Facial Expression Recognition Improvement through an Appearance Features Combination

Taoufik Ben Abdallah[1], Radhouane Guermazi[2] and Mohamed Hammami[3]
[1]*Faculty of Economics and Management, Sfax University, Tunisia*
[2]*College of Computation and Informatics, Saudi Electronic University, Kingdom of Saudi Arabia*
[3]*Faculty of Sciences, Sfax University, Tunisia*

Keywords: Facial Expression Recognition, Local Binary Pattern, Eigenfaces, Controlled Environment, Uncontrolled Environment.

Abstract: This paper suggests an approach to automatic facial expression recognition for images of frontal faces. Two methods of appearance features extraction is combined: Local Binary Pattern (LBP) on the whole face region and Eigenfaces on the eyes-eyebrows and/or on the mouth regions. Support Vector Machines (SVM), *K* Nearest Neighbors (*K*NN) and MultiLayer Perceptron (MLP) are applied separately as learning technique to generate classifiers for facial expression recognition. Furthermore, we conduct to the many empirical studies to fix the optimal parameters of the approach. We use three baseline databases to validate our approach in which we record interesting results compared to the related works regardless of using faces under controlled and uncontrolled environment.

1 INTRODUCTION

The Automatic Facial Expression Recognition (AFER) is becoming an increasingly important research filed due to its wide range of applications such as intelligent human computer interaction, educational software, etc. Generally AFER is performed by three steps: face tracking, feature extraction and expression classification. The second step-that consist in extracting features from the appropriate facial regions- is the most important to build robust facial expression recognition system. Three main approaches directions can be distinguished (Huang, 2014): (1) Geometric feature-based approaches, (2) Appearance feature-based approaches and (3) Hybrid feature-based approaches.

Geometric feature-based approaches generally used fidical points located on the face for calculating geometric displacement and/or geometric rules (distances, angles, etc.). For example, Porwat Visutsak (Visutsak, 2013) calculates Euclidean distances between eight fiducial points manually located on a neutral face and their corresponding located on an expressive face to construct a geometric movement vector. Sanchez *et al.* (Sánchez et al., 2011) initially establish a set of fiducial points on a neutral face and then move them according to the facial expressions, using the optical flow in order to extract geometric movement features between a pair of consecutive frames.

Appearance feature-based approaches mainly describe the changes in texture on a face by wrinkles, bulges and furrows. Many features defined in the literature are used in the recognition of facial expressions such as the Gabor filters (Deng et al., 2005), the Histogram of Oriented Gradient (HOG) (Donia et al., 2014; Ouyang et al., 2015), the Scale-Invariant Feature Transform (SIFT) (Soyel and Demirel, 2010), the Haar-like features (Yang et al., 2010), etc. Recently, some researchers have suggested the use of Local Binary Pattern (LBP) (Shan et al., 2009; Mliki et al., 2013; Chao et al., 2015; Happy, 2015) and its variants as Mean Based weight Matrix (MBWM) (Priya and Banu, 2012), Pyramid of Local Binary Pattern (PLBP) (Khan et al., 2013) and Completed Local Binary Pattern (CLBP) (Cao et al., 2016).

Hybrid feature-based approaches combine a set of geometric and appearance features. Zhang *et al.* (Zhang et al., 2014) combine the geometric features extracted through a set of distances between fiducial points located automatically by ASM and appearance features calculated through SIFT transformations. Wan *et al.* (Wan and Aggarwal, 2014) calculate metric distance based on a set of shape and texture features. As for shape, they extract the coordinates of

Abdallah, T., Guermazi, R. and Hammami, M.
Facial Expression Recognition Improvement through an Appearance Features Combination.
DOI: 10.5220/0006288301110118
In *Proceedings of the 19th International Conference on Enterprise Information Systems (ICEIS 2017) - Volume 3*, pages 111-118
ISBN: 978-989-758-249-3

68 fiducial points located on a face using the Constrained Local Model (CLM) (Saragih et al., 2009). As regards texture, they apply the Gabor Filter on different face regions and ACP for reducing the texture vector dimensionality.

The majority of the approaches proposed in the literature have used small dimensions sub-regions (8×8 pixels,16×16 pixels, etc) to extract appearance feature (Mliki et al., 2013; Khan et al., 2013; Zhang et al., 2014). Therefore, our main contribution is to set up an approach via using the whole face and/or the large dimensions sub-regions describing the face parts(the mouth, the eyes and/or the eyebrows). This approach is based on features extracted from the Local Binary Pattern (LBP) method on the whole face and the Eigenfaces method on the eyes-eyebrows and/or the mouth parts. An extensive experimental study of the combination of these parts is provided in order to find the best one.

The remainder of this paper is organized as follows. Section 2 presents the proposed approach. Section 3 depicts the experimental results. A conclusion is drawn and perspectives are forecasted in Section 4.

