Structured Learning and Prediction in Facial Emotion Classification and Recognition

Khalid Ounachad, Mohamed Oualla, Abdelghani Souhar, Abdelalim Sadiq

Abstract: Structured prediction methods have become, in recent years, an attractive tool for many machine-learning applications especially in the image processing area as in customers satisfaction prediction by using facial recognition systems, in criminal investigations based on face sketches recognition, in aid to autistic children and so. The main objective of this paper is the identification of the emotion of the human being, based on their facial expressions, by applying structured learning and perfect face ratios. The basic idea of our approach is to extract the perfect face ratios from a facial emotion image as the features, this face emotional images are labeled with their kind of emotions (the seven emotions defined in literature). For this end, first we determined sixty-eight landmarks point of image faces, next we applied a new deep geometric descriptor to calculate sixteen features representing the emotional face. The training and the testing tasks are applied to the Warsaw dataset: The Set of Emotional Facial Expression Pictures (WSEFEP) dataset. Our proposed approach can be also applied in others competitor facial emotion datasets. Based on experiments, the evaluation demonstrates the satisfactory performance of our applied method, the recognition rate reaches more than 97% for all seven emotions studied and it exceeds 99.20% for neutral facial images.

Keywords: Supervised learning, Structured Learning, prediction, Facial Emotion Recognition, Perfect Face Ratios, Emotional Facial Expression, WSEFEP dataset.

I. INTRODUCTION

Emotion is one of the most important semantic attributes of face. People can immediately recognize the emotion of a person. So, it is an easy task for humans to classify emotion but challenging task for machines. Emotion classification can be used as part of a face recognition process. The facial emotion recognition area (FER) became an enormously important field in the computer vision topic and also in the artificial intelligence topic owing to its utility in classic functional aspects of customer's purchasing behavior, human-computer interaction, identification of a person in the process of detection of the Criminal by using face sketches [1][2] in the criminology. In this paper We are interested in the detection of the Criminal in the criminology domain.

Facial Emotion Recognition systems (FERS) [3] aim to

recognize an emotion in a dataset of photos or sketches

images [4][5] (Fig.1). In the case of sketches, the task is particularly difficult: There doesn't exist a large dataset of training facial emotion image sketches and classifying 63
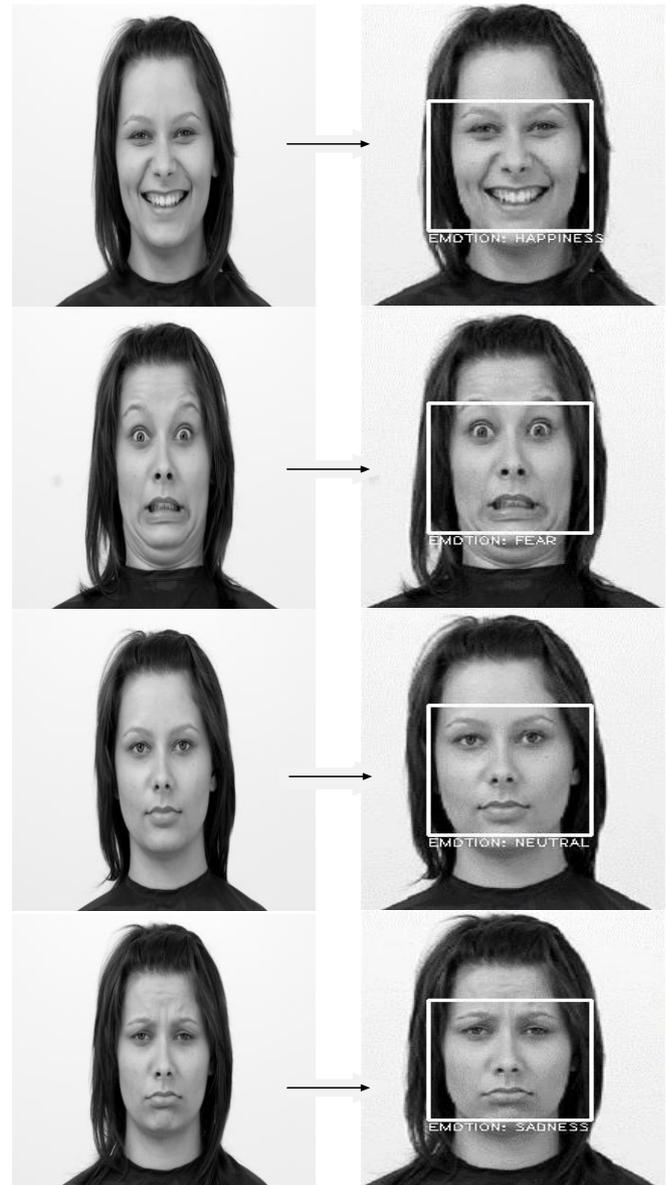
* Correspondence Author

Ounachad*, Ibn Tofail University, Faculty of sciences, Kenitra, Morocco. khalid.ounachad@uit.ac.ma

Oualla, Faculty of sciences and technology, Moulay Ismail University, Errachidia, Morocco. mohamedoualla76@gmail.com

Souhar, Ibn Tofail University, Faculty of sciences, Kenitra, Morocco. houssouhar@gmail.com

Sadiq, Ibn Tofail University, Faculty of sciences, Kenitra, Morocco. a.sadiq@uit.ac.ma

