International Journal of Intelligent Information Technologies
Volume 17 • Issue 3 • July-September 2021

Performance Evaluation of Machine Learning for Recognizing Human Facial Emotions

Alti Adel, LRSD Laboratory, Department of Computer Science, Faculty of Sciences, University of Ferhat Abbas SETIF-1, Sétif, Algeria

https://orcid.org/0000-0001-8348-1679

Ayeche Farid, LMETR - E1764200, Optics and Precision Mechanics Institute, University of Setif 1, Algeria

https://orcid.org/0000-0002-1556-1156

ABSTRACT

Facial expression recognition is a human emotion classification problem attracting much attention from scientific research. Classifying human emotions can be a challenging task for machines. However, more accurate results and less execution time are still the issues when extracting features of human emotions. To cope with these challenges, the authors propose an automatic system that provides users with a well-adopted classifier for recognizing facial expressions in a more accurate manner. The system is based on two fundamental machine learning stages, namely feature selection and feature classification. Feature selection is realized by active shape model (ASM) composed of landmarks while the feature classification algorithm is based on seven well-known classifiers. The authors have used CK+ dataset, implemented and tested seven classifiers to find the best classifier. The experimental results show that quadratic classifier (DA) provides excellent performance, and it outperforms the other classifiers with the highest recognition rate of 100% for the same dataset.

KEYWORDS

Active Shape Model, Generalized Procrust Analysis, Human Facial Emotions, Machine Learning, Quadratic Classifier

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 human 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

DOI: 10.4018/IJIIT.2021070105

standard set of seven emotion 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 (Ekman, 2004). Some researchers consider the neutral face as a seventh expression (Tian, Kanade & Cohn, 2001; Michel & Kaliouby, 2003; Chuang & Shih, 2006; Barman, & Dutta, 2019). Despite their powerful benefits that provide in human-computer interaction systems, classifying human emotions can be a challenging task for machines. However, subpar accuracy and less execution time are there still the mins issues when extracting features of human emotions.

Generally, FER systems can 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. It will evaluate the performance and accuracy of seven classifies among the most common algorithms in FER, K-Nearest Neighbor (KNN), Naıve 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, showing how driving new suitable classifiers for those exploring automated emotion recognition via machine learning.

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 system on human emotions has been developed is the Facial Action Coding System (FACS) created by (Paul Ekman et al. 1994). This system based on human observations and manual labeling process. FACS can serve many researchers, particularly those with a psychological background and lack of concentration. It 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 & Yacoob, 1997) 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 is considered as a major objective to identify or detect human emotions within an appropriate classifier. According to (Kobayashi & Hara, 1997), 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.

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 in (Kong, 2019) 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 more complex features.

In (Chuang, & Shih., 2016), authors adopted 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 (Littlewort et al. 2002; Bartlett et al., 2003; Kotsia et al., 2007; Lei et al., 2008; Lucey et al., 2010).

The authors in (Yuan, Wu, & Zhang, 2013) proposed model based Local Binary Models (LBP) with Principal Component Analysis (PCA) to obtain more accurate facial features by improving the local and holistic facial characteristics in a merged manner. (Chang, Feris, & Turk, 2006) 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.

In order to examine in more detail variations in facial expressions over time, new expression recognition technique based-Gabor's wavelets have been proposed (Valstar, & Pantic, 2006). Due to the limited temporal segmentation of facial gestures in spontaneous facial behavior recorded in real contexts, (Torre et al., 2007) 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 in (Bhattacharjee et al., 2010) proposed a model to classify facial expressions based on fuzzy rules. The results shows that fuzzy rules predicts better non-linear overlapping classes than the Multi-Layer Perceptron (MLP) (Khanam, Shafiq, & Akram, 2008) 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, Werner, & Gross, 2008). 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 (Khanam, Shafiq, &Akram, 2008) 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.

In (Cheon, & Kim, 2009) proposed a model based on local Gabor to extract local dynamic characteristics and to detect facial action units in real-time. Moreover, authors in (Happy, & Routray, 2014) 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. In (Chen et al., 2014) authors focused on their efforts on detecting respective deformation characteristics of facial expressions by exploiting the characteristics of Histogram Oriented Gradients (HoG) of facial components.

