Scientific Visualization, 2018, volume 10, number 5, pages 1 - 12, DOI: 10.26583/sv.10.5.01

# The Performance Improvement on 4-FB Face using Random Forest Classifier

B. Nassih[1], A. Amine[2], M. Ngadi[3], D. Naji[4], N. Hmina[5]

LGS, National School of Applied Sciences, Ibn Tofail University,
B.P. 241, university campus, Kenitra, Morocco

[1] ORCID: 0000-0002-2389-4380, nassih.bouchra@univ-ibntofail.ac.ma
[2] ORCID: 0000-0002-6806-3513
[3] ORCID: 0000-0002-4632-0531
[4] ORCID: 0000-0002-3279-0241
[5] ORCID: 0000-0002-7739-4875

**Abstract**

The growing demand in the field of security leads to the development of interesting approaches in face classification. For this reason, we bring a new method based on the extracted features of the four Frequency Blocks (4-FB) and Random Forest (RF) to classify faces and non faces, thus, we have used the fusion of three extracted features based on Discrete Cosines Transform (DCT), Uniform Local Binary Pattern (ULBP), and Histogram of Selected Regions (HSR), using RF classifier then, compared to other classifiers such as ID3, C4.5, K Nearest Neighbor (KNN), and Neural Network (NN). Firstly, we have used each descriptor separately, secondly, we have combined the three descriptors in pairs and in trinomial, to evaluate the performance of the proposed method. Our new approach is to apply RF classifier only on local features, where, exhibits pertinent information about the face image as eyes, mouth, and nose which displays by four Frequency Blocks (4-FB). The performance of our proposed method was evaluated on our created database named BOSS. Moreover, RF classifier combined with DCT and ULBP provided better results with 4-FB compared to other used methods which exceeds 99%.

**Keywords:** Face Classification, HSR, DCT, ULBP, RF, ID3, C4.5, KNN, NN, Feature Extraction, Feature Fusion.

## 1. Introduction

Recently there has been much progress in face classification due to an important role in various applications. For instance, a security system that allows access only to a people who are members of a certain group or a surveillance system that can give an alert to law enforcement agencies of the presence of a person who belongs or has a link with an international terrorist group. Each of these applications relies on the integration of a face classification system. This article is devoted to the problem of face classification [1] based on feature extraction merged by feature selection and classification task. Therefore, the objective of the feature extraction procedure is to extract invariant features representing the face information. In addition, feature selection is a global problem in machine learning and pattern recognition. It reduces the number of features and removes redundant data that helps to improve accuracy. Further, there are many difficulties and challenges in face classification, for example, the huge variations in facial expression, lighting conditions, beards, mustaches and glasses impact on the face. Thus, all these factors can influence the classification process to differentiate faces and non faces. For this reason, we have created our own facial database with complex lighting conditions and we tested it on a variety of descriptors and classifiers. In the other hand, the key challenges for improving face classification performance is finding and combining efficient and discriminative information of face image which presented by only 4-FB based on varied methods of feature extraction and classification process.

In this paper, we have proposed a new method to classify faces (faces vs non-faces) using classical descriptors which are DCT, ULBP, and HSR, with the varied classification process as RF, ID3, C4.5, KNN, and NN in order to extract the locale features across a reduced area face presented by 4-FB which we work in a smaller feature space instead of the whole image.

# 2. Related work

In the last decade, face classification has attracted keen attention, while many approaches have been proposed in this area. Concerning feature extraction methods, there are various algorithms like, DCT [2],[3], Discrete Wavelet Transform (DWT) [4], [5], [6], as well as, Local Binary Patterns (LBP) [7], Spacial Local Binary Patterns [8], which used it to generate a series of ordered LBP histograms for capturing spatial information, LBP combined with Histograms of Oriented Gradient (HOG) as a fusion descriptor [9], improved LBP [10], Local Gradient Patterns (LGP) and binary HOG [11] as local transform features for face detection, Scale Invariant Feature Transform (SIFT) as locale features [12], too, in [13], the authors kept all initial SIFT key points as features and detected the key points described by a partial descriptor on a large scale, and HOG which the authors extract HOG descriptors from a regular grid [14]. Regarding feature selection step, there are many algorithms in the literature which used in this context, such as, Genetic Algorithms (GAs) [2], where the authors have been selected the optimal feature subset, as well Principal Component Analysis (PCA) for dimensionality reduction [15], firefly algorithm [16], Particle Swarm Optimization (PSO) was used to select a subset of features that effectively represents pertinent information extracted for better classification [17]. In classification tasks, many approaches have been used as a Random Forest (RF) [18], [19], [20], K Nearest Neighbor (KNN) [21], Neural Network (NN) as presented in [22], then a Support Vector Machine (SVM) as a supervised learning algorithm [23], and Adaboost [24], where these methods become a popular technique for classification problems.

