



Performance Evaluation of Machine Learning for Recognizing Human Facial Emotions

Farid Ayache¹, Adel Alti^{2,3*}

¹ LMETR Laboratory Optics and Precision Mechanics Institute, University of SETIF-1, Setif 19000, Algeria

² Department of Management Information Systems, College of Business & Economics Qassim University, P.O. Box 6633, Buraidah, 51452, Saudi Arabia

³ LRSD Laboratory Computer Science Department University of SETIF-1 Setif 19000, Algeria

Corresponding Author Email: alti.adel@univ-setif.dz

<https://doi.org/10.18280/ria.340304>

ABSTRACT

Received: 25 March 2020

Accepted: 7 May 2020

Keywords:

human facial emotions, active shape model, machine learning, Generalized Procrust Analysis, quadratic classifier

Facial Expression Recognition is a human emotion classification problem that attracted much attention from scientific research. Classifying human emotions can be a challenging task for machines. However, more accurate results and less execution time are there still the main issues when extracting features of human emotions. To cope with these challenges, we propose an automatic system that provides users with well-adopted classifier for recognizing facial expressions more accurately. The system consists of two fundamental machine-learning stages, namely, feature selection and feature classification. Feature selection is performed using Active Shape Model (ASM) composed of landmarks while the feature classification has examined seven well-known classifiers. We have used CK+ dataset, implemented and tested seven classifiers to find the best classifier. Experimental results showed that Quadratic classifier provides excellent performance and outperforms other classifiers with the highest accuracy of 92.42% on the same dataset.

1. INTRODUCTION

Facial Expression Recognition (FER) systems over the past several decades have attracted much attention from scientific research. FER has proven several benefits and showed great success in computer vision due to their major importance in various areas of our daily life such as Human-Machine Interface (HCI), automatic psychological analysis, the security and surveillance field in particular airports, robotic education to offer a better learning experience by having a better understanding of the feelings of students and online learning systems to estimate the criminal tendency and security of the conductor.

Facial expressions are one of those things, which are of great importance to humans in social communication, as they tend to convey emotions, energies, and expressions without using words. The human face is capable of generating thousands of facial expressions. Machine learning approaches to FER all require a set of training image examples, each labeled with a single emotion category. A standard set of six emotions classification is Anger (AN), Disgust (DI), Fear (FE), Happiness (HA), Sadness (SA), and Surprise (SU) as 'atomic expressions'. These six expressions are unique among different races, religions, cultures, and groups [1-3]. Some researchers consider the neutral face as a seventh expression [4-6]. Despite their powerful benefits that provide in human-computer interaction systems, classifying human emotions can be a challenging task for machines. However, higher classification accuracy and less execution time are there still the main issues when extracting features of human emotions.

Generally, FER systems may be categorized into two fundamental approaches, namely geometric-based and texture-based approaches. Each one has advantages and

drawbacks. In this paper, we mainly address the challenge of searching, in a novel classification way, for an appropriate classifier from the existing classifiers recognizing facial expressions. By novel, we mean (i) considering relevant features beyond what is explicitly shown in the faces images, (ii) realizing comparative machine learning to human facial emotions including two new classification techniques, the first one is *one vs others* and the second is *test vs training*. Our goal is to provide adequate classifiers helping users to reduce their execution time and increase the recognition rate.

This paper presents an automatic landmarks extraction module enhanced with Active Shape Model. The evaluation of performance and accuracy of seven classifiers among the most common algorithms in FER, K-Nearest Neighbor (KNN), Naive Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT), Quadratic classifier (DA), Random Forest (RF) and Multi-Layer Perceptron (MLP) to predict facial expression's class in a CK + dataset. The proposed approach driving new suitable classifiers for exploring automated emotion recognition via machine learning.

2. RELATED WORK

There are many attempts by researchers to classify human emotions. However, high accuracy is still the main issue when classifying human emotions. The first work on human emotions has been done is the Facial Action Coding System (FACS) developed by Ekman et al. [7] based on human observations and manual labeling process. It noted that FACS could serve many researchers, particularly those with a psychological background and lack of concentration. The proposed system also extracts many facial features using

action coding. However, relying on action coding technique may not identify the faces' expression in a more precise manner due to the traditional coding techniques. For that reason, new local paramedical representations of movements have been proposed by Black and Yacoob [8] to transmit the information to an appropriate classifier. Due to the important procession time of such technique, users are not able to identify the huge face's features.

Ensuring increased accuracy and less execution time are considered as a major objective to identify or detect human emotions within an appropriate classifier. According to the study [9], better-classification of human emotions is an important task in exploiting hidden features, the recognition rate can be improved. From this idea, the authors proposed an approach that uses Artificial Neural Networks (ANN) to predict facial expressions recognition performance using facial image data obtained by a CCD camera. The ANN can classify complex discriminating faces for human facial expression recognition (e.g. anger, disgust, fear, happiness, sadness, and surprise). This work is limited to identifying a small range of human facial expressions and neglecting reduced feature information. The Artificial Neural Networks (ANN) is also used in the works [10-14].

Besides, as we know, human expressions in an HCI are several and can be detected using many classifiers that require large and deep neural networks with a significant processing time to identify such expressions. As an improvement of neural network, Kong [15] developed a deep Convolutional Neural Networks (CNN) for the recognition of facial expressions. The work consists of two connected channels; the first channel contains the input extracted eyes while the second consists of one input exhibiting the mouth. The collected information from the two channels converges into a fully connected layer, which is used to learn global information from these local characteristics and is then used for classification. The major lack of this approach is the complexity level and the computational time that increases with every additional layer in the purpose of extracting complex features.

Chuang and Shih [16] adopted an Independent Component Analysis (ICA) to extract facial features. They used SVM (Support Vector Machines) to detect and predict facial expressions. Such kind of classifier offers high accuracy rate for some facial expressions. Furthermore, feature extraction and selection are also crucial for facial expression recognition, which is not well represented in this work. Some other approaches exploiting SVM classifier are also proposed and can be found in the researches [17-21].

