Recognizing Human Emotions from Facial Images by Landmark Triangulation: A Combined Circumcenter-Incenter-Centroid Trio Feature-Based Method

Avishek Nandi, Paramartha Dutta and Md Nasir

Abstract Human emotion reflected in facial expression is generated by coordinated operation of muscular movement of facial tissue which associates with the emotional state of the human subject. Facial expression is one of the most significant non-articulated forms of social communication and it is highly adopted by scientific community for the purpose of automated emotion analysis. In the present scope, a triangular structure is induced with three points, viz., circumcenter, incenter, and centroid are considered as the geometric primitive for extraction of relevant features. Information extracted from such features is utilized for the purpose of discrimination of one expression from another using MultiLayer Perceptron (MLP) classifier in images containing facial expressions available in various benchmark databases. Results obtained by applying this method found to be extremely encouraging.

1 Introduction

Expression recognition in human face has much significance in the field of affective computing from intelligent emotional assistance, interest estimation for e-learning system to criminal tendency estimation [25]. The human face is capable of generating 7,000 different kinds of facial expression, only Six of these expressions are recognized by behavioral scientist Ekman at el, namely, Anger (AN), Disgust (DI), Fear (FE), Happiness (HA), Sadness (SA), and Surprise (SU) as "atomic expressions" [14, 15]. They also proved that these six expressions are unique among different

A. Nandi (✉) · P. Dutta · Md Nasir
Visva-Bharati University, Santiniketan 731235, India
e-mail: avisheknandi10@gmail.com

Md Nasir
e-mail: nasir.vb@gmail.com

Fig. 1 Reference image for each of the basic emotions of faces

races, religions, cultures, and groups [13]. These six basic expressions are portrayed in Fig. 1.

The universality of the basic facial expressions and emotion is argued in [6, 11, 22]. Ekman et al. tried to settle the issue of universality dichotomy in their research facial expression across different cultures in [14] but they also mentioned that the universality phenomena of basic emotions are only valid when subjects are having strong emotional behavior. In the strong emotional condition, the facial expressions are not masked by the cultural influences and the expressions are static across religion, sex, race, culture, and educational background. But according to Russell, the facial expressions are sensitive to the context of the situation in which expressions are displayed [26]. But our research follows the universality of facial expressions which is argued by the Ekman and well accepted in the research community. This paper is organized as follows, apart from the first introductory section, Sect. 2 contains a survey of relevant literature. Section 3 emphasizes the motivation and contribution of our work. Section 4 discusses about the methodologies along with the essential flow diagram and algorithm. Landmark points generated by Active Appearance Model (AAM), the landmark selection and triangulation formation are emphasized in Sects. 4.1, 4.2 and 4.3, respectively. The circumcenter-incenter-centroid trio feature descriptor, MultiLayer Perceptron (MLP), and Classification Learning are discussed in Sects. 4.4, 4.5 and 5, respectively. Results and comprehensive analysis with respect to CK+, JAFFE, MMI, and MUG databases are reported in Sects. 6 and 7, respectively. Section 8 figures overall conclusions.

2   Literature Survey

Some recent studies of automatic facial expression recognition include the following researches. The work of Happy and Routray introduced salient facial patches to differentiate one expression to another; they also have used their own learning free landmark localization method for detection of facial landmarks in a robust and autonomous way [17]. In contrast to that Almaev and Valstar used local Gabor patterns to extract local dynamic features for detecting facial action units in a real-time manner in [2]. Another significant work of Yuan et al. introduces a hybrid combination of Local binary Patterns (LBP) with Principal Component Analysis (PCA) features which describes local and holistic facial features in a fused manner [30]. The Histogram oriented Gradients (HoG) features of facial components also explored by Chen et al. which are used to detect deformation features respective to facial expressions [9]. Martin et al. used Active Appearance Model (AMM) features of gray scale and edge images to achieve a greater robustness in a varying light condition [23]. The work of Cheon and kim introduce a differential AAM feature which is the computation of the Directed Hausdorff Distance (DHD) in between neutral face image and Excited face image with K-Nearest Neighbor (KNN) classifier [10]. Another interesting work by Barman and Dutta also uses AAM features to compute shape and distance signature along with statistical features to boost up expression recognition performance [3]. They also have extended their work of distance signature generated out of AMM landmark points with newly introduced texture the signature generated out of salient facial patched localized by AAM landmarks along with stability index [4, 5]. Most of the facial expression recognizer uses MultiLayer Perceptron (MLP), Radial Basis Function Network(RBF), Support Vector Machine (SVM) for the purpose of classification of facial expressions [8, 16, 19]. In comparison with MLP the fuzzy MLP classifier performs better because it can identify decision surfaces in case of nonlinear overlapping classes, whereas an MLP is restricted to crisp boundaries only [7].

3   Motivation and Contribution

The methods mentioned previously have a major drawback that it does not effectively classify facial expressions in a robust manner. To overcome this issue, we proposed a circumcenter-incenter-centroid trio feature as a more accurate shape descriptor in this context which is able to recognize facial expressions in a person independent manner. The circumcenter-incenter-centroid trio feature shows that it inherently captures the person independent information and also ensures good accuracy in different groups and ages of people.

The contribution of the present article has the following points of merits.

– An effective triangulation formation on the face is proposed.
– A novel feature descriptor based on circumcenter, incenter, and centroid of a face triangulation is introduced.
– A good person independent expression recognition with distance slope measure of circumcenter-incenter-centroid trio feature is achieved.

4 Methodology

The flow of the computation involves image space and feature space operations as depicted in Fig. 2. The algorithm with steps of computations is given below.

Algorithm 1: Proposed Training Algorithm for Facial Expression Recognition

---
1 INPUT: A set of N expression images of size $n \times m$.
2 STEP1: Extract the Landmark points using AAM fitting algorithm.
3 STEP2: Select the salient landmark points $L = [(x_1, y_1), (x_2, y_2), ..., (x_L, y_L)]$
4 STEP3: Generate the set of Triangulation T using Landmark points, where
  $T = [(x_i, y_i), (x_j, y_j), (x_k, y_k)]$ and $i, j, k \in [1, L]$.
5 STEP4: for t_i in T Compute three triangle center
  $p_{ijk} = [(p_1, p_2), (p_3, p_4), (p_5, p_6)]$
6 Compute distance triplet $D_{ijk} = [d_1, d_2, d_3]$ and slope triplet $S_{ijk} = [s_1, s_2, s_3]$
  as
7 $d_1 = dist((p_1, p_2), (p_3, p_4))$
8 $d_2 = dist((p_3, p_4), (p_5, p_6))$
9 $d_3 = dist((p_1, p_2), (p_5, p_6))$
10 $s_1 = slope((p_1, p_2), (p_3, p_4))$
11 $s_2 = slope((p_3, p_4), (p_5, p_6))$
12 $s_3 = slope((p_1, p_2), (p_5, p_6))$
13 STEP5: Combine distance D and slope S to get enhanced feature set.
14 STEP6: Divide the dataset in 70%,15% and 15% ratio for the purpose of
  training, validation and testing.
15 STEP7: Train MLP classifier with training data and use validation data to
  overcome overfitting.
16 STEP8: Test the Classifier on the basis of test database.
17 OUTPUT: Classified expression level for input image database.

4.1 Active Appearance Model (AAM)

AAMs are shape predictor model works by optimizing the geometric parameters of deformable bodies of a class of shape models [12]. We have used AAM by
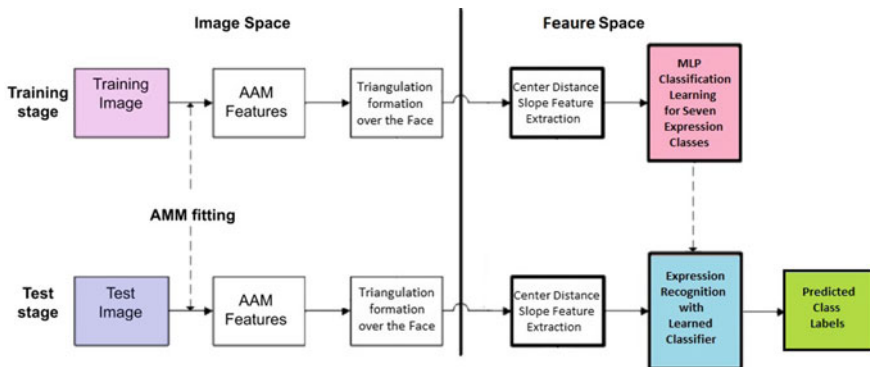
Fig. 2 Image space and feature space flow of computation

Tzimiropoulo et al for generating face description by inducing sixty-eight landmark points on the face [27, 28]. Python and Dlib implementation of [27] can be found on "https://github.com/davisking/dlib-models". We have used Python implementation of [27] for our research.

A deformable shape object can be expressed as $S = [(x_1, y_1), ..., (x_L, y_L)]^T$, a L element vector comprising L landmark coordinate points (x_i, y_i), $\forall i = 1, ..., L$. The AAM model is trained with a manually annotated set of N training images $I_1, I_2, ...I_N$ where each image consisting of L landmark points. The training of AAM has four steps of computation.

1. First holistic features are extracted using the $F()$ operator, i.e. $F(I_i)$, $\forall_i = 1, ..., T$.
2. Warping of extracted features from candidate image I_i according to the reference shape by $W()$ operator, i.e. $F(I_i)(W(s_i))$, $\forall_i = 1, ..., N$. Where vector shape parameters are defined as $a = [a_1, a_2, ..., a_n]^T$.
3. The Warped images are vectorized as $a_i = F(I_i)(W(s_i))$, $\forall i = 1, ..., N$ where $a_i \in \mathbb{R}^{M \times 1}$.
4. Finally, PCA is computed on the extracted vectors generating

$$\{\bar{a}, U_a\} \tag{1}$$

where $\bar{a}$ is mean appearance vector and U_a is orthonormal basis eigenvectors. $a_c = \bar{a} + U_a c$ is the new appearance model instance where $c = [c_1, c_2, ..., c_m]$ are appearance vector parameters.

4.2 AAM Landmark Selection

The landmarks generated by AMM describe geometrical positions and the shape of facial components. We have selected 21 principal landmark points as Barman and Dutta identified these points as salient landmark points in their research [3]. These points are mainly corner points and mid-points of facial components. The selected principal landmark points are as follows—two corner points and one midpoint on both eyebrows, two corner points of eyes and two middle points of eyelids, two corner points on nostril and one on middle of nose, four points on outer lips region and four on inner lips region as depicted on Fig. 3a. We left out outer points intentionally because those points are very less sensitive toward expressional changes.

4.3 Triangulation

The triangulation structure is formed by fixing three pivot points from the set of 21 points $\gamma = [(x_1, y_1), (x_1, y_1), ..., (x_{21}, y_{21})]$. Triangulation is formed using the formula (2).

$$\delta = < \sigma_i, \sigma_j, \sigma_k > \tag{2}$$

where σ_i, σ_j and σ_k are $\in \gamma$ and $\sigma_i \neq \sigma_j \neq \sigma_k$. The possible number of triangulation using 21 points is $\binom{21}{3} = 1330$. Figure 3b depicts the formed triangulation constituting principal landmark points.

It is needless to mention that the shape information of the triangulations is sensitive toward the expressional variation on the face. The shape of a triangle is highly correlated to geometrical positions of different types of centers. We have considered three types of classical triangle center as centroid, incenter, and circumcenter for this work. Triangle centers have this property that they are immune to similarity transformation (rotation, reflection, and translation) so only the change of shape is reflected irrespective of size and position of the triangle.

4.4 Circumcenter-Incenter-Centroid Trio

We have considered three types of triangle centers as Centroid, Incenter, and Circumcenter of to form the triangulation.

– **Centroid** of a triangle is the "Center of Gravity" point of the triangle and also the single intersection point of three lines bisecting in middle point of each side.
– **Incenter** of a triangle is the meeting point of three angel bisector and also the center of incircle of the triangle.
– **Circumcenter** of a triangle is the point of congruence of three side normal and also the center of circumcircle encompassing the triangle.
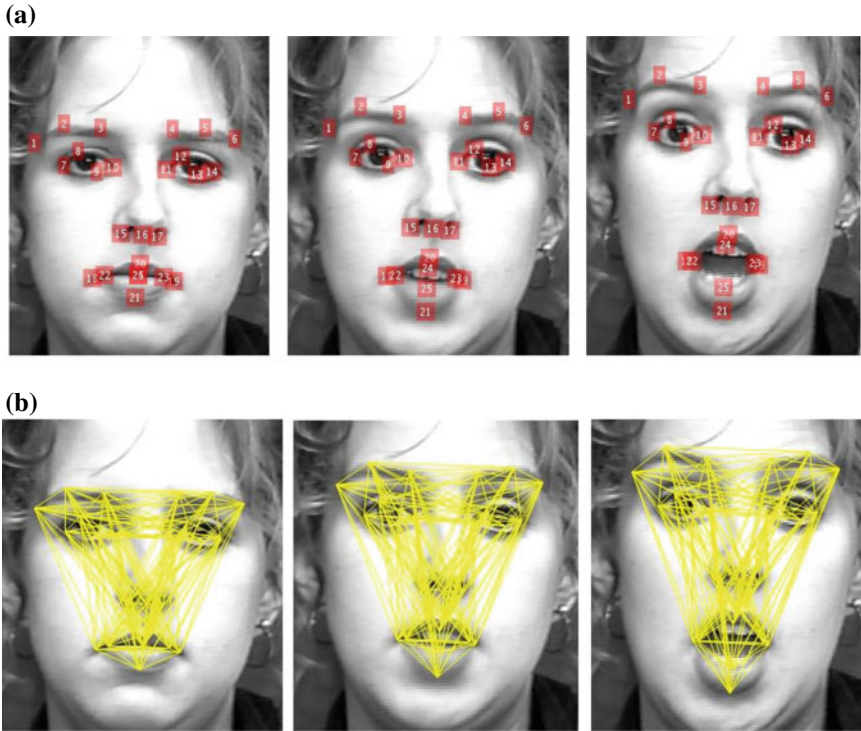
(a)

(b)

Fig. 3 Principal Landmark selection and Triangle formation Example. **a** Principal landmark points plotted on Face. **b** Triangulation formation

Incenter and centroid of a triangle always remain inside the triangle whereas circumcenter may get outside for obtuse triangles. The features we have considered are three types of distance and three types of slope vectors originating from centroid-incenter, centroid-circumcenter, and incenter-circumcenter pair of the centers of a triangle. These three distances and three slope features are combined to make a six-feature set and which are computed for each triangle in the triangulation set δ of Eq. 2.

4.5 MultiLayer Perceptron

MultiLayer perceptron is a feedforward neural network having at least three layers one input layer, one or more hidden layer, and one output layer. MLP is learned with a backpropagation algorithm which uses supervised methods of learning. MLP uses nonlinear synaptic activation function to learn the properties of high dimensional data. The sigmoid $y(v_i) = \frac{1}{(1+e^{-v_i})}$ and hyperbolic tangent $y(v_i) = tanh(v_i)$ activation function is most popular activation function in the literature. Layers of MLP are

interconnected with a weight matrix W where w_{ij} is the weight connecting i'th node of current layer to the j'th node of the following layer.

Backpropagation learning of MLP is done by changing the connection weight after processing each single input vector on basis of error computed in the output layer. The error is formulated as $e_j(n) = d_i(n) - y_j(n)$ where d is target value and y the value computer by the MLP and this error is backpropagated to output layer to input layer. Connection weights w_{ij} is adjusted to minimize the error $\eta(n) = \frac{1}{2} \sum e_j^2(n)$. The change of weight is computed using Eq. 3

$$\Delta w_{ji}(n) = -\eta \frac{\partial \mathcal{E}(n)}{\partial v_j(n)} y_i(n) \tag{3}$$

η is the learning rate and y_i output of the previous neuron. The output layer weight are updated using the formulae (4) from Chap. 4 of the book [18].

$$-\frac{\partial \mathcal{E}(n)}{\partial v_j(n)} = e_j(n)\phi'(v_j(n)) \tag{4}$$

The derivative of activation function is ϕ'. The hidden layer weights are updated using the formulae (5) from Chap. 4 of the book [18].

$$-\frac{\partial \mathcal{E}(n)}{\partial v_j(n)} = \phi'(v_j(n)) \sum_k -\frac{\partial \mathcal{E}(n)}{\partial v_k(n)} w_{kj}(n) \tag{5}$$

5   Classification Learning

These three-distance feature and three-slope feature are combined and learned with MultiLayer Perceptron (MLP) with scale conjugate backpropagation learning for classification of expressions to six different atomic expression classes [24]. We have used MATLAB implementation of [24] is used for the training the model for the purpose of pattern recognition and classification learning. The 70% of dataset is used for training purposes, 15% of the dataset is used for testing purposes and the rest of the 15% of the dataset is used for validation checking. Validation dataset used to overrule the overfitting of MLP classifier at the time of training.

6   Results

We have tested the proposed machine with four well-known expression database, the CK+ [20], JAFFE [21], MMI [29] and MUG [1] database. The computing environment we have used is Intel(R) Core(TM) i3-3217U CPU @ 1.80 GHz with 4 GB RAM. Dlib Shape Predictor Model implementation library for 68 landmark points

of AAM for Python can be found on GitHub link "https://github.com/davisking/dlib-models" [27] used for landmark extraction. The AAM model achieved an 100% detection accuracy in CK+, JAFFE, MUG database and 99.2% in MMI database.

6.1 Extended Cohn-Kanade (CK+) Database

The CK+ expression dataset is a combination of posed and spontaneous expressions having 327 image sequences from neural to peak expression. This dataset is consist of six basic expressions Anger (AN), Disgust (DI), Fear (FE), Happiness (HA), Sadness (SA), Surprise (SU) and also Contempt (CO) expression. Only the peak expressions are selected form this database. The proposed MLP classifier achieved a 99% overall accuracy in this database. We have also computed N-Fold accuracy for training, testing, validation, and overall dataset reflected in Fig. 4. Some example images with their true class level and predicted class levels are added for reference in Table 1.

6.2 Japanese Female Facial Expression (JAFFE) Database

The JAFFE database is a posed expression database containing 213 images of 10 different female adults showing six atomic facial expressions and there is no occlusion like spectacle, hair falling on face in JAFFE dataset. This dataset is consist of six basic expressions Anger (AN), Disgust (DI), Fear (FE), Happiness (HA), Sadness (SA), Surprise (SU) and also Neutral (NE) expression. In this dataset, the system obtained an overall precision of 97.18%. We have also computed N-Fold accuracy
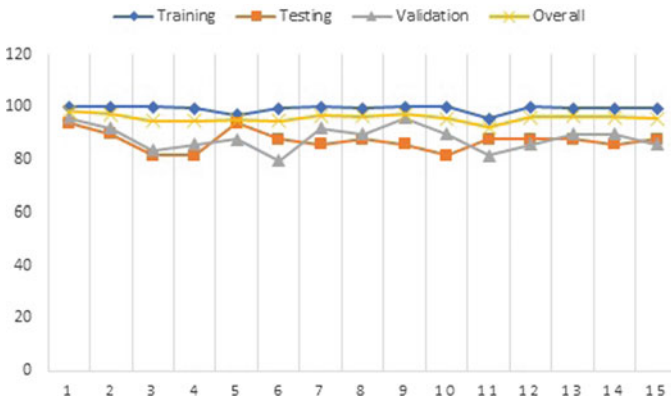

Fig. 4 N-Fold training, testing, validation, and overall accuracy of CK+ database [20]

Table 1 Visual representation of sample images of CK+ database with expression levels predicted by our system

Input image			
True expression level	Anger	Disgust	Fear
Expression level predicted by our system	Anger	Disgust	Fear
Verdict	Truly classified	Truly classified	Truly classified
Input image			
True expression level	Happiness	Sadness	Surprise
Expression level predicted by our system	Happiness	Anger	Surprise
Verdict	Truly classified	Misclassified	Truely classified

for training, testing, validation, and overall dataset reflected in Fig. 5. Some example images with their true class level and predicted class levels are added for reference in Table 2.

6.3 MMI Database

MMI is a posed expression dataset having multiple phases of data collection. The Phase-III of the MMI dataset contains 400 images with single Facial Action Unit (FAU) coded expression levels we have manually annotated 222 of them to six basic expression levels Anger (AN), Disgust (DI), Fear (FE), Happiness (HA), Sadness (SA) and Surprise (SU). Effective learning of the system with 222 manually annotated expressions results in an overall 96.87% accuracy. We have also computed N-Fold accuracy for training, testing, validation, and overall dataset reflected in Fig. 6. Some example images with their true class level and predicted class levels are added for reference in Table 3.
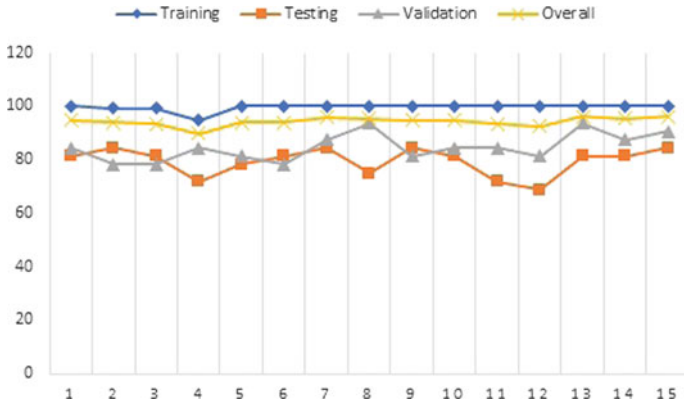
Fig. 5 N-Fold accuracy plot for training, testing, validation and overall images of JAFFE database [21]

Table 2 Visual representation of sample images of JAFFE database with expression levels predicted by our system

Input image			
True expression level	Anger	Disgust	Fear
Expression level predicted by our system	Anger	Disgust	Anger
Verdict	Truly classified	Truly classified	Misclassified
Input image			
True expression level	Happiness	Sadness	Surprise
Expression level predicted by our system	Happiness	Sadness	Surprise
Verdict	Truly classified	Truely classified	Truely classified
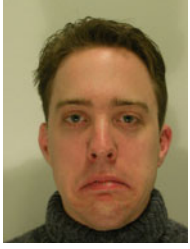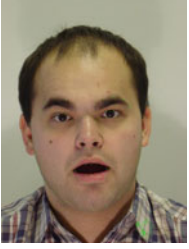
Fig. 6 N-Fold accuracy plot for training, testing, Validation, and overall images of MMI database [29]

Table 3 Visual representation of sample images of MMI database with expression levels predicted by our system

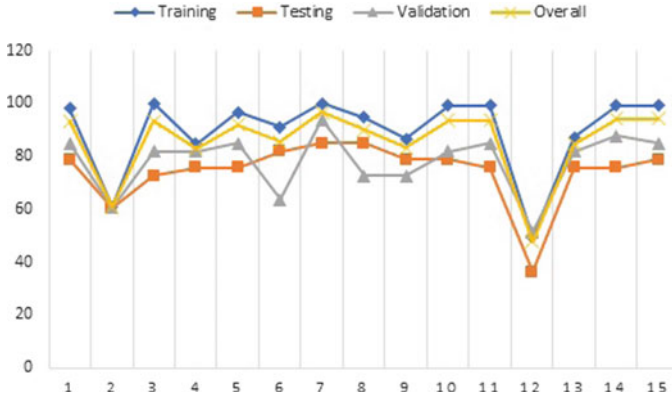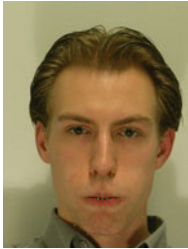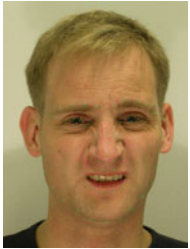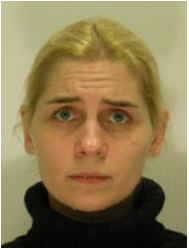| Input image			
True expression level	Anger	Disgust	Fear
Expression level predicted by our system	Anger	Anger	Fear
Verdict	Truly classified	Misclassified	Truly classified
Input image			
True expression level	Happiness	Sadness	Surprise
Expression level predicted by our system	Happiness	Sadness	Surprise
Verdict	Truly classified	Truely classified	Truely classified

Fig. 7 N-Fold accuracy plot for training, testing, Validation, and overall images of MUG database [1]

6.4 Multimedia Understanding Group (MUG) Database

MUG is a mixed expression dataset with posed and spontaneous expressions containing 401 images from 26 subjects. This dataset is consist of six basic expressions Anger (AN), Disgust (DI), Fear (FE), Happiness (HA), Sadness (SA), Surprise (SU), and also Neutral (NE) expression. Our system accomplishes an 97.26% of overall accuracy in MUG dataset. We have also computed N-Fold accuracy for training, testing, validation, and overall dataset reflected in Fig. 7. Some example images with their true class level and predicted class levels are added for reference in Table 4.

7 Discussions

The four different types of performance measure of training, validation, testing, and overall accuracy are shown in table no Table 5. The confusion matrix of the CK+, JAFFE, MMI, and MUG is presented in Tables 8, 9, 10 and 11, respectively. The MLP classifier shows very good recognition rate in CK+ database with 99.08% overall accuracy and the Table 6 shows that Anger, Contempt, Disgust, Happiness, and Surprise expressions are recognized with 100% precision. In the JAFFE dataset overall accuracy is 97.18% obtained and the Table 6 shows that the Neutral, Surprise, Anger, and Disgust expressions show 100% recognition rate. The MMI dataset shows 100% precision in Happiness expression with an overall 96.87% accuracy. The MUG database shows 100% Recognition rate for Disgust, Fear, Happiness, Sadness, and Surprise expression. Good accuracy in all four datasets implies that the system

Table 4 Visual representation of sample images of MUG database with expression levels predicted by our system

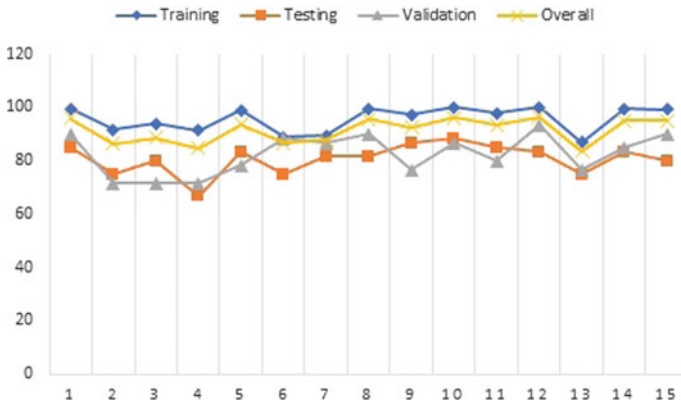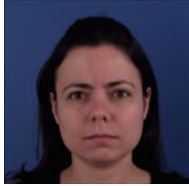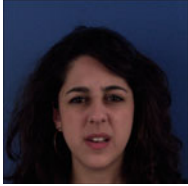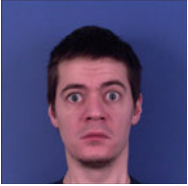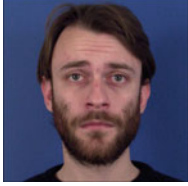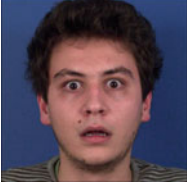| Input image			
True expression level	Anger	Disgust	Fear
Expression level predicted by our system	Fear	Disgust	Fear
Verdict	Misclassified	Truly classified	Truly classified
Input image			
True expression level	Happiness	Sadness	Surprise
Expression level predicted by our system	Happiness	Sadness	Surprise
Verdict	Truly classified	Truely classified	Truely classified

Table 5 Accuracy comparison table with training, testing, validation, and overall precisions

	CK+	JAFFE	MMI	MUG
Training	100	100	100	100
Validation	100	96.8	90.91	93.33
Testing	93.88	84.38	87.88	88.33
Overall	99.08	97.18	96.87	97.26

learns the expressions of a human face image in a person's independent manner. The circumcenter-incenter-centroid trio feature efficiently grabs the expression related cues showing effective and efficient learning of the system. Comparison with other machine learning technique ate also portrayed in Table 7.

Table 6 Recognition rate for each individual expressions for all four-benchmark database

Database	Anger	Contempt	Disgust	Fear	Happiness	Sadness	Surprise	Neutral	Overall
CK+	100	100	100	96	100	92.9	100	–	99.08
JAFFE	96.7	–	100	96.9	93.5	93.5	100	100	97.18
MMI	96.8	–	95.8	91.6	100	96.4	97.4	–	96.87
MUG	96.5	–	100	100	100	100	100	84	97.26

Table 7 Comparison between different machine learning techniques for different benchmark databases

Machine leaning methods	Features	CK+	JAFFE	MMI	MUG
Multilayer perceptron (MLP)	Circumcenter-Incenter-Centroid trio	**99.08**	**97.18**	**96.87**	**97.26**
Support vector machine (SVM)	Circumcenter-Incenter-Centroid trio	89.00	69.50	70.30	77.80
Linear discrimnant analysis (LDA)	Circumcenter-Incenter-Centroid trio	85.50	61.50	66.60	68.30
K-Nearest neighbor (KNN)	Circumcenter-Incenter-Centroid trio	77.40	71.80	66.70	82.30
MultiLayer perceptron (MLP)	Distance and texture signature from [4]	90.60	88.70	56.10	73.00

Table 8 CK+ confusion [20]

	AN	CO	DI	FE	HA	SA	SU
AN	**45**	0	0	0	0	0	0
CO	0	**18**	0	0	0	0	0
DI	0	0	**59**	0	0	0	0
FE	0	0	0	**24**	1	0	0
HA	0	0	0	0	**69**	0	0
SA	1	1	0	0	0	**26**	0
SU	0	0	0	0	0	0	**83**

Table 9 Confusion JAFFE [21]

	NE	HA	SA	SU	AN	DI	FE
NE	**30**	0	0	0	0	0	0
HA	1	**29**	1	0	0	0	0
SA	1	0	**29**	0	0	0	1
SU	0	0	0	**30**	0	0	0
AN	1	0	0	0	**29**	0	0
DI	0	0	0	0	0	**29**	0
FE	0	0	1	0	0	0	**31**

Table 10 MMI confusion [29]

	AN	DI	FE	HA	SA	SU
AN	**60**	0	0	0	1	1
DI	1	**23**	0	0	0	0
FE	1	0	**22**	1	0	0
HA	0	0	0	**45**	0	0
SA	0	0	1	0	**27**	0
SU	1	0	0	0	0	**38**

Table 11 MUG confusion [1]

	NE	AN	DI	FE	HA	SA	SU
NE	**21**	3	0	0	0	0	1
AN	1	**55**	1	0	0	0	0
DI	0	0	**71**	0	0	0	0
FE	0	0	0	**47**	0	0	0
HA	0	0	0	0	**87**	0	0
SA	0	0	0	0	0	**48**	0
SU	0	0	0	0	0	0	**66**

8 Conclusions

Overwhelming degree of accuracy on different benchmark databases such as CK+, JAFFE, MMI, and MUG vindicates the effectiveness and efficiency of the proposed method with judiciously chosen geometric feature set for the said purpose (Table 12).

Table 12 Confusion matrix of JAFFE, CK+, MMI, and MUG databases

	AN	CO	DI	FE	HA	SA	SU
AN	**45**	0	0	0	0	0	0
CO	0	**18**	0	0	0	0	0
DI	0	0	**59**	0	0	0	0
FE	0	0	0	**24**	1	0	0
HA	0	0	0	0	**69**	0	0
SA	1	1	0	0	0	**26**	0
SU	0	0	0	0	0	0	**83**

Table 8: CK+ Confusion [20].

	NE	HA	SA	SU	AN	DI	FE
NE	**30**	0	0	0	0	0	0
HA	1	**29**	1	0	0	0	0
SA	1	0	**29**	0	0	0	1
SU	0	0	0	**30**	0	0	0
AN	1	0	0	0	**29**	0	0
DI	0	0	0	0	0	**29**	0
FE	0	0	1	0	0	0	**31**

Table 9: Confusion JAFFE [21].

	AN	DI	FE	HA	SA	SU
AN	**60**	0	0	0	1	1
DI	1	**23**	0	0	0	0
FE	1	0	**22**	1	0	0
HA	0	0	0	**45**	0	0
SA	0	0	1	0	**27**	0
SU	1	0	0	0	0	**38**

Table 10: MMI Confusion [29].

	NE	AN	DI	FE	HA	SA	SU
NE	**21**	3	0	0	0	0	1
AN	1	**55**	1	0	0	0	0
DI	0	0	**71**	0	0	0	0
FE	0	0	0	**47**	0	0	0
HA	0	0	0	0	**87**	0	0
SA	0	0	0	0	0	**48**	0
SU	0	0	0	0	0	0	**66**

Table 11: MUG Confusion [1].