The remainder of this paper is organized as follows: Section 2, presents the selection of 4-FB, Section 3, gives an overview of the used methods of ULBP, HSR and DCT as a means of feature extraction algorithms. In Section 4, a brief description of several classifier methods is given. The proposed approach is presented in Section 4, experimental results of the proposed technique along with a comparative analysis are presented and are discussed in Section 5. Finally, we draw conclusions and we give avenues for future work in Section 6.

# 3. The selection of 4 regions of the face (4-FB)

The aim of 4-FB for the feature selection algorithm is to select a subset of the extracted features that minimizing the classification error and improve the execution time. For this reason, we proposed a feature selection approach for face classification based on RF to focus on local and significant features. In the entire face image, the chosen blocks represent left eye, right eye, nose and mouth respectively. To that end, we divided the entire image into 9 blocks, afterwards, we selected manually only the \$4\$ blocks to keep the useful information and avoid the unnecessary ones. Thus, our proposed method selects the most important part of information in the face (eyes, nose, mouth). In this respect, we reduce the number of operations, and the running time. The precision of this phase significantly impacts the performance of the next phases as long as we work in a smaller feature space presented by 4-FB.

# 4. The extraction of features using several feature extraction techniques

Feature extraction can be considered as the key of face classification. The extracted features contained the relevant information from the input data (face image). Further, feature extraction can be defined as the process where a geometrical or a vectorial model is obtained by gathering important characteristics of the face. After feature extraction part, redundant data have to be discarded. The choice of the feature set is a very important and critical task in the case of classification and detection problem. Thus, we have chosen DCT, ULBP and HSR as a means of feature extraction to generate features and to combine them. Later, we will give a brief overview of these descriptors to characterize features in this context.

## 4.1. Histogram of Selected Regions

First, Histogram is a graph which represents frequency of the data, moreover, it has numerous utilizations in image processing. The first one is the examination of the image where we can anticipate around an image by simply taking a gander at its histogram. The second one is brightness purposes. It has a wide application in: image brightness, equalization of image and thresholding that predominately used in computer vision. In our work, we use the Histogram of Selected Regions as a feature extraction method for only the significant information on the face image like eyes, nose and mouth as impacts the Fig. 1, which presents clearly the four selected regions by mesh applied to the face image.

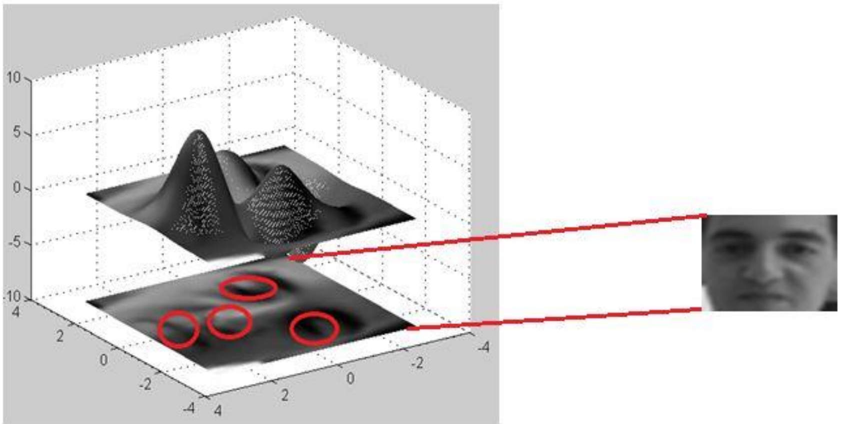

Fig.1. Mesh of face image on only 4-FB

## 4.2. Uniform Local Binary Pattern

LBP is a good technique used frequently in facial analysis and provided outstanding results in many problems relating to face and activity analysis [25]. The LBP was first introduced by Ojala et al [26], who exhibits the high discriminative power of this operator for texture classification. An extension of the original operator was made in [27] and called uniform patterns. The idea behind the Uniform LBP is to detect local characteristic textures in image like, spots, line ends, edges and corners. Through its recent extensions, the ULBP operator has been made into a really powerful measure of image texture, showing excellent results in terms of accuracy and computational complexity in many empirical studies. Moreover, the ULBP is resistant to lighting effects in the sense that they are invariant to monotonic gray-level transformations, and it has been shown to have a high discriminative power for texture classification [26].

Face image can organize as a composition of micro-patterns which can be effectively detected by the ULBP descriptor. In [28], the authors divided the face image into several regions, The ULBP histograms extracted from each region are concatenated into one histogram called Histogram features which presents the features of the entire face image.

The choice of parameters of ULBP is essentially based on the concept of neighborhood. In this sense, two parameters are used, namely: the number of neighboring pixels to be analyzed (N) and the radius of the circle on which these neighbors are located (R). Generally, these two parameters are set empirically so, that N is set to eight (8) and R, which corresponds to small values, is often set to 1. As long as this descriptor uses the notion of binary comparison, it makes it possible to combine a good image description with an ease of calculation. For this reason, we set N to 8 and R to 1.

