

Facial Emotion Recognition Using NLPCA and SVM

Chirra Venkata Rami Reddy^{1,2*}, Uyyala Srinivasulu Reddy³, Kolli Venkata Krishna Kishore¹

¹ Department of Computer Science & Engineering, VFSTR, Guntur, India

² Research Scholar, Department of Computer Applications, national institute of technology, Tiruchirappalli, India

³ Machine Learning & Data Analytics Lab, Department of Computer Applications, national institute of technology, Tiruchirappalli, India

Corresponding Author Email: chvrr58@gmail.com

<https://doi.org/10.18280/ts.360102>

Received: 15 November 2018

Accepted: 7 January 2019

Keywords:

Gabor wavelet, Haar wavelet, PCA, NLPCA, SVM

ABSTRACT

The aim of this present work is to achieve better accuracy of facial emotion recognition and classification with limited training samples under varying illumination. A method (involving two versions) for achieving high accuracy with limited samples is proposed. Global and local features of facial expression images were extracted using Haar Wavelet Transform (HWT) and Gabor wavelets respectively. Dimensionalities of extracted features are reduced using Nonlinear principal component analysis (NLPCA). Concatenated and weighted fusion techniques have been employed for fusing the global and local features. To recognize and classify six emotions (joy, surprise, fear, disgust, anger, and sadness) from facial expressions a Support Vector Machine was used. The proposed method is evaluated on Extended Cohn-Kanade dataset. The average recognition rates of 97.3 % and 98 % are achieved with the two versions of the proposed method, providing better recognition accuracy compared with the existing methods.

1. INTRODUCTION

Facial expressions are very useful for expressing emotions and intentions and primarily for interacting with other people. New facial expression recognition systems have been implemented in several fields, including psychology, computer graphics, consumer neuroscience, media testing & advertisement, psychotherapy, medicine, and transportation security. Psychologists were the first among other people who realized and investigated the importance of facial expressions and the need for recognizing emotions through computer applications.

In this context, Mehrabian [1] explained that the communication of feelings involves conveying 55 % of the total message through facial expressions, 38% through para language like intonations, and only 7 % through linguistic languages. In the late 70s, Ekman [2] conducted the initial studies on facial expressions and demonstrated facial expressions like happiness, fear, anger, sadness, disgust, surprise, and neutrality through his experiments. These seven expressions comprise the basic facial expressions of human beings.

In 2013, Du [3] presented 21 expressions called compound expressions, which include 6 basic expressions and 15 derived from the basic expressions. According to him, compound expressions include neutrality, anger, sadness, happiness, fear, disgust, surprise, happy and surprised, sad and fearful, happy and disgusted, sad and angry, sad and disgusted, sad and surprised, fearful and surprised, fearful and angry, fearful and disgusted, angry and surprised, angry and disgusted, disgusted and surprised and hatred. In general, facial expression classification comprises of 3 steps: To detect faces, to extract features, and recognize facial expressions [4]. Face will be

detected using a face detector in the first step. Facial features are extracted from either static images or frames in a video in second step and a feature vector is formed. In the third step, a classifier classifies the expressions.

For extracting features geometric-based extraction methods [5] and appearance-based extraction methods [6] are used. The former methods extract location and shape-related metrics from the eyes, eyebrows, mouth, and nose. Conversely, appearance-based feature extraction methods extract skin appearance or facial features. Facial features can be extracted from entire face or specific regions in a face by these methods. Feature extraction methods in facial expression classification systems generate high-dimensional data. Dimensionality reductions through linear or nonlinear methods are applied to reduce feature dimensions.

Appearance-based methods such as Gabor wavelet [7], LBP [8], pyramid Local Binary Pattern, LDA [6] or ICA [6], PCA [6], project facial images into a lower dimensional subspace. Geometric-based techniques include the multi-resolution active shape model, canny edge detection, and AAM.

High-accuracy facial expression recognition can be achieved through vigorous feature extraction techniques. Most classifiers fail to achieve accurate recognition rates with insufficient features. The extraction of unique features that represent facial expressions is critical to enhancing the accuracy of feature extraction techniques. In automatic facial expression recognition and classification systems, classifiers play a crucial role. The classifier recognizes as well as classifies facial expressions.

There are several classification algorithms such as SVM, least-square method, two nearest neighbor, random forest, decision tree, Naive Bayes, distance ratio-based classifier, and the Hidden Markov Model. Today, most researchers use deep

learning as well as ML algorithms for extraction and classification of features. For efficient facial expression recognition and classification, deep neural networks [9] are also used.

In this study, appearance-based facial features are extracted using the Gabor wavelet and Haar wavelet feature extraction methods. Gabor wavelets exhibit high-discriminative power. Dimensionality of features will be reduced by NLPCA. SVM act like a classifier to classify the expressions.

Section 2 explains the literature survey whereas section 3 describes about proposed method. Feature extraction techniques are presented in section 4. In section 5, dimensionality reduction technique has been explained whereas section 6 deals with fusion techniques. Section 7 describes the classifier and the proposed algorithm is described in section 8. Section 9 describes experimental results as well as conclusion.