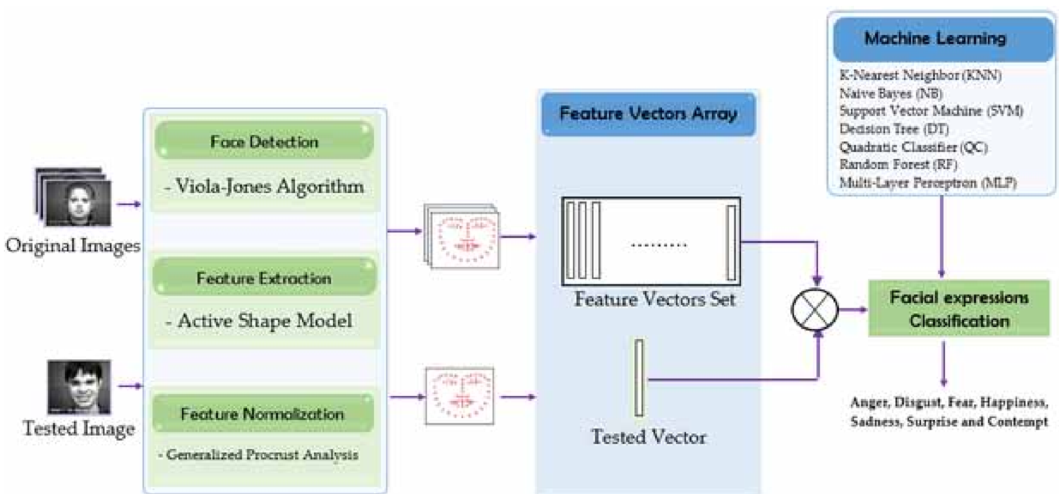
In (Barman, & Dutta, 2017) proposed model based Active Appearance Model (AAM) to enhance expression recognition performance. Then, the shape and distance signatures as well as statistical functionalities are the input data of the learning model. (Barman, & Dutta, 2019) present a core lightweight ontology 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 Appearance Model (AAM) 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.

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 its accuracy. Each machine learning has advantages and drawbacks. However most of them suffer from the poor accuracy and high execution time in most facial images with less 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 landmark and select the best classifier model with high accuracy and precision, reduced execution time that suits the facial expressions field. The general architecture of proposed framework is described in Figure 1.

Figure 1. General architecture of the proposed Framework



The Face Detection Stage

The *face detection* stage locates the face in input image using Viola-Jones algorithm. 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 (Viola, & Jones, 2001).

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 an optimal edges for a given object through geometric transformations. We used and adapted this algorithm for generating 68 landmarks to obtain a more accurate results. It was included in our framework after the face detection stage. The face is represented by a feature vector consisting of $\boldsymbol{n}$ point coordinates. The ASM model is trained with a set of 2D points mentioned manually form training images where each image consists of n references points:

$$q = \left[\left(x_1, y_1\right), \left(x_2, y_2\right), \ldots. \left(x_n, y_n\right)\right]^T, i = 1 \ldots n \tag{1}$$

The feature extraction process is described as the following steps:

Step 1: Computed the mean of data as follows:

$$\overline{q} = \frac{1}{q}\sum_{i=1}^{m} q_i \tag{2}$$

Step 2: Align the shapes after landmarking the corresponding points using the Procrustes Algorithm. The sum of instances is computed and reduced to the mean of each shape by:

$$min\sum_{i=1}^{m}\left(q - q_i\right)^2 \tag{3}$$

Step 3: Compute the covariance of data aligned shapes in the face image:

$$S = \frac{1}{m-1}\sum_{i=1}^{m}\left(q - q_i\right)\left(q - q_i\right)^T \tag{4}$$

Step 4: Compute the eigenvectors 'U_i' and eigenvalue *'i'* of the covariance matrix sorted in a descending order by eigenvalue size and remove the small eigenvalues while maintaining most of them (98%).We approximate any instance of the shape, including training examples, by projecting onto the first *'t'* eigen vectors:

$$q = q + \sum_{i=1}^{t} b_i * u_i \tag{5}$$

Step 5: The weight vector b is computed and identified shape's feature:

$$b = \left[b_1, b_2 \ldots b_t\right]^T \tag{6}$$

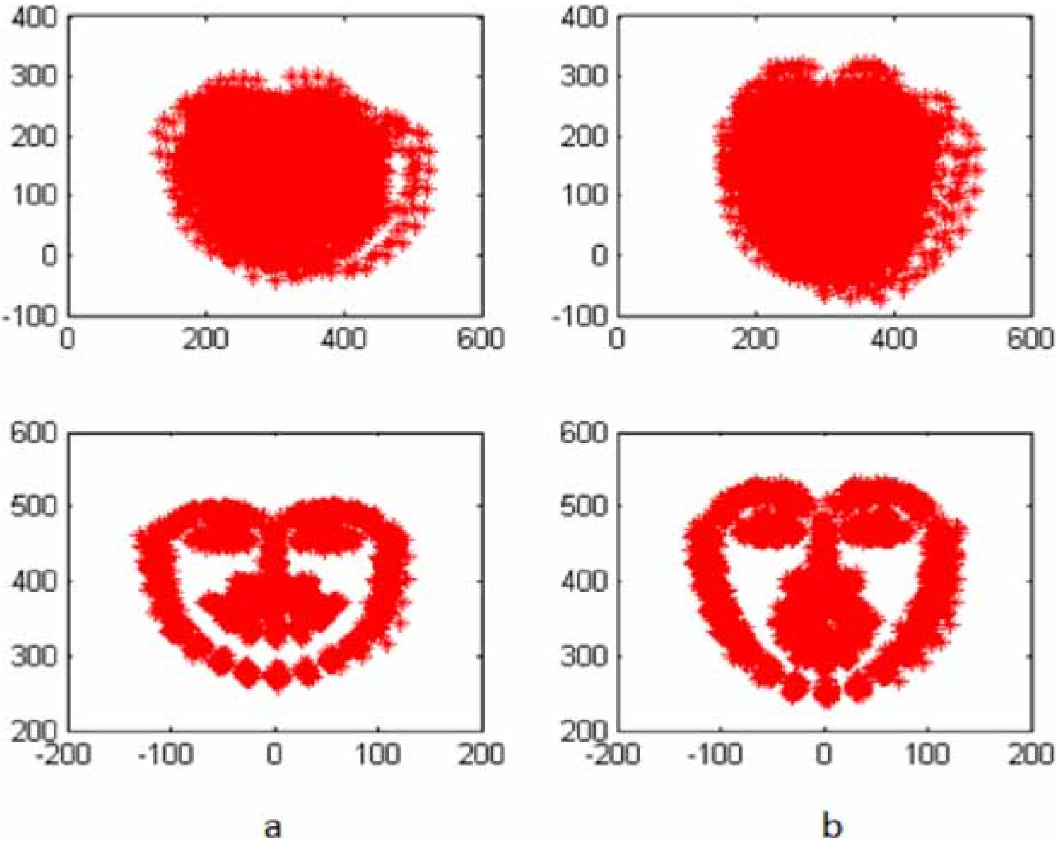
Step 6: Finally varying the two weights b_i enables us to explore the allowable variations in the shape.

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 informations.

Algorithm 1: Feature extraction process.

Figure 2. Extracted shapes before and after applying GPA on facial expressions, a: Happy, b: surprise




```
Input: m shapes' vectors.
Output: Optimal shape model.
Algorithm
Compute the mean of data q̄ of m shapes' vectors;
Compute and reduce the mean of each shape;
Compute the covariance S of data aligned shapes;
Compute the eigenvectors U_i and eigenvalue i of the covariance
matrix;
b=0;
Repeat
   Generate point's models according to Eq.5;
   Compute the translation, rotation and scaling parameters
   Update points models according to Eq.6;
Until all geometric parameters become stable
Return Optimal shape model.
```

This will significantly increase and improve the recognition rate for the proposed system. We used Generalized Procrust Analysis (GPA) proposed by (Gower., 1975) and enhanced in (Berge., 1977). Figure 2 shows an example of shapes before and after applying the GPA. Figure 3 shows an example of result of our system for extracting.

