International Journal of Electrical and Computer Engineering (IJECE)
Vol. 13, No. 4, August 2023, pp. 4052~4067
ISSN: 2088-8708, DOI: 10.11591/ijece.v13i4.pp4052-4067

Robust face recognition using convolutional neural networks combined with Krawtchouk moments

Yassir El Madmoune, Ilham El Ouariachi, Khalid Zenkouar, Azeddine Zahi
Laboratory of Intelligent Systems and Application, Faculty of Sciences and Technology, University Sidi Mohamed Ben Abdellah, Fez, Morocco

Article Info

Article history:

Received Jul 16, 2022
Revised Oct 25, 2022
Accepted Nov 6, 2022

Keywords:

Convolutional neural networks
Face recognition
Image moments
Krawtchouk moments
Noisy conditions

ABSTRACT

Face recognition is a challenging task due to the complexity of pose variations, occlusion and the variety of face expressions performed by distinct subjects. Thus, many features have been proposed, however each feature has its own drawbacks. Therefore, in this paper, we propose a robust model called Krawtchouk moments convolutional neural networks (KMCNN) for face recognition. Our model is divided into two main steps. Firstly, we use 2D discrete orthogonal Krawtchouk moments to represent features. Then, we fed it into convolutional neural networks (CNN) for classification. The main goal of the proposed approach is to improve the classification accuracy of noisy grayscale face images. In fact, Krawtchouk moments are less sensitive to noisy effects. Moreover, they can extract pertinent features from an image using only low orders. To investigate the robustness of the proposed approach, two types of noise (salt and pepper and speckle) are added to three datasets (YaleB extended, our database of faces (ORL), and a subset of labeled faces in the wild (LFW)). Experimental results show that KMCNN is flexible and performs significantly better than using just CNN or when we combine it with other discrete moments such as Tchebichef, Hahn, Racah moments in most densities of noises.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Yassir El Madmoune
Laboratory of Intelligent Systems and Application, Faculty of Sciences and Technology, University Sidi Mohamed Ben Abdellah
Fez, Morocco
Email: yassir.elmadmoune@usmba.ac.ma

1. INTRODUCTION

Face recognition is an essential aspect of biometric technologies [1], [2]. that has received significant attention due to the fast development of technology such as digital cameras [3], the Internet [4], and mobile devices [5], as well as the rising desire for security [6]–[9]. Face recognition offers various benefits over other biometric systems, including natural, non-intrusive, and simple. However, face recognition has become one of the most challenging pattern recognition problems, owing the wide range of lighting circumstances, face expression, head size, pose variation, complex background, motion blurring, noisy conditions, and other environmental variables that could reduce recognition performance [10].

Three primary categories can be used to categorize the various approaches that have been utilized for face recognition: holistic approaches [11]–[21], feature-based approaches [22]–[30] and hybrid approaches [31]–[34]. In early 1990, researchers in face recognition field started using holistic approaches, i.e., facial detection systems use the entire face region as input to accomplish face recognition. In this approach, we find two sub-categories of techniques: the first one is based on linear methods like Eigenfaces principal component analysis (PCA) [11], [12], Fisherfaces linear discriminative analysis (LDA) [13], [14], independent component analysis (ICA) [15], discrete wavelet transform (DWT) [16] and discrete cosine transform (DCT) [17]. The second technique is based on non-linear methods such as Kernel PCA (KPCA) [18], kernel linear discriminant analysis (KLDA) [19], Gabor-KLDA [20], and CNN [21]. In the first decade of the 21st century, studies have focused on feature-based approaches, and could possibly be separated into two distinct types: local appearance-based techniques that consider the facial image as a collection of discrete vectors with low dimensions and focus on crucial parts of the face like the nose, mouth, and eyes to create additional information and make face recognition easier. Local binary pattern (LBP) [22], histogram of oriented gradients (HOG) [23], correlation filters (joint transform correlator (JTC) [24], VanderLugt correlator (VLC) [25]) and discrete orthogonal moments (DOM) [26] are the most methods used in this sub category. In the second sub-category, keypoints-based techniques are utilized to detect particular geometric characteristics based on the geometry of the facial features (e.g., the space between the eyes, the circumference of the head) using algorithms like scale-invariant feature transform (SIFT) [27] and descriptors like speeded-up robust features (SURF) [28], binary robust independent elementary features (BRIEF) [29] and fast retina keypoint (FREAK) [30]. In early 2010, the face recognition community focused on hybrid approaches that combine local and subspace features to maximize the strengths of both types of approaches which could provide enhanced performance in face recognition systems, such as Gabor wavelet and linear discriminant analysis (GW-LDA) [31], multi-sub-region-based correlation filter bank (MS-CFB) [32], CNNs and stacked auto-encoder (SAE) [33], advanced correlation filters and Walsh LBP (WLBP) [34], Figure 1 shows a brief organization of the previous mentioned approaches. Recently, deep learning (DL) and more specifically convolution neural networks (CNN) is the most commonly methodology used for extracting features in face recognition, it has significant advantages due to its learning ability, generalization, and robustness [14], [15]. Deep and extensive neural networks have demonstrated remarkable performance with massive training datasets and the computing capacity of graphical processing units (GPUs); It could generate the fundamental feature representation of data and create high-level features from the low-level pixels.

Ding *et al.* [35] presented the noise resistant network (NR-network), a deep learning network-based system that extracts low-level and high-level face characteristics using a multi-input structure; they used a downscaling approach to reduce the resolution of their dataset in order to accelerate the processing, focusing on facial recognition in noisy conditions. However, basic design and massive pooling operations are lost certain facial features. As a result, such a system will not be able to recognize faces in noisy environments. Meng *et al.* [36] presented a deep CNN with sub-networks for denoising and recognizing faces under noise; unlike traditional approaches, which train the two sub-networks separately, this method trains them together; hence, it requires more time. Wu *et al.* [37] proposed a light CNN framework based on three networks that reduce the computational costs and the number of parameters to train a 256-D compact embedding from enormous face data with several noisy labels, Ma *et al.* [38] introduce a robust local binary pattern (LBP) guiding pooling (G-RLBP) mechanism to enhance the accuracy of CNNs models while effectively reducing the noise effects.

Dimensionality reduction and feature extraction are essential parts of any facial recognition system. Despite the fact that face images have a high dimensionality despite their small size, which leads to a significant amount of computational time, complexity, and memory occupation; the performance of any classifier is mainly determined by the good discriminating features included inside the face image [39], [40]. In this sense, the presence of noisy training data can harm the ultimate performance of trained convolutional neural networks. Although a recent research demonstrated that deep CNNs work well even on noisy samples with sufficient clean data [41], this conclusion is not always applicable in face recognition. Experimental tests indicate that noisy data appears to reduce the performance of trained face recognition CNNs [42]. To overcome these constraints and improve performance, another feature extraction technique that can deal with noise must be used.

Orthogonal moments are robust in the presence of image noise and have a near-zero redundancy measure in a feature set. In this respect, 2D DOM that are based on the Krawtchouk polynomials [43] has the ability to extract local features from any region of interest in an image in addition to the global feature extraction capability. Apostolidis and Papakostas [44] showed that using Krawtchouk moments as an image local descriptor and a watermarking attack can affects the accuracy of deep learning models when it applied in medical images. Amakdouf *et al.* [45] came up with quaternion radial Krawtchouk moments that could be useful in the field of color image analysis by showing a good representation capability and robustness to different noises. Hassan *et al.* [46] demonstrated that invariant quaternion Krawtchouk moments are more effective than continuous orthogonal moments at representing images and showed more stability against the translation, rotation, and scaling transformation. Rahman *et al.* [47] introduced a new method for face recognition in which sparse representation of face images is created by selecting a discriminatory set of orthogonal Krawtchouk moments. Following the considerations presented above, the Krawtchouk DOM are investigated for grayscale face image recognition. The essence of our suggested model is to employ Krawtchouk moments as a fast and accurate object descriptor; the whole face shape moments could be computed and fed it as input layer to a convolutional neural network, the robustness of our proposed model is tested on small and large

size databases with the presence of two types of noise and compared with CNN combined with others 2D DOM and without them. The main contributions of this study are summarized as follows:
- A new architecture named Krawtchouk moments convolutional neural networks (KMCNN), defined by Krawtchouk orthogonal polynomials, is introduced for the first time in this paper.
- A robust face recognition approach against various types of noises is proposed.
- An application of the suggested KMCNN model for face reconstruction and recognition is presented.

The remainder of this paper is structure as follows. Section 2 a brief review of 2D Krawtchouk orthogonal moments and the process of creating image moments. Section 3 describes the proposed KMCNN model and its architecture. The databases are considered in section 4. Experiments and results details are also conducted to evaluate the KMCNN compared with CNN only and its combination with other 2D orthogonal moments in this section. Section 5 concludes this paper.