2. LITERATURE SURVEY

There have been considerable studies on expression classification of facial images; however, results have been unsatisfactory under varied illumination, pose variation, and occlusions. Deng et al. [10] used a Gabor filter for extracting features and implemented PCA and LDA for feature reduction. Experiments performed on the JAFFE database yielded 97% accuracy.

Praseeda Lekshmi et al. [11] implemented an RBF neural network to classify four facial expressions using the JAFFE database. Kishore et al. [12] proposed a method which is constructed from the fusion of facial features. Gabor wavelets and DCT were implemented to retrieve the features and PCA was implemented to reduce dimensionality. The Hybrid Emotional Neural Network was used as a classifier to classify emotions. Experiments performed on the Cohn-Kanade (CK) dataset yielded 98.5% of Recognition rate.

For recognizing the facial expressions and classification a feature learning-based pixel difference representation was proposed by Sun et al. [13]. A DFM called as discriminative feature matrix was first generated from original images and a discriminative feature dictionary was constructed. DFD served as the input to vertical 2D linear discriminate analysis in the direction (V-2DLDA). To classify the images, a nearest neighbor classifier was used. The proposed method was evaluated on a CK+ dataset and yielded an accuracy of 91.87 %.

Uddin et al. [9] stated a depth camera-based novel approach. Local directional rank histogram pattern (LDRHP) and local directional pattern (LDP) was used to retrieve the features. KPCA was used for dimensionality reduction. Finally, the convolutional neural network was used for classification.

RamiReddy et al. [8] proposed a method which is constructed from the combination of local and global features. DLBP and DCT were used to get global and local features from non-dynamic face expression images respectively. For classification RBF neural network was used. Evaluation was performed using the Cohn-Kanade dataset, which yielded an accuracy of 97 %.

Kumar et al. [14] proposed a framework depending on getting the instructive regions of a face image. In the feature extraction stage, LBP features were extracted and then Procrustes analysis was used to model the reference image. Features were finally extracted only from selected regions. An

SVM was used as a classifier to classify the expressions. Experiments performed on the MUG database yielded an accuracy of 98 %.

To recognize and classify facial expressions in video sequences, Kamarol et al. [15] proposed an approach based on appearance features. At first step, Viola-Jones face detector was used for detecting faces from frames. The Spatio temporal texture map (STTM) algorithm was then used to model the facial features. SVM classifier was implemented to classify the features. The accuracy achieved was approximately 95.37 on the CK+ database.

Siddiqi et al. [16] stated an approach based on localized features. SWLDA was implemented to find the features from images which contain expressions of a face. SWLDA used partial F-test values to choose localized features from the frames. An HCRF model was used for recognition. Four datasets were used for performance evaluation of the proposed approach and yielded an accuracy of 96.37 %.

Qayyum et al. [17] Proposed an approach based on a stationary wavelet transform. Features were extracted using the stationary wavelet transform. The horizontal and vertical sub-bands of a stationary wavelet transform contain muscle movement information of facial expressions. DCT was applied on these sub-bands for feature dimensionality. A Feed Forward Neural Network (FFNN) with back propagation algorithm was implemented to recognize and classify the expressions. Average accuracies of 96.61% and 98.83% were achieved on extended Cohn-Kanade and JAFFE datasets, respectively.

Ding et al. [18] stated a method for recognizing and classifying facial expressions in videos. A peak expression frame from the video was detected using DLBP. Logarithm-Laplace (LL) domain was used to obtain accurate facial features. The Taylor expansion theorem was applied to extract the features and a Taylor feature map was created, which was given as an input to Taylor Feature Pattern (TFP) for extracting accurate features. CK and JAFFE datasets were used for testing.

3. PROPOSED SYSTEM

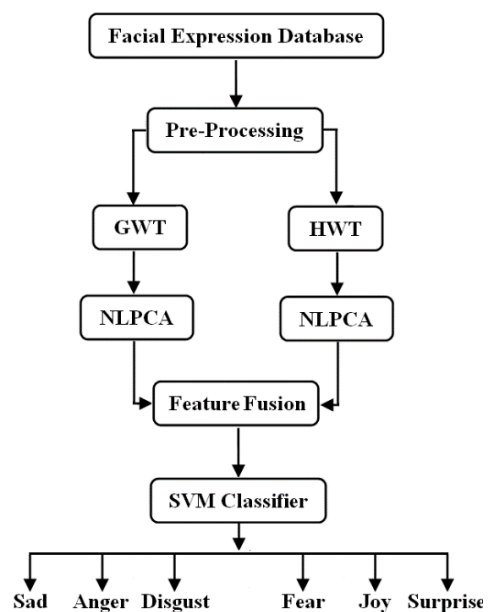


Figure 1. Architecture of the Proposed method

The architecture of facial expression recognition and classification system is outlined in Figure 1. Global and local features are extracted using Haar wavelets and Gabor filters respectively. Because the features extracted are high in dimensionality, the required time complexity for training the model is higher. Dimensionality reduction is typically achieved through PCA. However, in the proposed method, NLPCA is implemented for dimensionality reduction to improve the performance of the model. PCA detects linear correlations in a feature set, whereas NLPCA finds both linear and nonlinear correlations. Information content is equally distributed among the components in NLPCA. The reduced features from both extraction methods are fused to improve accuracy.