### 4.3. Discrete Cosine Transform

The DCT is a predominant tool which introduced by Ahmed et al in the early seventies [29]. It helps to isolate the image into parts (or spectral sub-groups) of varying significance as for the image visual quality. The DCT is like the Discrete Fourier Transform, because, it changes a signal or an image from the spatial domain to the frequency domain. As known in image, most of the energy is concentrated in the lower frequencies, so if we transform an image into its frequency components and neglect the high frequency coefficients, we can reduce the amount of data needed to describe the image without sacrificing too much image quality. For calculating the DCT descriptor, we divide the image into 9 blocks of 8*8 pixels. But in our work, we took only the four blocks, which represent two eyes, nose and mouth, then we compute in each block the DCT descriptor. In short, the length of the DCT feature vector is 256, because, we have 64 coefficients in each block of 8*8 pixels.

## 5. The classification using several different classifiers

As regards the classification process, various image features are organized as input data into categories, for instance, faces and non faces, which used various classification methods as ID3, C4.5, KNN, NN, and RF. The classification algorithms typically employ two phases of processing: training and testing. Next, we will expose a brief overview of the use classification process.

### 5.1. ID3

The ID3 algorithm is among the Decision Tree implementations developed by Ross Quinlan, [30]. The ID3 is a supervised learning algorithm [31], which constructs a decision tree from a constant group of models. Afterwards, based on the resulting tree, we arrange the future examples. This algorithm by and large utilizes nominal characteristics for grouping with no missing values. It constructs a tree focused on the information (information gain) got from the training instances and after that, it utilizes the same to arrange the test datum. Further, the ID3 algorithm selects the best attribute based on the concept of entropy and information gain for developing the tree. One impediment of ID3 is that it is excessively delicate to highlights with substantial quantities of qualities. The essential parameter of ID3 is the entropy which makes it possible to find the most significant parameters in order to measure the heterogeneity of the node.

### 5.2. C4.5

The C4.5 is an improved version of the ID3 algorithm, it takes into account the numerical attributes as well as the missing values. The algorithm uses the function of entropy gain combined with a Split Info function to evaluate the attributes at each iteration. The advantage of using entropy for the ID3 or C4.5 algorithm is that these two algorithms operate for symbolic data, whether for categorical variables (such as colors) or discrete numeric variables. Nevertheless, among the disadvantages of the both methods is that the efficiency of the learning and the relevance of the model produced, remain dependent on the continuous variables which must be discretized before the implementation of the algorithm.

The C4.5 algorithm is used as a parameter the function of the entropy gain combined with a Split Info function to evaluate the attributes for each iteration.

## 5.3. K Nearest Neighbors

The K Nearest Neighbors is a non-parametric method for information grouping [32], then, it is a straightforward technique that stores every accessible case and characterizes new cases in light of a likeness measure. In the training phase, the KNN is relatively fast and simple [33]. Moreover, the instance of KNN is grouped by a larger part vote of its neighbors, with the case being attributed to the class most basic among its K closest neighbors estimated by a separation work by a distance function. In the event that K = 1, at that point the case is basically allotted to the class of its closest neighbor.. Among the parameters to be optimized we have K which represents the number of nearest neighbors used in the classification and the distance metric. After a series of tests, we set K to 7 and we chose the Euclidean distance as the distance parameter.

## 5.4. Neural Network

Neural Network (also called an Artificial Neural Network (ANN)) is an artificial system made up of virtual abstractions of neuron cells. Focused on the human cerebrum, Neural Networks are described in terms of their depth, including how many layers they have between input and output, or the model's so-called hidden layers. They can also be described by the number of hidden nodes, the model has or in terms of how many inputs and outputs each node has. Variations on the classic neural-network design allow various forms of forward and backward propagation of information among tiers. The Table 1 presents the adopted parameters of the neural network structure.

Table 1. The adopted parameters of BPNN

| *Number of layers* | *3* |
|---|---|
| Activation function | sigmoid |
| Learning Rate | 0.1 |
| Number of hidden layer | 1 |
| Number of epochs | 1000 |
| Number of neural per hidden layer | 100 |

## 5.5. Random Forest (RF)

The general method of random decision forests was introduced by Ho [34], the introduction of RF proper was first appeared in [35], which describes a method of building a forest of uncorrelated trees. Further, the RF is an ensemble learning method that grows many classification trees [36]. To classify an object from an input vector, the input vector is put down each of the trees in the forest. Then, the ensemble learning methods can be divided into two main groups: Bagging and boosting. In bagging, models are fitted in parallel where successive trees do not depend on previous trees. Each tree is independently built using bootstrap sample of the dataset. A majority vote determines prediction. The RF adds an additional degree of randomness to bagging [35]. Although, each tree is constructed using a different bootstrap sample of the dataset, the method by which the classification trees are built is improved. The RF predictor is an ensemble of individual classification tree predictors. For every perception, every individual tree votes in favor of one class and the woods predicts the class that has the majority of votes. One of the important properties of RF is their convergence with a sufficient number of trees, therefore they avoid over-learning. In addition, they are able to deal naturally with a large-scale problem based on the important variables of the problem [37].