Retrieval Number: C6421029320/2020©BEIESP
DOI: 10.35940/ijeat.C6421.049420

152

Published By:
Blue Eyes Intelligence Engineering
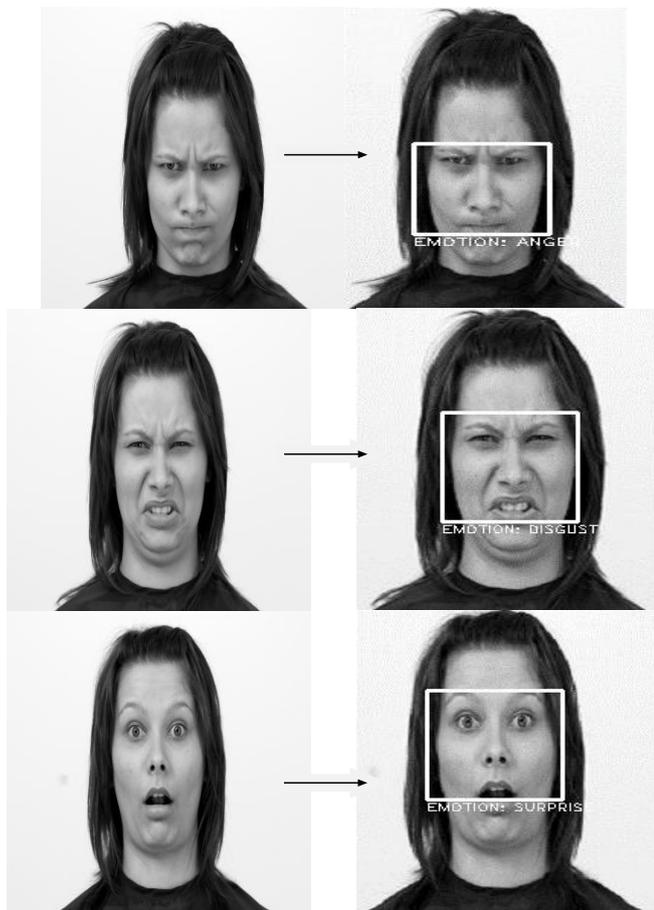*& Sciences Publication*

Fig. 1. An Input image photo to the structured learning system and the Outputs of the application, with the kind of the emotion superimposed over the subjects 'photo faces to indicate the detected emotion. Column 1 the input images and column 2 their output facial emotion images emotion can be also difficult.

The matching process is a general concept used to fully describe the task of the significant find of correspondences between two items by using some distances, that has to be accomplished for recognition features extracted or for a more complex comparison system. In a facial emotion recognition system matching is used to recognize the right kind of emotion using the precomputed features.

We tend to look for a mapping function f define as: $f: X \rightarrow Y$ and the aims is to map an input X to an output Y. When we modeling by using machine learning [8][9][10] methods, the form of the input X is varied, the output Y is a number or a label, such as the various methods of classification, regression, and distribution density estimation. In Structured Learning, Y is no longer limited to a number, but can be a complex structured object, such as an image, an image region, a label sequence [11], or parse tree and so on.

The principal objective of this article is originally to nominate a new facial emotion recognition approach based on structured learning and also the face ratios which allows a fast recognition of a facial emotion. The elementary and original idea is composed into two parts: the first one is to extract the perfect face ratios (PFR) for a query facial emotion and also to extract the PFR for each facial emotion in the dataset as the features; the second part is to consider the kind of emotion as the labels and to calculate the probe kind emotion of the query

face by using structured learning. In our method, we use the facial landmarks point and we extract sixteen features for to computed perfect face ratios.

This paper is separately organized as follows. In section 2 a related works and our made major contribution, in section 3 a background information about structured learning concept and about the perfect face ratios concept, specially their formalisms and their definitions. In section 4 holds in the proposed architecture of our approach, the full process is explained in depth. The dataset and the experimental results are explained clearly and concisely in section 5. Finally, in the last section, we impart a conclusion and some perspectives.

II. RELATED WORKS

In relatively recent years there has been a considerable and enormous research in methods for learning structured prediction, as well as in their application in many areas such as in the natural language processing domain and in the computer vision area, especially in image segmentation [12], in recognizing hand-written digits [13] and in OCR Letter sequence recognition [14].

A. Image Segmentation

To make concrete the abstract concepts discussed in [12, the authors tackle the problem by using the challenging task of foreground–background image segmentation. In the foreground– background image segmentation step, they turn over a natural image and they decided for each pixel whether it serves as the foreground object or as the background object. To attain this end, for each pixel i they proposed and defined one dual output variable $y_i \in \{0,1\}$, if i belongs to the foreground $y_i = 1$, else $y_i = 0$. Accordingly, the entire observed image will be represented by a single observation variable $x \in X$.

B. Recognizing Hand-Written Digits

In [13] the authors predict human handwriting digit. The task is to classify the image into one of the 10 digits 0-9. They give an example showing how the structured learning can be used to recognize images of hand-written digits. The test score is 0.97.

C. OCR Letter Sequence Recognition

In OCR Letter sequence recognition [14] the authors use structured learning to predict human handwriting word. The task is to classify the image into one of the 26 characters a-z. They give an example illustrates the results of optical character recognition. The example is taken from [13]. The test score is 0.71-0.78.

Our work is inspired in a significant part by the recent and successful works that showed that proportionately simple features be eligible for to grant to an extremely good performance in a structured learning-based framework. In our contribution, we propose a novel approach based on structured-learning-and-prediction for facial emotion classification and recognition. We clarify the structured learning-based proceeding supplementary to use only simple features.

Training a facial emotion on these simple features that can accurately predict the emotion of the query facial emotion is the key technical contribution of our paper. Obviously based on the experiments, this technic achieves our goal.

III. STRUCTURED LEARNING AND PERFECT FACE RATIOS

Structured prediction or structured learning is a cover term for the techniques of supervised machine learning that concerns predicting structured objects, moderately than scalar discrete values or real values [15], It is an explicit generalization of the usual prototype of supervised machine learning approaches, regression and classification [15]. All of these can be a reflection of discovering and finding a function that optimizes and reduces some loss and deficit over a training set or a training group. The differences concern the type of functions that are extensively used and concern also the failures and losses. In classification paradigm, the destination domain are discrete class labels, and the failure is for the most part the 0-1 loss, i.e. counting and check in order the misclassifications. In regression paradigm [15], the destination domain is the real numbers, and the failure is mostly mean squared error. In structured prediction paradigm, both the destination domain and the failure are more or less random and haphazard. This means the target is not to predict a label or a number, but in some way greatly too complicated object like related to a sequence or a graph. We purpose to define a function Discriminant F. Given the discriminant function, given an input value x, select the value Y that maximizes F as the output.