4. FEATURE EXTRACTION

4.1 Haar wavelet

The Haar wavelet is a transform that includes multi-layer decomposition. The Haar wavelet transform is implemented for the extraction of global attribute from a face image by decomposing image of face into wavelet coefficients. Decomposed wavelet coefficients consist of four portions, each with a quarter of the authentic image area as presented in Figure 2. LA (All Components), LH (Horizontal Component), LV (Vertical Components) and LD (Diagonal Component). We have used LA coefficients to obtain features from facial expression images.

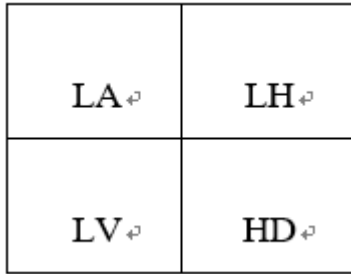


Figure 2. Haar wavelet decomposition

4.2 Gabor wavelet

The extraction of local features like eyes, mouth, cheeks, and nose, involves finding attributes in a localized manner. Local attributes can be pulling out using Gabor wavelets [7] in frequency and spatial domain. The Gabor wavelet transformation is represented by

$$G_{o,s}(x,y) = (o^2 + s^2) \div (2\sigma^2) * \ell^{-((o^2+s^2)*(x^2+y^2)) \div (2\sigma^2)} * [\ell^{i*(ox+sy)} - \ell^{(-\sigma^2 \div 2)}] \quad (1)$$

where o is the orientation and s is the scale, which are obtained by the following equations.

$$o = k_m \div f^n * \cos(pi * m \div 8) \quad (2)$$

$$s = k_m \div f^n * \cos(pi * m \div 8) \quad (3)$$

$$k_m = pi \div 2, f = \sqrt{2} \quad (4)$$

Here, $f = \sqrt{2}$ is the spacing factor between the kernels and k_m is the frequency domain.

The Figure 3 shows the real part of the Gabor kernels at 4 scales and 5 orientations and their magnitude.

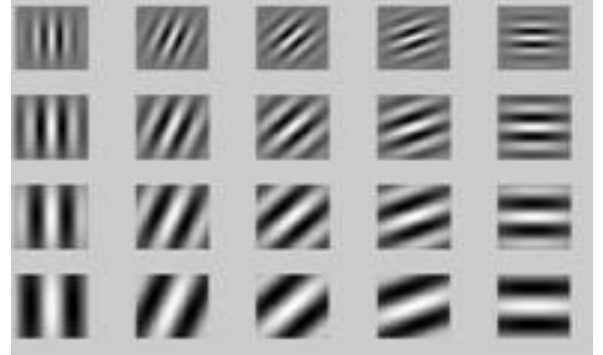


Figure 3. Real part of the Gabor kernels

In the proposed method, Gabor wavelets with five orientations and four scales are used. Gabor kernels are convolved with the face image to extract facial features using the below relationship.

$$O_{m,n} = I(x,y) * G_{u,v}(x,y) \quad (5)$$

$I(x,y)$ is a gray-level face image, and $G_{u,v}(x,y)$ is a Gabor kernel.

The Figure 4 shows the Gabor wavelet representation of an image.



Figure 4. Gabor wavelet representation of an image

By concatenating all features using orientation selectivity, spatial frequencies, and spatial localities, a feature vector is generated. Because Gabor kernels are complex representations, a high-dimensionality feature vector is generated.

5. DIMENSIONALITY REDUCTION

5.1 NLPCA

The feature vectors that are generated using Gabor and Haar wavelets are high in dimensionality so consequently it requires more training time. Therefore, transformation of higher

dimensional feature vectors to a lower dimensional sub-space using a nonlinear transformation function is required. The nonlinear generalization of PCA is a NLPCA [19]. Therefore NLPCA is implemented to decrease the dimensionality of feature vector. It projects the principal components from straight lines to nonlinear curves. Nonlinear principal components can be computed using a neural network referred to as an auto-encoder.

The edifice of the NLPCA network model is displayed in Figure 5. Neurons in the input layer and the output layer are equivalent to optimum features selected from Gabor wavelets. Hidden neurons are placed to the left and the right of the output layer. The bottleneck layer is an encoding layer. A nonlinear function maps the high-dimensional input space with the K-dimensional bottleneck space. Reverse transform mapping is performed from the bottleneck space to the original space represented by the outputs \hat{x} to ensure minimum error. Five-layer architecture with a node ratio of 1: 20 in terms of input layer dimensions is used in the compression layer to obtain an optimum feature number (dimensions).