2 METHODOLOGY

Our approach is essentially based on the process of Knowledge Discovery from Databases (KDD). In fact, we distinguish three major steps. The first one is designed to prepare data that consist in presenting each image by a set of features. The second one is considered for building classifiers. The last one validates the obtained classifiers (Figure 1). Note that each input image is assumed to contain only one face. The details of the data preparation and supervised learning steps are provided in the following sections.

2.1 Data Preparation

In this stage, two steps are distinguished:(i) the detection of face parts and (ii) features extraction.

2.1.1 Detection of Face Parts

For detecting the sub-regions face parts, we use the Viola and Jones' algorithm (Viola and Jones, 2001). This algorithm was widely used for face detection. It is still reliable and ready to use in a lot of image processing software. In our work, we use this algorithm to detect the eyes sub-region. After that, we increase the size of the sub-region detected so that it includes the eyebrows. The mouth detection was very low. To obtain more accurate results, we only used the lower part of the face upon mouth processing (Figure 2). For any image, first the whole face is detected then we crop the obtained image to only consider the lower part (35% of the height). Then, 25% of the left side columns and 25% of the right side ones are removed. We detect the mouth part from the resulting image. This improvement led to an increase of nearly 40% in the rate of correct detection of the mouth.

Following the detection of face parts, images are converted to grayscale level, resized to 140×140 pixels resolution for faces, 40×90 pixels resolution for eyes-eyebrows and 30×50 pixels resolution for mouth, and then preprocessed by histogram equalization to reduce lighting conditions effects.

2.1.2 Features Extraction

Two renowned methods are used to extract features vector namely Eigenfaces based on PCA and LBP. We refer to each feature vector as $\vec{F}^{method}_{application}$, with the method either PCA or LBP, and the application either the face, the eyes-eyebrows or the mouth. Our contribution focuses on studying how the fusion of the vectors $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{eyes}$, $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{mouth}$, and $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{mouth} \wedge \vec{F}^{PCA}_{eyes}$ can improve the facial expression recognition. As a matter of fact, we combine the features generated by LBP on the whole face and by PCA on the eyes and eyebrows and/or the mouth.

Unlike the majority of related works that apply LBP on uniform regions, we use LBP only on the whole face by adjusting its parameters to define fewer number of features and to avoid regions selection steps. However, the choice of using PCA on eyes-eyebrows and/or mouth regions is done according to the research work of Daw-Tung Lin (Lin, 2006) which shows that applying PCA on face parts is more important than applying PCA on the whole face.

Eigenfaces (Sirovich and Kirby, 1978) is to convert the pixels of an image into a set of features through a multivariate statistical study based on PCA.

Formally, we use M images in the training set and each image, noted X_i with $i = 1,2,...,M$, is a 2-dimensional array sized $l \times c$ pixels. An image X_i can be converted into one-dimensional array of D pixels ($D = l \times c$). Define the training set of M images by $X = (X_1, X_2, ..., X_M) \subset \Re^{D \times M}$. The covariance matrix Γ is defined as follow (equation 1):

$$\Gamma = \frac{1}{M} \sum_{i=1}^{M} (X_i - \bar{X})(X_i - \bar{X})^T \tag{1}$$

Where $\Gamma \subset \Re^{D \times D}$ and $\bar{X} = \frac{1}{M} \sum_{i=1}^{M} X_i$ refers to the mean image of the training set.
According to Γ, we calculate the k eigenvectors

