Real Time Automated Facial Expression Recognition App Development on Smart Phones

Humaid Alshamsi [1] and Veton Kepuska [1]
Department of Electrical and Computer Engineering
Florida Institute of Technology
Melbourne, USA
alshamsih2015@my.fit.edu

Hongying Meng [1,2]
Department of Electrical and Computer Engineering
Brunel University London
UB8 3PH,UK

Abstract— **Automated facial expression recognition (AFER) is a crucial technology to and a challenging task for human computer interaction. Previous methods of AFER have incorporated different features and classification methods and use basic testing approaches. In this paper, we employ the best feature descriptor for AFER by empirically evaluating the feature descriptors named the Facial Landmarks descriptor and the Center of Gravity descriptor. We examine each feature descriptor by considering one classification method, such as the Support Vector Machine (SVM) method, with three unique facial expression recognition (FER) datasets. In addition to test accuracies, we present confusion matrices of AFER. We also analyze the effect of using these feature and image resolutions on AFER performance. Our study indicates that the Facial Landmarks descriptor is the best choice to run AFER on mobile phones. The results of our study demonstrate that the proposed facial expression recognition on a mobile phone application is successful and provides up to 96.3% recognition accuracy.**

Keywords—Machine Learning; Automated Facial Expression Recognition; Facial Landmarks; Support Vector Machine; Mobile Computing.

I. INTRODUCTION

Automated facial expression recognition (AFER) has a great potential for improving our life quality. For instance, the AFER system is useful in medical applications, such as when surgical staff need to assess a patient's behavior, either during or after surgery. The AFER system could also be used as a security check method at airports to find out whether a traveler exhibits any significant changes in emotional state; or when a suspect is under criminal investigation. A robotic teacher could also possibly offer a better learning experience by having a clearer understanding of student feeling.

Designing a facial expression recognition (FER) system that runs on smart phone technology is difficult and challenging due to the huge number of images needed for testing.Storage of images, light conditions, sunglasses and masks, and head poses also pose challenges in smart phone application. Over the years, researchers have proposed various techniques for computer recognition of human facial expressions. Previous work has typically focused on the proposal of basic feature descriptors and classification methods for FER. In contrast, in this paper, we aim to identify the best feature descriptors by performing an extensive comparison study.

AFER technologies are versatile in that they have recently been applied to various disciplines, from psychology to smart phone use [12]. The relationship between embedded devices and the future vision of computerization has benefitted from this technology. FER carries with it immense possibilities for future use. Humans have the ability to communicate without words; facial expressions do the work of telling the viewer how a person feels in terms of energy and emotion, among other things, without speaking. Darwin studied the biology of facial expressions and attested to their significance to the survival of the human race. There are 6 basic, universal emotions conveyed via facial expression, according to Ekman [8], and they range from fear, sadness, disgust, and anger, to happiness and surprise. Facial expressions convey the deepest of human emotions. Per Ekman, the basic expressions just mentioned are actually the building blocks of all facial expressions, and thus, sometimes facial expressions are hard to discern and figure out; the 43 muscles of the face form expressions with the processes of animation and protraction.

II. RELATED WORKS

FER inquiry has become more important in recent years, partly due to the versatility of its application in areas such as computer animation, identification and recognition capabilities, examination of facial aspects, and simulation necessities [1, 2]. As previously mentioned, humans instinctively read facial expressions to elucidate emotion. Previously, many interpretations of human emotion have been made. The 6 basic facial expressions per Ekman are hallmarked as useful for research. Ekman and cohort's Facial Action Coding System (FACS) delineates all of the 6 expressions possible in a static image [3]. The FACS is universally recognized as valid to research, particularly for use in psychology applications. A manual labeling process built on human perception is key to the system, significantly different from the common research practice of classifying facial expressions in an automated fashion.

FER systems have historically been created using videos and static imagery. The first local paramedic motion system used to classify information accordingly was used by Yacoob and Black [4]. Created by Chang, a low-dimensional manifold system to parse out facial expression via representation, trailing, and identification [5] used the embedded Lipschitz system of alignment of facial features for construction of the manifold.

Chang incorporated Gabor wavelets to pay attention to changes in facial expression which were then examined in the temporal space, attributed to Valstar and cohort, to delineate expressions [6]. Templates of fluctuating intensity were used by Kumano and colleagues to study some of the more uncommon expressions of the face, thus showing the differences between classification templates [7]. Found not to be as useful, the Adaboost classifier and SVMs were not used as often after Kumano's research [8]. Following this, the graphic spectral methods of Torre were then used [9], and Zeng's addition to the research encompassed an understanding of the separation of unemotional and emotional facial expressions [10]. It is commonly understood that humans exhibit unemotional tendencies sometimes simply as a product of habit. Contemporary research has focused on the recognition of emotion in the wild, as demonstrated by the EmotiW challenge (2013[13], 2014[14] and 2015[15]). For these annual "challenges," FER systems are created to sense emotions from brief videos and static images in order to ascertain use in real world deployments. Some concerns regarding the challenges are: 1) the inputs are connected to distinct labels which are related to comprehensive and broad expressions, which is a method that is acknowledged to be not conducive to reproducing actual and real states of emotion, and 2) the images studied are only affectations of posed countenances from movies. Incidentally, real-time use of these approaches on portable devices is not possible.