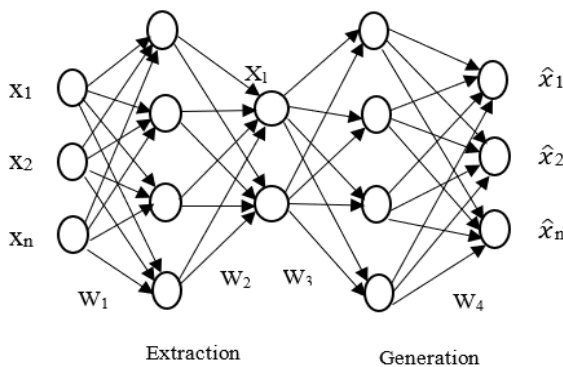


Figure 5. Nonlinear principal component analysis neural network

6. FUSION

6.1 Concatenated fusion

In this phase, the features extracted from both methods are concatenated for increasing the accuracy of the model. Normalization is implemented so that features share a common scale.

$$S = (S_1 + S_2) \quad (6)$$

6.2 Weighted fusion

The weighted summation method is applied to combine or fuse normalized features. Weights are set based on feature ranking. Feature ranking is computed by determining the variance between a feature vector and the mean of the feature set. Based on this view, weights are assigned to the features. The weighted summation is given as

$$S = \sum_{n=0}^N w_n s_n \quad (7)$$

7. SUPPORT VECTOR MACHINE CLASSIFIER

A SVM is a ML algorithm that is simple to implement and provides good generalization performance, and with a little tuning, the same algorithm can solve a variety of problems. The SVM starts training on two sets of vectors in an n-dimensional space and determines the hyper plane that maximizes the margin between two close points in the training set, which are known as support vectors. Further calculations involve only support vectors.

The kernel function does operations in an input space, but not in attribute space, which exhibits higher dimensionality. The kernel function maps attributes of the input space to the attribute space which decreases computational complexity.

A legitimate inner product in the attribute space is represented by a kernel. Input space does not separate training set linearly, but in the case of the attribute space, the training set will be linearly separable. This is known as the “Kernel trick” Proposed by Shenetal [14-15].

The following are the various SVM kernels:

- (1). Gaussian radial basis function:

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad (8)$$

- (2). Exponential radial basis function:

$$K(x, x') = \exp\left(-\frac{\|x - x'\|}{2\sigma^2}\right) \quad (9)$$

8. PROPOSED SYSTEMALGORITHM

In this study, four methods are proposed, namely PCA_SVM with concatenated fusion (Proposed method-1 with CF), PCA_SVM with weighted fusion (Proposed method-1 with WF), NLPCA-SVM with concatenated fusion (Proposed method-2 with CF), and NLPCA-SVM with weighted fusion (Proposed method-2 with WF). In all the methods, Gabor and Haar wavelets are used to extract global features as well as local features from facial expression images and an SVM is used as a classifier to recognize six emotions.

In the PCA_SVM with a concatenated fusion method, PCA was implemented to decrease the dimensionality of features. A concatenated fusion technique was implemented for fusing global and local features.

In the PCA_SVM with a weighted fusion method, PCA was implemented to decrease the dimensionality of features. A weighted fusion technique was implemented for fusing global and local features.

In the NLPCA_SVM with a concatenated fusion method, NLPCA was implemented to decrease the dimensionality of features. A concatenated fusion technique was implemented for fusing global and local features.

In the NLPCA_SVM with a weighted fusion method, NLPCA was implemented to decrease the dimensionality of features. A weighted fusion technique was implemented for fusing global and local features.

1. Training samples: <classes, expressions>I(x, y)=I
<no of classes, no of samples, 6 expressions>(x, y)

2. (i) Let Gabor wavelet transform is $G_{u,v}(x, y)$
(ii) $O_{m,n} = I(x, y) * G_{u,v}(x, y)$
(iii) Feature set: $X_i = O_{m,n}$
3. **Case I:** Feature extraction using PCA

$$(i) \bar{x} = \frac{1}{M} \sum_{i=1}^M x_i$$

$$(ii) \omega_i = x_i - \bar{x}$$

$$(iii) c = \frac{1}{M} \sum_{n=1}^M \omega_n \omega_n^T = BB^T$$

$$(iv) C: u_1, u_2, \dots, u_N$$

$$(v) \text{Feature vector } V1 = (u_1, u_2, \dots, u_N)$$

Case II: Feature extraction using NLPKA

$$(i) (\text{pc, net, network}) = \text{nlpca}(X_i, n_c)$$

$$(ii) \text{pc_new} = \text{nlpca_get_components}(\text{net}, V1)$$

$$(iii) V1 = \text{nlpca_get_data}(\text{net}, \text{pc_new})$$

Here n_c is no of nonlinear components from dataset

4. Let X_i is a feature set
 $X_i = H_n[I(x, y)]$;
where H_n is Haar Wavelet transform
5. Repeat step 3 on the feature set which is generated in step 4 and form feature vector $V2$.
6. **Case I:**
(i) Train set= $V1+V2$
(ii) Model= svmtrain (Train set, Group);
(iii) Repeat step 1 to step5 on test sample: $T(x, y)$
(iv) svmclassify(model, $T(x, y)$)
Case II:

- (i) Train set= $S = \sum_{n=0}^N w_n s_n$
- (ii) Model=svmtrain(Train set, Group);
- (iii) Repeat step 1 to step5 on test sample: $T(x, y)$
- (iv) svmclassify(model, $T(x, y)$)