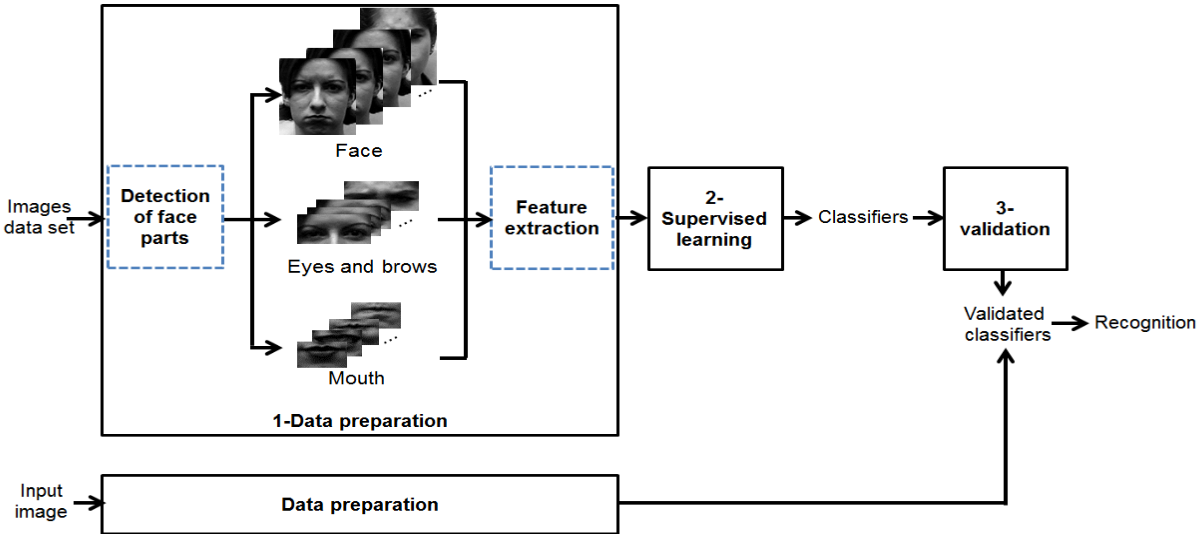

Figure 1: Approach proposed for facial expression recognition.

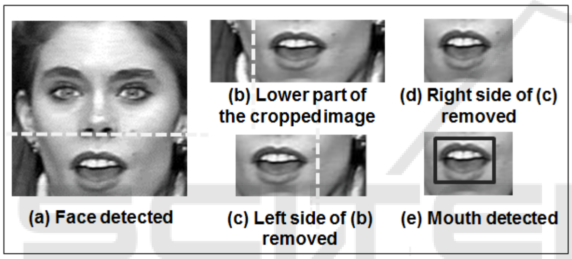

Figure 2: Mouth detection.

(i.e. eigenfaces) corresponding to k largest non-zero eigenvalues ($k << M$). Each image X_i is projected into the eigenfaces space to obtain the features learning vector $\vec{F}^{PCA}_{application} \subset \Re^{M \times k}$.

The features of each of the test images Y_i is calculated by projecting the mean-subtracted image $Y_i - \bar{X}$ on the eigenfaces space.

Local Binary Pattern was introduced by Ojala *et al.* (Ojala et al., 1996) as an effective solution to texture description based on the comparison of the luminance level of a pixel with its neighbors' levels. This operator is used to code each pixel of an image (named center pixel) into grayscale by thresholding its neighborhood with its value. We used the extended version of LBP (Ojala et al., 2002) which spreads the neighboring p pixels in a circular shape with a radius R, indicating the distance between the center pixel and its neighbors. The size of the feature vector $\vec{F}^{LBP}_{face}$ is equal to 2^p values.

To guide the LBP optimal parameters choice, we set the number of neighbors to a value often used in the literature which is equal to 8 (Priya and Banu, 2012; Saha and Wu, 2010; Shan et al., 2009), and conduct an empirical study by calculating the global recognition rate according to R.

2.2 Supervised Learning

We used three supervised learning techniques for expression-classification: Support Vector Machines (SVM) (Vapnik, 1995), K Nearest Neighbors (KNN) (Duda et al., 2000) and MultiLayer Perceptron (MLP) (Malsburg, 1961). For SVM, we apply two kernels: the polynomial kernel and the Gaussian kernel of Radial Basis Function (RBF). Likewise, we have opted for the technique "one versus one" which, for as m classes problem, generates $\frac{m \cdot (m-1)}{2}$ binary models. For KNN, we have applied the Euclidean distance for calculating similarity between an input test image and the set of learning images. For MLP, we have defined a single hidden layer and seven neurons in the output layer corresponding to the seven facial expressions to classify. Then, we have applied the backpropagation algorithm to adjust the synaptic weights and the sigmoid activation function defined as $f(x) = \frac{1}{1+exp(-x)}$.

3 EXPERIMENTAL RESULTS

The purpose of the conducted experiments is evaluating the performance of the approach proposed. In the present work, we distinguish three sets of experiments. The first experiment series is primarily dedicated to determine LBP, PCA, SVM and MLP parameters. The second experiment series aims to evaluate our approach under controlled environment through three learning techniques i.e. SVM (polynomial kernel, RBF kernel), KNN and MLP. In these series of experiment, we consider the universal representation (joy, surprise, disgust, sadness, anger and fear), proposed by Ekman (Ekman, 1972), with the addition of the representation of neutrality for facial expression recognition, using frontal faces. Therefore, we used three-baseline datasets separately: Japanese Female Facial Expression (JAFFE) (Lyons et al.,) (213 images), Cohn Kanade (CK) (Kanade et al., 2000) (606 extracted images) and Fedutum (Wallhoff,) (231 extracted images). We apply 10-fold cross-validation to evaluate the classifiers so that the faces used in the learning phase do not contribute to the test phase. The third experiment series considered an image dataset that encompasses JAFFE, CK and Feedtum together to construct a classifier, treating several types of persons under uncontrolled environment.