III. FACIAL EXPRESSION RECOGNITION

In our research, the Facial Landmarks method was used to extract features following the pre-processing step, which detected and subsequently tracked facial movement with the Hausdorff Distance method of detection [20]. Feature extraction and classification were the essential basis of the method. We then analyzed the histogram values in contrast to the training set's stored values using an algorithm based on SVM. One of the most-known databases for videos of facial expressions is the Cohn-Kanade (CK+) database [19], which we used along with the Karolinska Directed Emotional Faces (KDEF) [16] and Japanese Female Facial Expression (JAFFE) Datasets [17]. Action Units (AUs) were actuated to best characterize the most prevalent of facial expressions in the recordings.

Feature Descriptors: The Haar descriptor [27] is one of the most popular descriptors for use in FER [28], [29]. The Gabor descriptor [30] is also another common feature descriptor used for FER [31], [32], [33], [34], [35]. The BRIEF descriptor [36] and HOG descriptor have been examined for use in FER as well [37], [38], [39]. In this study, we evaluate the performance of a relatively new feature descriptor called the facial landmarks descriptor [20] in facial expression recognition. To the best of our knowledge, we are the first to consider the facial landmarks descriptor in real time FER development on a mobile phone platform.

Classification Methods: The SVM classifier [33], [34] is the most common classifier that has been applied to FER [40], [33], [34], [41], [42]. Several studies have reported that SVM with a linear kernel produce similar test outcomes when compared to a radial basis function (RBF) kernel [33], [43], [34]. The K-Nearest Neighbors (K-NN) classifier has also recently been used for FER [44]. The results of their application have been very accurate with very few errors to date. The databases are used for the testing of the proposed system and the application is built on an Android platform and can run on any Android platform.

A. Overview of the System

In Fig. 1, we see a block diagram showing the structure of real-time FER on a mobile phone platform. Extraction of features and classification of those features are the two essential steps in the system. The stage encompassing feature extraction includes the preprocessing step wherein an arrangement of images at 8 frames/second from a video recorder is examined to detect the facial space and to normalize properties in order to light the image correctly. The Facial Landmarks feature extractor was used in this stage. A SVM algorithm was then employed in the feature classification step, which is largely acknowledged in image processing as a strong nonparametric arrangement classifier.
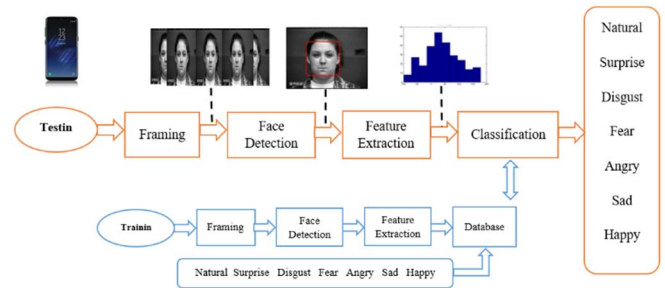

Fig. 1. Overview of structure of the facial expression recognition.

B. Feature Extraction

Facial landmarks are points on specific parts of the facial image which indicate, for instance, the location of the nose, the eyes, the brows, and the mouth within an image. These points are tracked to follow the facial muscles' movements in time. If all facial landmarks are considered as a connected graph, we assume that the density of the graph differs in each facial expression; the pre-trained facial landmark detector inside the Dlib library is used to estimate the location of 68 (x, y) coordinates that map the facial structures on the face. After the detection of the facial landmarks, the "Center of Gravity" (COG) of all landmarks is calculated (Formula [1,2]) and is shown in Fig. (2,3). Graphs are very useful mathematical tools that can provide a wealth of information regarding the interrelationships of spatial points, in this case, of the facial landmarks. In order to extract features from these facial landmarks, spectral graph analysis is used, through which a characteristic vector, depicting areas of density in a graph, is extracted. The face offset correction is calculated by taking the tip of the nose and finding the angle for the nose as shown in Fig. (4) (formals from [3-8]).

$$X_{COG} = \frac{1}{68} \sum_{i=1}^{68} x_i \qquad (1)$$

$$Y_{COG} = \frac{1}{68}\sum\nolimits_{i=1}^{68} y_i \qquad (2)$$

Where X_{COG} is x-coordinate of COG and Y_{COG} is y-coordinate of COG.

$$x_{relative\ i} = x_i - X_{COG} \qquad (3)$$

$$y_{relative\ i} = y_i - Y_{COG} \qquad (4)$$

$$EUC_i = \sqrt{(x_i - X_{COG})^2 + (y_i - Y_{COG})^2} \qquad (5)$$

$$\beta_{nose} = \tan^{-1}\left(\frac{y_{28} - y_{31}}{x_{28} - x_{31}}\right) \qquad (6)$$

$$\theta_i = \tan^{-1}\left(\frac{y_i - Y_{COG}}{x_i - X_{COG}}\right) - \beta_{nose} \qquad (7)$$

$$i = 1,2,...., 68$$

$$F = \{x_{relative\ i}, y_{relative\ i}, EUC_i, \theta_i\}i = 1^{68} \qquad (8).$$

Where EUC_i is Euclidean distance between point i and central point, β_{nose} is the angle of nose, point-28 top of the nose and point-31 tip of the nose. θ_i defines relative angle and is corrected when the face is not perfectly horizontal, and feature vector length is 272.