In Structured Learning paradigm, there are mainly two tasks involved [15]: Inference Task and Learning Task. Learning Task is based on a training set $(x1, y1)$, $(x2, y2)$, ..., (xn, yn), learn a model, which model is often a parametric model, so learning process is to determine the parameters of the process. Inference task is to predict an optimal output Y according to a given input X after training the model. Structured Learning has a variety of methods. One of the most important methods is the method based on the probability graph model. Another type of method is Large Margin Method for Structured Learning, which is structured SVM, is an extension of the traditional SVM method.

A. A Formalism

Having functions that generate arbitrary objects appear a taste hard to manipulate. There is one very necessary fundamental formula at the kernel of structured prediction. This formula enables us to predict the haphazard complex outputs, while we can say how suitable a given output y is with the input x.

$$y^* = arg \max_{y \in Y} f(x, y) \qquad (1)$$

x: The input,

Y: Set of all possible outputs and

f: function of the compatibility that indicate how well y be appropriate the input x.

y*: The element of Y that maximizes the compatibility. It be entitled the Prediction for x,

The all possible outputs set Y is commonly and commonly a finite set but very large one (like graphs set, like sentences set in the English language, also like all images of a given resolution). To find the argmax in the above equation by in-depth search is wherefore out of the question. To maximize over y efficiently we have to restrict ourselves to the function f. The most usual method for creating such f is based on energy functions or use the conditional random fields (CRFs). In real use of structured learning and prediction paradigm, this approach inaugurates fundamentally three technical questions:

− The first challenge: How to Choose a parametric form of the function f?

The second challenge is to solve the equation?

$$\underset{y \in Y}{\text{argmax}}\, f(x, y) \qquad (2)$$

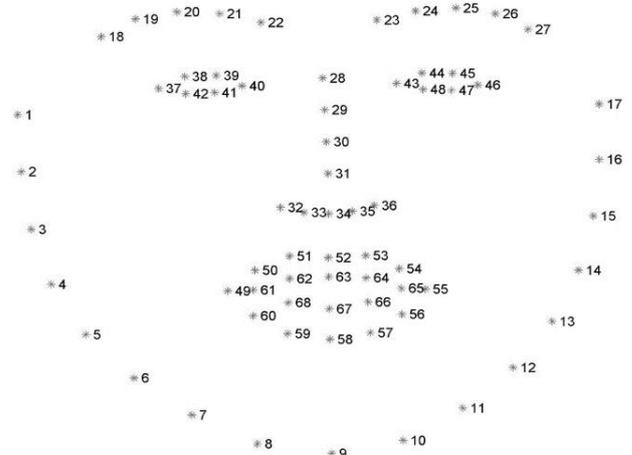Fig. 2.The 68_face_landmarks[4] :"The Mouth" is the list of points [49, 68],"The Right Eyebrow" is the list of points [18, 22], "The Left Eyebrow" is the list of points [23, 27],"The Right Eye" is the list of points [37, 42],"Left Eye" is the list of points [43, 48], "The Nose" is the list of points [28, 3] and "The Jaw" is the list of points [1, 17] .

− How to minimize a loss of the function f by Learning the parameters for the same function f?

The function f is defined as a linear function of some parameters w (weights) with another function named a joint feature function which allows to joint x and y:

$$f(x, y) = w^T joint\_feature(x, y) \qquad (3)$$

The probe output prediction is given as:

$$y^* = \underset{y \in Y}{\text{argmax}}\, w^T joint\_feature(x, y) \qquad (4)$$

w: are the weights (parameters) that are learned from the input data,

joint feature: is another function defined by the user-specified structure of the model used in structured learning paradigm.

The formula of the joint feature function is given by the Models included also in structured learning paradigm. y is defined as a discrete vector, and all most models used in present a pairwise decomposition of the energy of the function f over all entries of y, the joint feature function can be:

$$f(x, y) = w^T joint_{feature}(x, y) =$$
$$\sum_{i \in V} w_i^T joint_{feature}(x, y_i) +$$
$$\sum_{(i,j) \in E} w_{i,j}^T joint_{feature_{i,j}}(x, y_i, y_j) \quad (5)$$

Relatively to the equation (5)

V are a set of all nodes corresponding to the all entries of y and E are a set of margins between these nodes.

The particular form of $joint\_feature(x, y_i)$ and $joint_{feature_{i,j}}(x, y_i, y_j)$ depends on the model used.

B. Perfect Face Ratios

The region of interest related to each key point is described using special descriptors. Generally, the result is $v \in R^d$, so v is a d-dimensional vector describing the physical region properties; that is, texture or shape among others properties are described. It seems reasonable to select those detectors and descriptors that are not affected by transformations. They will perform better in complex scenarios due to their robustness against transformations.

To calculate the proportions of the face (face ratios), we used some different distances, exaltedly twenty distances. All of them are inspired by Cannons work, these measures are also used during a perfect face conception [5]. We define the following distances as:

Definition: Let Pi be the names of a face landmarks point [6][7] for each index i in the discrete interval [1,68] or [0,67] as illustrated on Fig. 3. (the x_coordinate of each point pi is designate by pi.x and the y_coordinate of the same point pi is designated by pi.y).

The ratios of the distances taken as the scientific measuring guns of the beauty in a perfect face are calculated. We can choose:

$$X=(x_i) \quad i \in \{1, …, n\}$$
$$Y=(y_i) \quad i \in \{1, …, n\}$$

And we can choose the Euclidean distance is define by:

$$distance(X,Y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \qquad (6)$$

Then, the sixteen ratios are x1 to x16 defined as:

$$x1 = \frac{\sqrt[2]{(P42.x - P45.x)^2 + (P42.y - P45.y)^2}}{\sqrt[2]{(P0.x - P16.x)^2 + (P0.y - P16.y)^2}} \qquad (7)$$

$$x2 = \frac{\sqrt[2]{(P36.x - P39.x)^2 + (P36.y - P39.y)^2}}{\sqrt[2]{(P0.x - P16.x)^2 + (P0.y - P16.y)^2}} \qquad (8)$$

$$x3 = \frac{\sqrt[2]{(P47.x - P54.x)^2 + (P47.y - P54.y)^2}}{\sqrt[2]{(P4.x - P13.x)^2 + (P4.y - P13.y)^2}} \qquad (9)$$