3.1 First Experiment Series: LBP, PCA, SVM and MLP Parameters

Our approach has several parameters that can affect the efficiency of our classifier. Among these parameters, we identify: R the radius representing the distance between a center pixel and its neighbors in the LBP method, the number k of eigenfaces in PCA, the SVM parameters which are γ and C, and the number of neurons in the hidden layer. To find these parameters, we conduct four empirical studies.

The first shows the recognition rate variation of LBP features calculated on the whole face, using an SVM (polynomial kernel), according to the radius R based on 10 cross validation. The study of the radius R was performed over an interval ranging from 1 to 20 as the majority of related works use values less than 16. Moreover, the results are unstable outside the selected range which makes the optimization of R very difficult. The choice of the optimal radius was made according to the maximum recognition rate of the classifier generated by $\vec{F}^{LBP}_{face}$ as vector features and SVM (polynomial kernel) as learning technique. Figure 3 shows the results obtained, using different values of R for the three defined databases.

The best value of R for the JAFFE and CK databases is equal to 16 where the recognition rate reached 92.49% and 87.79% respectively. Similarly, $R = 17$ provides the best performance for the Feedtum database. However, the results found by $R = 16$ is also interesting and close to that found with $R = 17$. So, in our work, we fix $R = 16$ for all experiments performed and whatever the database used. Note that the choice of a large radius R is argued by the use of LBP on high dimensions region (whole face).

The second study shows the variation in the number k of eigenfaces according to 10 cross validation obtained by three classifiers: the first is generated by $\vec{F}^{PCA}_{eyes}$, the second is generated by $\vec{F}^{PCA}_{mouth}$, and the third is made up of $\vec{F}^{PCA}_{eyes} \wedge \vec{F}^{PCA}_{mouth}$. SVM with a polynomial kernel is applied in these experiment series. Choosing the best number of eigenvectors was achieved according to the highest average recognition rate provided by the three suggested classifiers. Figure 4 presents the recognition rate of each classifier for each value of k using JAFFE database. The optimal number of k eigenfaces is equal to 30 in which the three classifiers record the best recognition rate, varying beteween 84.95 and 87.79%. This study is performed likewise, using CK and Feedtum databases in that we find almost $k = 30$ is the optimal number of eigenfaces. Note that we can find performances that slightly exceed the recognition rate obtained for k values of above 50. Meanwhile, the parallel increase of the complexity of the classifier makes the performance gain negligible as compared to the loss in terms of complexity.

The third study is reserved to determine the optimal values of the SVM (RBF kernel) parameters: the variable γ, which allows us to change the size of the kernel and the constant C, which reduces the number of fuzzy observations. This is done by a study of the variation in the global recognition rate of 10 cross validation according to γ and C on the test phase, using three combinations of features i.e. $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{eyes}$, $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{mouth}$ and $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{mouth} \wedge \vec{F}^{PCA}_{eyes}$. We varied γ from 0.01 to 1.12 (step=0.03) and C from 1 to 8 (step=1). The optimal values are $C = 3$ and $\gamma = 0.1$ for each combination and for each database.

Finally, regarding the MLP, we used a single hidden layer perceptron and a sigmoid function to activate the hidden and the output layers. Obviously, the number of neurons of the input layer is the number of features, the number of neurons of the output layer is equal to 7 corresponding to the seven universal facial expressions and the number of neurons in the hidden layer was found by the empirical study. This study is performed by calculating the recognition rate of 10 cross validation for each classifier ($\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{eyes}$, $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{mouth}$ and $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{mouth} \wedge \vec{F}^{PCA}_{eyes}$) and for each database according to the number of neurons in the hidden layer. Generally, the ideal number of neurons in the hidden layer is 10 for all classifiers and databases used.

3.2 Second Experiment Series

In this section, experiments are expressed in terms of 10 cross validation. They were carried out, using SVM (with a polynomial or RBF kernel), KNN, and