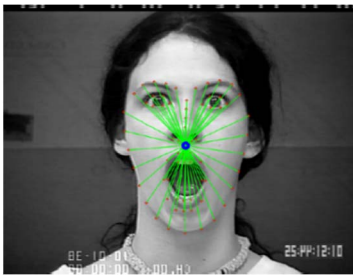Fig. 2.   Facial landmarks and COG using CK+ dataset.

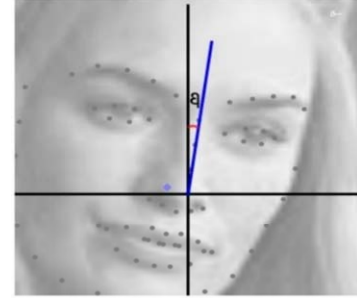Fig. 3.   Line mapping between COG and facial landmarks.

Fig. 4.   Face offset correction.

C. Classification Method

Because SVMs are widely believed to be able to obtain optimal separation between classes, they are largely used for a wide range of pattern delineation problems. For this research, we used SVMs by training them to classify facial expressions according to the facial characteristics we wanted to include. The high-dimension area is separated by the SVM.

Vapnik and associates [18] recently promoted SVMs as decidedly effective in pattern recognition functions. An SVM has the ability to locate the hyperplane that divides the largest portion of points of similar class on similar flanks, when given a series of points from two classes. SVM then optimizes the range from the hyperplane to either category. Vapnik [18] classified this as an Optimal Separating Hyperplane (OSH). The purpose of this hyperplane is to diminish the chance of wrongly allocating the samples in the training set as well as the silent examples from the test set. Researchers are currently trying to figure out how to expand the binary classification capabilities of SVMs [21] to include resolutions for multiclass classification issues. Proposed methods include the construction of a classification system wherein binary classifiers are incorporated into one multiclass classifier, while others focus on all classes together. The expense of solving multiclass issues within the parameters of large-scale problems has redirected research away from same.

The support vector machine [22], given labelled training data

$$D = \{(x_i, y_i)\}_{i=1}^l, x_i \in X \subset R^d, y_i \in Y = \{-1, +1\} \quad (9)$$

constructs a maximal margin linear classifier in a high dimensional feature space, $\phi(x)$ defined by a positive definite kernel function, $k(x, x')$, specifying an inner product in the feature space,

$$\phi(x).(x') = k(x, x') \qquad (10).$$

A common kernel is the Gaussian radial basis function (RBF)

$$k_{RBF}(x, x') = e^{-\gamma\|x - x'\|^2} \qquad (11).$$

Where $\gamma > 0$, kernel parameters to be decided by the users. Where linear classifier is defined as

$$h_{w,b}(x) = g(w^T x + b) \qquad (12).$$

Where w is weights and b is bias.

The discriminant function implemented by a support vector machine is given by

$$f(x) = \left\{ \sum_{i=1}^{l} \alpha_i y_i k(x_i, x) \right\} + b \qquad (13)$$

To find the optimal coefficients α of this expansion it is sufficient to maximize the functional,

$$w(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} y_i y_j \alpha_i \alpha_j k(x_i, x_j) \qquad (14)$$

which can be considered here as a cost function for optimization.

The popular SVM guide [26] suggests the following setting to train a kernel classifier:

1. Scale each feature to an interval like [-1; +1].

2. Use Gaussian kernel. Choose γ; that gives the highest cross validation (CV) accuracy.

3. Obtain the model w using the selected γ.

IV. SYSTEM EVALUATION

The proposed system was first evaluated in a public dataset. It will demonstrate its efficiency on facial expression recognition and validate the method.

A. Facial Expression Datasets

We studied the 3 main FER datasets: the extended CK+ Facial Expression Dataset [23], [24], the JAFFE dataset [17], and the KDEF dataset [16].

1) CK+ Dataset

One of the most universally accepted FER databases, the CK+ database contains facial images of 210 adults from 18-50 years of age, 69% of whom were female. An instructor helped to guide the study's participants through an array of requested expressions; 327 cycles were classified as containing at least 1 of 7 distinct facial expressions. These 7 expressions are: angry, contempt, disgust, fear, happy, sad, and surprise. Every cycle begins with a neutral facial countenance and resolves with a more specified expression. For our purposes, we ruled out the facial expression connoting "contempt," and instead targeted

the 6 basic expressions [25]. The "neutral" image was chosen as the first frame of each sequence in our study and the last frame consisted of 1 of the basic 6 expressions. In total, we employed 636 facial pictures for our research, comprised by number of images in this way: angry-45, disgust-59, fear-25, happy-69, sad-28, and surprise-83. Sample images from the CK+ dataset are shown in Fig. 5, wherein angry is represented by the color red, happy is orange, fear is yellow, neutral is green, sad is blue, surprise is indigo, and disgust is violet.