The Feature Vector Construction and Classification Stage

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

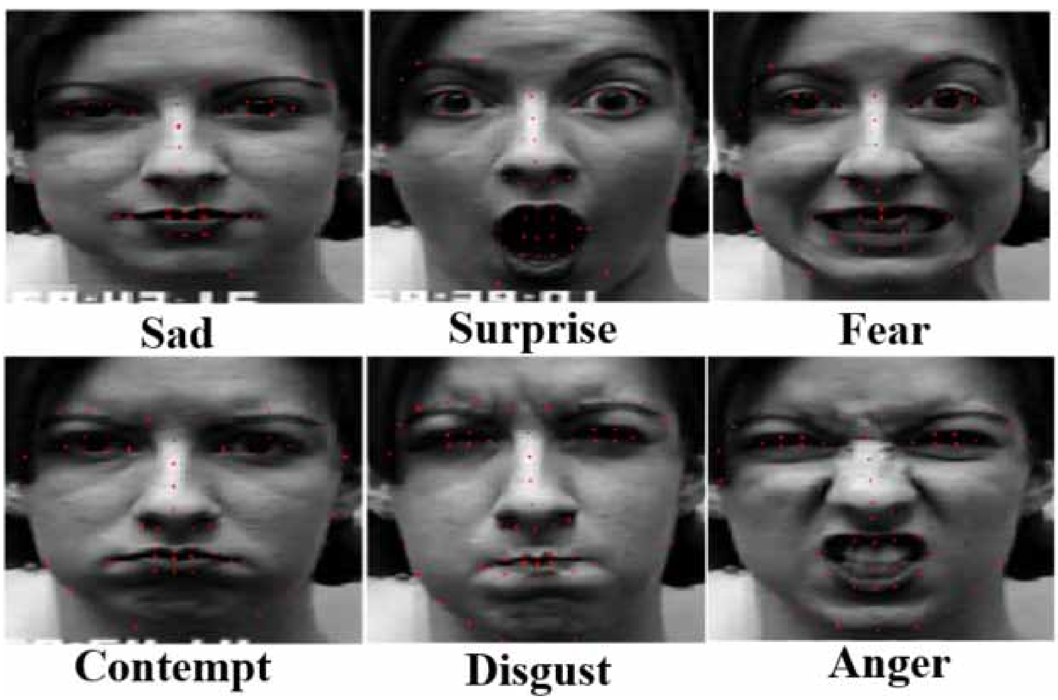
$$F_{i,j} = \left(f_1; f_1; \ldots; f_n\right)\ f = \left(x; y\right) \tag{7}$$

where:

- ***i*** and j denote the i^{th} frame defined for the j^{th} facial expression;
- ***n*** is the number of extracted marks;
- **f = (x, y)** is the Cartesian coordinate frame ***F***. In the presented work, we set n = 68.

Figure 3 shows an example of the landmarks extracted by our system on standard set of seven facial expressions. We conduct several experiments with the most common algorithm and compare its accuracy and execution time.

Figure 3. Samples of facial expression landmarks images result of our system in CK+ database



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):** A supervised machine learning algorithm. KNN performs better on patterns in small regions of an image, such as the curve of an eyebrow.
- **Naive Bayes (NB):** A learning algorithm based on the Byes theorem. Naive Bayes enables to classify a set of observations according to rules determined by the algorithm itself.
- **Multiclass Support Vector Machine (SVM):** Supervised learning algorithms that analyze and classify data by specifying a sample class, or regression. They perform well when classifying human facial expressions with consistent head poses and illumination.
- **Decision Tree (DT):** A decision support tool based on shape-to-tree graph. A hierarchical representation of the data structure in the form of decision sequences (tests) is used to predict a class. Tests are performed in the internal nodes and decisions are made in the leaf nodes.
- **Quadratic Classifier (QC):** A more general version of the linear classifier and use a quadratic decision surface to separate the measurements of two or more classes of objects or events.
- **Random Forest (RF):** Exactly the same principle of the decision tree. It randomly selects observations and specific characteristics to build several decision trees and merge them to get more accuracy results and stable prediction.
- Multi-Layer Perceptron (MLP): Class of artificial neural networks and use dynamic temporal behavior when classifying an image. *Multi-Layer Perceptron* are highly connected networks of elementary processors that process more granular elements within an image, making them better at distinguishing between two similar emotion classifications.

DATASET AND TRAINING

Dataset

For our experiments, we used the extended Cohn-Kanade (CK+) dataset (Sebe et al., 2007). The CK+ has 326 images of peak facial expressions for seven emotion categories are anger (AN), contempt (CON), 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 best quality to consider these properties of good dataset. Two other dataset are used to evaluate three classifiers SVM, KNN and DA, the first is the ORL (Sebe *et al.,* 2007) dataset used for facial recognition while the second is JAFFEE database (Sebe et al., 2007) used for facial expression recognition. We used ORL and JAFFEE datasets to evaluate SVM, KNN and DA classification methods to recognize faces and facial expressions. SVM is applied to facial images in "one-vs-one" while KNN is applied with k equal 3.

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:

- **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 features extraction is compete, 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.

- **Test vs Training:** We conduct in-depth experiments on our dataset using seven classifiers in both test and training face images. The dataset with a total of 326 images was split into 2 parts: 292 *for training* and *70 for testing*. We train seven classifiers with varying pairs {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 remain from the training set. Finally, we evaluate the performance and accuracy of FER using seven classifiers in all face images in the testing set.