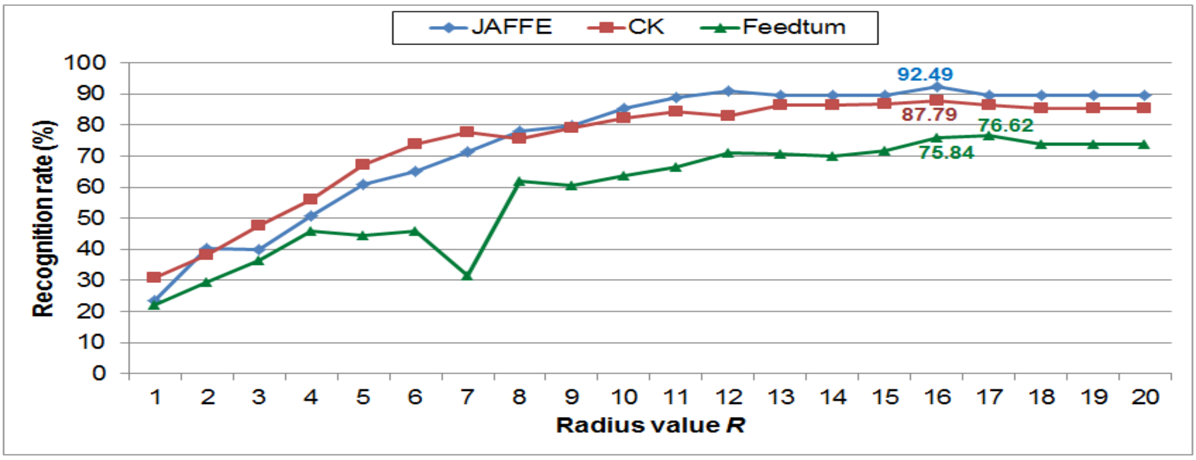

Figure 3: Recognition rate variation vs LBP radius value on different databases.

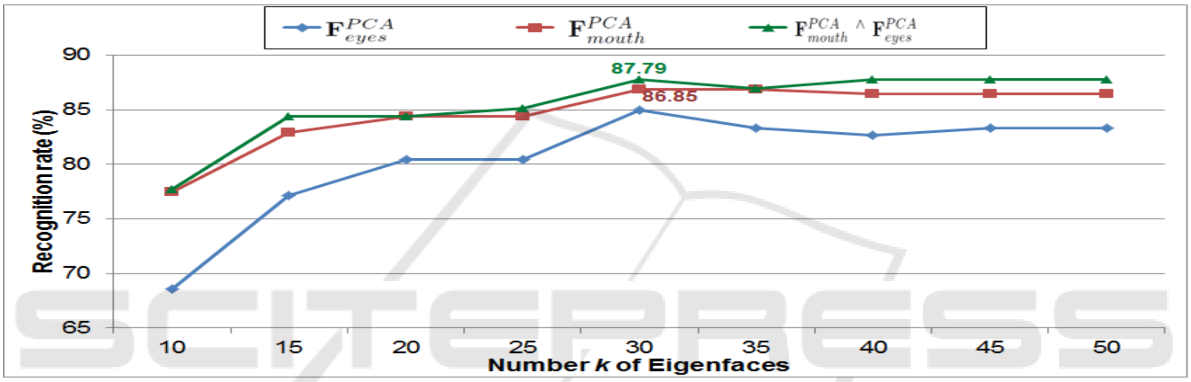

Figure 4: Recognition rate variation vs number of eigenfaces by SVM, KNN and MLP classifiers, using JAFEE database.

MLP. To illustrate the performance of our approach under controlled environment, we present the results separately for each database.

For JAFFE, Figure 5 illustrates the evaluation of our three features combination using SVM, KNN and MLP. The best recognition rate is 94.37%, using the classifier that combines LBP on the whole face with PCA on the mouth, and uses SVM (RBF kernel) as a learning technique. This classifier ($\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{mouth}$ / SVM with RBF kernel) outperforms the classifiers based on $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{eyes}$ or $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{eyes} \wedge \vec{F}^{PCA}_{mouth}$ irrespective of the learning technique used.

For CK, cross validation shows that our classifier generated by $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{mouth}$ as features vector and SVM (RBF kernel) as leaning technique outperforms all other classifiers with a recognition rate equal to 92.51% (Figure 6). On the other hand, the performance of the classifiers $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{eyes}$ / SVM (polynomial kernel), $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{eyes}$ / MLP and $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{eyes} \wedge \vec{F}^{PCA}_{mouth}$/ MLP deteriorates remarkably with recognition rate, varying between 86.13 and 88.79%.

Similarly, for Feedtum, our combination of LBP on the whole face and PCA on the mouth gives the best classifier, reaching 82.68% of facial expression recognition, using SVM with polynomial kernel or MLP (Figure 7). The decrease in Feedtum facial expression recognition is due to the fact that the facial expressions of this dataset are poorly generated which causes ambiguity. To conclude, the best classification obtained, using the three databases is that of $\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{mouth}$/ SVM with RBF kernel. In fact, the mouth region is more discriminating than the eye-eyebrows region for facial expression recognition.

To fairly compare our facial expressions recognition results to other related works, we suggest some works, using the same JAFFE database (213 images annotated by the six universal facial expressions and the expression of neutrality) under the same 10 cross validation strategy. Table 1 shows the performance comparison between our best classifier found and the existing approaches in terms of the features vector, the size of features vector and the recognition rate (Recog. rate). We can notice that our approach leads to the best result (94.37%) and reduce the number of features from 5379 features (Mliki et al., 2013) to 286 features.