9. EXPERIMENTAL RESULT

The Extended Cohn-Kanade (CK+) Facial Emotion Database is considered for evaluating the proposed methods. For validating the four proposed methods (two variations of each version), three datasets are created from the CK+ database with 900, 1200, and 1500 samples. Dataset-I contains 900 facial expression images that consist of six facial expressions taken from 15 people. Out of 900 face images, 450 are reserved for training and 450 are reserved for testing. Dataset-II contains 1200 facial expression images consisting of six facial expressions obtained from 20 people. Out of the 1200 images available, 600 are reserved for training and 600 are reserved for testing. Dataset-III contains 1500 facial expression images consisting of six facial expressions obtained from 25 people. Out of the 1500 face images, 750 are reserved for training and 750 are reserved for testing. The accuracy rates of all four proposed methods PCA_SVM with concatenated fusion (Proposed method-1 with CF), PCA_SVM with weighted fusion (Proposed method-1 with WF), NLPKA-SVM with concatenated fusion (Proposed method-2 with CF), and NLPKA-SVM with weighted fusion (Proposed method-2 with WF) are evaluated on the three datasets as given in Table 1

Table 1. Accuracy of proposed methods

Proposed Methods / Datasets	Dataset-I	Dataset-II	Dataset-III
Proposed method-1 with CF	97.33	97.16	96.4
Proposed method-1 with WF	98	97.5	96.53
Proposed method-2 with CF	97.33	97.16	96
Proposed method-2 with WF	98	97.66	96.93

Table 2. Performance evolution of each method on Dataset-I

training samples	testing samples	Proposed Method-1 with CF	Proposed Method-1 with WF	Proposed Method-2 with CF	Proposed Method-2 with WF
540	360	98.61	98.61	98.6	98.89
450	450	97.33	98	97.33	98
360	540	93.51	96.85	95.74	96.85

Table 3. Performance evolution of each method on Dataset-II

training samples	testing samples	Proposed Method-1 with CF	Proposed Method-1 with WF	Proposed Method-2 with CF	Proposed Method-2 with WF
720	480	98.96	98.5	98.33	98.95
600	600	97.16	97.5	97.16	97.66
480	720	93.61	95.41	95.14	95.83

Table 4. Performance evolution of each method on dataset-III

training samples	testing samples	Proposed Method-1 with CF	Proposed Method-1 with WF	Proposed Method-2 with CF	Proposed Method-2 with WF
900	600	98	98.33	98.16	98.33
750	750	96.4	96.53	96	96.93
600	900	94.22	95.33	95.33	95.44

Table 5. Accuracy rate of each expression for the Proposed methods on Dataset-I

Proposed Method/Expression	Surprise	Fear	Sad	Anger	Disgust	Joy
Proposed Method-1with CF	90.66	100	98.7	96	100	100
Proposed Method-1with WF	90.66	100	98.6	98.6	100	100
Proposed Method-2 with CF	89.33	100	98.7	98.6	97.33	100
Proposed Method-2 with WF	90.66	100	98.7	98.6	100	100

Table 6. Accuracy rate of each expression for the Proposed methods on Dataset-II

Proposed Method/Expression	Surprise	Fear	Sad	Anger	Disgust	Joy
Proposed Method-1with CF	91	99	98	95	100	100
Proposed Method-1with WF	91	100	99	96	99	100
Proposed Method-2 with CF	89	100	99	96	99	100
Proposed Method-2 with WF	91	100	98	97	100	100

Table 7. Accuracy rate of each expression for the proposed methods on Dataset-III

Proposed Method/Expression	Surprise	Fear	Sad	Anger	Disgust	Joy
Proposed Method-1 with CF	92	96.8	91.2	99.2	99	100
Proposed Method-1 with WF	90.83	96.66	95.83	92.59	96.66	100
Proposed Method-2 with CF	93.6	96.8	89.6	96	100	100
Proposed Method-2 with WF	91.66	97.5	95.83	93.33	96.66	100

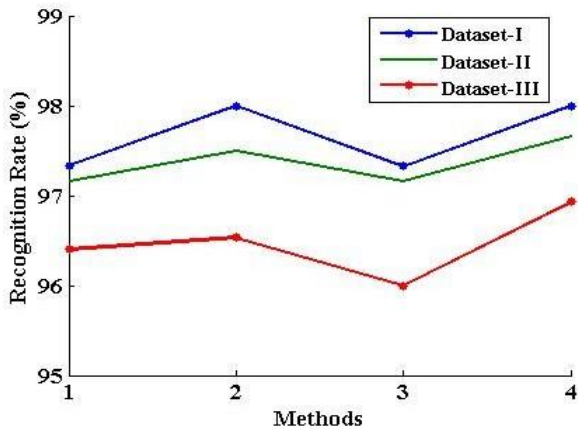


Figure 6. Comparison of recognition rates of four methods on Dataset-I, Dataset-II and Dataset-III

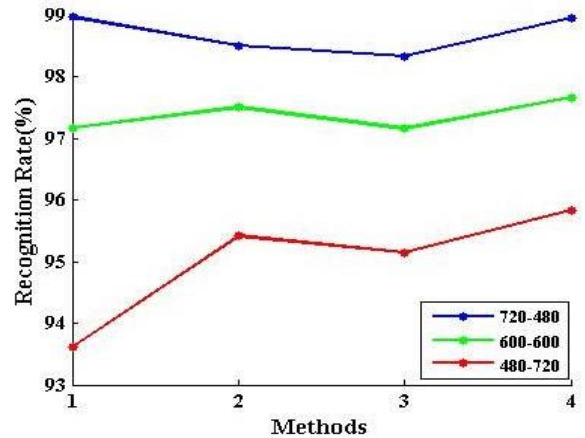


Figure 8. Comparison of recognition rates of the proposed methods on Dataset-II with different train and test samples