Evaluation Metrics

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

- Precision (Pre) is the number true positives over predicted positives. The formula for Precision (Pre) is:

$$Pre = \frac{TP}{TP + FP} \tag{8}$$

- Recall (Rec) to the number of true positives over actual positives. The formula for Recall (Rec) is:

$$Rec = \frac{TP}{TP + FN} \tag{9}$$

- F1score to evaluate a weighted average of Pre and Rec. It is an important factor based on weighted recall. The F1-score is computed as follows:

$$F1s = 2 * \frac{Pre * Rec}{Pre + Rec} \tag{10}$$

System Configuration

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.

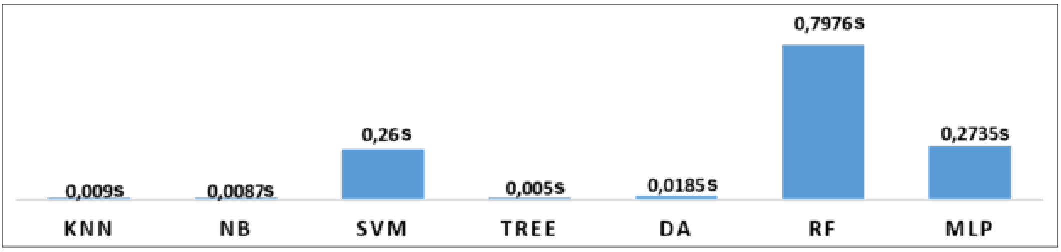
PERFORMANCE COMPARISONS

To achieve in-depth comparison 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.

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. Figure 4 show the execution time comparison for seven classifiers trained on CK+ dataset. For RF, KNN, and NB show superior improvement in terms of the execution time. The FR classifier yields lower execution time than other classifiers except for SVM that use linked test and training stages. The size of our features vector was smaller of 136 than other techniques and descriptors LBP, HOG… etc.

Figure 4. Execution time comparison between seven classifiers



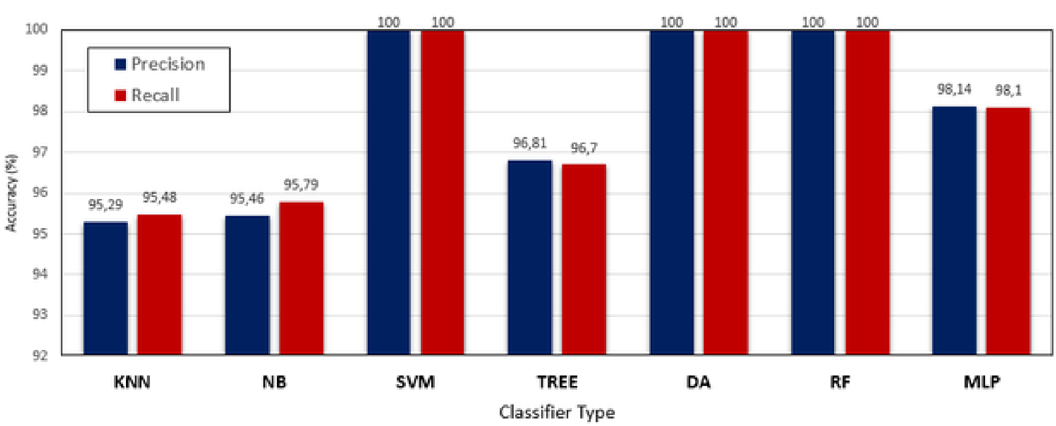
One vs Others

Table 1 shows the accuracy, precision, recall and f-score for the seven classifiers trained on CK+ dataset. As expected, SVM, DA and FR presents an ideal recognition rate 100% which prove 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	95.09
NB	96.32	95.79	95.13	95.46
SVM	**100.00**	**100.00**	**100.00**	**100.00**
TREE	97.55	96.70	96.94	96.81
DA	**100.00**	**100.00**	**100.00**	**100.00**
RF	**100.00**	**100.00**	**100.00**	**100.00**
MLP	98.16	98.10	98.19	98.14

Figure 5. Performance comparison between seven classifiers



The confusion matrix is used to evaluate the performance of seven classifiers on CK+ dataset. Table 2, 3, 4 and 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.