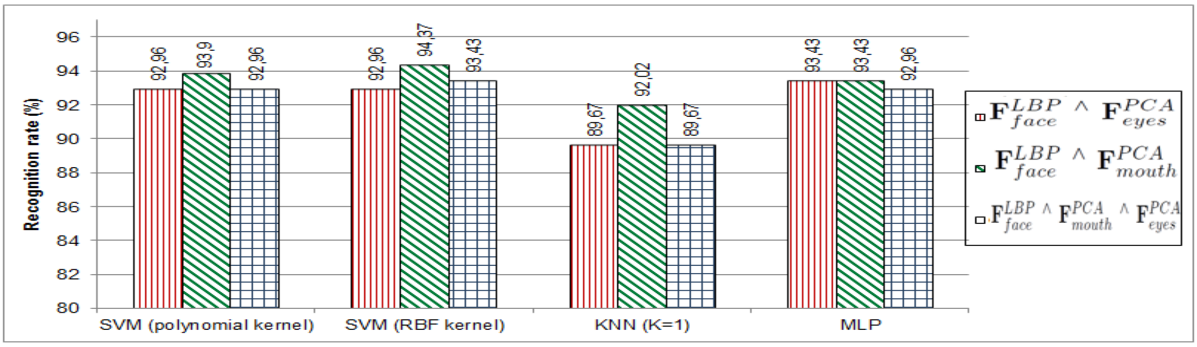

Figure 5: Experimental results by the four classifiers on JAFFE database.

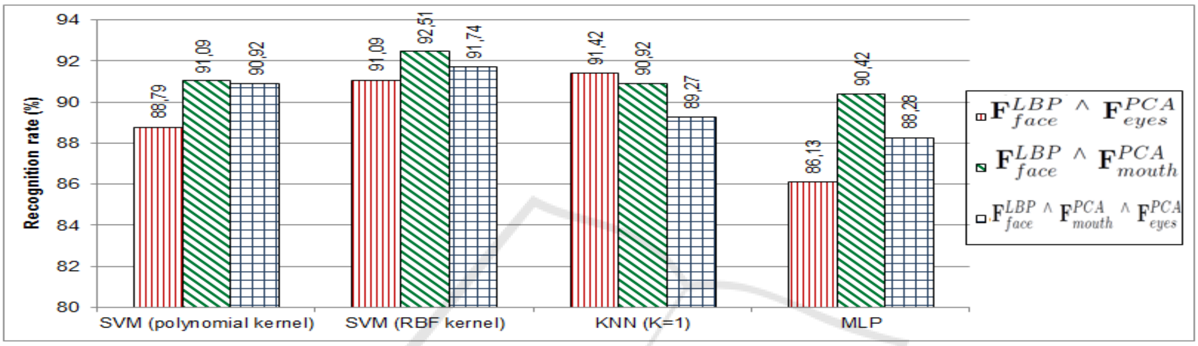

Figure 6: Experimental results by the four classifiers on CK database.

Table 1: Comparative evaluation of the proposed approach with the literature for facial expression recognition on JAFFE database.

Reference	Features vector	Features number	Recog. rate(%)
(Shan et al., 2009)	Boosted-LBP	-	81
(Priya and Banu, 2012)	MBWM	-	91.35
(S.Zhang et al., 2012)	LBP+LFDA	-	90.7
(Chen et al., 2012)	Shape features+ Gabor wavelet	-	83
(Mliki et al., 2013)	LBP	5379	93.89
(Chakrabartia and Duttab, 2013)	Eigenspaces	-	84.16
(Happy, 2015)	Uniform of LBP	-	91.8
Ours	$\vec{F}_{face}^{LBP} \wedge \vec{F}_{mouth}^{PCA}$	286	94.37

3.3 Third Experiment Series

Encouraged by the results of the proposed classifiers using 10 cross validation, we devote this section to evaluate the performance of our best combination of features ($\vec{F}_{face}^{LBP} \wedge \vec{F}_{mouth}^{PCA}$) under uncontrolled environment where the classifier must be able to recognize facial expressions of a person who has not necessarily belonged to the same environment and contributed to the learning process. The preparation of the classifier took into consideration images database that encompasses all the 1050 images of three datasets: JAFFE, CK and Feedtum (Table 2). Using LBP on the whole face and the PCA on the mouth records the best performance (51.88%).

For improving the results obtained, we applied two combination techniques of classifiers: (1) by majority vote and (2) by score learning.
The combination by majority vote consists in comparing the results of each classifier (SVM, *K*NN or MLP) in which the final decision corresponds to the class predicted by at least two classifiers. In case of conflict, we consider the prediction of the SVM classifier with an RBF kernel. Combining score learning is to seek a classifier based on the probabilities estimated for each class by each learning technique. We perform two score learning combination methods. The first one is labelled "Score_Tech" in which each of the four classifiers provides seven probabilities of belonging to seven facial expressions. The second one is labelled "Score_Desc" where two classifiers are used: one is based on LBP on the whole face and SVM (RBF kernel) as a learning technique, and the other is based on PCA on the eyes and eyebrows or mouth and SVM (RBF kernel) as a learning technique. Each