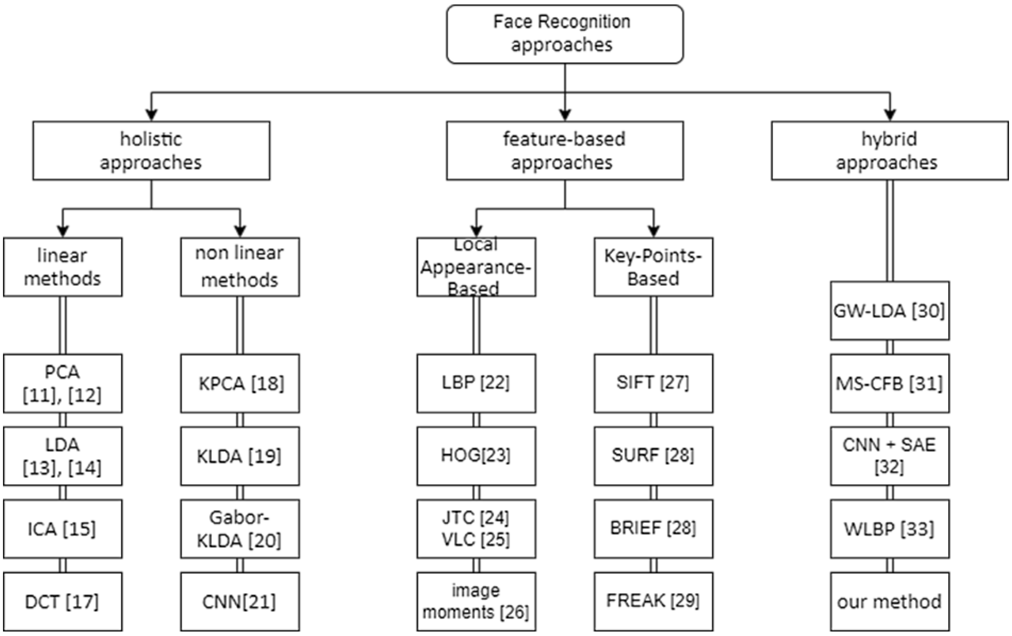


Figure 1. Summary of face recognition approaches

2. 2D KRAWTCHOUK MOMENTS

Krawtchouk moments are a set of orthogonal moments based on the discrete Krawtchouk polynomials defined over the coordinate image space. Their implementation does not involve any numerical approximation. In this section, we will give a brief formulation of 2D weighted Krawtchouk moments, including polynomials and describe their capacity to capture significant features from images with a significant dimensionality reduction.

2.1. Krawtchouk polynomials

The Krawtchouk polynomials were initially presented by Krawtchouk [48], and recently utilized by Yap *et al.* [49] image analysis fields. The orthogonality relation of the Krawtchouk discrete polynomials is given by (1).

$$\sum_{x=0}^{N-1} w_k(x;p,N)k_n(x;p,N)k_m(x;p,N) = \rho_k(n;p,N)\delta_{nm} \quad n,m = 1,\dots,N, \tag{1}$$

where $w_k(x;p,N)$ is the weighting function defined as (2):

$$w_k(x;p,N-1) = \binom{N-1}{x} p^x (1-p)^{N-1-x}, \tag{2}$$

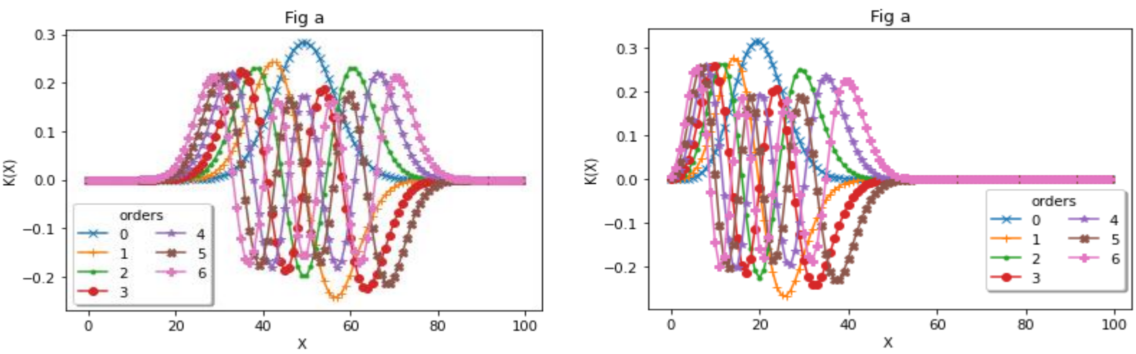
with the norm function is:

$$\rho_k(n;p,N-1) = (-1)^n \left(\frac{1-p}{p}\right)^n \frac{n!}{(1-N)_n}. \tag{3}$$

Using the definition above, Yap *et al.* [49] presents the recurrent formula by using the normalized Krawtchouk polynomials.

$$k_n(x;p,N-1) = A_n k_{n-1}(x;p,N-1) - B_n \bar{k}_{n-2}(x;p,N-1)$$
$$\bar{k}_0(x;p,N-1) = w_k(x;p,N-1)$$
$$\bar{k}_1(x;p,N-1) = w_k(x;p,N-1)\frac{(N-1)p-x}{\sqrt{(N-1)p(1-p)}} \tag{4}$$
$$with\ A_n = \frac{((N-1p-2(n-1)p+n-1-x)}{\sqrt{p(1-p)n(N-n)}}\ and\ B_n = \sqrt{\frac{(n-1)(N-n+1)}{(N-n)n}}$$

Figures 2(a) and (b) show the weighted Krawtchouk polynomials up to the 7th degree for *p*=0.5 and *p*=0.2, respectively. The graphs illustrate the impact of the localization parameter *p*, which permits the polynomials to be moved to the appropriate location.



(a) (b)

Figure 2. Weighted Krawtchouk polynomials up to the 7th degree for *N*=168, (a) *p*=0.5 and (b) *p*=0.2

2.2. Krawtchouk moments

In general, moments are defined as scalar values, that are consistent and efficient data descriptors [50]. They may be used to represent 1D signals like voice and 2D signals such as images without information redundancy in the moment set and to detect slight signal intensity variations [51]. For a two-dimensional signal with intensity function *f(x, y)* of size *N1×N2*, Krawtchouk moments ψ_{nm}can be defined as (5) [50], [52]:

$$\psi_{nm} = \&\sum_{x=0}^{N_1-1}\sum_{y=0}^{N_2-1}\bar{K}_n(x;p,N_1-1)\bar{K}_m(y;p,N_2-1)f(x,y)$$
$$n = \&0,1,\dots,M_1-1 \text{ and } m = 0,1,\dots,M_2-1, \tag{5}$$

where *M1* and *M2* are the maximum moment orders used to describe the intensity signal *f(x, y)*. To recover *f(x, y)* from Krawtchouk Moments, the (6) is used:

$$\hat{f}(x,y) = \sum_{n=0}^{M_1-1}\sum_{m=0}^{M_2-1}\bar{K}_n(x;p,N_1-1)\bar{K}_m(y;p,N_2-1)\psi_{nm}$$
$$x = 0,1,\dots,N_1-1 \text{ and } y = 0,1,\dots,N_2-1, \tag{6}$$

where $\hat{f}(x,y)$ is the reconstructed function, $\hat{f}(x,y) = f(x,y)$ when all moments are taken into account throughout the reconstruction process.

Figures 3(a) to (c) shows respectively a sample of an original face image from YaleB database [53], a noisy image when we mix the original image with 5% of salt and pepper and speckle noises. Figures 4 and 5 show respectively reconstructions of face images mixed with salt and pepper and speckle noises, where subfigures (a) to (j) show the reconstruction up to orders 168 from 10 to 160 with 20 increments by using 2D Krawtchouk moments and noisy images shown in Figures 3(b) and (c). We choose *p = 0.5* to obtain the highest representation; In the early stages, we can see a more striking resemblance between the noisy images and the reconstructed ones. This indicates that 2D Krawtchouk moments have the ability to extract more information from images in a smaller space, which means that instead of using the original picture, we may employ image moments to reduce dimensionality and extract meaningful features for classification.

(a) (b) (c)

Figure 3. sample of image with/without noise from YaleB database, (a) original image, (b) with 5% of salt and pepper noise, and (c) with 5% of speckle-noise

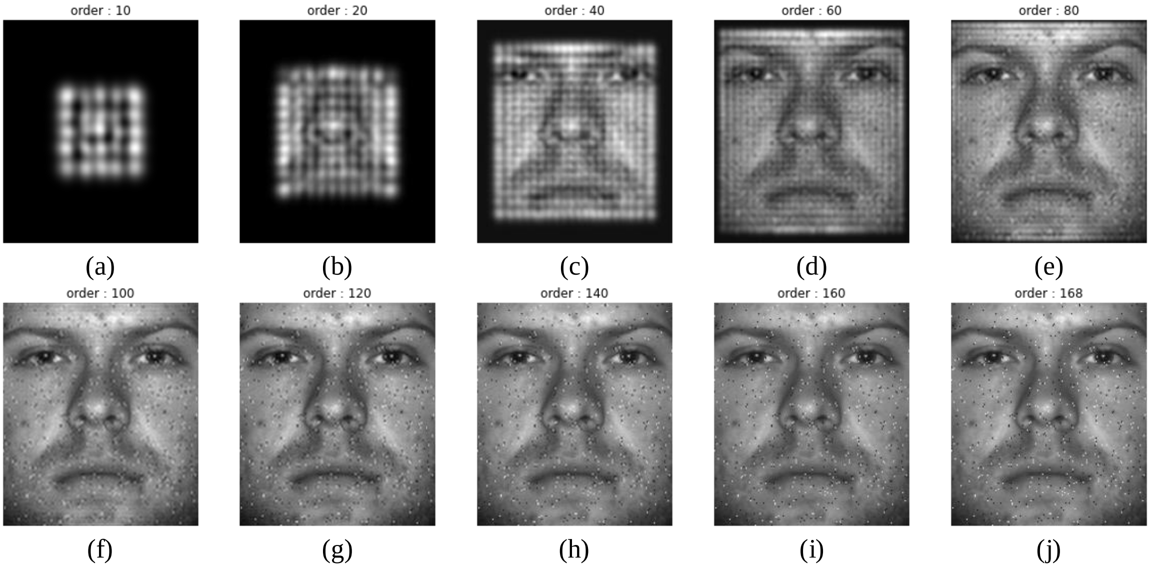


Figure 4. Reconstruction results using 2D Krawtchouk moments and an image from YaleB dataset mixed with 5% of salt and pepper noise, (a) to (j) from 10 to 160 with 20 increments

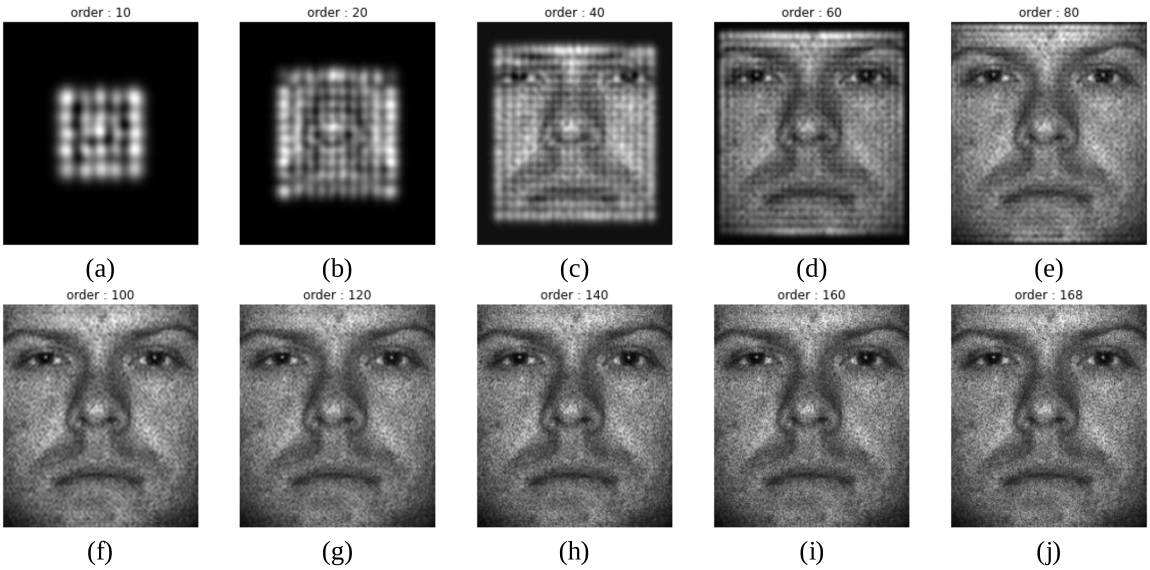


Figure 5. Reconstruction results using 2D Krawtchouk moments and an image from YaleB dataset mixed with 5% of speckle noise, (a) to (j) from 10 to 160 with 20 increments

3. PROPOSED MODEL

In this work, we presented a novel architecture for face recognition problems named KMCNN that combine the idea of orthogonal moments with the 2D CNN model, as shown in Figure 6. Indeed, duo to Krawtchouk moments property for representing face images in lower orders without redundancy, as demonstrated in the previous section; which facilitates the production of small 2D moments matrices that are inserted into a 2D convolutional neural network. Therefore, we get two benefits from this combination, the processing complexity is significantly reduced, and the computational speed is increased. Table 1 provides a summary of the principal model layers, and the suggested architecture design is organized as follows:

- 2D Krawtchouk Moment layer: As the first layer of the KMCNN, Krawtchouk discrete orthogonal moments compute the input image moments by using (5) and provide a matrix whose size is proportional to the moment order value. This layer optimizes the image representation and decreases the processing dimension significantly. The subsequent 2D convolutional layer is then given this matrix of moments.
- 2D Convolution layer: the purpose behind this layer is to recognize the presence of a set of features in the moment matrix rather than the original image, by the use of 2D convolution operators. The output activation value $a(i,j)^L$ at position (i,j) is calculated by (7).

$$a(i,j)^L = f\left(\sum_{x=i}^{i+N-1}\sum_{y=j}^{j+N-1}\sum_{s=0}^{S-1} W_{s,x,y}M_{s,x,y} + b^L\right) \tag{7}$$

where the matrix of moments M convolves with the L^{th} filter with a size of $N\times N$, S is the number of input channels, W is the weight matrix with size (C, N, N), *i,j* are the indices of the output position, *x,y* are the indices of the input position. f is the activation function.

- Activation functions ReLU: The output feature maps from the convolution layer are given a non-linear transformation when they are sent through the activation layer. By transforming the data into a non-linear format, it facilitates the identification of complex features that cannot be explained using a linear combination of a regression technique. The most regularly used non-linear functions are sigmoid and hyperbolic tangent; in this work, the rectified linear unit (ReLU) defined by $f(x)=max(0, x)$ is used, since it improves the non-linearity, avoids network saturation and speeds up training time networks [54]–[56].
- Batch normalization: Allows each convolutional layer to learn more independently. This layer normalizes the output of the preceding layers to enhance their learning process and prevent overfitting and divergence risks [57].
- 2D Max-pooling layer: A pooling layer is typically added following the convolution layers, to decrease the size of the feature maps. Consequently, the number of network parameters, as a result the computation time is accelerated and the chance of the model falling in overfitting is diminished.
- Fully connected layer: Is the last layer in our proposed KMCNN, this layer performs a linear combination on the data received from the preceding layers, and then applies the softmax function to produce the probability of each class as a new output vector.

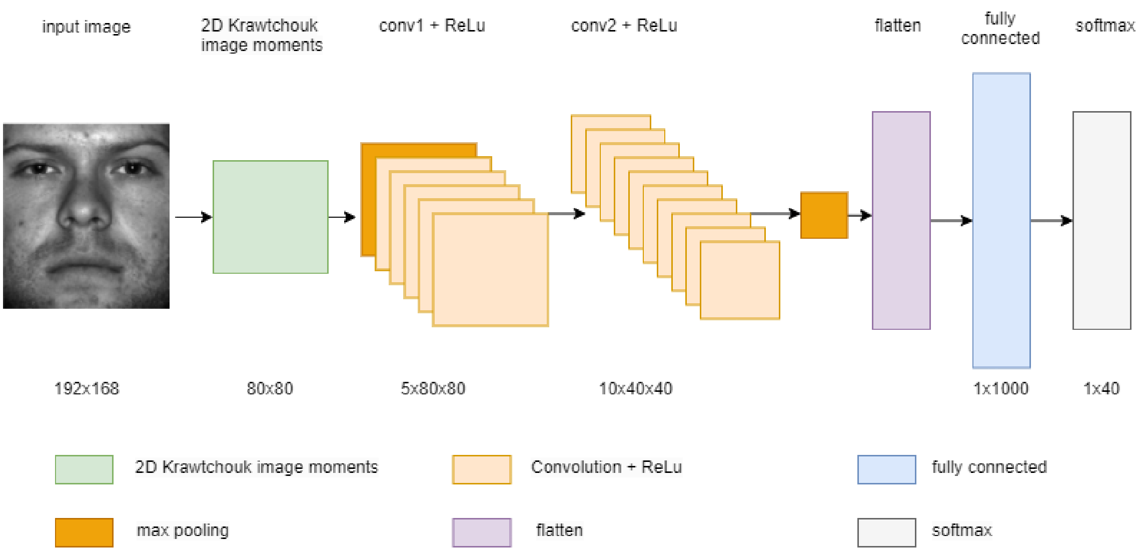


Figure 6. 2D KMCNN architecture

Table 1. Specifics of the suggested model.

Layer	Purpose	Filter	Number of filters	Activation
0	Input image	-	-	NxN
1	Moments layer	-	-	nxn
2	Conv+ReLU	3x3	5	$nxnx5$
3	MaxPooling	2x2	-	$n/2xn/2 * 5$
4	Conv+ReLU	3x3	10	$n/2xn/2 * 10$
5	MaxPooling	2x2	-	$n/4xn/4 * 10$
6	Fully connected	-	-	1,000
7	Softmax	-	-	number of subjects

4. EXPERIMENTS AND RESULTS

This section presents details about experiments conducted on face images using the proposed model KMCNN and provides a thorough description of the databases. The experiments are divided into two parts, the first one was conducted in free noise environment where the second was performed with presence of noise. Additionally, this section discusses the recognition accuracies obtained.

4.1. Experiments

In this sub-section, we evaluated the classification performance of the proposed model by carrying out a number of relevant experiments using YaleB extended database [53], our database of faces (ORL) database [58] and a subset of 10 classes from labeled faces in the wild (LFW) database [10]. By randomly dividing each database into 70% for training and 30% for testing, the efficacy of the suggested model is examined, the results are properly compared to several 2D orthogonal moment-based approaches. All experiments were conducted in the cloud using Google Collaboratory with 2.20 GHz, Intel(R) Xeon(R) CPU, NVIDIA-SMI GPU and 13 GB of RAM. The evaluation of the recognition rate of the suggested model with/without noise is structured around three primary comparisons:

- First, a comparison of the accuracy of Tchebichef, Krawtchouk, Hahn, and Racah moments as an input layer of the suggested CNN architecture was conducted using YaleB extended, ORL, and a subset of LFW database without any presence of noise. A comparison with existing methods is also presented.
- Second, we compared our suggested model KMCNN against CNN only, in noisy environments using two forms of noise (salt and pepper and speckle)
- Third, we have compared our proposed model with CNN combined with other 2D discrete orthogonal moments and Krawtchouk combined with pre-trained VGG16 model [59] in the same noisy environments.

In addition, we used different densities of noise to test our model in noisy environments, by varying the salt and pepper noise densities from 1% to 5% and Speckle noise by varying the variance value from 0.1 to 0.5 and a fixed the mean at 0.

4.2. Datasets

In the course of those experiments, three face image databases are utilized, in order to investigate the recognition rate performance. The two first databases contain grayscale images, whereas the third provide red, green and blue (RGB) images that have been transformed to grayscale format. The selected databases are as follows:

- The extended YaleB database: comprises 16,128 pictures of 28 individuals in 9 different positions and 64 lighting settings. This database follows the same data structure of the YaleB Database. In contrast to the original YaleB database consisting of 10 participants, the extended database was originally revealed by Lee *et al.* [53]. Since we are not concentrating on position variation, only the frontal face image of each subject with 64 different illuminations will be selected, totaling 2,432 images from 38 distinct subjects. Manual alignment, cropping, and resizing to 168 by 192 pixels is performed on every image used in the experiments. Figure 7 depicts a selection of facial image instances.
- The ORL database [58] consists of 400 images in total, including 40 persons with 10 unique image (4 females and 36 males), the images were captured at various periods, changing the illumination, face gestures (open/closed eyes, smiling/not smiling), and facial characteristics (glasses/no glasses). Figure 8 shows that all of the images were taken with the people standing up and facing forward against a black background. The dimensions of each image are 70 by 80 pixels, and there are 256 levels of gray for each individual pixel.
- The LFW database [10] includes 13,233 face images gathered from the internet. This collection contains 5,749 identities for 1,680 individuals with two or more images. In this work we choose a subset of 10 classes of people that have the most available images with total of 1456 images and the dimensions of each image are 240 by 240 pixels. Figure 9 shows an example of images used.

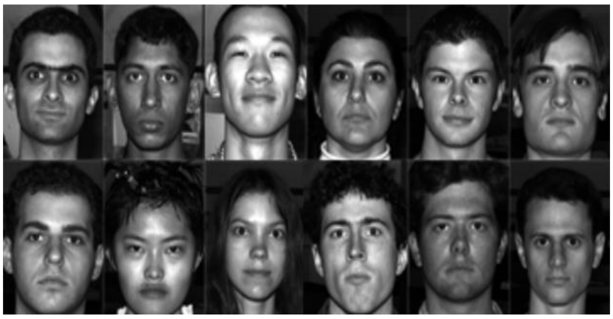


Figure 7. Examples of images from YaleB database

Figure 8. Examples of images from ORL database

Figure 9. Examples of images from a subset of LFW database

4.3. Results

4.3.1. Face recognition with free noise

Experiment 1: comparison between orthogonal moments.

As mentioned in the first sub-section, we begin our experiments by analyzing the classification performance of the suggested CNN architecture using original images from YaleB [53], ORL [58], and LFW [10], and compared the results with CNN combined with Tchebichef moments [60], Hahn moments [61], Racah Moments [62], the corresponding classification accuracy results using the databases mentioned before started from lower orders to the maximum order are listed and summarized in Tables 2 to 4, each column in the tables represents the performance in terms of accuracy of the suggested CNN architecture combined with a different type of 2D orthogonal moments.

Based on results from Table 2, the greatest score is achieved at the order (168,168) using Krawtchouk moments as an input layer on YaleB database with 92.03% of accuracy, followed by Tchebichef moments with a precision of 87.98% at the order (140,140), Hahn moments with 86.36% of accuracy at the order (160,160) and Racah moments with 82.99%. Nevertheless, we can see that Krawtchouk moments give interesting results starting from order 60 by surpassing 90% in terms of accuracy. However, As shown in Table 3, the fusion of CNN and Krawtchouk moments does not surpass other fusions of CNN with 2D discrete moments when we tested it on small-size face images from ORL database. Table 4 clearly shows that the suggested KMCNN outperforms the rest of models based on other 2D discrete orthogonal moments; we can clearly notice that the combination of Krawtchouk moments with convolutional neural networks gives 74.30% in terms of accuracy at order 20 when we test it on original images of LFW database. As a conclusion achieved from Tables 2 to 4, we can say that face image recognition by using CNN and Krawtchouk moments as input layer was significantly improved compared with results obtained using CNN combined with Tchebichef, Racah and Hahn moments.

Experiment 2: comparison with the state-of-the-art methods.

In order to illustrate the effectiveness of the suggested model, the classification accuracy is compared with the state-of-the-art approaches for face recognition. Table 5 shows the comparative analysis of the recognition rate, between the suggested KMCNN and the other approaches for the extended Yale B and the ORL databases. Each row from the table shows the method and the corresponding accuracy achieved, whereas the last row represents the accuracy of the KMCNN.

Table 2. Classification accuracies for different orders using Tchebichef, Krawtchouk, Hahn and Racah moments tested on YaleB database

Order	Tchebichef moments	Krawtchouk moments	Hahn moments	Racah moments
10	43.45	28.74	47.23	38.86
20	71.52	66.93	71.65	68.82
40	85.15	89.47	82.18	80.43
60	85.34	90.41	84.34	82.45
80	84.88	90.55	84.34	**82.99**
100	84.21	90.95	80.56	82.45
120	82.18	91.22	82.72	81.37
140	**87.98**	90.95	81.24	80.29
160	85.42	91.76	**86.36**	78.40
168	85.69	**92.03**	85.96	77.59

Table 3. Classification accuracies for different orders using Tchebichef, Krawtchouk, Hahn and Racah moments tested on ORL database

Order	Tchebichef moments	Krawtchouk moments	Hahn moments	Racah moments
10	56.1	28.05	63.41	47.56
20	91.46	78.05	93.9	87.8
30	95.12	96.34	93.9	95.12
40	96.34	**97.56**	**97.56**	96.34
50	**98.78**	95.12	97.56	96.34
60	97.56	95.12	97.56	**97.56**
70	96.34	95.12	97.56	96.34

Table 4. Classification accuracies for different orders using Tchebichef, Krawtchouk, Hahn and Racah moments tested on LFW database

Order	Tchebichef moments	Krawtchouk moments	Hahn moments	Racah moments
10	41.44	56.16	38.36	41.44
15	39.73	68.84	39.04	34.93
20	44.86	**74.32**	36.99	40.41
25	46.23	72.26	37.67	40.41
30	50.68	70.55	42.81	40.75
40	53.08	64.04	47.95	47.6
60	60.96	59.93	52.05	50.0
80	**64.38**	56.16	56.16	56.85
100	63.7	54.11	**58.22**	55.82
150	61.99	50.34	57.53	56.16
200	63.36	54.79	58.22	**57.53**
250	63.7	52.4	57.53	56.16

Table 5. the comparison between the state-of-the-art methods and the KMCNN for the YaleB and ORL databases

YaleB		ORL	
Methods	*Accuracy*	*Methods*	*Accuracy*
LSP [63]	85.6	PCA [18]	93.91
POEM [64]	90.5	2DHOG [65]	97
LBP [66]	78.6	SIFT [67]	97
GENet [68]	84.21	SURF [69]	88
Gabor [70]	87.19	HOG + ConvNet [71]	95.5
KMCNN	92.03	KMCNN	97.56

To validate the efficacy of our suggested approach, we compared it to other methods were used the Extended YaleB and ORL databases. Following the comparison procedure, it is obvious that our methodology exceeds the methods indicated above in terms of recognition rate. Thus, we may assume that our KMCNN has the potential to be very effective in a wide range of computer vision applications.

4.3.2. Face recognition with noise

The second experiment was performed on the same databases, but instead of using original images, we compared our model KMCNN with the proposed CNN architecture using noisy images. Each column of Tables 6 to 11 illustrates the precision of the suggested model in terms of accuracy employing various salt and pepper and speckle noise degradations, starting from 1% to 5%. Each row represents results obtained from each order starting from lower orders to bigger ones, except the last row that shows the accuracy of the proposed CNN architecture without using Krawtchouk moments.

According to the results shown in Tables 6, 8, and 10, the KMCNN obtained good classification rates for various degradations of salt and pepper noise, beginning at order 40 when evaluated in YaleB, particularly when the accuracy was 88.93% even with 5% of noise. We also remark that results of the KMCNN are more accurate than those of CNN only, even if samples are under 5% of the same noise using ORL database. The high robustness of the KMCNN can be noticed when we compare it with CNN taking (as input) noisy images from LFW database; it achieves an accuracy between 71.58% and 73.97% compared with CNN that not even surpassing 68% in terms of accuracy.

Taking into account the speckle noise classification rate values shown in Tables 7, 9, and 11, it can be clearly observed that the KMCNN provides the highest classification rates with YaleB database, particularly when the accuracy was 90.82% even with a variance of σ=0.4, and it performs better than CNN. By using noisy images from ORL database the KMCNN surpasses 96%, while CNN only did not even surpass 93%. Alternatively, we notice also that the KMCNN gives interesting accuracies in very low orders using LFW database.

Considering the results depicted in Tables 6 to 11, it is evident that the accuracy values increase with the order of the moments up to an optimal order; after that it starts to decrease, but what is important is that the best results are obtained in lower orders and they are better than the results obtained by CNN. From this fact, we may deduce that the KMCNN is highly noise-tolerant, which is necessary for face recognition in noisy environments. Hence the KMCNN confirms the fact presented in [42], indicating that face recognition using CNN is intolerant to noisy conditions.

Table 6. Classification accuracy using Krawtchouk moments and YaleB database in noise-free and salt and pepper noisy environment

Krawtchouk moments + CNN	Free Noise	Salt and Pepper noise				
		1 %	*2 %*	*3 %*	*4 %*	*5 %*
Order	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*
10	28.74	25.91	25.91	27.39	24.83	25.10
20	66.93	65.45	63.02	64.23	60.59	61.26
40	89.47	77.59	83.13	85.56	82.45	84.07
60	90.41	88.66	87.85	87.98	87.58	86.63
80	90.55	89.87	90.41	89.60	89.33	88.93
100	90.95	88.93	89.33	89.47	88.79	87.98
120	91.22	89.87	87.71	87.71	86.90	86.36
140	90.95	89.06	87.17	87.04	86.36	84.48
160	91.76	87.71	85.96	85.56	85.15	83.80
168	92.03	87.98	84.88	84.34	84.21	81.91
CNN	94.90	81.64	80.56	81.78	80.16	80.16

Table 7. Classification accuracy using Krawtchouk moments and YaleB database in noise-free and speckle noisy environment

Krawtchouk moments +CNN	Free Noise	Speckle noise				
		1 %	*2 %*	*3 %*	*4 %*	*5 %*
Order	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*
10	28.74	28.60	28.07	29.28	27.93	29.28
20	66.93	68.28	66.35	65.72	65.58	67.47
40	89.47	89.74	89.20	88.93	88.66	86.90
60	90.41	68.63	90.01	86.90	88.79	89.47
80	90.55	91.09	91.63	89.47	86.77	89.60
100	90.95	91.22	90.01	88.52	90.28	89.33
120	91.22	91.63	90.41	81.51	90.82	90.28
140	90.95	91.09	89.87	90.41	87.31	90.55
160	91.76	90.68	90.41	89.74	90.55	90.41
168	92.03	90.55	89.87	90.41	88.79	90.14
CNN	94.90	93.65	87.58	90.41	82.32	91.22

In the last experiment, we compared our KMCNN model with other models based on CNN combined with 2D orthogonal moments like Tchebichef moments [60], Hahn moments [61], Racah moments [62] and Krawtchouk moments combined with pre-trained VGG16 model [59] using the same noisy conditions presented in the previous experiment. The accuracy results of the noisy images from YaleB, ORL and LFW databases for the KMCNN and prementioned models are respectively shown in Figures 10 to 15, a descriptive legend is given in Figure 16.

Table 8. Classification accuracy using Krawtchouk moments and ORL database in noise-free and salt and pepper noisy environment

Krawtchouk moments + CNN	Free Noise	Salt and Pepper noise				
		1 %	*2 %*	*3 %*	*4 %*	*5 %*
Order	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*
10	28.05	30.49	23.17	28.05	30.49	29.27
20	78.05	79.27	78.05	79.27	74.39	75.61
30	96.34	93.9	96.34	96.34	95.12	92.68
40	97.56	98.78	97.56	96.34	96.34	95.12
50	95.12	90.24	93.9	96.34	91.46	85.37
60	95.12	93.9	91.46	96.34	91.46	89.02
70	95.12	92.68	93.9	90.24	86.59	85.37
CNN	94.30	91.86	90.24	91.05	91.86	90.24

Table 9. Classification accuracy using Krawtchouk moments and ORL database in noise-free and speckle noisy environment

Krawtchouk moments + CNN	Free Noise	Speckle noise				
		1 %	*2 %*	*3 %*	*4 %*	*5 %*
Order	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*
10	28.05	30.49	28.05	25.61	35.37	28.05
20	78.05	78.05	78.05	79.27	79.27	75.61
30	96.34	96.34	96.34	95.12	91.46	96.34
40	97.56	97.56	97.56	96.34	96.34	95.12
50	95.12	93.9	95.12	92.68	93.9	95.12
60	95.12	95.12	95.12	95.12	95.12	93.9
70	95.12	95.12	95.12	95.12	91.46	92.68
CNN	94.30	92.68	92.68	91.86	91.86	91.05

Table 10. Classification accuracy using Krawtchouk moments and LFW database in noise-free and salt and pepper noisy environment

Krawtchouk moments + CNN	Free Noise	Salt and Pepper noise				
		1 %	*2 %*	*3 %*	*4 %*	*5 %*
Order	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*
10	56.16	55.82	54.79	54.45	54.45	54.11
15	68.84	70.55	70.55	68.49	69.18	71.92
20	74.32	73.97	71.58	74.66	71.92	72.26
25	72.26	72.95	71.58	70.21	68.84	73.29
30	70.55	69.18	70.89	68.15	69.52	69.86
40	64.04	70.55	67.12	68.15	67.81	68.49
60	59.93	60.27	58.22	56.51	54.79	58.56
80	56.16	55.48	57.88	53.42	51.71	52.74
100	54.11	52.4	52.4	53.08	53.77	55.48
150	50.34	50.0	47.6	48.97	47.95	46.92
200	54.79	49.32	46.23	46.92	45.21	44.86
250	52.4	47.26	44.52	45.21	43.84	44.86
CNN	77.80	63.84	66.59	67.73	67.96	65.44

Table 11. Classification accuracy using Krawtchouk moments and LFW database in noise-free and speckle noisy environment

Krawtchouk moments + CNN	Free Noise	Speckle noise				
		1 %	*2 %*	*3 %*	*4 %*	*5 %*
Order	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*	*Accuracy*
10	56.16	55.82	57.19	57.53	54.11	54.79
15	68.84	69.18	68.84	70.55	70.21	69.18
20	74.32	72.26	73.29	71.58	71.23	71.92
25	72.26	73.29	71.23	72.6	71.23	71.58
30	70.55	69.18	70.21	70.21	70.21	70.89
40	64.04	65.07	68.15	66.44	66.1	69.18
60	59.93	57.19	60.27	58.9	57.19	56.16
80	56.16	56.85	54.79	56.51	55.82	55.14
100	54.11	55.48	54.45	54.45	53.42	54.11
150	50.34	50.34	52.74	51.37	53.42	51.03
200	54.79	53.77	50.0	49.32	46.92	47.6
250	52.4	54.45	51.37	46.23	46.23	48.29
CNN	77.80	72.99	72.86	71.85	71.21	71.39

Examining the given results in the aforementioned figures, the proposed KMCNN achieved the greatest recognition performance for the four classifiers on the three datasets. In fact, the depicted graphs all demonstrate the same general trend, where the recognition rate values increase by increasing the order of the noisy image moments up to an optimal order, then start to decrease. The obtained results indicate that the KMCNN offers a better strategy to handle noise compared to the combination of CNN with other 2D discrete orthogonal moments. Perhaps this is due to our suggested KMCNN is able to accurately reflect global features by employing discrete orthogonal polynomials with a near-zero redundancy measure in a feature set, as well as their robustness against the effects of noise.

Comparing the results with an architecture that use Krawtchouk moments with VGG16 [59] as pre-trained convolutional neural networks, the KMCNN gives interesting accuracies. This is probably due to the flexibly of the proposed CNN to take different dimension as input layer, however using pre-trained CNNs like VGG16 requires a fixed input shape which lead to the necessity of resizing the image moment and transform it to RGB format. As a result, the capacity of our architecture to represent appropriate features for face recognition was proved. Finally, based on the results depicted in Figures 10 to 15, the proposed KMCNN has reached very satisfactory recognition accuracies, even in a noisy environment, also, it might have a great utility in real-world applications against this type of noise.

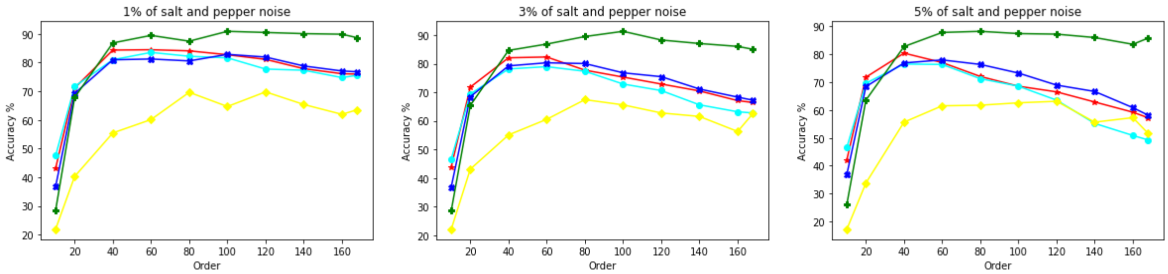


Figure 10. Classification accuracy for different orders using 2D discrete orthogonal moments moments+CNN and Krawtchouk+VGG16 in noisy conditions with salt and pepper and YaleB database

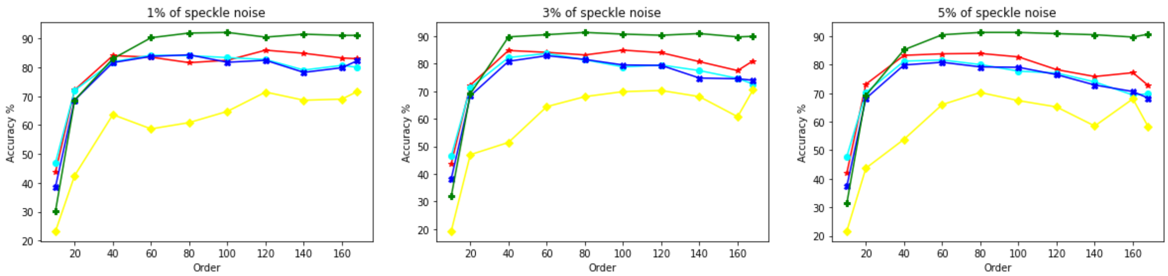


Figure 11. Classification accuracy for different orders using 2D discrete orthogonal moments+CNN and Krawtchouk moments+VGG16 in noisy conditions with speckle and YaleB database

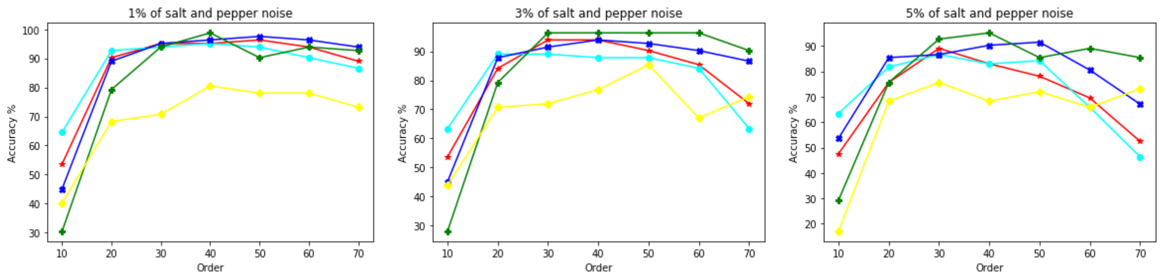


Figure 12. Classification accuracy for different orders using 2D discrete orthogonal moments+CNN and Krawtchouk moments+VGG16 in noisy conditions with salt and pepper and ORL database

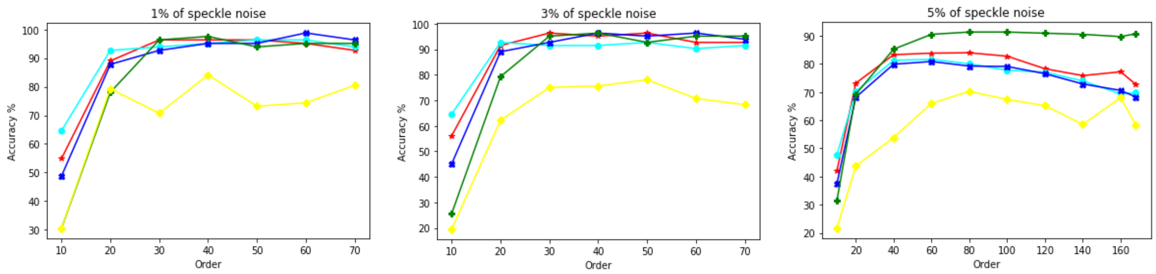


Figure 13. Classification accuracy for different orders using 2D discrete orthogonal moments+CNN and Krawtchouk moments+VGG16 in noisy conditions with speckle and ORL database

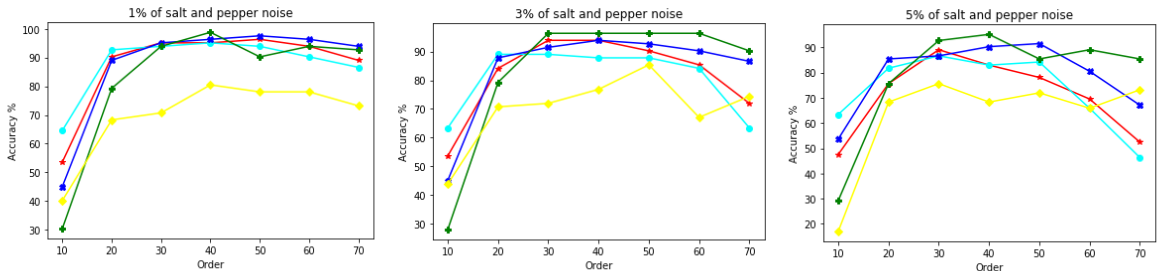


Figure 14. Classification accuracy for different using 2D discrete orthogonal moments +CNN and Krawtchouk moments+VGG16 in noisy conditions with salt and pepper and LFW database

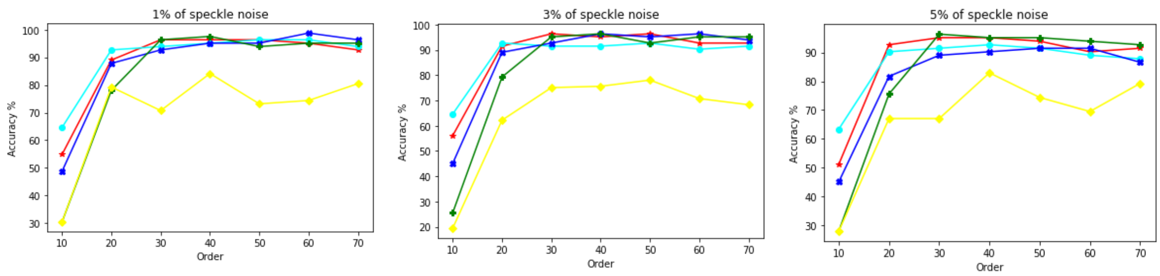


Figure 15. Classification accuracy for different orders using 2D discrete orthogonal moments+CNN and Krawtchouk moments+VGG16 in noisy conditions with speckle and LFW database

Figure 16. A clear legend for Figures 10 to15 presented above

5. CONCLUSION

In this paper, we have suggested a novel face recognition approach that can tolerate deformations produced by two forms of noise: salt and pepper and speckle. The suggested model is founded on the combination of features extracted by the calculation of Krawtchouk moments and convolutional neural networks. Applying Krawtchouk moments on images produced various feature vectors that were then fed into CNN's input layer. The proposed model performed well on small-sized face images (70×80) from the ORL database, large-sized face images (168×192) from the YaleB database, and images (240×240) from the LFW database. The experimental results demonstrated that the suggested model enhanced the accuracy of face recognition with noisy images and surpassed CNN alone and when we combined it with 2D discrete moments like Tchebichef Hahn and Racah significantly. For future works, we plan to further examinate the robustness of the proposed model using different types of noise. We also plan to extend our model to improve the accuracy of 3D noisy face images.

ACKNOWLEDGEMENTS

The authors thankfully acknowledge the Laboratory of Intelligent Systems and Applications (LSIA) for his support to achieve this work.

REFERENCES

[1] A. K. Jain, A. A. Ross, and K. Nandakumar, "Face recognition," in *Introduction to Biometrics*, Boston, MA: Springer US, 2011, pp. 97–139.

[2] A. Ono, "Face recognition with Zernike moments," *Systems and Computers in Japan*, vol. 34, no. 10, pp. 26–35, Sep. 2003, doi: 10.1002/scj.10414.

[3] Y. V. Lata *et al.*, "Facial recognition using eigenfaces by PCA," *International Journal of Recent Trends in Engineering*, vol. 1, no. 1, pp. 587–590, 2009.

[4] P. B. Balla and K. T. Jadhao, "IoT based facial recognition security system," in *2018 International Conference on Smart City and Emerging Technology (ICSCET)*, Jan. 2018, pp. 1–4, doi: 10.1109/ICSCET.2018.8537344.

[5] L. M. Mayron, "Biometric authentication on mobile devices," *IEEE Security & Privacy*, vol. 13, no. 3, pp. 70–73, May 2015, doi: 10.1109/MSP.2015.67.

[6] M. Owayjan, A. Dergham, G. Haber, N. Fakih, A. Hamoush, and E. Abdo, "Face recognition security system," in *New Trends in Networking, Computing, E-learning, Systems Sciences, and Engineering*, 2015, pp. 343–348.

[7] D.-L. Wu, W. W. Y. Ng, P. P. K. Chan, H.-L. Ding, B.-Z. Jing, and D. S. Yeung, "Access control by RFID and face recognition based on neural network," in *2010 International Conference on Machine Learning and Cybernetics*, Jul. 2010, pp. 675–680, doi: 10.1109/ICMLC.2010.5580558.

[8] J. R. Barr, K. W. Bowyer, P. J. Flynn, and S. Biswas, "Face recognition from video: a review," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 26, no. 05, Aug. 2012, doi: 10.1142/S0218001412660024.

[9] M. C. de Pinho, N. M. Ribeiro, and F. R. Gouveia, "Automatic detection of human faces in uncontrolled environments: Identification of direction and movement," in *6th Iberian Conference on Information Systems and Technologies (CISTI 2011)*, 2011, pp. 1–7.

[10] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," in *Workshop on faces in'Real-Life'Images: detection, alignment, and recognition*, 2008, pp. 1–11.

[11] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, Jan. 1991, doi: 10.1162/jocn.1991.3.1.71.

[12] H. J. Seo and P. Milanfar, "Face verification using the LARK representation," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 4, pp. 1275–1286, Dec. 2011, doi: 10.1109/TIFS.2011.2159205.

[13] K. Simonyan, O. Parkhi, A. Vedaldi, and A. Zisserman, "Fisher vector faces in the wild," in *Proceedings of the British Machine Vision Conference 2013*, 2013, pp. 8.1--8.11, doi: 10.5244/C.27.8.

[14] B. Li and K.-K. Ma, "Fisherface vs. Eigenface in the dual-tree complex wavelet domain," in *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Sep. 2009, pp. 30–33, doi: 10.1109/IIH-MSP.2009.322.

[15] M. Annalakshmi, S. M. M. Roomi, and A. S. Naveedh, "A hybrid technique for gender classification with SLBP and HOG features," *Cluster Computing*, vol. 22, no. S1, pp. 11–20, Jan. 2019, doi: 10.1007/s10586-017-1585-x.

[16] Z.-H. Huang, W.-J. Li, J. Shang, J. Wang, and T. Zhang, "Non-uniform patch based face recognition via 2D-DWT," *Image and Vision Computing*, vol. 37, pp. 12–19, May 2015, doi: 10.1016/j.imavis.2014.12.005.

[17] Z. Sufyanu, F. S. Mohamad, A. A. Yusuf, and M. B. Mamat, "Enhanced face recognition using discrete cosine transform," *Engineering Letters*, vol. 24, no. 1, pp. 52–61, 2016.

[18] J. H. Shah, M. Sharif, M. Raza, and A. Azeem, "A survey: Linear and nonlinear PCA based face recognition techniques," *International Arab Journal of Information Technology*, vol. 10, no. 6, pp. 536–545, 2013.

[19] S. R. Arashloo and J. Kittler, "Class-specific kernel fusion of multiple descriptors for face verification using multiscale binarised statistical image features," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2100–2109, Dec. 2014, doi: 10.1109/TIFS.2014.2359587.

[20] A. Vinay, V. S. Shekhar, K. N. B. Murthy, and S. Natarajan, "Performance study of LDA and KFA for Gabor based face recognition system," *Procedia Computer Science*, vol. 57, pp. 960–969, 2015, doi: 10.1016/j.procs.2015.07.493.

[21] S. Lawrence, C. L. Giles, Ah Chung Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," *IEEE Transactions on Neural Networks*, vol. 8, no. 1, pp. 98–113, Jan. 1997, doi: 10.1109/72.554195.

[22] P. Khoi, L. Huu, and V. Hoai, "Face retrieval based on local binary pattern and its variants: A comprehensive study," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 6, pp. 249–258, 2016, doi: 10.14569/IJACSA.2016.070632.

[23] M. Karaaba, O. Surinta, L. Schomaker, and M. A. Wiering, "Robust face recognition by computing distances from multiple histograms of oriented gradients," in *2015 IEEE Symposium Series on Computational Intelligence*, Dec. 2015, pp. 203–209, doi: 10.1109/SSCI.2015.39.

[24] C. S. Weaver and J. W. Goodman, "A technique for optically convolving two functions," *Applied Optics*, vol. 5, no. 7, pp. 1248–1249, Jul. 1966, doi: 10.1364/AO.5.001248.

[25] A. V Lugt, "Signal detection by complex spatial filtering," *IEEE Transactions on Information Theory*, vol. 10, no. 2, pp. 139–145, Apr. 1964, doi: 10.1109/TIT.1964.1053650.

[26] J. S. Rani, D. Devaraj, and R. Sukanesh, "A novel feature extraction technique for face recognition," in *International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007)*, 2007, pp. 428–435, doi: 10.1109/ICCIMA.2007.141.

[27] L. Lenc and P. Král, "Automatic face recognition system based on the SIFT features," *Computers & Electrical Engineering*, vol. 46, pp. 256–272, Aug. 2015, doi: 10.1016/j.compeleceng.2015.01.014.

[28] G. Du, F. Su, and A. Cai, "Face recognition using SURF features," in *MIPPR 2009: Pattern Recognition and Computer Vision*, Oct. 2009, pp. 593–599, doi: 10.1117/12.832636.

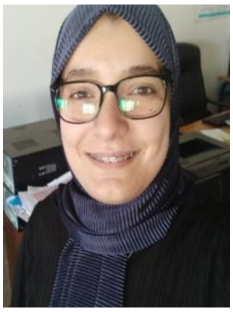
[29] M. Calonder, V. Lepetit, M. Ozuysal, T. Trzcinski, C. Strecha, and P. Fua, "BRIEF: Computing a local binary descriptor very fast," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, pp. 1281–1298, Jul. 2012, doi: 10.1109/TPAMI.2011.222.

[30] A. Alahi, R. Ortiz, and P. Vandergheynst, "FREAK: Fast retina keypoint," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2012, pp. 510–517, doi: 10.1109/CVPR.2012.6247715.
[31] A. A. Fathima, S. Ajitha, V. Vaidehi, M. Hemalatha, R. Karthigaiveni, and R. Kumar, "Hybrid approach for face recognition combining Gabor Wavelet and Linear Discriminant Analysis," in *2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS)*, Nov. 2015, pp. 220–225, doi: 10.1109/CGVIS.2015.7449925.
[32] Y. Yan, H. Wang, and D. Suter, "Multi-subregion based correlation filter bank for robust face recognition," *Pattern Recognition*, vol. 47, no. 11, pp. 3487–3501, Nov. 2014, doi: 10.1016/j.patcog.2014.05.004.
[33] C. Ding and D. Tao, "Robust face recognition via multimodal deep face representation," *IEEE Transactions on Multimedia*, vol. 17, no. 11, pp. 2049–2058, Nov. 2015, doi: 10.1109/TMM.2015.2477042.
[34] F. Juefei-Xu, K. Luu, and M. Savvides, "Spartans: Single-sample periocular-based alignment-robust recognition technique applied to non-frontal scenarios," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 4780–4795, Dec. 2015, doi: 10.1109/TIP.2015.2468173.
[35] Y. Ding, Y. Cheng, X. Cheng, B. Li, X. You, and X. Yuan, "Noise-resistant network: a deep-learning method for face recognition under noise," *EURASIP Journal on Image and Video Processing*, vol. 2017, no. 1, Dec. 2017, doi: 10.1186/s13640-017-0188-z.
[36] X. Meng, Y. Yan, S. Chen, and H. Wang, "A cascaded noise-robust deep CNN for face recognition," in *2019 IEEE International Conference on Image Processing (ICIP)*, Sep. 2019, pp. 3487–3491, doi: 10.1109/ICIP.2019.8803443.
[37] X. Wu, R. He, Z. Sun, and T. Tan, "A light CNN for deep face representation with noisy labels," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2884–2896, Nov. 2018, doi: 10.1109/TIFS.2018.2833032.
[38] Z. Ma, Y. Ding, B. Li, and X. Yuan, "Deep CNNs with robust LBP guiding pooling for face recognition," *Sensors*, vol. 18, no. 11, Nov. 2018, doi: 10.3390/s18113876.
[39] T. S. Arulananth, B. Manjula, and M. Baskar, "Human position tracking and detection using geometric active contours," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, Jul. 2020, pp. 509–512, doi: 10.1109/ICIRCA48905.2020.9182825.
[40] M. Jahangir Alam, T. Chowdhury, and M. Shahzahan Ali, "A smart login system using face detection and recognition by ORB algorithm," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 20, no. 2, pp. 1078–1087, Nov. 2020, doi: 10.11591/ijeecs.v20.i2.pp1078-1087.
[41] D. Rolnick, A. Veit, S. Belongie, and N. Shavit, "Deep learning is robust to massive label noise," *Prepr. arXiv.1705.10694*, 2017.
[42] F. Wang *et al.*, "The devil of face recognition is in the noise," in *Computer Vision – ECCV 2018*, 2018, pp. 780–795.
[43] Pew-Thian Yap, R. Paramesran, and Seng-Huat Ong, "Image analysis by Krawtchouk moments," *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1367–1377, Nov. 2003, doi: 10.1109/TIP.2003.818019.
[44] K. D. Apostolidis and G. A. Papakostas, "Digital watermarking as an adversarial attack on medical image analysis with deep learning," *Journal of Imaging*, vol. 8, no. 6, May 2022, doi: 10.3390/jimaging8060155.
[45] H. Amakdouf, A. Zouhri, M. EL Mallahi, and H. Qjidaa, "Color image analysis of quaternion discrete radial Krawtchouk moments," *Multimedia Tools and Applications*, vol. 79, no. 35–36, pp. 26571–26586, Sep. 2020, doi: 10.1007/s11042-020-09120-0.
[46] G. Hassan, K. M. Hosny, R. M. Farouk, and A. M. Alzohairy, "New set of invariant quaternion Krawtchouk moments for color image representation and recognition," *International Journal of Image and Graphics*, vol. 22, no. 04, Jul. 2022, doi: 10.1142/S0219467822500371.
[47] S. M. M. Rahman, T. Howlader, and D. Hatzinakos, "On the selection of 2D Krawtchouk moments for face recognition," *Pattern Recognition*, vol. 54, pp. 83–93, Jun. 2016, doi: 10.1016/j.patcog.2016.01.003.
[48] M. Krawtchouk, "On interpolation by means of orthogonal polynomials," *Memoirs Agricultural Inst. Kyiv*, vol. 4, pp. 21–28, 1929.
[49] P. T. Yap, P. Raveendran, and S. H. Ong, "Krawtchouk moments as a new set of discrete orthogonal moments for image reconstruction," in *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No.02CH37290)*, 2002, pp. 908–912, doi: 10.1109/IJCNN.2002.1005595.
[50] R. Mukundan, P. Raveendran, and W. A. Jassim, "New orthogonal polynomials for speech signal and image processing," *IET Signal Processing*, vol. 6, no. 8, pp. 713–723, Oct. 2012, doi: 10.1049/iet-spr.2011.0004.
[51] K.-H. Thung, R. Paramesran, and C.-L. Lim, "Content-based image quality metric using similarity measure of moment vectors," *Pattern Recognition*, vol. 45, no. 6, pp. 2193–2204, Jun. 2012, doi: 10.1016/j.patcog.2011.12.001.
[52] S. H. Abdulhussain, A. R. Ramli, S. A. R. Al-Haddad, B. M. Mahmmod, and W.A. Jassim, "On computational aspects of Tchebichef polynomials for higher polynomial order," *IEEE Access*, vol. 5, pp. 2470–2478, 2017, doi: 10.1109/ACCESS.2017.2669218.
[53] K.-C. Lee, J. Ho, and D. J. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 684–698, May 2005, doi: 10.1109/TPAMI.2005.92.
[54] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *ICML'10: Proceedings of the 27th International Conference on International Conference on Machine Learning*, 2010, pp. 807–814.
[55] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," *Prepr. arXiv.1505.00853*, 2015.
[56] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249–256.
[57] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *32nd International Conference on Machine Learning, ICML 2015*, 2015, vol. 1, pp. 448–456.
[58] "The database of faces," *AT&T Laboratories Cambridge*. 2001, Accessed: Mar. 10, 2022. [Online]. Available: https://cam-orl.co.uk/facedatabase.html.
[59] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, Sep. 2014.
[60] R. Mukundan, S. H. Ong, and P. A. Lee, "Image analysis by tchebichef moments," *IEEE Transactions on Image Processing*, vol. 10, no. 9, pp. 1357–1364, 2001, doi: 10.1109/83.941859.
[61] P.-T. Yap, R. Paramesran, and S.-H. Ong, "Image analysis using Hahn moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 11, pp. 2057–2062, Nov. 2007, doi: 10.1109/TPAMI.2007.70709.
[62] H. Zhu, H. Shu, J. Liang, L. Luo, and J.-L. Coatrieux, "Image analysis by discrete orthogonal Racah moments," *Signal Processing*, vol. 87, no. 4, pp. 687–708, Apr. 2007, doi: 10.1016/j.sigpro.2006.07.007.
[63] Y. Jiang, Y. Wu, W. Li, L. Wang, and Q. Liao, "Log-domain polynomial filters for illumination-robust face recognition," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 504–508, doi: 10.1109/ICASSP.2014.6853647.
[64] Ngoc-Son Vu and A. Caplier, "Enhanced patterns of oriented edge magnitudes for face recognition and image matching," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1352–1365, Mar. 2012, doi: 10.1109/TIP.2011.2166974.

[65] M. M. Abdelwahab, S. A. Aly, and I. Yousry, "Efficient web-based facial recognition system employing 2DHOG," *Prepr. arXiv.1202.2449*, 2012.
[66] Y. Wang, Z. Xu, W. Li, and Q. Liao, "Illumination-robust face recognition with block-based local contrast patterns," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2017, pp. 1418–1422, doi: 10.1109/ICASSP.2017.7952390.
[67] Y. Feng, X. An, and X. Liu, "The application of scale invariant feature transform fused with shape model in the human face recognition," in *2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, Oct. 2016, pp. 1716–1720, doi: 10.1109/IMCEC.2016.7867511.
[68] Y. Gan, T. Yang, and C. He, "A deep graph embedding network model for face recognition," in *2014 12th International Conference on Signal Processing (ICSP)*, Oct. 2014, pp. 1268–1271, doi: 10.1109/ICOSP.2014.7015203.
[69] A. Vinay, D. Hebbar, V. S. Shekhar, K. N. B. Murthy, and S. Natarajan, "Two novel detector-descriptor based approaches for face recognition using SIFT and SURF," *Procedia Computer Science*, vol. 70, pp. 185–197, 2015, doi: 10.1016/j.procs.2015.10.070.
[70] T. M. Abhishree, J. Latha, K. Manikantan, and S. Ramachandran, "Face recognition using Gabor filter based feature extraction with anisotropic diffusion as a pre-processing technique," *Procedia Computer Science*, vol. 45, pp. 312–321, 2015, doi: 10.1016/j.procs.2015.03.149.
[71] K. Chaturvedi and D. K. Vishwakarma, "Face recognition in an unconstrained environment using ConvNet," in *Proceedings of the 2020 2nd International Conference on Big Data Engineering and Technology*, 2020, pp. 67–71, doi: 10.1145/3378904.3378905.

BIOGRAPHIES OF AUTHORS

Yassir El Madmoune He received a M.S. degree in Intelligent Systems and Networks, from the Faculty of Science and Technology, University of Sidi Mohammed Ben Abdellah, Fez, Morocco in 2018. He is currently pursuing his Ph.D. degree in Computer Science at the Faculty of Science and Technology of Fez. His research interests include pattern recognition and computer vision. Email: yassir.elmadmoune@usmba.ac.ma.



Ilham El Ouariachi She received a M.S. degree in Intelligent Systems and Networks, from the Faculty of Science and Technology, University of Sidi Mohammed Ben Abdellah, Fez, Morocco in 2016. He is currently pursuing his Ph.D. degree in Computer Science at the Faculty of Science and Technology of Fez. His research interests include pattern recognition and computer vision. Email: ilham.elouariachi@usmba.ac.ma.

Khalid Zenkouar He received a Ph.D. degree in image analysis from Faculty of Science, University Sidi Mohamed Ben Abdellah, Fez, Morocco in 2006. Now he is a professor of the Department of computer engineering, Faculty of Science and Technology Fez Morocco. He is a member in the LSIA Laboratory (Laboratory of Intelligent Systems and Application). His current research interests include image analysis, machine intelligence and pattern recognition. Email: khalid.zenkouar@usmba.ac.ma.

Azeddine Zahi Received his PhD degree in 1997 in Computer Sciences from Mohammed V University in Rabat. He is a research professor at Sidi Mohamed Ben Abbdellah University of Fez since 1995. His research interests include fuzzy data mining, heuristics and automatic learning. Email: azeddine.zahi@usmba.ac.ma.