$$x4 = \frac{\sqrt[2]{(P39.x - P42.x)^2 + (P39.y - P42.y)^2}}{\sqrt[2]{(P0.x - P16.x)^2 + (P0.y - P16.y)^2}} \qquad (10)$$

$$x5 = \frac{\sqrt[2]{(P45.x - P16.x)^2 + (P45.y - P45.y)^2}}{\sqrt[2]{(P0.x - P16.x)^2 + (P0.y - P16.y)^2}} \qquad (11)$$

$$x6 = \frac{\sqrt[2]{(P36.x - P0.x)^2}}{\sqrt[2]{(P0.x - P16.x)^2 + (P0.y - P16.y)^2}} \qquad (12)$$

$$x7 = \frac{\sqrt[2]{(P45.y - P26.y)^2}}{\sqrt[2]{(P8.x - P33.x)^2 + (P8.y - P33.y)^2}} \qquad (13)$$

$$x8 = \frac{\sqrt[2]{(P27.y - (P27.y + |P33.y - P27.y|)^2}}{\sqrt[2]{(P8.x - P27.x)^2 + (P8.y - (P27.y + |P33.y - P27.y|))^2}} \qquad (14)$$

$$x9 = \frac{\sqrt[2]{(P33.x - P27.x)^2 + (P33.y - P27.y)^2}}{\sqrt[2]{(P8.x - P27.x)^2 + (P8.y - (P27.y + |P33.y - P27.y|))^2}} \qquad (15)$$

$$x10 = \frac{\sqrt[2]{(P33.x - P8.x)^2 + (P33.y - P8.y)^2}}{\sqrt[2]{(P8.x - P27.x)^2 + (P8.y - (P27.y + |P33.y - P27.y|))^2}} \qquad (16)$$

$$x11 = \frac{\sqrt[2]{(P42.x - P39.x)^2 + (P42.y - P39.y)^2}}{\sqrt[2]{(P35.x - P31.x)^2 + (P35.y - P31.y)^2}} \qquad (17)$$

$$x12 = \frac{\sqrt[2]{(P54.x - P48.x)^2 + (P54.y - P48.y)^2}}{\sqrt[2]{((P36.x + |P39.x - P36.x|) - (P42.x + |P45.x - P42.x|))^2 + (P36.y - P42.y)^2}} \qquad (18)$$

$$x13 = \frac{\sqrt[2]{(P57.x - P8.x)^2 + (P57.y - P8.y)^2}}{\sqrt[2]{(P8.x - P27.x)^2 + (P8.y - (P27.y + |P33.y - P27.y|))^2}} \qquad (19)$$

$$x14 = \frac{\sqrt[2]{\left(P27.x - (P42.x + |P45.x - P42.x|)\right)^2 + \left(P42.y - (P27.y + |P33.y - P27.y|)\right)^2}}{\sqrt[2]{(P27.x - P8.x)^2 + (P27.y + |P33.y - P27.y| - P8.y)^2}} \qquad (20)$$

The ratios x15 and x16 are detailed in the approach section.

Let X=(xi) / i $\in$ {1…16} and Y=(yj) / j $\in$ {0,1,2,3,4,5,6}. We consider X the vector of facial emotion features, X= (x1, x2, x3, x4, x5, x6, x7, x8, x9, x10, x11, x12, x13, x14, x15, x16) and Y the vector of emotions (each emotion is one from the seven-emotions defined in literature) of a facial emotion, Y=(0,1,2,3,4,5,6).

IV. APPROACH

The Fig.3 details and presents a general overview on our originally proposed structured learning-based framework for facial emotion recognition. In the online case of the Facial Emotion Recognition System (FERS), turn over a facial emotion, we extracted sixteen features in this steep, the series of all measured characteristics construct a vector of float values, this vector is treated as though be
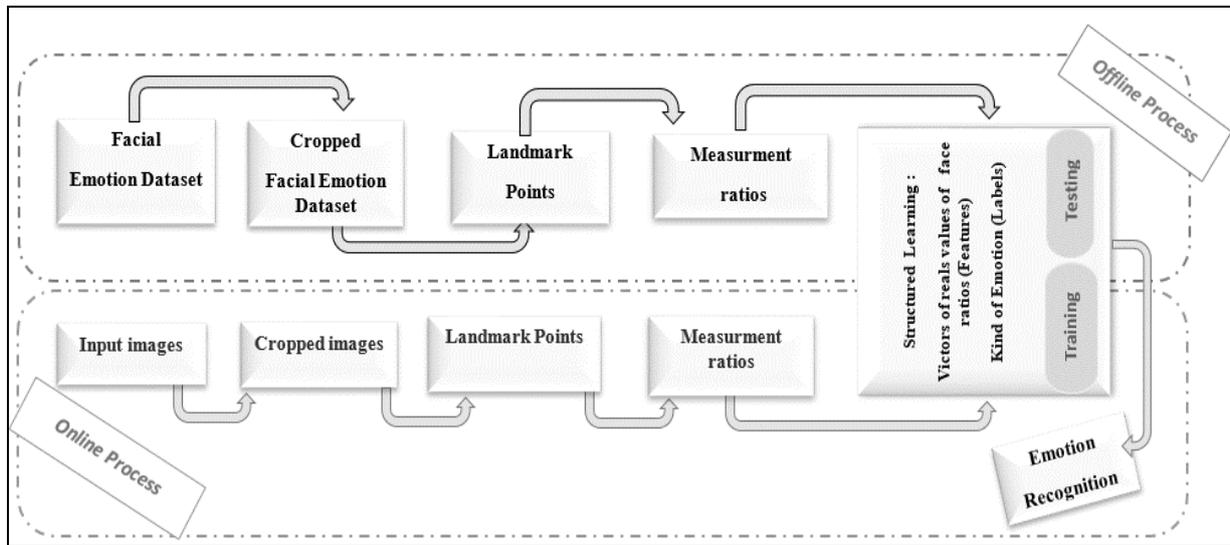
Fig.3. **A summary capture of our proposed framework based on structured learning with human face ratios for facial emotion recognition according to the model "The 68_face_landmarks".**

the identifier of the facial emotion from which the values have been elicited and calculated. In the offline case of the Facial Emotion Recognition scheme, certain measurements based on same distances are effectuated for each facial emotion in the facial emotion photos dataset.