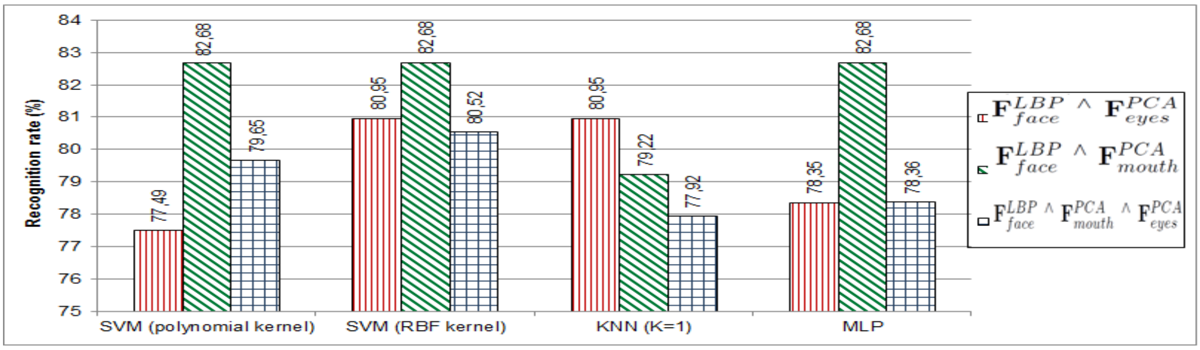

Figure 7: Experimental results by the four classifiers on Feedtum database.

classifier provides seven probabilities corresponding to the seven belonging facial expressions. We opted for the use of SVM with an RBF kernel to construct the classifier based on the estimated probabilities for each facial expression. Table 3 illustrates the results of different classifier combination techniques, using $\vec{F}^{LBP}_{face}$ and $\vec{F}^{PCA}_{mouth}$ features.

Table 2: Experimental results under uncontrolled environment.

Combination features	**Recog. rate(%)**			
	SVM (Poly)	SVM (RBF)	*KNN* (*K*=1)	MLP
$\vec{F}^{LBP}_{face} \wedge \vec{F}^{PCA}_{mouth}$	46.88	**51.88**	46.88	49.38

Table 3: Comparison of different combination techniques.

	Majority vote	Score_Tech	Score_Desc
Recog. rate(%)	52.6	60.31	70.63

The best classifier is achieved by score learning ("Score_Desc" method) reaching 70.63% of recognition rate. Unlike the combination of classifiers by majority vote that has not resulted in meaningful increases in the recognition of facial expressions under uncontrolled environment, the "Score_Desc" method allowed a significant improvement of 18.75% compared to the results obtained without combination.

4 CONCLUSION AND PERSPECTIVES

In this work, we propose an approach to the recognition of the six universal expressions and the neutrality on artificially taken still images with front side faces under controlled and uncontrolled environment. Four major contributions of this work could be enumerated. Firstly, we improve the mouth detection, using the Viola and Jones program on the lower part of the face automatically located. Secondly, we achieve several empirical studies to find the optimum parameters of the approach proposed. Then, we demonstrate that considering the whole face and the mouth together can improve the facial expression recognition rate. Finally, we improve facial expression recognition under uncontrolled environment according to a combination of classifiers based on score learning.

As perspectives. we can test other combination methods such as the score weight and the transferable belief model in order to improve the performance of the approach proposed under uncontrolled environment. We may also explore a variety of images that display faces captured at a natural environment(spontaneous expressions, face poses, etc.).

REFERENCES

Cao, N., Ton-That, A., and Choi, H. (2016). An effective facial expression recognition approach for intelligent game systems. *International Journal of Computational Vision and Robotics*, 6(3):223–234.

Chakrabartia, D. and Duttab, D. (2013). Facial expression recognition using Eigenspaces. In *CIMTA'13: International Conference on Computational Intelligence, Modeling Techniques and Applications*, volume 10, pages 755–761. ELSEVIER.

Chao, W., Ding, J., and Liu, J. (2015). Facial expression recognition based on improved Local Binary Pattern and class-regularized locality preserving projection. *Journal of Signal Processing*, 2:552–561.

Chen, L., Zhoua, C., and Shenb, L. (2012). Facial expression recognition based on SVM in E-learning. In *CSEDU'12: International Conference on Future Computer Supported Education*, volume 2, pages 781–787.

Deng, H., Jin, L., Zhen, L., and Huang, J. (2005). A new facial expression recognition method based on Local Gabor Filter Bank and PCA plus LDA. *International Journal of Information Technology*, 11(11):86–96.

Donia, M., Youssif, A., and Hashad, A. (2014). Spontaneous facial expression recognition based on

Histogram of Oriented Gradients descriptor. *Journal of Computer and Information Science*, 7(3):31–37.