Fig. 5. CK+ Facial Expression Dataset.

2) JAFFE Dataset

For our purposes, we used all 213 of the 6 basic expressions, including neutral, found in the JAFFE database. 10 Japanese females facially expressed the following number of emotions: angry-30, fear-32, happy-31, neutral-30, sad-31, and surprise-30. JAFFE sample images are depicted in Fig. 6, wherein red represents angry, orange is happy, yellow is fear, green is neutral, blue is sad, indigo is surprise, and violet is disgust.
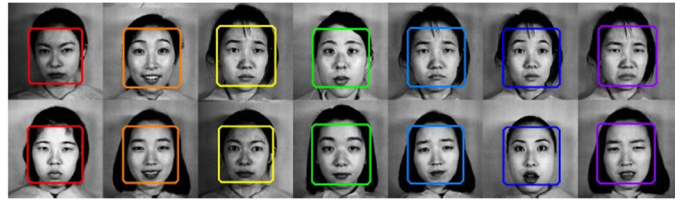

Fig. 6. JAFFE Facial Expression Dataset.

3) KDEF Dataset

For this study, we considered the frontal images only of 70 male and 70 female actors found in the KDEF database. In the KDEF database, the actors, from 20-30 years of age, were photographed displaying 5 different angles; their faces were free of any intrusive elements such as facial hair, jewelry, makeup, or glasses. We studied a total of 980 facial expression representations, consisting of the following numbers: angry-140, disgust-140, fear-140, happy-140, neutral-140, sad-140, and surprise-140. Sample representations from the KDEF set are found in Fig. 7, which shows angry as red, happy as orange, fear as yellow, neutral as green, sad as blue, surprise as indigo, and disgust as violet.

Fig. 7.   KDEF Facial Expression Dataset.

B. Performance

We analyzed the Facial Landmarks feature extraction performance and the SVM algorithm. Our testing results comprise findings from 3 separate databases. The tables that follow illustrate our results. Table I demonstrates the system results of the Cohn and Kanade (CK+) dataset which contains 636 images from 123 subjects, where 327 sequences are labeled with one of seven facial expressions, namely (angry, disgust, fear, happy, sad, surprise, and contempt). Table II demonstrates the system results of the Japanese Female Facial Expression (JAFFE) dataset which contains 213 images of 7 facial expressions posed by 10 Japanese female models. Table III demonstrates the system results of the Karolinska Directed Emotional Faces (KDEF) dataset which contains 980 images of 7 facial expressions.

TABLE I.   THE CONFUSION MATRIX USING PROPOSED METHOD ON CK+ DATABASE. THE NUMBERS ARE GIVEN IN % ACC.

	N	F	S	H	S	A	D
Natural (N)	100	0	0	0	0	0	0
Fear (F)	0	100	0	0	0	0	0
Surprise(S)	0	0	100	0	0	0	0
Happy (H)	0	0	0	100	0	0	0
Sad (S)	0	0	0	0	87.5	0	12.5
Angry (A)	0	0	0	0	0	92.3	7.6
Disgust (D)	0	5.88	0	0	0	0	94.1

TABLE II.   THE CONFUSION MATRIX USING PROPOSED METHOD ON JAFFE DATABASE. THE NUMBERS ARE GIVEN IN % ACC.

	N	F	S	H	S	A	D
Natural (N)	100	0	0	0	0	0	0
Fear (F)	11.1	88.9	0	0	0	0	0
Surprise(S)	11.1	0	88.9	0	0	0	0
Happy (H)	0	0	0	88.9	11.1	0	0
Sad (S)	0	0	0	0	100	0	0
Angry (A)	0	11.11	0	0	0	88.9	0
Disgust (D)	0	0	0	0	12.5	0	87.5

TABLE III.   THE CONFUSION MATRIX USING PROPOSED METHOD ON KDEF DATABASE. THE NUMBERS ARE GIVEN IN % ACC.

	N	F	S	H	S	A	D
Natural (N)	88.1	4.7	0	0	0	2.3	4.7
Fear (F)	0	90.5	4.7	2.3	0	2.3	0
Surprise(S)	4.7	0	90.4	4.7	0	0	0
Happy (H)	2.4	0	0	92.8	2.3	2.3	0
Sad (S)	0	4.7	2.3	2.3	90.4	0	0
Angry (A)	2.3	2.3	0	2.3	0	92.8	0
Disgust (D)	0	2.3	4.7	0	2.3	0	90.4

V. SYSTEM EVALUATION

A. Feature Extraction

In this section, we outline the method used to extract image features and the portion of code that correlates to the extraction of features. We stored the features we extracted in order to create a database for testing. When employing the central part of the code, 70% of the training images are used and 30% of the test images are used. The feature extraction module then takes over once those images are uploaded.