Table 2. Confusion Matrix of CK+ Dataset using SVM, DA and RF

	Anger	Contempt	Disgust	Fear	Happiness	Sandiness	Surprise
Anger	**100**	0.00	3.39	0.00	0.00	0.00	0.00
Contempt	0.00	**100**	0.00	3.22	0.00	3.57	0.00
Disgust	2.22	0.00	**100**	0.00	3.33	0.00	0.00
Fear	0.00	5,56	0.00	**100**	0.00	3.57	0.00
Happiness	0.00	0.00	3.13	6.45	**100**	0.00	0.00
Sandiness	2.22	5,56	6.25	3.20	0.00	**100**	0.00
Surprise	0.00	5,56	0.00	0.00	0.00	0.00	**100**

Table 3. Confusion Matrix of CK+ Dataset using KNN

	Anger	Contempt	Disgust	Fear	Happiness	Sandiness	Surprise
Anger	**95.56**	0.00	3.39	0.00	0.00	0.00	0.00
Contempt	0.00	**94.44**	0.00	3.22	0.00	3.57	0.00
Disgust	2.22	0.00	**98.30**	0.00	3.33	0.00	0.00
Fear	0.00	5,56	0.00	**92.00**	0.00	3.57	0.00
Happiness	0.00	0.00	3.13	6.45	**100**	0.00	0.00
Sandiness	2.22	5,56	6.25	3.20	0.00	**89.29**	0.00
Surprise	0.00	5,56	0.00	0.00	0.00	0.00	**98.78**

Table 4. Confusion Matrix of CK+ Dataset using NB

	Anger	Contempt	Disgust	Fear	Happiness	Sandiness	Surprise
Anger	**93.33**	0.00	5,08	0.00	0.00	0.00	0.00
Contempt	0.00	**94.44**	0	4	0.00	0.00	0.00
Disgust	8,89	0.00	**93.22**	0.00	0.00	0.00	0.00
Fear	0.00	0.00	0.00	**92.00**	0.00	7,14	0.00
Happiness	0.00	0.00	0.00	0.00	**100.00**	0	0.00
Sandiness	0.00	0.00	0.00	0.00	0.00	**100.00**	0.00
Surprise	0.00	5,56	0.00	4	0.00	0.00	**97.56**

Table 5. Confusion Matrix of CK+ Dataset using TREE

	Anger	Contempt	Disgust	Fear	Happiness	Sandiness	Surprise
Anger	**100**	0.00	0.00	0.00	0.00	0.00	0.00
Contempt	0.00	**100**	0.00	0.00	0.00	0.00	0.00
Disgust	2.22	0	**98.30**	0.00	0.00	0.00	0.00
Fear	2.22	5.56	0.00	**92.00**	0.00	0.00	0.00
Happiness	2.22	0.00	0.00	0.00	**98.55**	0.00	0.00
Sandiness	4.45	0.00	0.00	0.00	1.45	**89.29**	0.00
Surprise	0.00	5.56	0.00	0.00	0.00	0.00	**98.78**

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 classifier come in the last.

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 show the accuracy results of the proposed approach for ten pairs. Moreover, while analyzing the results, we obtain that the DA classifier present very interesting accuracy results. SVM, NB, RF and KNN classifiers with good results can promote facial expression recognition performance. Then will come TREE and NN as the last place.

We evaluated the efficiency of three classifiers SVM, KNN and DA in terms of accuracy rate. Table 8 shows the recognition rate of each classifier. The results show that the SVM method gives a recognition accuracy equal to 98.50% although the recognition accuracy with KNN is still good, 94.50% and DA with 98.45%. As it can be observed that the highest precision for facial recognition is recorded for DA. This is due to the classification power of DA.

We also evaluated the accuracy rate of three classifiers SVM, KNN and DA on JAFFEE facial expression database. Table 9 shows the accuracy rates of each classification method in the JAFFEE database. We noticed that SVM classifier achieves a higher accuracy rate than KNN and DA of 84.28% in the database of facial expressions JAFFEE.

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

Table 6. Confusion Matrix of CK+ Dataset using MLP

	Anger	Contempt	Disgust	Fear	Happiness	Sandiness	Surprise
Anger	**100**	0.00	0.00	0.00	0.00	0.00	0.00
Contempt	2.22	**94.44**	0.00	0.00	0.00	0.00	0.00
Disgust	4.44	0.00	**96.61**	0.00	0.00	0.00	0.00
Fear	0.00	0.00	0.00	**100**	0.00	0.00	0.00
Happiness	0.00	0.00	1.69	0.00	**95.65**	3.57	1.22
Sandiness	0.00	0.00	0.00	0.00	0.00	**100**	0.00
Surprise	0.00	0.00	0.00	0.00	0.00	0.00	**100**

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

select best machine learning model and achieve more accurate results. The summary of the work is described below:

- 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. The proposed work identified the best facial expressions classifier with experimental testing and evaluation of every classifier. In Figures 4, 5, 6, 7 and 8 and Table 1, 2, 3 and 4, the performance comparison was done using considered metrics such as total accuracy

