

[22] M. Haghighat, S. Zonouz, M. A. Mottaleb, "CloudID: Trustworthy cloud-based and cross-enterprise biometric identification," *Expert Systems with Applications*, Vol. 42, No. 21, pp. 7905-7916, 2015.

[23] L. I. Smith, "A tutorial on principal components analysis," Cornell University, USA, 2002.

[24] M. J. Lyons, S. Akemastu, M. Kamachi and J. Gyoba, "Coding Facial Expressions with Gabor Wavelets", *3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 200-205, 1998.

[25] M. Brescia, *Multi Layer Perceptron Trained by Quasi Newton Learning Rule*, Naples, 2012 [Online]. Available: http://dame.dsf.unina.it/documents/DAME_MLPQNA_Model_Mathematics.pdf. [Accessed: 07- Jun- 2018].

[26] Brescia et al., *Publications of the Astronomical Society of the Pacific*, Vol. 126, No. 942, pp. 783-797, 2014.

[27] V. Kůrková, "Kolmogorov's theorem and multilayer neural networks", *Neural Networks*, Vol. 5, No. 3, pp. 501-506, 1992.

[28] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, San Francisco, CA, June 13-18, pp. 94-101, 2010.

[29] Q. Rao, X. Qu, Q. Mao and Y. Zhan, "Multi-pose facial expression recognition based on SURF boosting," *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, Xi'an, pp. 630-635, 2015.

[30] V. Mavani, S. Raman and K. P. Miyapuram, "Facial Expression Recognition Using Visual Saliency and Deep Learning," *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, Venice, pp. 2783-2788, 2017.