B. Testing of the System

The training database we developed was then tested by using the test images; we needed to interrogate if the database was formulated in the right way and that the system worked correctly. We created code in MATLAB because it was necessary to test the system with a classifier made to analyze the training images against the test images. The code was implemented on the CK+, JAFFE, and KDEF datasets. The code used to test, along with the SVM classifier, follows, as do the results of the testing.

C. Testing the proposed method

The Facial Landmarks descriptor's import is in its superior rate of recognition and its ability to compute in a complex manner. We used a derivation of the Landmarks + COG descriptor, the Landmarks + COG -64 descriptor, which is a 64-byte descriptor, for our purposes. It is illustrated in [5, 7] that a 64-byte descriptor is satisfactory when the goal is to produce good matches.

The face is the area of focus and as such is the main input for feature extraction. The area was changed to grayscale and Gaussian smoothing [5, 7] was administered on the grayscale image. A Gaussian low-pass filter with kernel of 2 [5, 7] and a window diameter of 8 x 8 [5] was used to filter the image.

We pre-smoothed the patch of the Landmarks + COG descriptor, which operates on a pixel scale, to filter out noise pollution, thereby stabilizing the descriptors and increasing repeatability. Figs. [9, 10] show the filtration result of MATLAB reproduction. The colored pie charts outline the following facial aspects: yellow is fear, red is disgust, green is angry, blue is surprise, black is sad, and gray is happy.
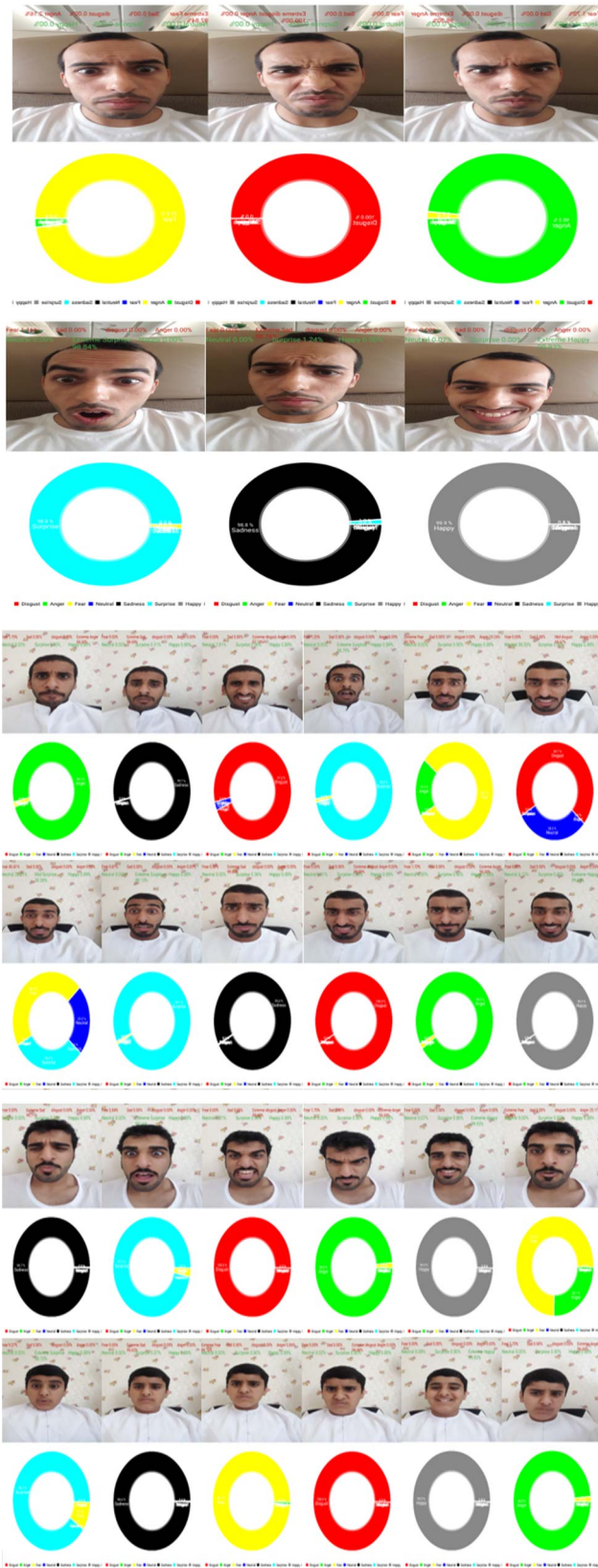
Fig. 8.   Output of automated facial expressions recognition module.

Fig. 8 shows simple results of testing using a smart phone platform and represents the percentages of facial expression

recognition detected in real time, where navy blue, yellow, red, green, royal blue, black, and gray pie charts represent natural, fear, disgust, angry, surprise, sad, and happy facial expression recognition respectively.