References

1. Aifanti N, Papachristou C, Delopoulos A (2010) The mug facial expression database. In: Proceedings of the 11th international workshop on image analysis for multimedia interactive services WIAMIS 10. IEEE, pp 1–4
2. Almaev TR, Valstar MF (2013) Local gabor binary patterns from three orthogonal planes for automatic facial expression recognition. In: Proceedings of the 2013 humaine association conference on affective computing and intelligent interaction. IEEE, pp 356–361
3. Barman A, Dutta P (2017) Facial expression recognition using distance and shape signature features. Pattern Recogn Lett
4. Barman A, Dutta P (2019) Facial expression recognition using distance and texture signature relevant features. Appl Soft Comput 77:88–105
5. Barman A, Dutta P (2019) Influence of shape and texture features in facial expression recognition. IET Image Process.
6. Bente G, Krämer NC, Eschenburg F (2008) Is there anybody out there. Mediated interpersonal communication, pp 131–157

7. Bhattacharjee D, Basu DK, Nasipuri M, Kundu M (2010) Human face recognition using fuzzy multilayer perceptron. Soft Comput 14(6):559–570
8. Boughrara H, Chtourou M, Amar CB, Chen L (2016) Facial expression recognition based on a mlp neural network using constructive training algorithm. Multimed Tools Appl 75(2):709–731
9. Chen J, Chen Z, Chi Z, Fu H (2014) Facial expression recognition based on facial components detection and hog features. In: Proceedings of the international workshops on electrical and computer engineering subfields, pp 884–888
10. Cheon Y, Kim D (2009) Natural facial expression recognition using differential-aam and manifold learning. Pattern Recogn 42(7):1340–1350
11. Collier G, Collier GJ (2014) Emotional expression. Psychology Press
12. Cootes TF, Edwards GJ, Taylor CJ (2001) Active appearance models. IEEE Trans Pattern Anal Mach Intell 6:681–685
13. Ekman P (1992) An argument for basic emotions. Cogn Emot 6(3–4):169–200
14. Ekman P (2004) Emotions revealed. BMJ 328(Suppl S5):0405184
15. Ekman P, Friesen WV (1971) Constants across cultures in the face and emotion. J Personal Soc Psychol 17(2):124
16. Er MJ, Wu S, Lu J, Toh HL (2002) Face recognition with radial basis function (RBF) neural networks. IEEE Trans Neural Netw 13(3):697–710
17. Happy S, Routray A (2014) Automatic facial expression recognition using features of salient facial patches. IEEE Trans Affect Comput 6(1):1–12
18. Haykin SS et al (2009) Neural networks and learning machines/Simon Haykin. Prentice Hall, New York
19. Kotsia I, Pitas I (2006) Facial expression recognition in image sequences using geometric deformation features and support vector machines. IEEE Trans Image Process 16(1):172–187
20. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I (2010) The extended cohn-kanade dataset (ck+): a complete dataset for action unit and emotion-specified expression. In: Proceedings of the 2010 IEEE computer society conference on computer vision and pattern recognition-workshops. IEEE, pp 94–101
21. Lyons M, Akamatsu S, Kamachi M, Gyoba J (1998) Coding facial expressions with gabor wavelets. In: Proceedings of the third IEEE international conference on automatic face and gesture recognition. IEEE, pp 200–205
22. Manstead AS, Fischer AH (2002) Beyond the universality-specificity dichotomy. Cogn Emot 16(1):1–9. https://doi.org/10.1080/0269993014000103
23. Martin C, Werner U, Gross HM (2008) A real-time facial expression recognition system based on active appearance models using gray images and edge images. In: Proceedings of the 2008 8th IEEE international conference on automatic face & gesture recognition. IEEE, pp 1–6
24. Møller MF (1993) A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw 6(4):525–533
25. Pham TH, Philippot P (2010) Decoding of facial expression of emotion in criminal psychopaths. J Pers Disord 24(4):445–459
26. Russell J (1997) Reading emotions from and into faces: resurrecting a dimensional-contexual perspective. The psychology of facial expression, pp 295–320
27. Sagonas C, Antonakos E, Tzimiropoulos G, Zafeiriou S, Pantic M (2016) 300 faces in-the-wild challenge: database and results. Image Vis Comput 47:3–18
28. Tzimiropoulos G, Pantic M (2013) Optimization problems for fast AAM fitting in-the-wild. In: Proceedings of the IEEE international conference on computer vision, pp 593–600
29. Valstar M, Pantic M (2010) Induced disgust, happiness and surprise: an addition to the mmi facial expression database. In: Proceedings of the 3rd international workshop on EMOTION (satellite of LREC): corpora for research on emotion and affect, p 65
30. Yuan L, Wu CM, Zhang Y (2013) Facial expression feature extraction using hybrid PCA and LBP. J China Univ Posts Telecommun **20**(2), 120–124 (2013)