We mention that our system has two parallel proceedings. In both of them, the first step is to read the input facial emotion images. In our chosen images database, the standard size is 200x250. The second step is to convert the all images into Gray level, they are resized or cropped into 200x250 pixels too. The method projected in our approach is based on Haar-cascades algorithm, we use the same process for detecting the faces.

Once the photo image is acquired and the preprocessing mode is done, the most relevant features are to be extracted in order to recognize the facial emotion. Afterwards, the descriptors are used to describe the region around the previously detected positions. To select the most important points in an image, usually denoted as key points, detectors are used. Usually, the landmarks points are located in regions of the image with high contrast, edges or corners. Note that the landmark point itself does not provide an ought information, thus a region of interest around the landmark point is defined. Shape_predictor_68_face_landmarks is used in our approach. For more clarity see Fig. 3.

In details, our system has two modes, in both them, the input facial emotion image and all facial emotion photos of the dataset are normalized and transformed into a Gray level, they are resized and cropped into 200x250 pixels. These dimensions are chosen: It's the proposed default choice of the datasets used and it's also the dimensions used in the related works. To localize and to detect faces in each facial emotion image, the Viola and Jones algorithm [16] [17] is used for. In the major first step of our system we pretreat and we normalize and assign all facial emotion images in the offline phase. To achieve this goal, they have been converted into a gray level images and they are all cropped to 200x250 pixels.

In parallel, we applied the same technique to online mode. After this step, we projected the famous algorithm of viola

and jones to find and to detect faces into emotional images. In the major second step, we also locate the 68_ landmarks points in each face. Our geometric descriptor used these 68 points as a parameter. This descriptor allows to extract an identity of each face via a certain calculation based on the ratios of the perfect human face. The system calculates and generates twenty harmonious distances, a vector will be dedicated to group these distances in order. This vector represents the input facial emotion image. It is a real proportionality with any other similar vector. As detailed in last section, the ratios xi are float fractions of these distances. All ratios are calculated and have been stored in vectors. The major considerable steps of our algorithm can be sum up along these lines:

Step 1: To pretreat and cropped all the facial emotion images.

Step 2: To detect the face and to locate and draw the landmark points in the facial emotion photos.

Step 3: To calculate the distances and the ratios x1 to x14 as defined in structured learning and perfect face ratios section.

Step 4: To calculate S: the surface of the nose. We defined S as:

$$x_{15} = S \qquad (21)$$

Step 5: in each detected face, let:

E= The tragus center of the ear
I= Center inter eyebrow
F= Center of the front
C= The chin
N= the Dorsum in the nose

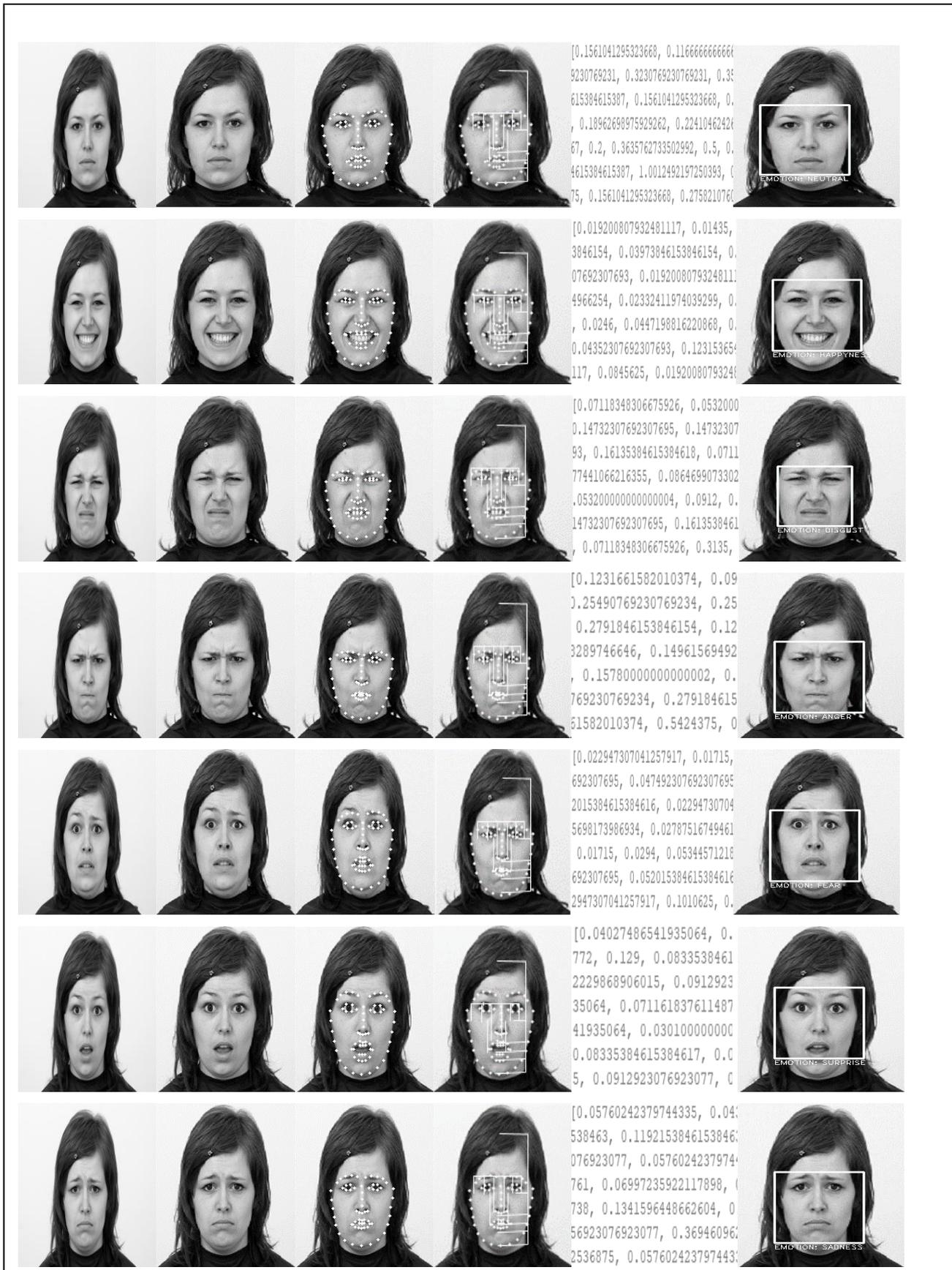To calculate the last ratio X16, we use the following