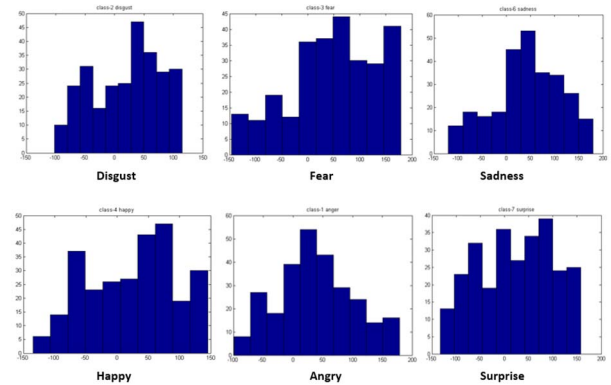Fig. 9.   Example from CK+ dataset used to extract the feature using Matlab.

Fig. 10. Histogram of Facial Landmarks descriptors using CK+ dataset.

We then plotted a histogram of the Facial Landmarks descriptors values to see them in a graphed-out way. As necessary, the amount of bins is variable.

In Fig. 10, a histogram shows that when the color blue is not found in any of the bins of the 10, this indicates the absence of an element for that particular bin. If a column lies at a height of 1, then 1 element in that column lies in the specified bin. Correspondingly, a column reaching a height of 2 then indicates the presence of 2 elements in that bin. This pattern follows for the remainder of the figure. Varied values are inherent in a binary test that contains 2 points selected randomly, so in different sequences of experimentation of the software, different values are inevitably found; ultimately, the same image is depicted regardless.

D. Testing summary

Finally, we did the work of appraising the performance of the project. We had to determine in what ways a mobile phone platform could support the requirements of FER. We tested the project through its varied stages of study.

We first evaluated each function irrespective of the other. We used a minute sample of data to test the code, at times building

another dataset, to establish accuracy of the algorithm and correctness of lines of code. We established a manual procedure to correlate the code result with the target result whereby a numeric model helped to sum up the data.

Secondly, we tested each stage after completion of that particular stage. We then integrated all of the stages and did a comprehensive test of the total performance of the system. We found at variable stages that certain algorithms which we thought would be useful were in fact not and we had to begin again. For instance, the literature speaks of the SVM algorithm in successful terms, but our attempts using the SVM algorithm were not succesful. We found that a large amount of multiclass features in need of training caused the failure. We then focused on working with the MATLAB tool and the Android platform to reprogram and retest the application. We finally found a solution after multiple tries.

Lastly, the whole of the project was checked for successful attainment of goals. The most important step, this ensured that our algorithm was capable enough to achieve its goal of being able to discern facial emotional expressions on a mobile phone platform.
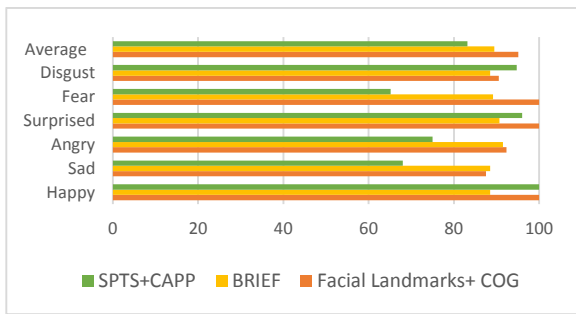

Fig. 11. *Performance comparison on Facial Landmarks+ COG, BRIEF and SPTS+CAPP features.*

VI. CONCLUSION AND DISCUSSION

The comparison between the previous studies and the proposed study is provided in (Table 4). The previous methods and the proposed methods were applied on three databases (CK+, JAFFE & KDEF). However, the difference between the performance of the proposed methods and the previous methods depended on factors such as: in the previous studies, accuracy is extracted from the highest performance of expressions, while in the proposed system the overall accuracy is calculated by dividing the number of correct predictions on the total number of images. In previous studies, all images in the databases were not used, while in the proposed system, all images in the databases were used.

Although the proposed method was successfully implemented on a smart phone platform and a satisfictory performance was achieved, the work can be further improved by adding new features such as a cloud computing server. It will be very interesting to try to run the system on the cloud, showing the facial expressions in real time.

In this paper, the proposed methods proved their effectiveness through achieving high performance. However, we documented some suggestions that may lead to improvements in the proposed system in the future, to its performance, and it's quality assessment as well, such as :

- Instead of using the proposed system to predict emotions from a facial image, a possible suggestion is to use this system to improve face recognition systems which can be used in different security models such as criminal detection.

- To make this application more accurate, some features need to be enhanced and more focused, such as the eyebrows, frowning, and dilation of pupils. The eyes are the core determiners of expression after the mouth in facial detection, so the dots can be placed at the start of the eyebrow, the middle, and the end.

TABLE IV. ACCURACY COMPARISON BETWEEN THE SYSTEMS.