Figure 6. Accuracy comparison between seven classifiers

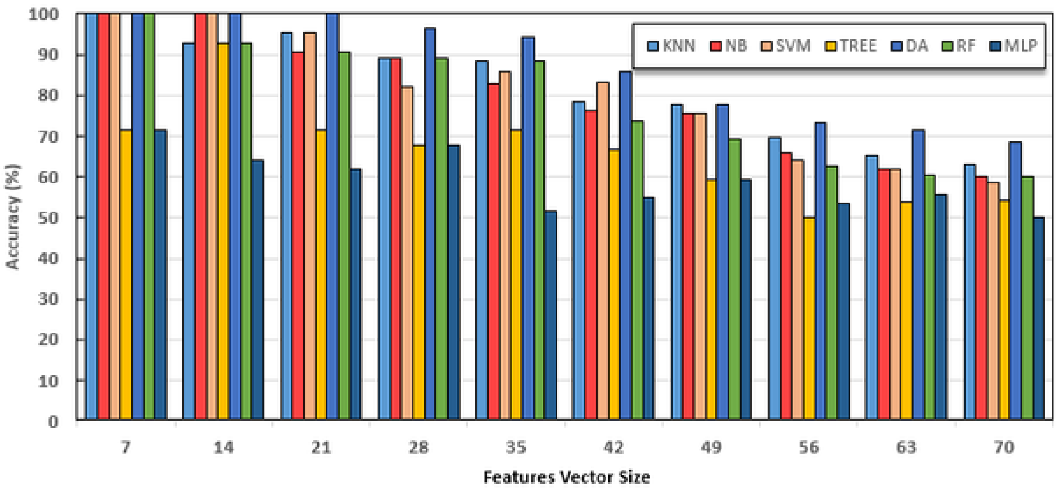


Figure 7. F-score comparison between seven classifiers

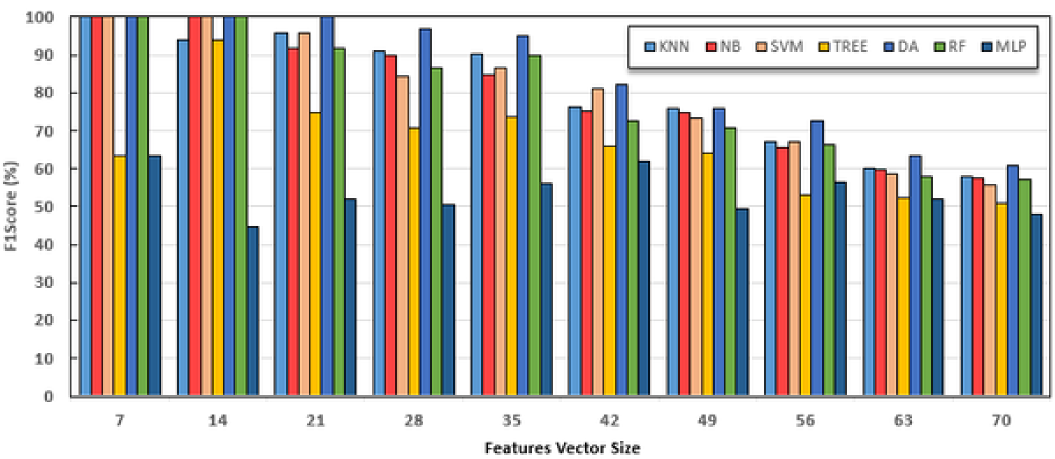


Table 8. Precision comparison between three classifiers using one vs one on ORL dataset

Classifier	Accuracy
SVM	98.50
KNN	94.50
DA	**98.45**

Table 9. Precision comparison between three classifiers using one vs one on ORL dataset

Classifier	Accuracy
SVM	**84.28**
KNN	72.83
DA	71.42

(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.

- 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** features.

CONCLUSION

In this paper, we have proposed fast and efficient simple facial expression recognition model. It is based on the extraction of landmarks to simulate the geometric shapes of facial expressions. They are defined through Active Shape Model (ASM). Our proposed approach is concretely validated and tested with two new classification approaches 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 more 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

Barman, A., & Dutta, P. (2019). Facial expression recognition using distance and texture signature relevant features. *Applied Soft Computing*, *77*, 88–105. doi:10.1016/j.asoc.2019.01.011

Barman, A., & Dutta, P. (2019). Influence of shape and texture features on facial expression recognition. *IET Image Processing*, *13*(8), 1349–1363. doi:10.1049/iet-ipr.2018.5481