Fig.4. The FER process of our approach report, Column1: the input facial emotion from the dataset. Column 2: cropped emotional images. Column 3: extraction of the 68 face landmark points. Column 4: generation of sixteen features and Column 5: the vector calculated from the facial emotion images; it permits to recognized the kind of emotion. The output reaction of FER system is the inquiry kind of emotion of the input facial emotions.

angles:

$$\theta 1 = \overline{FEI} = \frac{arc\ \cos(EF^2 + EI^2 - FI^2)}{2 * EF * EI} \quad (22)$$

$$\theta 2 = \overline{IEN} = \frac{arc\ cos(EN^2 + EI^2 - NI^2)}{2 * EN * EI} \quad (23)$$

$$\theta 3 = \overline{NEC} = \frac{arc\ cos(EN^2 + EC^2 - NC^2)}{2 * EN * EC} \quad (24)$$

The last ratio is defined as:

$$x_{16} = \frac{\theta 1 + \theta 2}{\theta 3} \quad (25)$$

$$x_{16} = \frac{\dfrac{arc\ cos(EF^2 + EI^2 - FI^2)}{2 * EF * EI} + \dfrac{arc\ cos(EN^2 + EI^2 - NI^2)}{2 * EN * EI}}{\dfrac{arc\ cos(EN^2 + EC^2 - NC^2)}{2 * EN * EC}} \quad (26)$$

Step 6: Create the vectors X and Y as:

X=xi,/ i ∈ {1,…,16}   the features and y=yi, / i ∈ {0,1,2,3,4,5,6} the labels.

We choose the integer value 0 to represent the anger emotion, the integer value 1 to represent the disgust emotion, the integer value 2 to represent the fear emotion, the integer value 3 to represent the happy emotion, the integer value 4 to represent the neutral emotion, the integer value 5 to represent the surprise emotion and the integer value 6 to represent the sad emotion.

Step 7:   Training and testing the data with structured learning algorithm.

Step 8: The recognition step: The output result is the kind of emotion prediction.

The classification algorithm used in the authors' approach is structured learning. They use prediction for recognition the probe output emotion of face image. The features are the ratios of some selected distances and the labels are the emotion, the authors applied the same steps for online process scenario in their FER (Facial Emotion Recognition) system.

V. EXPERIMENTS AND REPORT

To try and prove the operational effectiveness of our proposed alternative method, we started by acquiring the dataset of emotional facial expression pictures. We used the WSEFEP [17] dataset, the Warsaw is the acronym of: Set of Emotional Facial Expression Pictures [18]. The WSEFEP contains those pictures that admitted the highest recognition attainment (e.g., precision with intended performance) from independent judges, it comprises 210 high quality photos of 30 persons.

In our approach the WSEFEP database is used in the train step and also in the test step. The Fig. 4 richly explains step by step the obtained report as we advance in the process of the framework already described previously.

The first step of the system is to pretreat and normalize all emotional face images in the offline phase. Table-I shows the percentage of facial emotion images used relative to the kind of emotion of the input photo images. To pretrain and

Table- I: The percentage of facial emotion images (used from WSEFEP database) relative to the kind of emotion of the input photo images

Emotion	Anger	Disgust	Fear	Happiness	Neutral	Sadness	Surprise
%	14.28	14.28	14.28	14.28	14.28	14.28	14.28

normalize all emotional faces images, they have been converted into the Gray level images and they are transformed to 200x250 pixels: The dimension 200x250 are chosen because is the default dimension in the dataset used in the other comparative methods. The same problem-solving techniques are thus applied in the online mode also.

The second step is to locate and detect faces in the input image photos. To accomplish this target, we used the distinguished algorithm of viola and jones. After that the second step product is used also to locate the 68-point landmarks in each input face. Our algorithm descriptor uses these 68 landmark points as a parameter. The aim of the chosen descriptor is to identify each facial emotion image along the computing and extraction of ratios deducted from human perfect face ratios. The extraction is based on some harmonious distances. We define a vector of float values   to group the fractions and ratios of these measurements. In order, the ratios of these distances have been stocked in the vector as already detailed in structured learning and perfect face ratios section.

The learning and the testing process of the proposed method is consistent with the standard cross-validation approach of common existing learning methods in the image recognition literature (numbers, letters, faces, objects). To do this, the dataset is divided into two sets, one is used for training while the entire remaining emotional facial images is used for testing. The parameters used for all the experiments were selected according to the parameters of the works carried out on the recognition of the figures, as well as on the recognition of the letters.

The classification was carried out with the SVM classifier and the last stage of the recognition allows to select the probable emotion or very close to the emotional facial image, the entry of the authors' FER system. The authors compare their new method with a facial emotion recognition method defined in [3]. In the work of [3] the basic idea of the experiment is to recognize the emotion of facial emotion by using the Fuzzy distance of Hamming (FHD) and ratios of a Perfect human Face (PFR); the input facial emotion is compared with seven referential faces suitable to each emotion. The referential faces used are created by some classical means. They are: arithmetic average, contra-harmonic average, harmonic average, geometric average and quadratic average.

TABLE-II illustrates the accruing match scores for our approach found on structured learning and Ratios of a Perfect human Face (PFR). It lists the accuracies of facial emotion recognition method by using the fuzzy distance of Hamming and ratios of perfect face and also, it

reports the accuracies by using our proposed method in this paper. Our algorithm proves that, the recognition rate reaches more than 98% for the anger, the fear, the sadness and the surprise emotions and the same algorithm prove also, that, the recognition rate reaches more than 97% for the disgust emotion,however it proves that, the recognition rate reaches more than 99% for the neutral and the happiness emotions. The algorithmic complexity of the approach based on structured learning and perfect face ratios is $O(n^2)$.