The construction of decision trees is based on the standard "Classification and Regression Trees" (CART) algorithm. This algorithm uses the Gini index as a parameter to determine which attribute should be generated. The basic principle of CART consists, therefore, in choosing the attribute whose Gini index is minimum after the separation [38]. The adopted parameters of RF classifier are presented in Table 2.

Table 2. The parameters of the RF classifier

| | |
|---|---|
| The number of trees | 500 |
| The number of nodes | 661 |
| the number of leaves | 331 |

## 6. Proposed approach: 4-FB Selection with RF

In this study, we explore a powerful method based on RF classifier with only 4-FB to discriminate between faces and non faces images under different kind of lighting conditions which presented by BOSS database. Based on BOSS and the 4-FB, two scenarios were applied to evaluate our proposed method. These two steps are as follows:

The DCT, HSR, and LBP descriptors are applied separately. Thus, the different feature vectors that result from the different operations of extracting attributes will be used as inputs of the following classifiers: C4.5, ID3, NN, KNN and RF. The Figure 2 describes the first procedure applied to the 4-FB of the BOSS database.

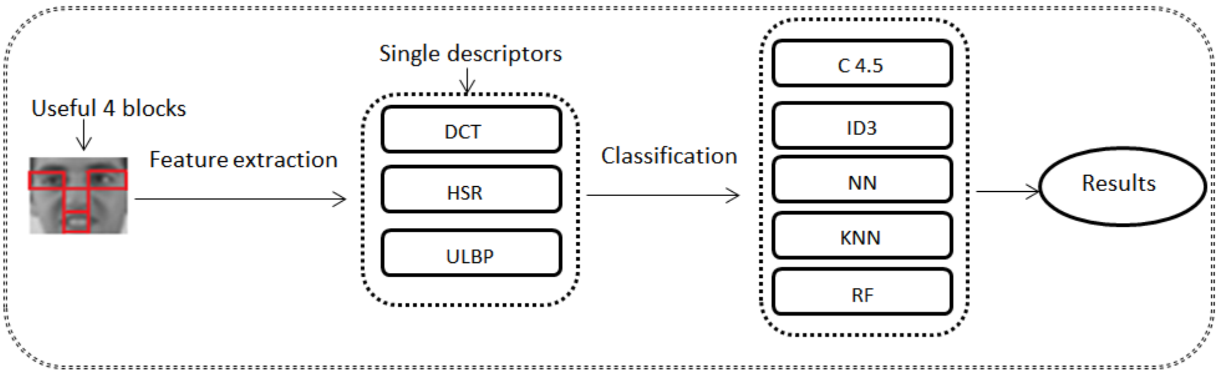


Fig.2. Scheme of our proposed approach using individual descriptors

Based primarily on the simple fusion of DCT, HSR and LBP descriptor characteristics, binomial and trinomial combinations were achieved by simple concatenations. The different feature vectors that result from the various concatenations of the feature vectors will be used as inputs of the following classifiers: C4.5, ID3, NN, KNN and RF in order to evaluate our approach. The diagram illustrated by Figure 3 summarizes the adopted approach.

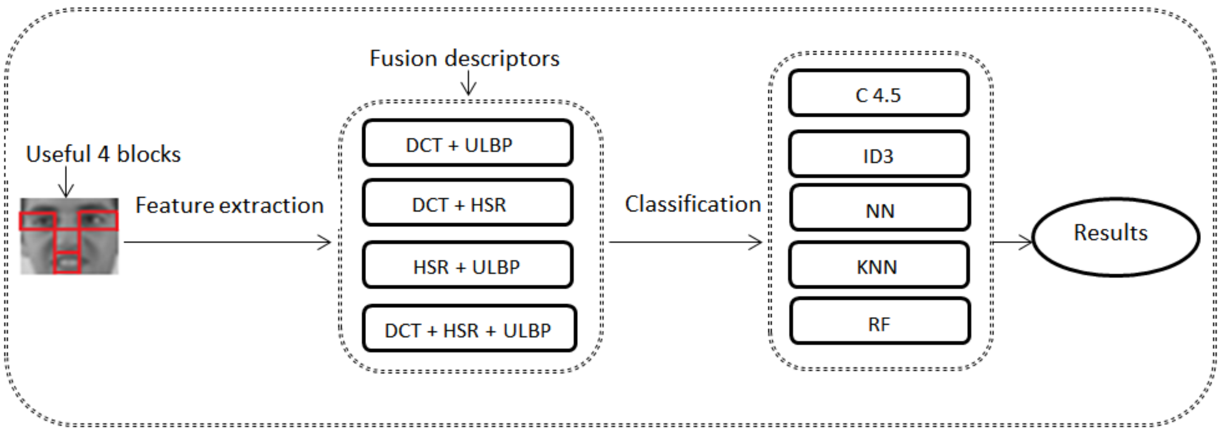


Fig.3. Scheme of our proposed approach using the fusion descriptors

# 7. Experimental Results

In this section, we will present the results of the adopted approach. Subsequently, we will evaluate the relevance of our approach on our own BOSS database based on the classification rate as evaluation criteria.