Systems	Accuracy (%)		
	CK+	JAFFE	KDEF
[19]	83.2	-	74.6
[45]	89.5	-	-
[46]	81.4	-	82.4
[47]	85.4	96.4	-
[48]	81.6	94.9	-
Proposed	**96.3**	**91.9**	**90.8**

REFERENCES

[1] Pantic, M., Pentland, A., Nijholt, A., Huang, T.S.: Human Computing and Machine Understanding of Human Behavior: A Survey, in Artificial Intelligence for Human Computing, 2007.

[2] Li, S.Z., Jain, A.K.: Handbook of Face Recognition, ISBN: 0-387-40595-X, 2005.

[3] Ekman, P.: Strong evidence for universals in facial expressions:a reply to Russell's mistaken critique. Psychol. Bull. 115(2), 268–287, 1994.

[4] Black, M. J., Yacoob, Y.: Recognizing facial expressions in image sequences using local parameterized models of image motion. Inter. Journal of CV, 25(1), pp. 23--48, 1997.

[5] Hu, C., Chang, Y., Feris, R., Turk. M.: Manifold based analysis of facial expression, Image and Vision Computing 24, pp. 60--614, 2006.

[6] Valstar, M.F., Pantic, M.: Fully automatic facial action unit detection and temporal analysis. Proceedings of IEEE Int'l Conf. Computer Vision and Pattern Recognition, 2006.

[7] Kumano, S., Otsuka, K., Yamato, J., Eisaku, S., Sata, Y.: Pose-Invariant facial expression recognition using variable intensity templates. Asian Conf. on Computer Vision, 2007.

[8] Bartlett, S., Littlewort, G., Frank, G., Lainscsek, C., Fasel, I., Movellan, J.: Fully automatic facial action recognition in spontaneous behavior. In Proc. Conf. Automatic Face & Gesture Recognition, pp. 223--230, 2006.

[9] Torre, F., Campoy, J., Ambadar, Z., Cohn, J.F.: Temporal Segmentation of Facial Behavior, International Conference on Computer Vision, October, 2007.

[10] Zeng, Z., Fu, Y., Roisman, G.I., Zhen, W.: Spontaneous emotional facial expression detection. Journal of Multimedia, 2006.

[11] Sakarkaya, M.; Yanbol, F.; Kurt, Z., "Comparison of several classification algorithms for gender recognition from face images," Intelligent Engineering Systems (INES), 2012 IEEE 16th International Conference on , vol., no., pp.97,101, 13-15 June 2012.

[12] Dieter Findling, Rainhard, and Rene Mayrhofer. "Towards Pan Shot Face Unlock". Int J of Pervasive Comp & Comm 9.3 (2013): 190-208. Web.

[13] Dhall, R. Goecke, J. Joshi, M. Wagner and T. Gedeon, Emotion Recognition In The Wild Challenge 2013, ACM ICMI 2013

[14] A. Dhall, R. Goecke, J. Joshi, K. Sikka and T. Gedeon, Emotion Recognition In The Wild Challenge 2014: Baseline, Data and Protocol, ACM ICMI 2014.

[15] Dhall, Abhinav et al. "Video And Image Based Emotion Recognition Challenges In The Wild: Emotiw 2015". ResearchGate. N.p., 2015. Web.

[16] Lundqvist, D., Flykt, A. and Ohman, A.: The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neu-roscience, Psychology section, Karolinska Institutet.

[17] Lyons, M., Budynek, J. and Akamatsu, S.: Automatic Classification of Single Facial Images, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.21, No.12, pp.1357–1362 (1999).

[18] C. Cortes and V. Vapnik, "Support-vector network," In Machine Learning, 1995, pp. 273-297.

[19] "The Extended Cohn-Kanade Dataset (CK+): A Complete Dataset For Action Unit And Emotion-Specified Expression". N.p., 2016.

[20] Gheorghita Ghinea, Rajkumar Kannan, and Suresh kannaiyan,"Gradient-Orientation-Based PCA Subspace for Novel Face Recognition"in IEEE Access, vol 2,2169-3536, 2014.

[21] G. Dai and C. Zhou, "Face Recognition Using Support Vector Machines with the Robust Feature," In Proc. of IEEE workshop on Robot and Human Interactive Communication, 2003, pp. 49-53.

[22] P. J. Phillips, "Support Vector Machines Applied to Face Recognition," In Advances in Neural Information Processing Systems 11, MIT Press, 1999, pp. 803-809.

[23] Kanade, T., Cohn, J. and Tian, Y.: Comprehensive Database for Facial Expression Analysis, Proc. 4th International Conference on Automatic Face and Gesture Recognition, pp.46–53 (2000).

[24] Lucey, P., Cohn, J., Kanade, T., Saragih, J., Ambadar, Z. and Matthews, I.: The Extended Cohn-Kanade Dataset (CK+): A Complete Dataset for Action Unit and Emotion-specified Expression, Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp.94–101 (2010).

[25] Ekman, P.: An Argument for Basic Emotions, Cognition and Emotion, Vol.6, pp.169–200 (1992).

[26] C.-W. Hsu, C.-C. Chang, and C.-J. Lin. A practical guide to support vector classification. Technical report, National Taiwan University, 2003.