Bartlett, M. S., Littlewort, G., Fasel, I., & Movellan, J. R. (2003, June). Real time face detection and facial expression recognition: development and applications to human computer interaction. In *2003 Conference on computer vision and pattern recognition workshop* (Vol. 5, pp. 53-53). IEEE.doi:10.1109/CVPRW.2003.10057

Bhattacharjee, D., Basu, D. K., Nasipuri, M., & Kundu, M. (2010). Human face recognition using fuzzy multilayer perceptron. *Soft Computing*, *14*(6), 559–570. doi:10.1007/s00500-009-0426-0

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. doi:10.1023/A:1007977618277

Chang, Y., Hu, C., Feris, R., & Turk, M. (2006). Manifold based analysis of facial expression. *Image and Vision Computing*, *24*(6), 605–614. doi:10.1016/j.imavis.2005.08.006

Chen, J., Chen, Z., Chi, Z., & Fu, H. (2014, August). Facial expression recognition based on facial components detection and hog features. In International workshops on electrical and computer engineering subfields (pp. 884-888). Academic Press.

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. doi:10.1016/j.patcog.2006.03.017

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).doi:10.1109/ICCV.2007.4408961

Ekman, P. (1994). *Strong evidence for universals in facial expressions: a reply to Russell's mistaken critique*. Academic Press.

Ghimire, D., & Lee, J. (2013). Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines. *Sensors (Basel)*, *13*(6), 7714–7734. doi:10.3390/s130607714 PMID:23771158

Gower, J. C. (1975). Generalized Procrustes Analysis. *Psychometrika*, *40*(1), 33–51. doi:10.1007/BF02291478

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. doi:10.1109/TAFFC.2014.2386334

Khanam, A., Shafiq, M. Z., & Akram, M. U. (2008, May). Fuzzy based facial expression recognition. In 2008 Congress on Image and Signal Processing (Vol. 1). Academic Press.

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* (Vol. 4, pp. 3732-3737).doi:10.1109/ICSMC.1997.633250

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. doi:10.1007/s00779-019-01238-9

Kotsia, I., Nikolaidis, N., & Pitas, I. (2007, April). Facial expression recognition in videos using a novel multi-class support vector machines variant. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07* (Vol. 2, pp. II-585). doi:10.1109/ICASSP.2007.366303

Littlewort, G., Fasel, I., Bartlett, M. S., & Movellan, J. R. (2002). Fully automatic coding of basic expressions from video. University of California, San Diego.

Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010, June). *The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In 2010 IEEE computer society conference on computer vision and pattern recognition-workshops*. IEEE.

Martin, C., Werner, U., & Gross, H. M. (2008, September). 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).doi:10.1109/AFGR.2008.4813412

Michel, P., & El Kaliouby, R. (2003, November). Real time facial expression recognition in video using support vector machines. In *Proceedings of the 5th international conference on Multimodal interfaces* (pp. 258-264). doi:10.1145/958432.958479

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. doi:10.1016/j.imavis.2016.01.002

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. doi:10.1016/j.imavis.2005.12.021

Ten Berge, J. M. (1977). Orthogonal Procrustes rotation for two or more matrices. *Psychometrika*, *42*(2), 267–276. doi:10.1007/BF02294053

The JAFFE Database. (n.d.). https://zenodo.org/record/3451524#.X3irdcIzbIV

The ORL Database. (n.d.). http://cam-orl.co.uk/facedatabase.html

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. doi:10.1109/34.908962 PMID:25210210

Valstar, M., & Pantic, M. (2006, June). Fully automatic facial action unit detection and temporal analysis. In *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)* (pp. 149-149).doi:10.1109/CVPRW.2006.85

Wang, M., Iwai, Y., & Yachida, M. (1998). Expression recognition from time-sequential facial images by use of expression change model. In *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 324-329).doi:10.1109/AFGR.1998.670969

Yuan, L. U. O., Wu, C. M., & Zhang, Y. (2013). Facial expression feature extraction using hybrid PCA and LBP. *Journal of China Universities of Posts and Telecommunications*, *20*(2), 120–124. doi:10.1016/S1005-8885(13)60038-2

Adel Alti obtained a Master's degree from the University of Setif (UFAS), Algeria, in 1998. He obtained a Ph.D. degree in software engineering from UFAS university of Sétif, Algeria, 2011. Right now, he is an associate professor, HDR at University of Sétif. He is a header of the Smart Semantic Context-aware Services research group LRSD. His areas of interest include mobility, ambient, pervasive, and ubiquitous computing, automated software engineering, mapping multimedia concepts into UML, semantic integration of architectural description into MDA platforms, context-aware quality software architectures and automated service management, Context and QoS. During his work he has published number of publications concerning these subjects.