## 7.1. Description of BOSS database

The Boss database is a new database of faces and non-faces. The most face images were captured in uncontrolled environments and situations, such as, illumination changes, facial expressions (neutral expression, anger, scream, sad, sleepy, surprised, wink, frontal smile, frontal smile with teeth, open / closed eyes,), head pose variations, contrast, sharpness and occlusion. Thus, the majority of individuals is between 18-20 years old, but some older individuals are also present with distinct appearance, hair style, adorns and wearing a scarf. The database was created to provide more diversity of lighting, age, and ethnicity than currently available landmarked 2D face databases. All images were taken in 26 ZOOM CMOS digital camera of full HD characteristics. The majority of images were frontal, nearly frontal or upright. The Figure 4 imparts some typical people images of BOSS database. We detect people faces in our BOSS database by using the cascade detected of Viola-Jones algorithm. All the faces are scaled to the size 30*30 pixels. This database contains 9,619 with 2,431 training images (with 771 faces and 1,660 non-faces) and 7,188 test images (178 faces and 7,010 non-faces). The face images stored in PGM format. The Figure 5 presents some typical detected face images of BOSS database. The BOSS database will be soon publicly available for research purposes, of various algorithms related to the face detection, classification, recognition and analysis.

Fig.4. Some people images of BOSS database

Fig.5. Some detected face images of BOSS database

## 7.2. Results

As a reminder, the first scenario is to apply the DCT, ULBP, and HSR descriptors separately to the 4-FB of the BOSS database. Then, the different feature vectors which result from the various operations of feature extraction task will used as inputs of varied classification methods which are C4.5, ID3, NN, KNN and RF. Thus, the results of the first scenario are presented in Table 3 and Figure 6.

Table 3: The performance of 4-FB based on individual descriptors combined with different classifiers

| Classifieur descripteur | Accuracy (%) | | | | |
|---|---|---|---|---|---|
| | C4.5 | ID3 | NN | KNN | RF |
| DCT | 76.18 | 64.82 | 86.80 | 80.18 | **100** |
| ULBP | 90.22 | 89.7 | 98.96 | 95.46 | **100** |
| HSR | 50 | 50 | 56.20 | 80.18 | **74.58** |

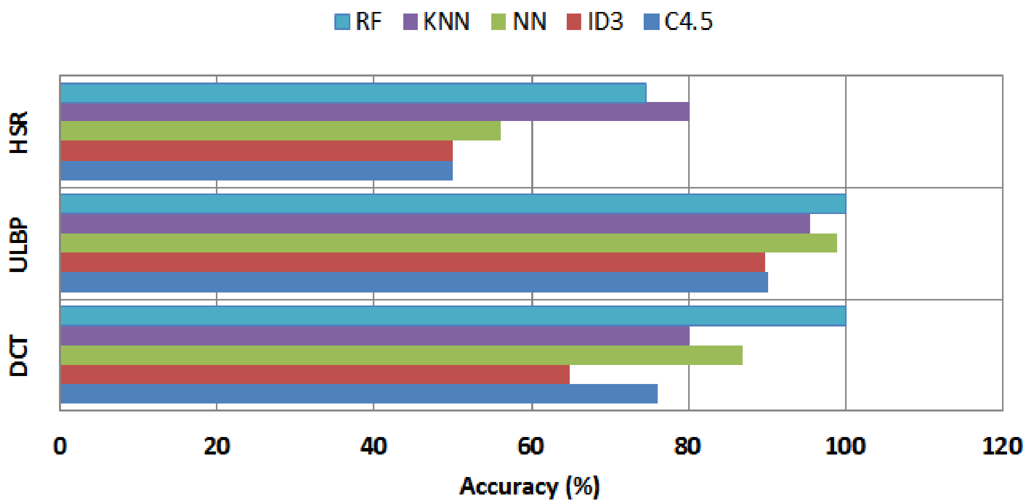


Fig. 6. The performance of RF using individual descriptors

The examination of the Table 3 and Fig. 6 have shown that the best results in terms of classification rate are generally obtained by applying the RF classifier with the DCT and ULBP descriptors, which we obtained respectively 100 % and 100 % in term of accuracy, against the extractor of HSR gets a very poor results with all the used classifiers compared to other descriptors. We can conclude that the best result was recorded by applying the RF classifier with the ULBP descriptor with 100% in term of classification rate.

The second scenario consists in carrying out the DCT-ULBP, DCT-HSR, HSR-ULBP and DCT-HSR-ULBP combinations based on the DCT, ULBP and HSR descriptors. The different feature vectors that will result will be used as inputs by varied classification methods which are: C4.5, ID3, NN, KNN and RF. Thus, the results of the second scenario are reported in the Tables 4, 5, the Figs 7 and 8.