[27] Papageorgiou, C., Oren, M. and Poggio, T.: A General Framework for Object Detection, Proc. 6th International Conference on Computer Vision, pp.555–562 (1998).

[28] Xu, C., Dong, C., Feng, Z. and Cao, T.: Facial Expression Pervasive Analysis Based on Haar-Like Features and SVM, Proc. The International Conference on E-business Technology and Strategy, pp.521–529 (2012).

[29] Jung, S., Kim, D., An, K. and Chung, M.: E cient Rectangle Feature Extraction for Real-time Facial Expression Recognition based on AdaBoost, Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.1941–1946 (2005).

[30] Gabor, D.: Theory of Communication, Journal of the Institution of Electrical Engineers, Vol.93, No.26, pp.429–457 (1946).

[31] Zhang, Z., Lyons, M., Schuster, M. and Akamatsu, S.: Comparison between Geometry-based and Gabor-wavelets-based Facial Expression Recognition using Multi-layer Perceptron, Proc. 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp.454–459 (1998).

[32] Gonzalez, I., Sahli, H., Enescu, V. and Verhelst, W.: Context-independent Facial Action Unit Recognition Using Shape and Gabor Phase Information, Proc. 4th International Conference on A ective Computing and Intelligent Interaction, pp.548–557 (2011).

[33] Wu, T., Bartlett, M. and Movellan, J.: Facial Expression Recognition Using Gabor Motion Energy Filters, Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp.42–47 (2010).

[34] Bartlett, M., Littlewort, G., Lainscsek, C., Fasel, I., Frank, M. and Movellan, J.: Fully Automatic Facial Action Recognition in Spontaneous Behavior, Proc. 7th International Conference on Automatic Face and Gesture Recognition, pp.223–228 (2006).

[35] Lee, C. and Shih, C.: Gabor Feature Selection for Facial Expression Recognition, Proc. International Conference on Signals and Electronic Systems, pp.139–142 (2010).

[36] Ozuysal, M., Calonder, M., Lepetit, V. and Fua, P.: Fast Keypoint Recognition Using Random Ferns, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.32, No.3, pp.448–461 (2010).

[37] Dalal, N. and Triggs, B.: Histograms of Oriented Gradients for Human Detection, Proc. IEEE Computer Society Conference on Computer Vi-sion and Pattern Recognition, pp.886–893 (2005).

[38] Dahmane, M. and Meunier, J.: Emotion Recognition using Dynamic Grid-based HoG Features, Proc. IEEE International Conference on Automatic Face and Gesture Recognition Workshops, pp.884–888 (2011).

[39] Orrite, C., Ganan, A. and Rogez, G.: HOG-Based Decision Tree for Facial Expression Classification, Proc. 4th Iberian Conference on Pat-tern Recognition and Image Analysis, pp.176–183 (2009).

[40] Shan, C., Gong, S. and McOwan, P.: Facial Expression Recognition based on Local Binary Patterns: A Comprehensive Study, Journal of Image and Vision Computing, Vol.27, No.6, pp.803–816 (2009).

[41] Littlewort, G., Bartlett, M., Fasel, I., Susskind, J. and Movellan, J.: Dynamics of Facial Expression Extracted Automatically from Video, Journal of Image and Vision Computing, Vol.24, No.6, pp.615–625 (2006).

[42] Nagi, G., Rahmat, R., Khalid, F. and Taufik, M.: Region-Based Facial Expression Recognition in Still Images, Journal of Information Processing Systems, Vol.9, No.1, pp.173–188 (2013).

[43] Bartlett, M., Littlewort, G., Frank, M., Lainscsek, C., Fasel, I. and Movellan, J.: Automatic Recognition of Facial Actions in Spontaneous Expressions, Journal of Multimedia, Vol.1, No.6, pp.22–35 (2006).

[44] Eleftheriadis, S., Rudovic, O. and Pantic, M.: Discriminative Shared Gaussian Processes for Multiview and View-Invariant Facial Expression Recognition, IEEE Trans. Image Processing, Vol.24, No.1, pp.189–204 (2015).

[45] H. Alshamsi, V. Këpuska, and H. Meng, "Real Time Facial Expression Recognition App Development on Mobile Phones," The

International Journal of Science & Technology, vol. 5, no. 10, pp. 507–512, Oct. 2016.

[46]   Liew, C. F., & T. Y. (aug. 2015). Facial Expression Recognition and Analysis: A Comparison Study of Feature Descriptors. IPSJ Transactions on Computer Vision and Applications, 7, 104-120. Retrieved August 31, 2015, from http://doi.org/10.2197/ipsjtcva.7.104

[47]   T. Jabid, M. H. Kabir, and O. Chae, ``Robust facial expression recognition based on local directional pattern,'' ETRI J., vol. 32, no. 5, pp. 784_794, 2010.

[48]   X. Zhao and S. Zhang, ``Facial expression recognition based on local binary patterns and kernel discriminant isomap,'' Sensors, vol. 11, no. 10, pp. 9573_9588, 2011.