To examine in more detail variations in facial expressions over time, new expression recognition technique based-Gabor's wavelets have been proposed [23]. Due to the limited temporal segmentation of facial gestures in spontaneous facial behavior recorded in real contexts, Torre's et al. [24] proposed spectral graph-based techniques to group similar shapes and appearance characteristics to certain geometric transformations. This study shows that even though the high recognition rate cannot be achieved through more general facial characteristics.

The ability to exploit the fuzzy techniques is used by researchers. The work [25] proposed a model to classify facial expressions based on fuzzy rules. The results show that fuzzy rules predict better non-linear overlapping classes than the Multi-Layer Perceptron (MLP) [26] that is limited only to sharp borders. However, this work does not offer the specification of complex facial expressions.

Some other approaches integrating Active Appearance Model (AMM) are also proposed by Martin et al. [27]. This approach used the characteristics of the gray-scale Active Appearance Model (AMM) and edge images to obtain greater robustness under variable lighting conditions. Besides, Cheon and Kim [28] proposed a differential AAM function based on directed Hausdorff distance (DHD) between the neutral face image and the excited face image with the K-Nearest Neighbor (KNN) classifier.

Almaev and Valstar [29] proposed a model based on local Gabor to extract local dynamic characteristics and to detect facial action units in real-time. Yuan et al. [30] proposed model-based Local Binary Models (LBP) with Principal Component Analysis (PCA) to obtain high accurate facial features by improving the local and holistic facial characteristics in a merged manner. Chang et al. [22] used a small subset of distinguishable facial expressions extracted from the human face. Once the facial expressions represented using a set of facial shapes, they are projected and aligned from three-dimensional space using an enhanced Lipschitz integration system. This work achieved lower accuracy is when classifying and smaller subsets with blended expressions. The main drawback found was that it needed to be extended with more facial expressions details extracting with multiple facial deformations. Happy and Routray [31] used facial patches to differentiate one expression from another. They also used their method for locating free landmarks and detecting facial landmarks robustly and autonomously. Chen et al. [32] focused on their efforts on detecting respective deformation characteristics of facial expressions by exploiting the characteristics of Histogram Oriented Gradients (HoG) of facial components. Barman and Dutta [33] proposed model-based Active Appearance Model (AAM) to enhance expression recognition performance. Then, the shape and distance signatures as well as statistical functionalities are the input data of the learning model. Barman et al. [34, 35] present a core lightweight ontology for remote signature which has been extended from AMM landmarks. This work enables the location of faces by landmarks AAM as well as the stability index.

Based on previous research results, we propose an approach of machine learning techniques using an enchanted Active Shape Model (ASM) to build facial features. Beside feature extraction, two new classification techniques, the first one is one vs others and the second is test vs training also considered to improve the machine-learning model.

3. PROPOSED FRAMEWORK

This section presents the machine learning techniques for facial expression recognition to solve defined problem above. We conduct experiments with other existing machine learning models and compare their accuracy. Each machine learning has advantages and drawbacks. However, most of them suffer from poor accuracy and high execution time in most facial images with fewer discriminant features.

To overcome that, we propose a system with a custom module called Active Shape Model (ASM) combined with seven most popular classifiers. This helps our system work well with input facial images that extracting more distinguishable facial landmarks and select the best classifier model with high accuracy and precision, reduced execution time that suits the facial expressions field. The architecture of the proposed framework is described in Figure 1.