The method of this paper helps to better recognize all facial emotions. This structured learning method has the neutral emotion rate as the best one with 99.22%, the second rate is for the happiness emotion with 99.10%, the third one is the rate of the both facial emotions: Anger and sadness with 98.99%, the fourth rate is 98.32% also for tow facial emotions: Fear and surprise, but the west rate is 97.75% for the disgust facial emotion.

The figure Fig.5 illustrates the comparison of accruing match scores between our approach and the used method in [3]. The x-axis shows the kind of emotions used and the y-axis shows their identification rates. The report clearly demonstrates the superiority of the used algorithm to recognize the kind's emotion of a facial emotion image, but the last rate of recognition is that of disgust emotion.

We tested our approach with seven different facial emotions by using structured learning and prediction method. This approach is based on features of perfect face ratios descriptor also. The cumulative match score proves the performance of the algorithm. The recognition rate reaches more than 99% in WSEFEP dataset.

However, to use our algorithm perfectly:

Hypothesis One: Large and labeled dataset is needed.

Hypothesis Two: A structured and multiple computing power are required for training.

Table-II: The Report of FHD _ PFR method and our approach based on Structured Learning and Average Face Ratios (AFR)

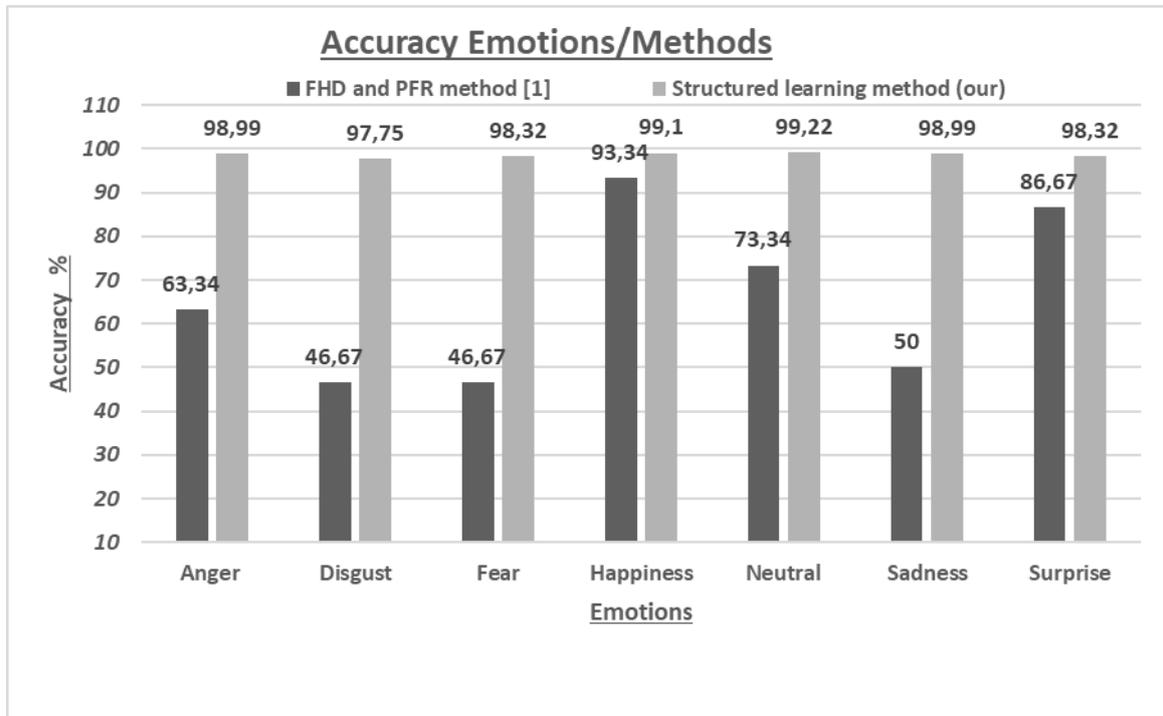| Methods / Type of Emotion	Anger	Disgust	Fear	Happiness	Neutral	Sadness	Surprise
FHD and PFR method	63.34	46.67	46.67	93.34	73.34	50	86.67
Structured learning method (our method)	98.99	97.75	98.32	99.1	99.22	98.99	98.32

Fig.5. A Comparative cumulative match scores scheme between our various facial emotion recognition approaches by using two different techniques (FHD_PFR method and Structured Learning method).

Hypothesis Three: Large memory, powerful operating system and efficient platform are demanded.

Hypothesis Four: The standardization of the input data, especially for the choice of the color image mode and for the choice of the dimensions to resizing each image is required.

Hypothesis Five: The face detector chosen proves its effectiveness but can be replaced by another more effective detectors.

Hypothesis Six: The number of points chosen was 68 while

there are more points landmarks that can be chosen.

Hypothesis Seven: The choice of the metric of rations measurements, especially that the face is a grid of pixels. whereas the distance of Euclid is very powerful for this type of plan. An adequate metric is needed.

Hypothesis Eight: In the recognition part, the choice of the classifier, the choice of his learner and his model is also an enormous challenge. A good learner and adequate model are demanded.

VI. CONCLUSION

In this paper, we proposed a new machine learning approach for a facial emotion recognition system. The method is based on structured learning and Ratios of a perfect Face. We determined and calculated same face ratios (exactly sixteen ones) based on the distances between all ratios of a perfect face as the features and the type of a face emotion as the labels. Our method is tested on WSEFEP dataset. Based on the experiments the report is very satisfactory, the experimental results show that our algorithm is effective and reliably recognized each emotion from neutral facial images. In the future work, we will optimize the number of the used features, we will compare our proposed method with the other machine learning algorithms and we will be applying also the same descriptor in other computer vision topics.