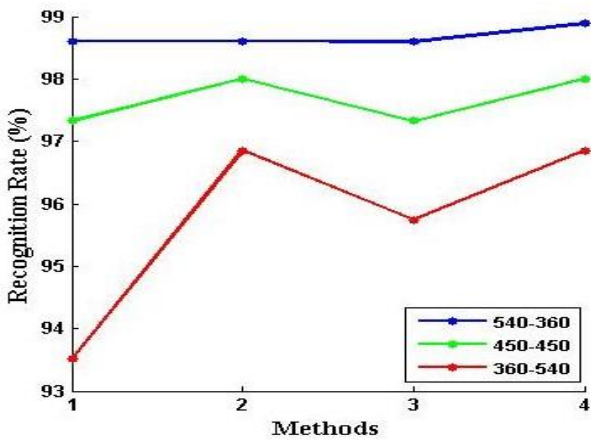


Figure 7. Comparison of recognition rates of four methods on Dataset-I with different train and test samples

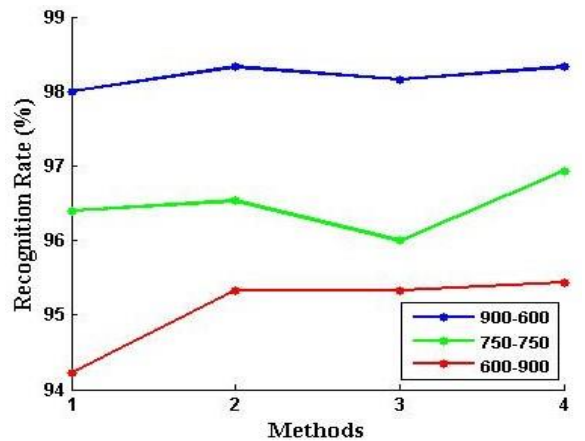


Figure 9. Comparison of recognition rates of the proposed methods on Dataset-III with different train and test samples

Table 8. Confusion matrix of proposed method-1 with CF on Dataset-I (Train samples=360, Test samples=540)

	Fear	surprise	Disgust	Joy	Anger	Sad
Fear	97.8			2.2		
Surprise	7.7	86.7		3.3	2.2	
Disgust			94			6
Joy	2.2			100		
Anger		1.1	1.1	2.2	88.9	4.4
Sad				6.7		93.3

Table 9. Proposed method-1 with CF confused matrix on Dataset-II (Train samples=480, Test samples=720)

	Fear	surprise	Disgust	Joy	Anger	Sad
Fear	97.5		2.5			
Surprise	7.5	87.5	0.8		2.5	1.7
Disgust			94.2		0.8	5
Joy	1.7			98.3		
Anger		0.8	0.8	2.5	90.9	5
Sad				6.7		93.3

Table 10. Proposed Method-1 with CF confused matrix on Dataset-III (Train samples=600, Test samples=900)

	Fear	surprise	Disgust	Joy	Anger	Sad
Fear	96		2.66	0.66	0.66	
Surprise	4.7	89.3	4	0.7		1.3
Disgust			98		2	
Joy	1.3			98.7		
Anger			1.3	0.7	96.7	1.3
Sad	1.3		2	8.7	1.3	86.7

Table 11. Proposed Method-1 with WF confused matrix on Dataset-I (Train samples=360, Test samples=540)

	Fear	surprise	Disgust	Joy	Anger	Sad
Fear	97.7			2.3		
Surprise	2.3	90	3.3	1.1		3.3
Disgust			100			
Joy				100		
Anger			1.1	2.2	96.7	
Sad				3.3		96.7

Table 12. Proposed Method-1 with WF confused matrix on Dataset-II (Train samples=480, Test samples=720)

	Fear	surprise	Disgust	Joy	Anger	Sad
Fear	96.7	0.8		2.5		
Surprise		90.8	2.5	5		1.7
Disgust			96.7		3.3	
Joy				100		
Anger			0.8	4.2	92.5	2.5
Sad				4.2		95.8

Table 13. Confusion matrix of Proposed Method-1 with WF on Dataset-III (Train samples=600, Test samples=900)

	Fear	surprise	Disgust	Joy	Anger	Sad
Fear	96	1.3	0.7	1.3	0.7	
Surprise	0.7	94	1.3	2		2
Disgust			96.7		2.6	0.7
Joy				100		
Anger			0.7	2	97.3	
Sad	4		2	4.7	1.3	88

Table 14. Confusion matrix of Proposed Method-2 with CF on Dataset-I (Train samples=360, Test samples=540)

	Fear	Surprise	Disgust	Joy	Anger	Sad
Fear	96.7			2.2	1.1	
Surprise	2.22	88.9	3.33	3.33		2.22
Disgust			97.8		2.2	
Joy				100		
Anger		1.1	1.1	4.44	93.33	
Sad				2.22		97.8

Table 15. Confusion matrix of Proposed Method-2with CF on Dataset-II (Train samples=480, Test samples=720)

	Fear	Surprise	Disgust	Joy	Anger	Sad
Fear	96.7	2.5			0.8	
Surprise		89.2	0.8	6.6		3.4
Disgust			97.5		2.5	
Joy				100		
Anger		0.8	0.8	4.2	91.7	2.5
Sad				4.2		95.8

Table 16. Confusion matrix of Proposed Method-2 with CF on Dataset-III (Train samples=600, Test samples=900)

	Fear	Surprise	Disgust	Joy	Anger	Sad
Fear	95.3	2.2	0.7	0.7	2.2	
Surprise		94.7	0.7	3.3	1.3	
Disgust			100			
Joy				100		
Anger		0.7	2.7	2	94.6	
Sad	4	3		6	0.7	87.3

Table 17. Confusion matrix of Proposed Method-2with WF on Dataset-I (Train samples=360, Test samples=540)

	Fear	Surprise	Disgust	Joy	Anger	Sad
Fear	100					
Surprise	2.22	81	3.33	3.33		1.11
Disgust			98.9	1.1		
Joy				100		
Anger			1.11	3.33	95.6	
Sad				3.33		96.67

Table 18. Confusion matrix of Proposed Method-2 with WF on Dataset-II (Train samples=480, Test samples=720)

	Fear	Surprise	Disgust	Joy	Anger	Sad
Fear	97.5			2.5		
Surprise		91.7	2.5	4.2		1.6
Disgust			96.7		3.3	
Joy				100		
Anger			1.6	3.4	93.4	1.6
Sad				4.2		95.8

Table 19. Confusion matrix of Proposed Method-2 with WF on Dataset-III (Train samples=600, Test samples=900)

	Fear	Surprise	Disgust	Joy	Anger	Sad
Fear	97.3		0.7	1.3	0.7	
Surprise	2	93.3	1.3	0.7		2.7
Disgust			98		2	
Joy				100		
Anger			2.7	1.3	96	
Sad	4		2	5.3	0.7	88

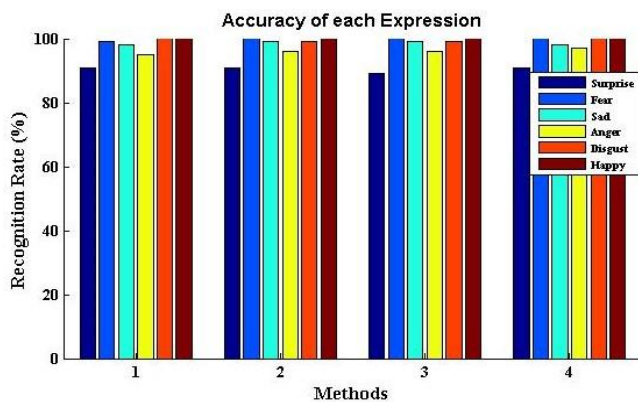


Figure 10. Comparison of six emotion rates of the proposed methods on Dataset-I

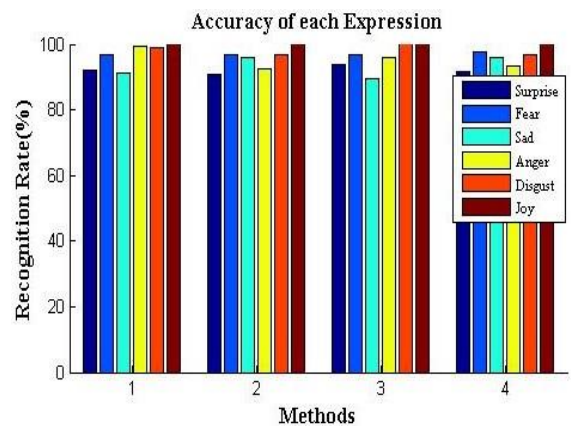


Figure 11. Comparison of six emotion rates of the proposed methods on Dataset-II

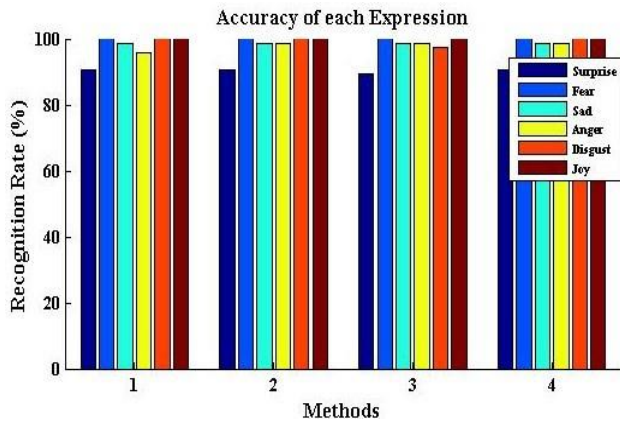


Figure 12. Comparison of six emotion rates of the proposed methods on Dataset-III

10. CONCLUSION

A novel method was proposed in this study based on the fusion of different feature sets extracted from input face images. A multi-feature system is an effective method compared to a uni-feature method to improve the recognition and classification accuracy of the FERCS. In the proposed methods, the Gabor wavelet method was adopted for extracting local features and the Haar wavelet was used for extracting global features. NLPCA was used for dimensionality reduction. Weighted fusion was then used for fusing the global and local features obtained from the facial expression samples. SVM was used as a classifier to recognize six emotions (joy, surprise, fear, disgust, anger, and sadness) from the images. The Extended Cohn-Kanade database was used in this study to evaluate the results of the proposed approach. The proposed method and its variations provided better results when compared with other existing methods tested on the CK+ database, as shown in Table 1. The future scope of this study involves the recognition of expressions under different head poses.

REFERENCES

- [1] Mehrabian DC. (1968). Communication without words. *Psychology Today* 2(4): 53-56. <https://dx.doi.org/10.4324/9781315080918-15>
- [2] Ekman P, Friesen WV, O'Sullivan M, Chan AYC, Diacoyanni-Tarlatzis I, Heider KG, Krause R, LeCompte WA, Pitcairn T, Bitti PER. (1972). Universals and cultural differences in facial expressions of emotion. *Journal of Personality and Social Psychology* 53(4): 712-717. <https://dx.doi.org/10.1037/0022-3514.53.4.712>
- [3] Du S, Tao Y, Martinez AM. (2014). Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences of the United States of America* 111(15): E1454-E1462. <https://dx.doi.org/10.1073/pnas.1322355111>
- [4] Pantic M, Rothkrantz LJM. (2000). Automatic analysis of facial expressions: the state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(12): 1424-1445. <https://dx.doi.org/10.1109/34.895976>
- [5] Deshmukh S, Patwardhan M, Mahajan A. (2016). Survey on real-time facial expression recognition techniques. *IET Biometrics* 5(3): 155-163. <https://dx.doi.org/10.1049/iet-bmt.2014.0104>
- [6] Struc V, Pavesic N. (2009): A case study on appearance based feature extraction techniques and their susceptibility to image degradations for the task of face recognition. *International Journal of Electrical and Computer Engineering* 3(6): 1351-1359. <https://doi.org/10.5281/zenodo.1071608>
- [7] Ramireddy CV, Kishore KVK. (2013). Facial expression classification using Kernel based PCA with fused DCT and GWT features. *IEEE International Conf. Computational Intelligence and Computing Research*, IEEE Press 1-6. <https://dx.doi.org/10.1109/ICCIC.2013.6724211>
- [8] Ramireddy CV, Kishore KVK, Bhattacharyya D, Kim TH. (2014). Multi-feature fusion based facial expression classification using DLBP and DCT. *International Journal of Software Engineering & Its Applications* 8(9): 55-68. <https://dx.doi.org/10.14257/ijseia.2014.8.9.05>
- [9] Uddin MZ, Khaksar W, Torresen J. (2017). Facial expression recognition using salient features and convolutional neural network. *IEEE Access* 5: 26146-26161. <http://dx.doi.org/10.1109/ACCESS.2017.2777003>
- [10] Deng HB, Jin, LW, Zhen LX, Huang JC. (2005). A new facial expression recognition method based on local gabor filter bank and PCA plus LDA. *International Journal of Information Technology* 11(11): 86-96. <https://dx.doi.org/10.1.1.102.3170>
- [11] PraseedaLekshmi V, Kumar MS. (2008). RBF based face recognition and expression analysis. *World Academy of Science Engineering and Technology* 2(6): 1175-1178. <https://doi.org/10.5281/zenodo.1072704>
- [12] Kishore KVK, Varma GPS. (2011). Hybrid emotional neural network for facial expression classification. *International Journal of Computer Applications* 35(12): 8-14. <https://doi.org/10.5120/4538-6420>
- [13] Sun Z, Hu ZP, Wang M, Zhao SH. (2017). Discriminative feature learning-based pixel difference representation for facial expression recognition. *IET Computer Vision* 11(8): 675-682. <https://doi.org/10.1049/iet-cvi.2016.0505>
- [14] Kumar S, Bhuyan MK, Chakraborty BK. (2016). Extraction of informative regions of a face for facial expression recognition. *IET Computer Vision* 10(6): 567-576. <https://doi.org/10.1049/iet-cvi.2015.0273>
- [15] Kamarol SKA, Jaward MH, Parkkinen J, Parthiban R. (2016). Spatiotemporal feature extraction for facial expression recognition. *IET Image Processing* 10(7): 534-541. <https://doi.org/10.1049/iet-ipr.2015.0519>
- [16] Siddiqi MH, Ali R, Khan AM, Park YT, Lee S. (2015). Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields. *IEEE Transactions on Image Processing* 24(4): 1386-1398. <https://doi.org/10.1109/TIP.2015.2405346>
- [17] Qayyum H, Majid M, Anwar SM, Khan B. (2017). Facial expression recognition using stationary wavelet transform features. *Mathematical Problems in Engineering* 9854050-1-9854050-9. <https://doi.org/10.1155/2017/9854050>

[18] Ding Y, Zhao Q, Li B, Yuan X. (2017). Facial expression recognition from image sequence based on LBP and Taylor expansion. IEEE Access 5: 19409-19419. <https://doi.org/10.1109/ACCESS.2017.2737821>

[19] Dong D, McAvoy TJ. (1996). Nonlinear principal component analysis-Based on principal curves and neural network. Computers & Chemical Engineering 20(1): 65-78. [https://doi.org/10.1016/0098-1354\(95\)00003-K](https://doi.org/10.1016/0098-1354(95)00003-K)