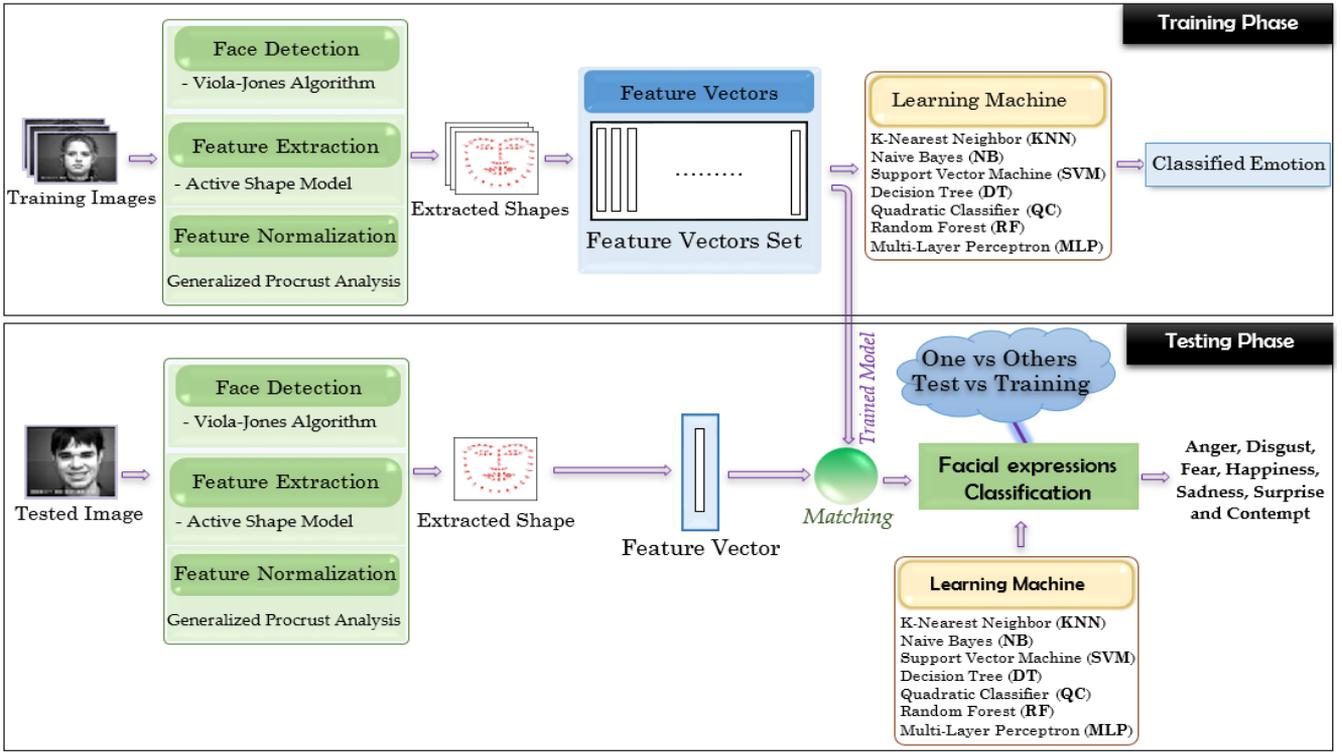


Figure 1. General architecture of the proposed framework

The proposed system consists of two fundamental successive phases, training and testing. During the training phase, the system extracts features and landmark points in all face images as shape vectors defined by Active Shape Model (ASM). These vectors will be aligned to eliminate the effect of translation, rotation and scaling using normalization algorithm. Likewise, the testing process will be extracted the features and landmark points in the test image and be compared with the trained one from the extracted vectors in the database to recognize its emotion classes.

3.1 The face detection stage

The *face detection* stage locates the face in input image using Viola-Jones algorithm [36]. This algorithm is often effective in solving the problem of complex background, brightness and it is insensitive to noise. It is defined through two main steps: the extraction of HAAR characteristics and the classification using Adaboost [36].

3.2 The feature extraction stage enhanced with active shape model

The *feature extraction* stage goes further in extracting the more discriminant facial landmarks of facial image. Often this means finding the local and global facial expressions features using the Active Shape Model (ASM), which can be most indicative of a particular class. This algorithm predicts optimal edges for a given object through geometric transformations.

We used and adapted this algorithm for generating 68 landmarks to obtain more accurate results. It was included in our framework after the face detection stage. The ASM model based on statistical analysis to determine the mean shape and its allowable variations in a set of shape images. A shape (q_i $i = 1..m$) representing a set of objects that consists of n Cartesian coordinates (x_j, y_j $j = 1..n$), each of which is a

particular landmark point. Landmark points are stacked into a shape vector as follows:

$$q_i = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{pmatrix}, \text{ for } i = 1..n \quad (1)$$

where, n is the number of landmarks.

The feature extraction process is described as the following steps (see Figure 2):

(1) Step. 1: Compute the mean shape \bar{q} of m shape vectors as follows:

$$\bar{q} = \frac{1}{m} \sum_{i=1}^m q_i \quad (2)$$

(2) Step. 2: Compute the covariance matrix S to capture the shape variability as follows:

$$S = \frac{1}{m-1} \sum_{i=1}^m (\bar{q} - q_i)(\bar{q} - q_i)^T \quad (3)$$

(3) Step. 3: Compute the eigenvector u_i and eigenvalue λ_i of the covariance matrix S . The eigenvector defines mode of variations of all points of the shape defined as follows:

$$S u_i = \lambda_i u_i \quad (4)$$

where, λ_i is the i th eigenvalue of the covariance matrix S and $U = (u_1 | u_2 | \dots | u_t)$ contains t eigenvectors of the covariance matrix S .

(4) Step. 4: We approximate any instance of the shape q by projecting onto the first t eigenvectors as follows:

$$q = \bar{q} + U \times b \quad (5)$$

where, $b = [b_1, b_2, \dots, b_n]^T$ denotes the weight vector which is identified as a feature of this instance of the shape. Varying the weights b_i enables us to explore the allowable variations in the shape. Thus, we find an optimal shape representation in the following steps.

- (1) **Step. 5:** Initialize the weighs vector b to zero.
- (2) **Step. 6:** Generate initial model instance using Eq. (5).
- (3) **Step.7:** Compute the translation, rotation and scaling parameters respectively (X_t, Y_t, s, θ) that best aligns the model instance.
- (4) **Step. 8:** Update the weight vector b using the following equation:

$$b = U^t(q - \bar{q}) \quad (6)$$

- (5) **Step. 9:** If weights or position parameters are changed go back to Step.4 Else Stop

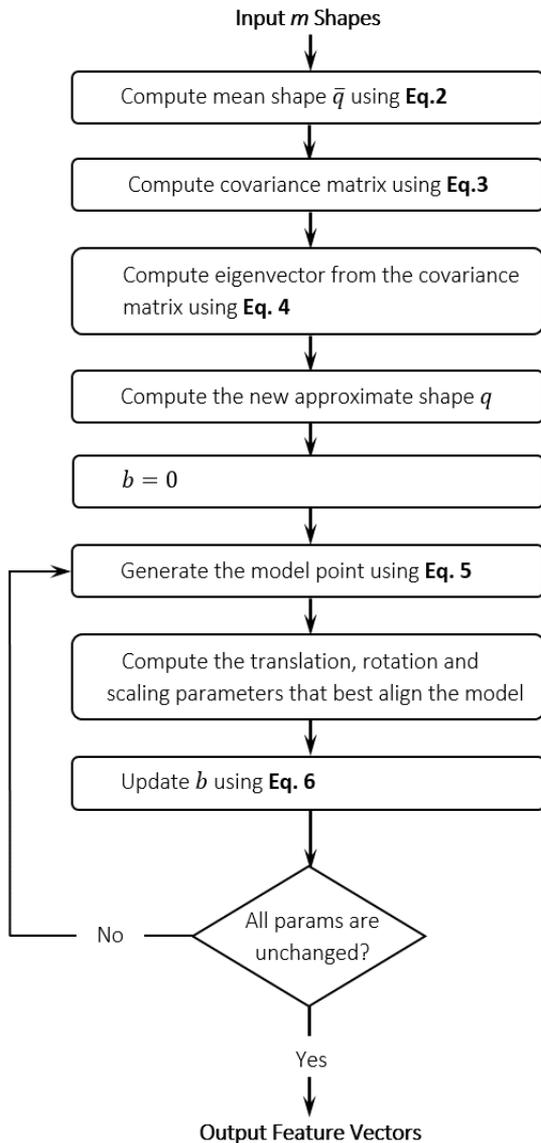


Figure 2. The flowchart of feature extraction stage

3.3 Feature normalization stage

After *feature extraction* stage, we proceed to the *feature normalization* stage. It is used to eliminate the effects of scale, rotation, and translation between all shapes using ASM model

to accentuate relevant facial information. This will significantly increase and improve the recognition rate for the proposed system. We used Generalized Procrust Analysis (GPA) proposed by Gower [37] and enhanced by Ten Berge [38]. Figure 3 shows an example of shapes before and after applying the GPA. Figure 3 shows an example result of our system for extracting.

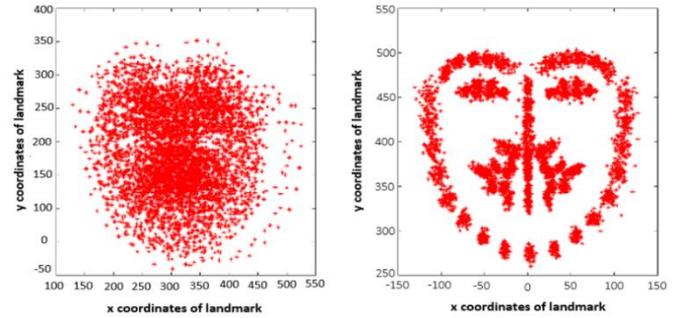


Figure 3. Shape normalization, a: before, b: after

3.4 Feature vector construction and classification stage

After the *feature normalization* stage, we will store the feature vector in a database. It contains all the features vectors that represent human emotions. Let F is the database containing the features vectors defined by:

$$F = (q_1 ; q_2 ; \dots ; q_m) \quad (7)$$

where, q_i is a shape vector defined using Eq. (1) and m is the number of shape images.

Figure 4 shows an example of the landmarks extracted by our system on standard set of six facial expressions. We conduct several experiments with the most common algorithm and compare its accuracy and execution time.

Our goal is to select the best algorithm from the most common algorithm that can be used in Facial Expression Recognition:

- K-Nearest Neighbor (KNN)
- Naive Bayes (NB)
- Multiclass Support Vector Machine (SVM)
- Decision Tree (DT)
- Quadratic classifier (DA)
- Random Forest (RF)
- Multi-Layer Perceptron (MLP).

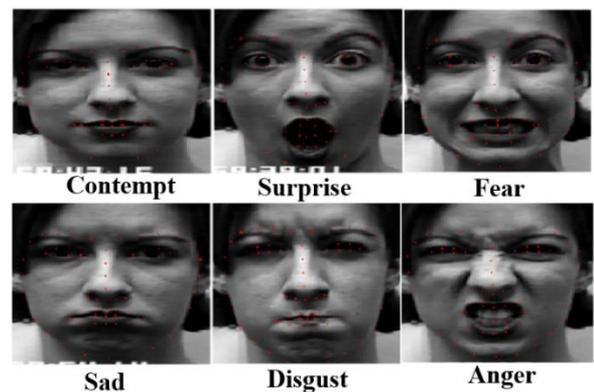


Figure 4. Samples of facial expression landmarks images result of our system in CK+ database

4. DATASET AND TRAINING

4.1 Dataset

For our experiments, we used the extended Cohn-Kanade (CK+) dataset [13]. The CK+ has 326 images of peak facial expressions for seven emotion categories are anger (AN), contempt (CO), disgust (DI), fear (FE), happiness (HA), sadness (SA), and surprise (SU), varying between posed and non-posed facial expressions of 210 adults. The images set ranging from 18 to 50 years old of age consisting of 69 female, 81, Euro-American, 13 Afro-American, and 6 other groups. The resolution of all images for training, verification and testing of size 256×256. It contains the following number sequences of individual anger expression (45), contempt (18), disgust (59), fear (25), happiness (69), sadness (28), and surprise (82). No subject has been collected with the same emotion more than once. We consider the best quality to consider these properties of good dataset. We applied the proposed ASM model to faces database [39, 40] for generating descriptions of face images with 68-landmarks.

4.2 Training

We consider two classification approaches *one vs others* and *test vs training* are the best way to conduct in-depth comparative study of seven classifiers among the most common algorithms in FER:

(1) One vs others: First, we take a face image from a database as a test image and the remaining images as the training set. By applying the training process of seven classifiers to original face images, new features of facial expressions are extracted. Once any feature extraction is completed, the same process is applied for all face images in the database. In this approach, testing process classified all face images and compared one by one to the rest of database. We evaluate the performance and accuracy of FER by each category and compare the effectiveness of each category versus other traditional categories.

(2) Test vs Training: We conduct in-depth experiments on our dataset using seven classifiers in both test and training face images. The dataset with 326 images was split into 2 parts: *training* and testing. We train seven classifiers with varying pairs of {test, training}. First, we take one face image of each class in the testing set then two test images of each class in the second step until all ten-face images of each class in the testing set and the remainder from the training set. Finally, we evaluate the performance and accuracy of FER using seven classifiers in all face images in the testing set.

4.3 Evaluation metrics

We opt to use 4 metrics for our framework's performance:

(1) Precision (Pre) to identify the number of emotions correctly classified among the classified ones. Precision output value between 0 and 1 with a close or equal value to 1 indicating better classification performance. The Pre-measure is defined as follows:

$$Pre = \frac{TP}{TP + FP} \quad (8)$$

(2) Recall (Rec) to identify the number of emotions correctly classified among the total number of expected

emotions. A best recall is expressed when Rec equal to 1. The formula for Recall (Rec) is:

$$Rec = \frac{TP}{TP + FN} \quad (9)$$

(3) F1score to evaluate a weighted average of Pre and Rec. This indirectly highlights missed expected classified emotions, which is an important factor based on weighted recall. F1-score reach best value at 1 (perfect Pre and Rec). It is computed as follows:

$$F1s = 2 * \frac{Pre * Rec}{Pre + Rec} \quad (10)$$

where,

- TP: the number of correctly classified emotions (true positives);
- FP: the number of incorrectly classified emotions (true negatives);
- TN: the number of correctly classified emotions that are not classified by the given approach (true negatives);
- FN: the number of incorrectly classified emotions that are not classified by the given approach (false negatives).

5. PERFORMANCE COMPARISONS

To achieve in-depth comparative study of seven classifiers in FER, it is necessary to compare them in terms of execution time, accuracy, precision, and F1-score through two new classification approaches one vs others and test vs training.

5.1 Execution time comparison

We have evaluated the execution time with the most common algorithms in FER. The execution time is the average response time needed to accomplish the testing stage of all classifiers except for SVM that use linked test and training stages. Figure 5 shows the execution time comparison for seven classifiers trained on CK+ dataset. For TREE, KNN, NB, and DA show superior improvement in terms of the execution time. The MLP and RF classifier yields lower execution time than other classifiers. The size of our features vector was smaller of 136 than other techniques and descriptors LBP, HOG... etc. The experiments are conducted on a computer with Intel Core i5-7500 CPU @3.4GHz, 32GB of RAM, GPU, and 1TB SSD hard disk. The framework is implemented with the C++builder.

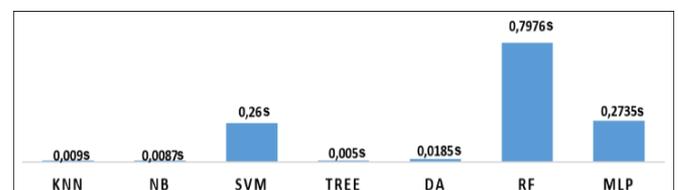


Figure 5. Execution time comparison between seven classifiers

5.2 One vs others

Table 1 shows the accuracy, precision, recall, and f-score for the seven classifiers trained on CK+ dataset. As expected,

SVM, DA, and FR present an ideal recognition rate of 100%, which proves their efficiency. We also observe that the recognition rates of KNN, NB, TREE, and NN are very satisfactory and they are very close to each other. We can notice that SVM, DA, and FR presents impressive accuracy results.

Table 1. Accuracy comparison between seven classifiers using *one vs others* approach

Classifier	Accuracy	Precision	Recall	F1Score
KNN	96.93	95.48	95.09	95.29
NB	96.32	95.79	95.13	95.46
SVM	100	100	100	100
TREE	97.55	96.70	96.94	96.81
DA	100	100	100	100
RF	100	100	100	100
MLP	98.16	98.10	98.19	98.14

Table 2. Confusion matrix of CK+ dataset using SVM, DA, and RF

	AN	CO	DI	FE	HA	SA	SU
AN	100	0.00	0.00	0.00	0.00	0.00	0.00
CO	0.00	100	0.00	0.00	0.00	0.00	0.00
DI	0.00	0.00	100	0.00	0.00	0.00	0.00
FE	0.00	0.00	0.00	100	0.00	0.00	0.00
HA	0.00	0.00	0.00	0.00	100	0.00	0.00
SA	0.00	0.00	0.00	0.00	0.00	100	0.00
SU	0.00	0.00	0.00	0.00	0.00	0.00	100

Table 3. Confusion matrix of CK+ dataset using KNN

	AN	CO	DI	FE	HA	SA	SU
AN	95.56	0.00	3.39	0.00	0.00	0.00	0.00
CO	0.00	94.44	0.00	0.00	0.00	3.57	0.00
DI	2.22	0.00	98.30	0.00	0.00	0.00	0.00
FE	0.00	5.56	0.00	92.00	0.00	3.57	0.00
HA	0.00	0.00	0.00	0.00	100	0.00	0.00
SA	2.22	5.56	0.00	4.00	0.00	89.29	0.00
SU	0.00	5.56	0.00	0.00	0.00	0.00	98.78

Table 4. Confusion matrix of CK+ dataset using NB

	AN	CO	DI	FE	HA	SA	SU
AN	93.33	0.00	5.08	0.00	0.00	0.00	0.00
CO	0.00	94.44	0.00	4.00	0.00	0.00	0.00
DI	8.89	0.00	93.22	0.00	0.00	0.00	0.00
FE	0.00	0.00	0.00	92.00	0.00	7.14	0.00
HA	0.00	0.00	0.00	0.00	100	0.00	0.00
SA	0.00	0.00	0.00	0.00	0.00	100	0.00
SU	0.00	5.56	0.00	4.00	0.00	0.00	97.56

Table 5. Confusion matrix of CK+ dataset using TREE

	AN	CO	DI	FE	HA	SA	SU
AN	100	0.00	0.00	0.00	0.00	0.00	0.00
CO	0.00	100	0.00	0.00	0.00	0.00	0.00
DI	2.22	0.00	98.30	0.00	0.00	0.00	0.00
FE	2.22	5.56	0.00	92.00	0.00	0.00	0.00
HA	2.22	0.00	0.00	0.00	98.55	0.00	0.00
SA	4.45	0.00	0.00	0.00	1.45	89.29	0.00
SU	0.00	5.56	0.00	0.00	0.00	0.00	98.78

The confusion matrix is used to evaluate the performance of seven classifiers on CK+ dataset. Table 2-5 show the confusion matrix on test sets using seven classifiers. From the observation of the confusion matrix, SVM, DA, and FR have

ideal matching values but also no overlap between facial expressions. As observed from the experiments above, the DA classifier is the best classifier in terms of execution time and accuracy results while KNN and NB classifiers come in the last.

Table 6. Confusion matrix of CK+ dataset using MLP

	AN	CO	DI	FE	HA	SA	SU
AN	100	0.00	0.00	0.00	0.00	0.00	0.00
CO	2.22	94.44	0.00	0.00	0.00	0.00	0.00
DI	4.44	0.00	96.61	0.00	0.00	0.00	0.00
FE	0.00	0.00	0.00	100	0.00	0.00	0.00
HA	0.00	0.00	1.69	0.00	95.65	3.57	1.22
SA	0.00	0.00	0.00	0.00	0.00	100	0.00
SU	0.00	0.00	0.00	0.00	0.00	0.00	100

5.3 Test vs training

After splitting the CK+ dataset according to test vs training classification model, we obtain the tested pairs: (7, 326-7), (14, 326-14)... (70, 326-70). Table 7 shows the accuracy results of the proposed approach for ten pairs. Moreover, while analyzing the results, we obtain that the DA classifier presents very interesting accuracy results. SVM, NB, RF, and KNN classifiers with good results can promote facial expression recognition performance. Then will come TREE and MLP as the last place.

5.4 Discussion

The main aim of the proposed work is to evaluate the performance of facial features-landmarks ASM based machine learning techniques by using the most common algorithms in FER. Our goal is to select best machine learning model and achieve high accurate results. The summary of the work is described below:

(1) Classification of facial expression is one of the main issues of computer vision can be a complex task for machines. Therefore, machine-learning techniques are needed to recognize emotional expressions and improve accuracy. In this study, an analytical framework is developed to identify the best classifier, and the landmarks ASM based classification was implemented and tested using the same dataset

(2) The proposed work identified the best facial expressions classifier with experimental testing and evaluation of every classifier. In Figures 5-7 and 8 and Table 2-5 and Table 6, the performance comparison was done using considered metrics such as total accuracy (Acc), Precision (Pre), Recall (Rec), and F1Score (F1s) by using each feature vector size separately. From Table 1, it is clear that the DA classifier is the best classifier.

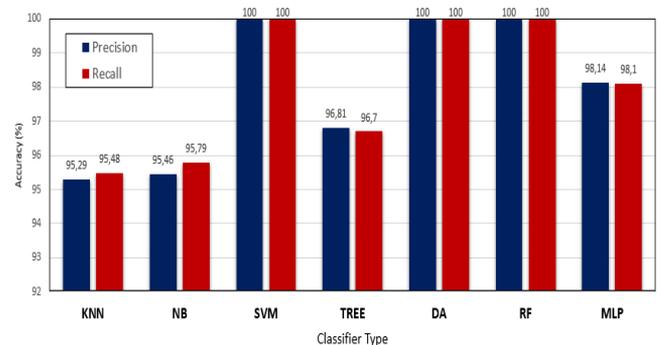


Figure 6. Performance comparison between seven classifiers

Table 7. Performance comparison between seven classifiers using *Test vs training* Approach

Size		7	14	21	28	35	42	49	56	63	70
KNN	Acc	100	92.85	95.23	89.28	88.57	78.57	77.55	69.64	65.07	62.85
	Pre	100	92.85	95.23	89.28	88.57	78.57	77.55	69.64	65.07	62.85
	Rec	100	95.23	96.42	92.38	91.83	74.14	74.28	64.76	55.71	53.84
	F1s	100	94.03	95.82	90.80	90.17	76.29	75.88	67.12	60.03	58.00
NB	Acc	100	100	90.47	89.28	82.85	76.19	75.51	66.07	61.90	60.00
	Pre	100	100	90.47	89.28	82.85	76.19	75.51	66.07	61.90	60.00
	Rec	100	100	92.85	90.71	86.59	74.48	74.28	64.76	57.56	55.13
	F1s	100	100	91.65	89.99	84.68	75.33	74.89	65.41	59.65	57.46
SVM	Acc	100	100	95.23	82.14	85.71	83.33	75.51	64.28	61.90	58.57
	Pre	100	100	95.23	82.14	85.71	83.33	75.51	64.28	61.90	58.57
	Rec	100	100	96.42	86.42	87.44	78.57	71.59	70.30	55.84	53.23
	F1s	100	100	95.82	84.23	86.57	80.88	73.49	67.15	58.71	55.77
TREE	Acc	71.42	92.85	71.42	67.85	71.42	66.66	59.18	50.00	53.96	54.28
	Pre	71.42	92.85	71.42	67.85	71.42	66.66	59.18	50.00	53.96	54.28
	Rec	57.14	95.23	78.57	74.04	75.79	65.60	69.82	56.25	51.12	48.14
	F1s	63.49	94.03	74.82	70.81	73.54	66.13	64.06	52.94	52.50	51.03
DA	Acc	100	100	100	96.42	94.28	85.71	77.55	73.21	71.42	68.57
	Pre	100	100	100	96.42	94.28	85.71	77.55	73.21	71.42	68.57
	Rec	100	100	100	97.14	95.91	78.57	74.28	71.90	57.14	54.76
	F1s	100	100	100	96.78	95.09	81.98	75.88	72.55	63.49	60.89
RF	Acc	100	100	90.47	85.71	88.57	71.42	69.38	67.85	58.73	60.00
	Pre	100	100	90.47	85.71	88.57	71.42	69.38	67.85	58.73	60.00
	Rec	100	100	92.85	87.14	91.15	73.94	72.02	64.76	56.72	54.76
	F1s	100	100	91.65	86.42	89.84	72.66	70.68	66.27	57.70	57.26
MLP	Acc	71.42	50.00	57.14	53.57	60.00	69.04	53.06	58.92	52.38	54.28
	Pre	71.42	50.00	57.14	53.57	60.00	69.04	53.06	58.92	52.38	54.28
	Rec	57.14	40.47	47.85	47.78	52.55	56.29	46.25	53.90	51.36	42.92
	F1s	63.49	44.73	52.08	50.50	56.02	62.02	49.42	56.30	51.86	47.94

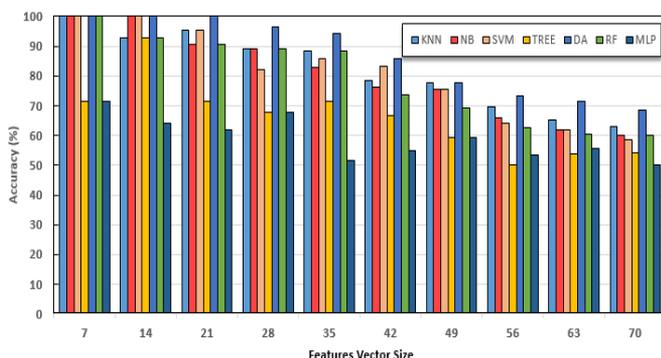


Figure 7. Accuracy comparison between seven classifiers

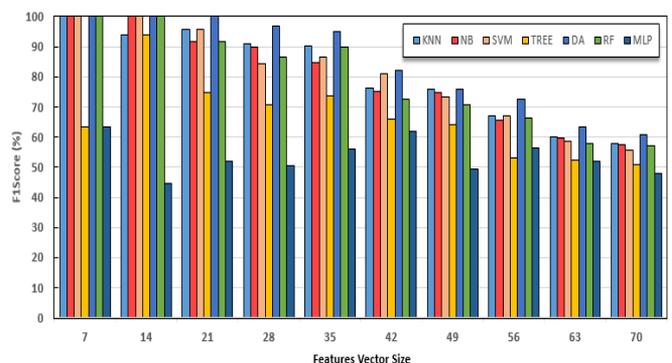


Figure 8. Fscore comparison between seven classifiers

(3) The proposed work considered two new classification processes one vs others and test vs training. It is shown in Table 1 and 7 with considered two classification processes. It is clearly indicated that DA classifier outperforms the other classifiers. This task of recognizing emotional expressions by extracting facial features-landmarks achieved at 0.00185 seconds using DA classifier with 136 size-features.

6. CONCLUSION

In this paper, we have presented fast and efficient simple facial expression recognition model. The model is based on extracting the landmarks to simulate the geometric shapes of facial expressions. They are defined through Active Shape Model (ASM). The proposed approach is validated and tested with new classification approach on the same dataset (i.e. CK+ dataset) using seven classifiers models among the most common algorithms in FER. Experimental results have shown that the Quadratic Analysis (DA) provides best accuracy

results while ensuring a lowest execution time. A comparison study using several classifiers shows that the recognition rate depending on three main factors: features extraction method, the classifier quality and the size of the learning set. A large learning dataset means high accurate results. Future work can focus on the study of other features extraction methods combined with deep learning machines using real-time data in FER systems.

REFERENCES

- [1] Ekman, P., Friesen, W.V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2): 124-129. <https://doi.org/10.1037/h0030377>
- [2] Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3-4): 169-200. <https://doi.org/10.1080/02699939208411068>
- [3] Ekman, P. (2004). *Emotions Revealed*. 2nd. Holt

- Paperbacks.
- [4] Tian, Y.I., Kanade, T., Cohn, J.F. (2001). Recognizing action units for facial expression analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2): 97-115. <https://doi.org/10.1109/34.908962>
- [5] Michel, P., El Kaliouby, R. (2003). Real time facial expression recognition in video using support vector machines. In *Proceedings of the 5th International Conference on Multimodal Interfaces*, pp. 258-264. <https://doi.org/10.1145/958432.958479>
- [6] Chuang, C.F., Shih, F.Y. (2006). Recognizing facial action units using independent component analysis and support vector machine. *Pattern Recognition*, 39(9): 1795-1798. <https://doi.org/10.1016/j.patcog.2006.03.017>
- [7] Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *APA PsycArticles*, 115(2): 268-287. <https://doi.org/10.1037/0033-2909.115.2.268>
- [8] Black, M.J., Yacoob, Y. (1997). Recognizing facial expressions in image sequences using local parameterized models of image motion. *International Journal of Computer Vision*, 25(1): 23-48. <https://doi.org/10.1023/A:1007977618277>
- [9] Kobayashi, H., Hara, F. (1997). Facial interaction between animated 3D face robot and human beings. In *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, Orlando, FL, USA, USA, 4: 3732-3737. <https://doi.org/10.1109/ICSMC.1997.633250>
- [10] Zhang, Z. (1999). Feature-based facial expression recognition: Sensitivity analysis and experiments with a multilayer perceptron. *International Journal of Pattern Recognition and Artificial Intelligence*, 13(6): 893-911. <https://doi.org/10.1142/S0218001499000495>
- [11] Fasel, B., Monay, F., Gatica-Perez, D. (2004). Latent semantic analysis of facial action codes for automatic facial expression recognition. In *Proceedings of the 6th ACM SIGMM International Workshop on Multimedia Information Retrieval*, pp. 181-188. <https://doi.org/10.1145/1026711.1026742>
- [12] Sebe, N., Lew, M.S., Sun, Y., Cohen, I., Gevers, T., Huang, T.S. (2007). Authentic facial expression analysis. *Image and Vision Computing*, 25(12): 1856-1863. <https://doi.org/10.1016/j.imavis.2005.12.021>
- [13] Sun, Y., Li, Z., Tang, C., Zhou, W., Jiang, R. (2009). An evolving neural network for authentic emotion classification. *2009 Fifth International Conference on Natural Computation*, Tianjin, pp. 109-113. <https://doi.org/10.1109/ICNC.2009.310>
- [14] Youssif, A.A., Asker, W.A. (2011). Automatic facial expression recognition system based on geometric and appearance features. *Computer and Information Science*, 4(2): 115. <https://doi.org/10.5539/cis.v4n2p115>
- [15] Kong, F. (2019). Facial expression recognition method based on deep convolutional neural network combined with improved LBP features. *Personal and Ubiquitous Computing*, 23(3-4): 531-539.
- [16] Chuang, C.F., Shih, F.Y. (2006). Recognizing facial action units using independent component analysis and support vector machine. *Pattern Recognition*, 39(9): 1795-1798. <https://doi.org/10.1016/j.patcog.2006.03.017>
- [17] Littlewort, G., Fasel, I., Bartlett, M.S., Movellan, J.R. (2002). Fully automatic coding of basic expressions from video. University of California, San Diego, San Diego, CA, 92093, 14.
- [18] Bartlett, M.S., Littlewort, G., Fasel, I., Movellan, J.R. (2003). Real time face detection and facial expression recognition: Development and applications to human computer interaction. *2003 Conference on Computer Vision and Pattern Recognition Workshop*, Madison, Wisconsin, USA, pp. 53-53. <https://doi.org/10.1109/CVPRW.2003.10057>
- [19] Kotsia, I., Nikolaidis, N., Pitas, I. (2007). Facial expression recognition in videos using a novel multi-class support vector machines variant. *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, Honolulu, HI, pp. II-585-II-588. <https://doi.org/10.1109/ICASSP.2007.366303>
- [20] Ghimire, D., Lee, J. (2013). Geometric feature-based facial expression recognition in image sequences using multi-class AdaBoost and support vector machines. *Sensors*, 13(6): 7714-7734. <https://doi.org/10.3390/s130607714>
- [21] Lucey, P., Cohn, J.F., 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. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, San Francisco, CA, pp. 94-101. <https://doi.org/10.1109/CVPRW.2010.5543262>
- [22] Chang, Y., Hu, C., Feris, R., Turk, M. (2006). Manifold based analysis of facial expression. *Image and Vision Computing*, 24(6): 605-614. <https://doi.org/10.1016/j.imavis.2005.08.006>
- [23] Valstar, M., Pantic, M. (2006). Fully automatic facial action unit detection and temporal analysis. In *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, pp. 149-149. <https://doi.org/10.1109/CVPRW.2006.85>
- [24] De la Torre, F., Campoy, J., Ambadar, Z., Cohn, J.F. (2007). Temporal segmentation of facial behavior. In *2007 IEEE 11th International Conference on Computer Vision*, pp. 1-8. <https://doi.org/10.1109/ICCV.2007.4408961>
- [25] Khanam, A., Shafiq, M.Z., Akram, M.U. (2008). Fuzzy based facial expression recognition. In *2008 Congress on Image and Signal Processing*, pp. 598-602. <https://doi.org/10.1109/CISP.2008.694>
- [26] Bhattacharjee, D., Basu, D.K., Nasipuri, M., Kundu, M. (2010). Human face recognition using fuzzy multilayer perceptron. *Soft Computing*, 14(6): 559-570. <https://doi.org/10.1007/s00500-009-0426-0>
- [27] Martin, C., Werner, U., Gross, H.M. (2008). A real-time facial expression recognition system based on active appearance models using gray images and edge images. In *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition*, pp. 1-6. <https://doi.org/10.1109/AFGR.2008.4813412>
- [28] Cheon, Y., Kim, D. (2009). Natural facial expression recognition using differential-AAM and manifold learning. *Pattern Recognition*, 42(7): 1340-1350. <https://doi.org/10.1016/j.patcog.2008.10.010>
- [29] Almaev, T.R., Valstar, M.F. (2013). Local Gabor binary patterns from three orthogonal planes for automatic

- facial expression recognition. 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, Geneva, pp. 356-361. <https://doi.org/10.1109/ACII.2013.65>
- [30] Yuan, L.U.O., Wu, C.M., Zhang, Y. (2013). Facial expression feature extraction using hybrid PCA and LBP. *The Journal of China Universities of Posts and Telecommunications*, 20(2): 120-124. [https://doi.org/10.1016/S1005-8885\(13\)60038-2](https://doi.org/10.1016/S1005-8885(13)60038-2)
- [31] Happy, S.L., Routray, A. (2014). Automatic facial expression recognition using features of salient facial patches. *IEEE transactions on Affective Computing*, 6(1): 1-12. <https://doi.org/10.1109/TAFFC.2014.2386334>
- [32] Chen, J., Chen, Z., Chi, Z., Fu, H. (2014). Facial expression recognition based on facial components detection and hog features. In *International Workshops on Electrical and Computer Engineering Subfields*, pp. 884-888.
- [33] Barman, A., Dutta, P. (2017). Facial expression recognition using distance and shape signature features. *Pattern Recognition Letters*. <https://doi.org/10.1016/j.patrec.2017.06.018>
- [34] Barman, A., Dutta, P. (2019). Facial expression recognition using distance and texture signature relevant features. *Applied Soft Computing*, 77: 88-105.
- [35] Barman, A., Dutta, P. (2019). Influence of shape and texture features on facial expression recognition. *IET Image Processing*, 13(8): 1349-1363. <https://doi.org/10.1049/iet-ipr.2018.5481>
- [36] Viola, P., Jones, M. (2001). Robust real-time object detection. *International Journal of Computer Vision*, 4(34-47): 4.
- [37] Gower, J.C. (1975). Generalized Procrustes analysis. *Psychometrika*, 40(1): 33-51. <https://doi.org/10.1007/BF02291478>
- [38] Ten Berge, J.M. (1977). Orthogonal Procrustes rotation for two or more matrices. *Psychometrika*, 42(2): 267-276. <https://doi.org/10.1007/BF02294053>
- [39] Sagonas, C., Antonakos, E., Tzimiropoulos, G., Zafeiriou, S., Pantic, M. (2016). 300 faces in-the-wild challenge: Database and results. *Image and Vision Computing*, 47: 3-18. <https://doi.org/10.1016/j.imavis.2016.01.002>
- [40] Tzimiropoulos, G., Pantic, M. (2013). Optimization problems for fast AAM fitting in-the-wild. 2013 IEEE International Conference on Computer Vision, Sydney, NSW, pp. 593-600. <https://doi.org/10.1109/ICCV.2013.79>