Table 4: The performance of 4-FB based on paired combinations of descriptors using different classification methods

| Classifieur Descripteur | Accuracy (%) | | | | |
|---|---|---|---|---|---|
| | C4.5 | ID3 | NN | KNN | RF |
| DCT-ULBP | 90.3 | 81.63 | 99.88 | 97.14 | 100 |
| DCT-HSR | 91.41 | 66.5 | 90.63 | 70.44 | 99.03 |
| HSR-ULBP | 93.85 | 86.22 | 98.3 | 91.32 | 99.03 |

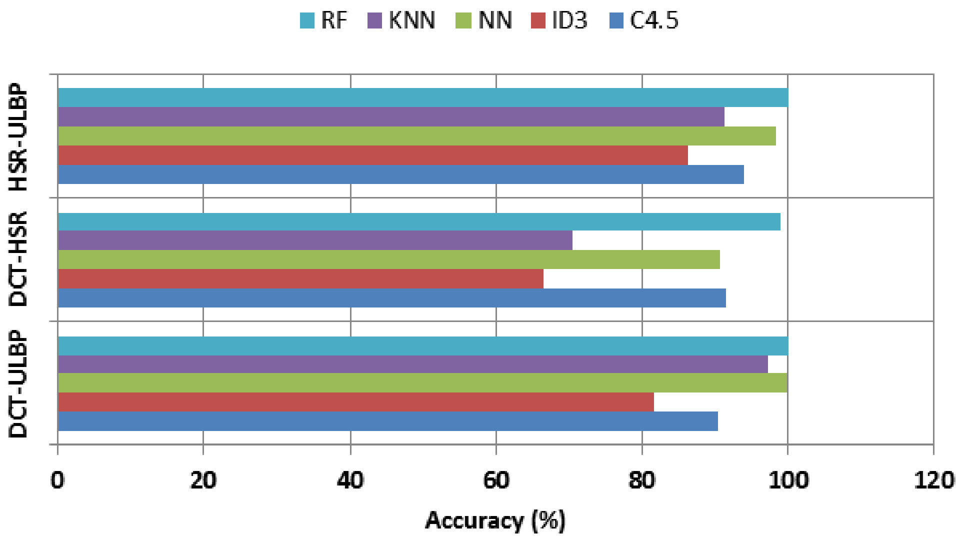


Fig. 7. The performance of RF using a paired combination of descriptors

Table 5: The performance of 4-FB based on trinomial combinations of descriptors using different classifiers

| Classifieur Descripteur | Accuracy (%) | | | | |
|---|---|---|---|---|---|
| | C4.5 | ID3 | NN | KNN | RF |
| DCT-HSR-ULBP | 91.41 | 66.07 | 90.45 | 70.44 | **99.03** |

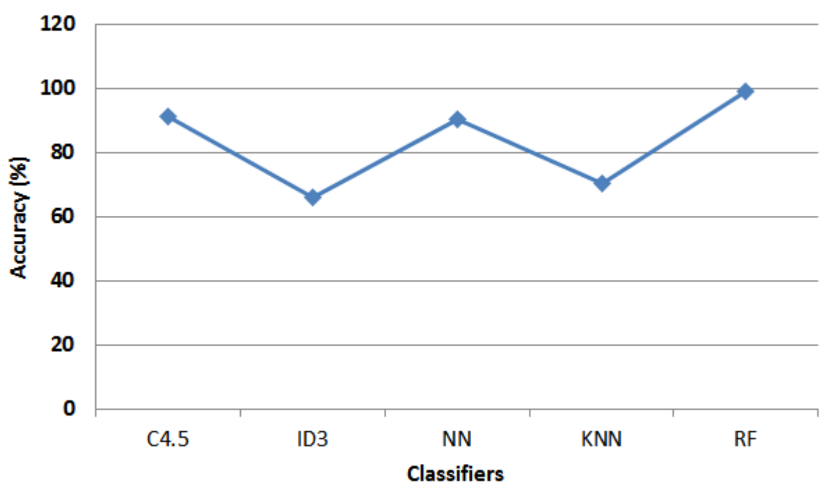


Fig. 8. The performance of RF using a trinomial combinations of descriptors

According to the Tables 4, 5, the Figs 7 and 8, the best results in term of classification rate are usually obtained by applying the RF classifier, which we obtained with the following descriptors combinations: DCT-ULBP, DCT-HSR, HSR-ULBP and DCT-HSR-ULBP an accuracy of order 100%, 99.03%, 99.97% and 99.03 % respectively. Then, we notice that the application of the HSR operator in combination with the RF classifier gave a good result in term of classification rate with the order of 99.97%. In addition, it appears that the performance of the combined extractors DCT-ULBP and HSR-ULBP with the RF classifier are generally close with a slight advance for DCT-ULBP. In addition, the performances of the combined descriptors of DCT-HSR and DCT-HSR-ULBP with the RF classifier are generally equal. At last, we can deduce that the application of the combination of the two descriptors DCT and ULBP with the RF classifier gives the best result in term of classification rate which we get 100%.

### 7.3. Comparison with the state-of-the-art methods

This section provides a comparative study of our approach and multiple well-known techniques recently published for face classification applied on the BOSS database. Refer to Table 7, the result clearly shows that the accuracy is best for the majority of algorithms based on BOSS database. For fair comparisons, it is clear that our approach of 4-FB+RF+DCT-ULBP produced the best classification rate compared to other published methods by 100% of classification rate which yields a significant improvement over the state-of-the-art methods.

Table 6. The performance of the other methods in the literature using BOSS database.

| Methods | Accuracy (%) | Research group |
|---|---|---|
| SLBP + NSVC | 99.99 | [39] |
| ULBP + NSVC | 99.47 | [39] |
| HOG + NSVC | 95.73 | [39] |
| DWT + NSVC | 91.90 | [39] |
| SLBP + DWT+ NSVC | 99.53 | [39] |
| SLBP + HOG + NSVC | 99.50 | [39] |
| ULBP + DWT + NSVC | 99.57 | [39] |
| **4-FB+RF+DCT-ULBP** | **100** | **Our approach** |
| PCA+SVM | 96.98 | [40] |
| Autoencoder+SVM | 97.52 | [40] |

## 8. Conclusions

In this paper, we tested our 4-FB approach on the BOSS database to further evaluate the relevance of BOSS and the performance of 4-FB using varied descriptors as DCT, ULBP and HSR, which involve pairing and trinomial descriptors to build more robust feature vectors using it later, as inputs of different classifiers which are RF, C4.5, ID3, NN and KNN. Our new approach is based on the 4-FB of the BOSS database combined with the RF classifier to discriminate between face and non-face. Then, we make a comparative study between RF and other classifiers including C4.5, ID3, NN and KNN to evaluate our approach. The experimental results on the BOSS database show that the proposed method based on 4-FB+RF+DCT-ULBP gets a promising results which exceeds 99% as classification rate compared to other used methods. Our future work includes applying the proposed technique to detect faces in the complex background.

## Acknowledgment

This research work is supported by the SAFEROAD project under contract No: 24/2017

## References


1. S. Zafeiriou, C. Zhang, and Z. Zhang, “A survey on face detection in the wild: past, present and future,” Computer Vision and Image Understanding, vol. 138, pp. 1-24, 2015.
2. Amine, A. el Akadi, M. Rziza, D. Aboutajdine, "GA-SVM and Mutual Information based Frequency Feature Selection for Face Recognition," infocomp Journal of Computer Science, vol. 8(1), 2009, pp. 20-29.
3. R. Chadha, P. P. Vaidya, M. M. Roja, "Face Recognition Using Discrete Cosine Transform for Global and Local Feature," Recent Advancements in Electrical, Electronics and Control Engineering, IEEE Xplore, 2011.
4. K. Ramesha, K. B. Raja, "Face Recognition System Using Discrete Wavelet Transform and Fast PCA," AIM Springer Berlin Heidelberg, vol. 147, 2011.
5. S. Sunitha, S. Katamaneni, P. M. Latha, "Efficient Gender Classification Using DCT and DWT," International Journal of Innovations in Engineering and Technology, vol. 7(1), 2016.
6. M. El Aroussi, A. Amine, S. Ghouzali, M. Rziza, D. Aboutajdine, "Combining DCT and LBP Feature Sets For Efficient Face Recognition," International Conference on Information and Communication Technologies: From Theory to Applications, 2008.
7. M. A. Rahim, M. N. Hossain, T. Wahid, M. S. Azam, "Face Recognition using Local Binary Patterns (LBP)," Global Journal of Computer Science and Technology Graphics and Vision, vol. 13(4), 2013.
8. J. Hu, P. Guo, "Spatial local binary patterns for scene image classification," International Conference on Sciences of Electronics, Technologies of Information and Telecommunications, 2013.
9. S. Paisitkriangkrai, Ch. Shen, J. Zhang, "Face Detection with Effective Feature Extraction", Computer Vision, 2010.
10. Jin, Q. Liu, X. Tang, and H. Lu, "Learning local descriptors for face detection", International Conference on Multimedia and Expo, 2005, pp.928-931.
11. Jun, I. Choi, and D. Kim, Senior Member IEEE, "Local Transform Features and Hybridization for Accurate Face and Human Detection," IEEE Transactions on Pattern Analysis and Machine Intelegence, vol. 35(6), 2013.
12. Y. Y. Wang, Z. M. Li, L. Wang, and M. Wang, "A Scale Invariant Feature Transform Based Method," International Journal of Information Hiding and Multimedia Signal Processing, vol. 4(2), 2013.
13. Geng, and X. Jiang, "Face recognition using sift features," International Conference on Image Processing, 2009, pp. 3277-3280.
14. O. Déniz, G. Bueno, J. Salido, F. De la Torre, "Face recognition using Histograms of Oriented Gradients," Pattern Recognition Letters, vol. 32(12), 2011, pp. 1598-1603.
15. S. S. Navaz, T. D. Sri and P. Mazumder, "Face Recognition Using Principal Component Analysis, and Neural Networks," International Journal of Computer Networking, Wireless and Mobile Communications, Vol. 3(1), 2013, pp. 245-256.
16. Ravindran, R. Aggarwal, N. Bhatia, A. Sharma, "Face Recognition using Evolutionary Approach," International Research Journal of Computers and Electronics Engineering, Vol. 3(1), 2015.
17. Seal, S. Ganguly, D. Bhattacharjee, M. Nasipuri and C. Gonzalo-Martin, "Feature Selection using Particle Swarm Optimization for Thermal Face Recognition," Advances in Intelligent Systems and Computing, Vol. 304, 2015.
18. Z. Kouzani, S. Nahavandi, K. Khoshmanesh, "Face Classification by a Random Forest," IEEE Region 10 Conference, 2007.
19. Kremic and A. Subasi, "Performance of Random Forest and SVM in Face Recognition," The International Arab Journal of Information Technology, vol. 13(2), 2016.
20. Salhi, M. Kardouchi, and N. Belacel, "Fast and Efficient Face Recognition System using Random Forest and Histogram of Oriented Gradients," Interna-

tional Conference of the Biometrics Special Interest Group, 2012, pp. 1-11.
21. J. P. Jose, P. Poornima, K. M. Kuma, "A Novel Method for Color Face Recognition Using KNN Classifier," International Conference on Computing, Communication and Applications (ICCCA), 2012.
22. M. Nandini, P. Bhargavi, G. R. Sekhar, "Face Recognition Using Neural Network," International Journal of Scientific and Research Publications, vol. 3(3), 2013.
23. Y. Liang, W. Gong, Y. Pan, W. Li, and Z. Hu, "Gabor Features-Based Classification Using SVM for Face Recognition," Lecture Notes in Computer Science, Vol. 3497, 2005, pp. 118-123.
24. Ying, M. Qi-Guang, L. Jia-Chen, G. Lin, "Advance and prospects of AdaBoost algorithm," Acta Automatica Sinica, vol. 39(6), 2013, pp. 745-758.
25. M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, "Computer Vision using Local Binary Patterns," Berlin, Germany: Springer-Verlag, 2011.
26. T. Ojala, M. Pietikainen, and D. Harwood, "A Comparative Study of Texture Measures with Classification Based on Feature Distributions," Pattern Recognition, vol. 29(1), 1996, pp. 51-59.
27. T. Ojala, M. Pietikainen and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE TPAMI, vol. 24(7), 2002, pp. 971-987.
28. T. Ahonen, IEEE, A. Hadid, M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 28(12), 2006.
29. T. N. N. Ahmed, and K. Rao. "Discrete Cosine Transform," IEEE Transactions Computers, vol. 23, 1974, pp. 90-94.
30. J.R. Quinlan, "Induction of Decision Trees," Machine Learning, vol. 1, 1986, pp. 81-106.
31. Shrivastava and V. Choudhary, "Comparison between ID3 and C4.5 in Contrast to IDS Surbhi Hardikar," VSRD-IJCSIT, Vol. 2(7), 2012, pp. 659-667.
32. Nigsch, A. Bender, B. v. Buuren, J. Tissen, E. Nigsch, J. B. Mitchell, "Melting point prediction employing k-nearest neighbor algorithms and genetic parameter optimization," Journal of Chemical Information and Modeling, vol. 46(6), 2006, pp. 2412-2422.
33. P. Hall, B. U. Park , R. J. Samworth , "Choice of neighbor order in nearest neighbor classification," Annals of Statistics, vol. 36(5), 2008, pp.2135-2152
34. T. K. Ho, "Random Decision Forests," International Conference on Document Analysis and Recognition, 1995, pp. 278-282
35. L. Breiman, "Random Forests," Machine Learning, vol. 45(1), 2001, pp. 5-32.
36. T. K. Ho, "The Random Subspace Method for Constructing Decision Forests," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20(8), 1998, pp. 832-844.
37. Biau, "Analysis of a random forests model," Journal of Machine Learning, 2012, pp. 1063-1095.
38. R. Kohavi and R. Quinlan, "Decision tree discovery," Handbook of Data Mining and Knowledge Discovery, 1999, pp. 267-276.
39. M. Ngadi, A. Amine, B. Nassih, H. Hachimi, and A. El-Attar, "The Performance of LBP and NSVC Combination Applied to Face Classification," Applied Computational Intelligence and Soft Computing, 2016.
40. Nassih, A. Amine, M. Ngadi, N. Hmina, "Non-Linear Dimensional Reduction Based Deep Learning Algorithm for Face Classification," International Journal of Tomography and Simulation, vol. 30, no. 4, 2017.