Duda, R., Hart, P., and Stork, D. (2000). *Pattern classification*. Library of Congress Cataloging-in-Publication Data.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. *University of Nebraska Press Lincoln*, 19.

Happy, S. L. (2015). Automatic Facial Expression Recognition using Features of Salient Facial Patches. In *IEEE Transactions on Affective Computing*, pages 511–518. IEEE.

Huang, X. (2014). *Methods for facial expression recognition with applications in challenging situations*. Phd thesis, University of OULU, INFOTECH OULU.

Kanade, T., Cohn, J., and Tian, Y. (2000). Comprehensive database for facial expression analysis. In *Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, pages 46–53. IEEE.

Khan, R., Meyer, A., Konik, H., and Bouakaz, S. (2013). Framework for reliable, real-time facial expression recognition for low resolution images. *Journal of Pattern Recognition Letters*, 34:1159–1168.

Lin, D. (2006). Facial expression classification using PCA and Hierarchical Radial Basis Function Network. *Journal of Information Science and Engineering*, 22:1033–1046.

Lyons, M., Kamachi, M., and J.Gyoba. The Japanese Female Facial Expression database (JAFFE). http://www.kasrl.org/jaffe.html.

Malsburg, C. (1961). *Frank Rosenblatt: principles of Neurodynamics: perceptrons and the theory of brain mechanisms*. Springer Berlin Heidelberg.

Mliki, H., Hammami, M., and Ben-Abdallah, H. (2013). Mutual information-based facial expression recognition. In *ICMV'13: ixth International Conference on Machine Vision*. Society of Photo-Optical Instrumentation Engineers (SPIE).

Ojala, T., Pietikainen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on feature distributions. *Journal of Pattern Recognition*, 29(1):51–59.

Ojala, T., Pietikainen, M., and Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with Local Binary Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987.

Ouyang, Y., Sang, N., and Huang, R. (2015). Accurate and robust facial expressions recognition by fusing multiple sparse representation based classifiers. *Journal of Neurocomputing*, 149:71–78.

Priya, G. and Banu, R. (2012). Person independent facial expression detection using MBWM and multiclass SVM. *International Journal of Computer Applications*, 55(17):52–58.

Saha, A. and Wu, Q. (2010). Facial expression recognition using curvelet based Local Binary Patterns. In *ICASSP'10: International Conference on Acoustics Speech and Signal Processing*, pages 2470–2473. IEEE.

Sánchez, A., Ruiz, J., Montemayor, A. M. A., Hernández, J., and Pantrigo, J. (2011). Differential optical flow applied to automatic facial expression recognition. *Journal of Neurocomputing*, 74(8):1272–1282.

Saragih, J., Lucey, S., and Cohn, J. (2009). Face alignment through Subspace Constrained Mean-Shifts. In *ICCV'09: International Conference on Computer Vision*.

Shan, C., Gong, S., and McOwan, P. (2009). A facial expression recognition based on Local Binary Patterns: a comprehensive study. *Journal of Image and Vision Computing*, 27(6):803–816.

Sirovich, L. and Kirby, M. (1978). Low dimensional procedure for characterization of human faces. *Journal of the Optical Society of America*, 4(3):519–524.

Soyel, H. and Demirel, H. (2010). Facial expression recognition based on discriminative Scale Invariant Feature Transform. *IEEE Electronics Letters*, 46(5).

S.Zhang, Zhao, X., and Lei, B. (2012). Facial expression recognition based on Local Binary Patterns and Local Fisher Discriminant Analysis. *WSEAS TRANSACTIONS on SIGNAL PROCESSING*, 8:21–31.

Vapnik, V. (1995). *The nature of statistical learning theory*. Springer-Verlag New York, Inc. New York, NY, USA.

Viola, P. and Jones, M. (2001). Rapid object detection using a Boosted Cascade of simple features. In *CVPR 2001: International Conference on Computer Vision and Pattern Recognition*, pages 511–518. IEEE.

Visutsak, P. (2013). Emotion classification through lower facial expressions using Adaptive Support Vector Machines. *Journal of Man, Machine and Technology*, 2(1):12–20.

Wallhoff, F. The FG-NET database with facial expressions and emotions. http://cotesys.mmk.e-technik.tu-muenchen. de/isg/content/feed-database.

Wan, S. and Aggarwal, J. (2014). Spontaneous facial expression recognition: a robust metric learning approach. *Journal of Pattern Recognition*, 47(5):1859–1868.

Yang, P., Liu, Q., and Metaxas, D. N. (2010). Exploring facial expressions with compositional features. In *CVPR'10: International Conference on Computer Vision and Pattern Recognition*, pages 2638–2644. IEEE.

Zhang, L., Tjondronegro, D., and V.Chandran (2014). Facial expression recognition experiments with data from television broadcasts and the Word Wide Web. *Journal of Image and vision computing*, 32:107–119.