REFERENCES

1. A. Akram, N.Wang , J. Li and X.Gao . "A Comparative Study on Face Sketch Synthesis". Vol. 6. IEEE Access. 2018.pp. 37084 - 37093.
2. T.Sini and T.C. Smita . "MindCam: An Approach for Sketch Based Image Retrieval". Vol. 8(2). International Journal of Information Systems and Computer Sciences. 2019. pp.67-71
3. MS.Bouzakraoui, S.Abdelalim and E.Nourddine. "Towards a framework for customer emotion detection". 2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)
4. K. Ounachad, M. Oualla, A. Souhar and A. sadiq. "Face Sketch Recognition: Gender Classification and Recognition". International Conference of Computer and Renewable Energies, ICCSRE'2018. To be published.
5. K. Ounachad, M. Oualla, A. Souhar and A. sadiq. "Fuzzy Hamming distance and Perfect Face Ratios based Face Sketch Recognition". IEEE 5th International Congress on Information Science and Technology. October 2018.
6. M. Sadiq, D. Shi, M. Guo and X. Cheng. "Facial Landmark Detection via Attention-Adaptive Deep Network". vol. 7. Journal of IEEE Access: Special section on Innovation and Application of Intelligent Processing of Data, Information and Knowledge as Resources in Edge Computing. 2019. pp.181041- 181050.
7. Y. Wu and Q. Ji. "Facial landmark detection: A literature survey". vol.127(2), Journal of Computer Vision. 2019. pp. 115–142.
8. S. Sra and R. Hosseini. "Algorithmic Advances in Riemannian Geometry and Applications". Advances in Computer Vision and Pattern Recognition. In: Minh H., Murino V. (eds). Geometric optimization in machine learning. Springer Cham .2016. pp. 73-91.
9. S. Thrun and L. Pratt. Learning to learn. Springer Science & Business Media.2012.
10. M. R. Boutell, J. Luo, X. Shen and C. M. Brown. "Learning multi-label scene classification". vol.37(9). International Journal of Pattern recognition, 2004. pp. 1757–1771.
11. S. Nowozin and C.H. Lampert. "Structured Learning and Prediction in Computer Vision". vol. 6(3–4). International Journal of Computer Graphics and Vision, 2011. pp. 185-365.
12. B. Taskar, C. Guestrin, and D. Koller. "Max-margin Markov networks". Vol.2. Neural Information Processing Systems Conference (NIPS03), December 2003.
13. P. Ekman and W. V. Friesen. (1987). "Universals and cultural differences in the judgements of facial expressions of emotion". Vol.53. Journal of Personality and Social Psychology.1987. pp.712 - 717.
14. A. Muller and S. Behnke. "Learning Structured Prediction in Python". Vol.5(1). Journal of Machine Learning Research. 2014. pp. 2055-2060.
15. P. Viola and M.J. Jones. (2001). "Robust Real-Time Face Detection". Vol.57. International Journal of Computer Vision.2004. pp.137–154.
16. M. Oualla, A. Sadiq and S. Mbarki. "comparative study of the methods using Haar-Like features". Vol.4(4). International journal of engineering sciences & research.2015. pp.35-43.
17. M. Olszanowski, G. Pochwatko, K. Kuklinski, M. Scibor-Rylski, P. Lewinski and R. K. Ohme,. "Warsaw set of emotional facial expression pictures: A validation study of facial display photographs. Frontiers in Psychology", vol.5.2015. Article 1516.
18. S. Stöckli, M. Schulte-Mecklenbeck, S. Borer and A. C. Samson. "Facial expression analysis with AFFDEX and FACET: A validation study". vol. 50(4). Journal of Behavior research methods, Psychonomic Society.2018. pp. 1446-1460.

AUTHORS PROFILE

Khalid Ounachad is a computer science professor at the preparatory classes for engineering schools, Moulay Youssef high school Rabat Morocco. He received his BS in software engineering from Sciences and technologies Faculty Moulay Ismail University, Errachidia in 2003, DESS degree in Information System engineering from the Faculty of Sciences Semlalia (FSSM) cadi ayyad University, Marrakech, Morocco in 2007. He is currently a PhD candidate in the Department of Computer Science in Faculty of Science, Ibn Tofail University, Kenitra, Morocco. His research interests include image processing, law investigation, multimedia information retrieval, machine learning, sentiments analysis and Artificial Intelligence.

Mohamed Oualla received his BS in Software Engineering from the Sciences and Technologies Faculty (FST), Moulay Ismail University, Errachidia in 1999, DESA in Computer Network and Telecommunication from the National School of Computer Science and Systems Analysis (ENSIAS), Mohammed V University, Rabat, Morocco in 2002 and PhD in 2017 from the Faculty of Science, Ibn Tofail University of Kenitra, Morocco. Currently, he is an Associate Professor in Computer Science with the Faculty of Science and Technologies (FST), Errachidia, Morocco. His research interests include image processing, pattern recognition, machine learning, big data and IoT. He has served as a reviewer for several international conferences and congresses.

Abdlghani Souhar is a full professor of computer science at Ibn Tofail University, Faculty of science Kenitra, Morocco. He received his M.S. degree in applied Mathematics in 1992, PhD degree in computer science in 1997 from Mohammed V University in Rabat-Morocco. His habilitation thesis concerned modelling complex systems as Arabic handwritten recognition and an intelligent system for generating mesh. His research interests include modelling complex systems, automatic processing of Arabic language, Artificial Intelligence and CAE/CAD. He has served as a reviewer for several international journals and conferences.

Abdelalim Sadiq received a B.S in software engineering from Sciences and technologies Faculty Moulay Ismail University, Errachidia in 1999, DESA degree in computer network and telecommunication from National School of Computer Science and Systems Analysis (ENSIAS) Mohammed V University, Rabat, Morocco in 2002 and Ph.D. degree in Computer Science from ENSIAS, Mohammed V University, Rabat, Morocco, in 2007. He is currently an associate professor in computer science Department of Sciences Faculty, Ibn Tofail University, Kenitra, Morocco and Team Leader Information System and Multimedia (SIM). His research interests include multimedia information retrieval and processing, sentiments analysis, IoT and data science. He has served as a reviewer for several international conferences and journals.

Retrieval Number: C6421029320/2020©BEIESP
DOI: 10.35940/ijeat.C6421.049420

160

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication