

# An Adaptive Face Recognition System Based on a Novel Incremental Kernel Nonparametric Discriminant Analysis

Arbia SOULA<sup>1\*</sup>, Salma BEN SAID<sup>1</sup>, Riadh KSANTINI<sup>2</sup> and Zied LACHIRI<sup>1</sup>

<sup>1</sup>Signal, Images and Information Technologies (LR-SITI-ENIT)  
National School of Engineers of Tunis (ENIT)  
University of Tunis El Manar,  
Tunis, TUNISIE

[e-mail: arbiasoula1@gmail.com]

<sup>2</sup>University of Windsor, Windsor, ON, Canada  
SUPCOM: Sécurité Numérique

[e-mail: ksontiniriadh@yahoo.fr]

\*Corresponding author: Arbia Soula

*Received September 15, 2018; revised October 6, 2018; accepted October 18, 2018;  
published April 30, 2019*

---

## Abstract

This paper introduces an adaptive face recognition method based on a Novel Incremental Kernel Nonparametric Discriminant Analysis (IKNDA) that is able to learn through time. More precisely, the IKNDA has the advantage of incrementally reducing data dimension, in a discriminative manner, as new samples are added asynchronously. Thus, it handles dynamic and large data in a better way. In order to perform face recognition effectively, we combine the Gabor features and the ordinal measures to extract the facial features that are coded across local parts, as visual primitives. The variegated ordinal measures are extracted from Gabor filtering responses. Then, the histogram of these primitives, across a variety of facial zones, is intermingled to procure a feature vector. This latter's dimension is slimmed down using PCA. Finally, the latter is treated as a facial vector input for the advanced IKNDA. A comparative evaluation of the IKNDA is performed for face recognition, besides, for other classification endeavors, in a decontextualized evaluation schemes. In such a scheme, we compare the IKNDA model to some relevant state-of-the-art incremental and batch discriminant models. Experimental results show that the IKNDA outperforms these discriminant models and is better tool to improve face recognition performance.

---

**Keywords:** Discriminant models, incremental learning, nonparametric discriminant analysis, Gabor filter, face recognition.

## 1. Introduction

Over the past decades, fresh research pathways, apropos of computer vision, abounded thanks to the face recognition paradigm. In actual fact, it has a plethora of applications ranging from human-computer smart interaction, surveillance, telecommunication, access control [1] to internet of things (IoT) [2]. Broadly speaking, face recognition is most esteemed for its facility to apperceive an individual's identity gleaned from his facial features. However, noticeable are the breakthroughs and improvements in the face recognition paradigm, a handful of loopholes do stand out [3] that are attributable to the wide intra-class facial change, such as, luminosity, pose, aging, expression, and the miniscule interclass dissimilarities, which are genuine factors that could affect the face recognition performance.

Into two divisible categories are face recognition methods classified [3]: Image-based methods use a vector that represents the entire face rather than the most significant facial features. Feature based methods try to extract features of the image and match them against the knowledge of the facial features.

Initially, Feature based methods that are developed for face recognition used mainly individual features of the faces, such as, mouth, eyes or nose to perform identification [4]. However, such methods did not lead to good results given the variability of poses, illumination, facial expression and the low amount of information used. Recently, however, a range of schemes are brought forward to overcome the difficulties of face recognition. In this respect, among the newfangled research works on face recognition based on filtering, dimension reduction and classification, we can mention:

Zhenhua Chai et al [5] proposed a new-fashioned local feature analysis formula, namely, Gabor Ordinal Measures (GOM), that combines Gabor features with ordinal measures, as a forceful formula to deal with intra-personal variations in face images and the inter-person similarity. Notwithstanding, this method is grounded on non-flexible Linear Discriminant Analysis (LDA) classifier. In effect, it deploys linear boundary to discriminate between facial classes. Their efficacy is lacking when employed on issues that necessitate nonlinear decision boundaries. Furthermore, the efficiency of LDA dwindles as the underlying class distribution is not normal.

Zhifeng Li et al [6] advanced a method for face recognition formulated on nonparametric discriminant analysis (NDA) and multiclassifier integration. A fresh formulation of scatter matrices, they have introduced, to extend the two-class NDA to multiclass NDA-based formula NSA (Nonparametric Subspace Analysis) and NFA (Nonparametric Feature Analysis). Indeed, they have formulated a dual NFA-based multiclassifier fusion framework by employing the overcomplete Gabor representation for face images to boost the recognition efficiency.

Recently, Arbia Soula et al presented in [7] and [8] a novel face recognition systems formulated on Gabor and ordinal filters for feature extrication, and on the Kernel Fisher Discriminant Analysis (KFD) and the Kernel Nonparametric Discriminant Analysis (KNDA), respectively, for dimension reduction and classification. In fact, both dimension reduction and classification methods are grounded on flexible non-linear facial classes' disjunction. More particularly, KNDA integrates the paramount nearby global variations data, to handle heteroscedastic face classes.

More recently, Convolutional Neural Network (CNN), with Convolution phase and fully connected layers, has been widely used for face features extraction and classification [9]. However, this model is not based on convex nor concave optimization problem. Thus,

finding the global optimal solution is not an easy task. For this reason, several research works have used Deep features to represent face images and then deployed Support Vector Machine (SVM) classifier to effectuate face recognition. Among them, we can mention:

Dat Tien Nguyen et al [10] presented an approach to optimize the level of security for face recognition systems based on a merging of deep learning and handcrafted features extricated from face images. In fact, this system is based on hybrid features descriptor, and uses CNN medium to extricate deep image attributes, and the multi-level local binary pattern (MLBP) formula to elicit the detail features of the skin from face images. Finally, SVM is made use of to put the image features into classes.

Typical employments for the existing face recognition techniques contend that all the data is given beforehand while learning is done altogether once. In this respect, such techniques are considered as batch learning methods. However, these systems still show some loopholes in several real-life functions where data are sequentially obtained, such as, video-based face recognition application. In fact, in a shifting scene, poses, luminescence and other image capture status may change quickly leading to accuracy degradation. In such contexts, it is unwise to re-train the classifiers on a continuous basis [11].

Because of this shortcoming, batch techniques necessitate a huge memory coupled with expensive time training for large datasets. Moreover, batch learning displays a decline in performance when data are not present up from the start. That's why, a novel learning formula is required.

In effect, incremental learning has been deemed to be more adequate than batch learning as new samples are added asynchronously, at different timestamps, or when dealing with a great deal of data. So far, many face recognition methods formulated on incremental learning and dimension reduction have been proposed in literature ([12] and [13]).

For instance, B. Raducanu et al. [12] introduced a novel formula to delineate subspace learning grounded on an incremental nonparametric discriminant analysis (INDA), where new examples can be added asynchronously, at different intervals. Then as new individuals are being added, we apply the incremental updates without the computation of the full scale within-class scatter matrix and between-class scatter matrix. The INDA deals with general data distributions in a proper manner, and it captures correctly the information between class boundaries. Thus, recently, we have used the INDA to build an incremental face recognition system in [13].

However, all the above mentioned incremental face recognition methods are based on non-flexible linear classifiers. In fact, they utilize linear boundaries to distinguish between face classes. Not up to standard is their efficacy when employed to difficulties that call for nonlinear decision boundaries or to problems as face datasets are nonlinearly separable. Furthermore, the efficiency of these methods dwindles as the underlying class distribution is not normal and face classes are heteroscedastic. Thus, these approaches are ineffecient for a variety of face recognition problems. Indeed, pose, luminescence and high variability in facial features induce high non-linearity and heteroscedasticity in the representation space.

As an alternative, Hsien-Ting Cheng et al. [14] presented an approach to multimodal person identity verification based on an incremental kernel Fisher's discriminant for training image dimension reduction, and relevant feature extraction. Support Vector Machine (SVM) is used for data fusion and classification. As IKFD is based on Kernels, it separates between face classes using flexible decision boundary. However, IKFD does not handle heteroscedastic data and assumes that face classes are normally distributed, which is not common in real world applications.

NaimulMefraz Khan et al. [11] proposed a video-based face recognition method formulated on adaptive sparse dictionary, to overcome the shifts in illumination, pose, occlusion and alignment in face images. In fact, it consists of dynamical updates with present probe image into the training matrix, using a novice confidence criterion and a Bayesian inference scheme to recognize faces from unconstrained videos.

Recently, Pawel Karczmarek et al. [15] presented a face recognition system based on Chain Code-Based Local Descriptor (CCBLD) for face features extraction, it consists to achieve Bag-of-Visual-Words method through the dictionary of chain-codes.

More recently, Lufian et al. [16] advanced an incremental face recognition method formulated on Deep learning, capable to update the classification model in which new samples can be added during use. In fact, this system is based on intelligent principle training method, namely, the S-DDL (self detection, decision and learning) employing incremental version of the Support Vectors Machine (SVM) algorithm in order to realize self-learning and enhance classification accuracy, while maintaining low computational time. The module of features extraction is based on CNN, and then incremental SVM is trained on the extracted features, so as to execute recognition. Although, incremental SVM makes no distribution assumption on the data, it can be misled by data spread, since its solution is being formulated on a finite number of support vectors.

As a matter of fact, in the present paper we are working out this problem by proposing an adaptive face recognition system based on a novel Incremental Kernel Nonparametric Discriminant Analysis (IKNDA). In fact, IKNDA is advantageous since it incrementally reduces data dimension and performs classification, while dealing with general data distributions. Moreover, unlike incremental SVM, it aptly distinguishes the structural information between class boundaries and is based on kernelization to perform flexible non-linear separation between face classes, thereby improving classification performance compared to face recognition systems based on classical incremental parametric models.

Also, our proposition integrates advantages of combining distinction of Gabor response with flexibility of ordinal filters [5]. In fact, it consists of several steps: First, multichannel Gabor filters are applied on the input image. Second, several ordinal measures are applied on obtained Gabor images and encoded to generate visual primitives in facial zones. Third, the spatial histograms of these primitives are concatenated into a feature vector whose size is brought to minimum using PCA. Finally, the novel IKNDA is further used to reduce dimension and classify feature vectors. The proposed adaptive face recognition system investigates the effectiveness of the IKNDA method and we showed that it is appropriate tool to deal with non-stationary learning environments.

The present is grouped as follows: In the next section we describe in details the main phases of our adaptive face recognition system, including the feature extraction phase and the recognition phase using the novel IKNDA. Also, the algorithm and comprehensive block diagram of the face recognition method are provided. In section III, a decontextualized evaluation is set into place to compare the IKNDA technique to the classical batch version of KNDA, as well as, to other relevant state-of-the-art incremental discriminant algorithms, on real datasets. Also, a contextualized comparative evaluation of the adaptive face recognition method based on IKNDA, is performed on several face datasets. Finally, in the last section, we provide concluding remarks and perspectives.

## 2. Face Recognition Method Phases

In this section, we give a full account of our face recognition method. This latter is made up of two main phases: Feature extraction and recognition using a novel Incremental Kernel Nonparametric Discriminant Analysis (IKNDA). Then, we provide the face recognition method algorithm and comprehensive block diagram.

### 2.1 Facial Feature Extraction

Feature extraction is a pivotal measure in face recognition operation. It consists in finding a specific representation of the data that can highlight relevant information, which would help in overcoming the human facial complications, such as, the light directions of imaging, differences of facial expression, variation of pose and aging.

Indeed, in our method we utilize a local feature analysis technique, namely, Gabor Ordinal Measures (GOM) for the representation of face features [5], which inherits the advantages of combining differences of Gabor features with forcefulness of some types of ordinal wavelets, as an auspicious answer to reduce intra-person similarities as well as maximize dissimilarity between persons. Thus, the 2D Gabor filters help to produce prominent local discriminating features that are appropriate for face recognition. The Gabor filters are expressed as follows [17]:

$$\psi_{\mu,v}(z) = \frac{\|k_{\mu,v}\|^2}{\sigma^2} e^{\left(\frac{\|k_{\mu,v}\|^2 \|z\|^2}{2\sigma^2}\right)} \left[ e^{ik_{\mu,v}z} - e^{-\frac{\sigma^2}{2}} \right] \quad (1)$$

Where  $\mathcal{V} \in \{0, \dots, 4\}$  and  $\mu \in \{0, \dots, 7\}$  are the scale and orientation of the Gabor wavelets, respectively, and  $z = (x, y)$  denotes the spatial position. The wave vector  $k_{\mu,v} = k_v e^{i\theta_\mu}$  is of a magnitude  $k_v = \frac{k_{max}}{\lambda^v}$ , where  $\lambda$  is the frequency ratio between filters and  $\theta_\mu = \frac{\pi\mu}{8}$ ,  $\theta_\mu \in [0, \pi]$ .

In practice, the Gabor filters decomposition and illustration of the facial image is the convolution of the image  $I$  with a set of Gabor kernels  $\psi_{\mu,v}(z)$ , defined as:

$$G_{\mu,v}(z) = I(z) \times \psi_{\mu,v}(z) \quad (2)$$

So, the Gabor filters response produced for a definite frequency and orientation in Equation (2), is a complex number, given by the following equation [18]:

$$G_{\mu,v}(z) = A_{\mu,v}(z) \cdot e^{i\theta(z)} \quad (3)$$

Where,  $A$  and  $\theta$  define the magnitude response and the phase of Gabor kernel at each image position  $z$ , respectively.

The complex Gabor filter is a powerful descriptor. In fact, it is a strong tool to characterize the image texture. Therefore, it can obtain the local region matching specific frequency, spatial locality and a definite orientation that are demonstrably discriminative and robust to expression changes and illumination.

The complex Gabor response can also be described with real part and imaginary part. Thus, we can obtain four features for each face image, which are phase, magnitude, real and imaginary Gabor feature images.

As for ordinal or multi-lobe differential filtering for ordinal feature extraction, they provide a richer representation of facial features and are well-conditioned to uniform noise.

Besides image space, ordinal features have the advantage of describing the neighboring relationship in various orientations and scales of Gabor images.

From a mathematical perspective, multi-lobe differential filter (MLDF) is formed by many positive and negative lobes, which allow the arrangement of dissociated image regions in intensity and feature aspects. In respect to intensity, it helps to determine relational nature of the average intensity values of two image zones. Whereas, on the feature account, qualitative details on the image features are calculated. Thus, the MLDF has the benefit of inertness to monotone illumination variation and forcefulness to noise. The ordinal values are distinctively “0” or “1” as the filtering results are, respectively, negative or positive. The MLDF can be presented with Gaussian Kernels as follows:

$$MLDF = c_p \sum_{i=1}^{N_p} \frac{1}{\sqrt{2\pi\delta_{pi}}} e^{\left[ \frac{-(x-\omega_{pi})^2}{2\delta_{pi}^2} \right]} c_n \sum_{j=1}^{N_n} \frac{1}{\sqrt{2\pi\delta_{nj}}} e^{\left[ \frac{-(x-\omega_{nj})^2}{2\delta_{nj}^2} \right]} \quad (4)$$

Where,  $\omega$  is the central position,  $\delta$  is the frequencies of 2D Gaussian filter,  $N_n$  is the sum of negative lobes and  $N_p$  is the sum of positive lobes and  $c_n$  and  $c_p$  are two constants.

Distinct ordinal feature representation techniques are administered to Gabor features, such as, Gabor magnitude, phase, real and imaginary, so as to record the forceful ordinal features in diverse directions.

Face image can be analyzed on two levels: local intensity level and local feature level. As for local intensity variation, it is insubstantial as facial skin would have the same intensity of reflection ratio. In response to such a limited role, ordinal measures derived from features level become more powerful as they have a manifest discriminatory power in face recognition. Consequently, the use of Gabor filter aims at getting a more discriminative feature as well as ameliorating the local details of face texture. As a pleasing result, the integration of feature Gabor images with ordinal filters leads to a better recognition rate.

As a result, the ordinal measures acquired from different components of Gabor images significantly expand the feature vector area of a face image. Thus, it is essential to integrate many binary codes in GOM facial features to find a characteristic texture parameter and minimize the length of GOM feature.

The Gabor Ordinal Measures (GOM) for facial feature description could be represented in the following algorithm:

---

Algorithm1. Face features extraction based on Gabor and Ordinal filters

---

1) We utilize a series of 2D Gabor filters constituted by five scales and eight orientations on each face pixel, so every single response of Gabor feature incorporate 40 items with a like dimension as the original face image, using the equations (1), (2) and (3) .

2) We employ four tri-lobe and four di-lobe ordinal wavelet using equation (4), with orientation values equal to  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  and  $135^\circ$  on Gabor response images in order to obtain ordinal measures.

3) We combine every eight Gabor Ordinal Measures for every pixel, given MLDF and scale into a visual code as the following model:

$$GOM - MAP - \sigma_v^i(x, y) = [GOM - \sigma_{0,v}^i(x, y), GOM - \sigma_{1,v}^i(x, y), \dots, GOM - \sigma_{7,v}^i(x, y)]$$

Where  $\sigma\{p: \text{phase}, m: \text{magnitude}, i: \text{imaginary}, r: \text{real}\}$  and  $GOM - \sigma_v^i(x, y)$  is the texture details derived at position  $(x, y)$  for the  $i - th$  ordinal measure at scale  $v$ .

4) Histograms representations are derived from every visual primitive. Then, they are joined to make the statistical apportionment of these parameters in a face image.

5) We utilize the Principal Component Analysis (PCA) to minimize feature vectors dimensions and maintain the appropriate information.

## 2.2 The Novel Incremental Kernel Nonparametric Discriminant Analysis (IKNDA)

In this part, we present the classical batch Kernel Nonparametric Discriminant Analysis (BKNDAs) as it is the basis of our incremental model. In fact, the incrementality consists of sequential updates of the KNDA-eigenspace representation. Then, we will describe the novel Incremental KNDA (IKNDA) in details.

### 2.2.1 Batch Kernel Nonparametric Discriminant Analysis (BKNDAs)

In this part, we depict the formulation of the batch Kernel Nonparametric Discriminant Analysis (BKNDAs) that is employed to reduce dimension and classify feature vectors. It introduces a nonparametric form of the within-class scatter matrix and rather than having a Gaussian distribution on the points of like classes, it normalizes the distances between every single point and their nearest neighbours, which in effect shows the advantage of the nearest neighbour rule.

We assume that we have  $C_b, b = 1, 2, \dots, L$  classes making an input space of  $N = \sum_{b=1}^L n_{C_b}$  examples, where each class  $C_b$  is formed by  $n_{C_b}$  samples in  $\mathbb{R}^M$ , namely,  $C_b = \{x_1^b, x_2^b, \dots, x_{n_{C_b}}^b\}$ .

The BKNDAs consists of two stages: it defines the nonlinear mapping and transforms the data samples into a superior dimensional feature space  $\mathcal{F}$ , where linear classification can be attained.

Let the function  $\varphi$  maps the classes  $C_b, b = 1, 2, \dots, L$  to higher-dimensional feature class  $\mathcal{F}_b = \{\varphi(x_i^b)\}_{i=1}^{n_{C_b}}, b = 1, 2, \dots, L$ , respectively.

Notwithstanding, if  $\mathcal{F}$  is of a very high dimension, this will be unfeasible to go straightforwardly into mapping. In this regard, the kernel trick [19] is applied to count the dot products of the higher-dimensional data rather than the examples themselves. In mathematical terms, it can be defined:

$$K(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle, \forall i, j \in \{1, 2, \dots, N\} \quad (5)$$

The decision function is described as follows:

$$y(x_i, \omega) = \sum_{i=1}^N f_i^x \omega_i + \omega_0 \quad (6)$$

Where  $(f_1^x, f_2^x, \dots, f_N^x): \mathcal{X} \rightarrow \mathcal{F}$  depicts a non-linear mapping from the input space to a feature space for the input variable  $x$ ,  $f_i^x = K(x, x_i), \forall i \in \{1, 2, \dots, N\}$  and  $\{\omega_i\}_{i=1}^N$  are the weights to be estimated.

BKNDAs is after an optimal subspace in order to attenuate the separability between classes. The latter is achieved by minimizing the within-class distance of the kernel feature classes all at a time, while also maximizing the between-class scatter matrix distance, through local

eigenvectors [20]. It has as objective the finding of the projection direction  $w$  in order to increase the objective function that follows:

$$J(\omega) = \frac{\omega^T S_{BK} \omega}{\omega^T S_{WK} \omega} \quad (7)$$

Where  $S_{WK}$  is the within-class scatter matrix and presented as follows:

$$S_{WK} = \sum_{b=1}^L K_b (I - 1_{nC_b}) K_b^T \quad (8)$$

where  $K_b, b = 1, 2, \dots, L$  are the kernel matrices for class  $C_b, b = 1, 2, \dots, L$ , respectively,  $I$  is the identity matrix and  $1_{nC_b}, b = 1, 2, \dots, L$ , are the matrices with all entries  $\frac{1}{nC_b}$ , respectively.

$S_{BK}$  is the between-class scatter matrix and defined as:

$$S_{BK} = \frac{1}{N} \sum_{b=1}^L \sum_{c=1, c \neq b}^L \sum_{i=1}^{nC_b} \psi_i(C_b, C_c) L_b(\varphi(x_i^c)) L_b(\varphi(x_i^c))^T \quad (9)$$

Where  $\psi_i$  are the weighting functions to nullify the outcome of items that are far from the boundary [21]. It is defined as follows:

$$\psi_i(C_b, C_c) = \frac{\min \left\{ d(\varphi(x_i^b), \varphi(xNN_{bi}^\kappa))^\gamma, d(\varphi(x_i^b), \varphi(xNN_{ci}^\kappa))^\gamma \right\}}{d(\varphi(x_i), \varphi(xNN_{bi}^\kappa))^\gamma + d(\varphi(x_i), \varphi(xNN_{ci}^\kappa))^\gamma} \quad (10)$$

Where  $\gamma$  is a control parameter which can range from zero to infinity, and  $d(\varphi(x_i^b), \varphi(xNN_{ci}^\kappa))$  is the Euclidean distance from  $x_i^b$  to its  $\kappa - NN$ 's from class  $C_c$  in the kernel space.

$$\left( L_b(\varphi(x_i^c)) \right)_j = K(x_i^b, x_j^b) - \left( M_c^\kappa(\varphi(x_i^c)) \right)_j, \forall i \in \{1, 2, \dots, nC_b\}, \forall j \in \{1, 2, \dots, N\} \quad (11)$$

Where  $M_c^\kappa(\varphi(x_i^c)) = \frac{1}{\kappa} \sum_{h=1}^{\kappa} \varphi(x_i^c)_{NN}(h)$  is the mean of the  $\kappa$  nearest neighbors and  $\varphi(x_i^c)_{NN}(h)$  define the  $h^{th}$  nearest neighbor of items  $x_i^c$  from class  $c$ . More precisely,  $\kappa$  is the free parameter which prescribes the number of neighbors up for consideration. Such a parameter should be upgraded for every single database. Eq. (11) represents the direction of the gradients of the corresponding class density functions in the feature space [22].

Expression (7) can be worked out by getting the leading eigenvectors of  $S_{WK}^{-1} S_{BK}$ . Since, the superior-dimensional feature space  $F$  is of size  $N$ . Moreover, numerical difficulty which given the matrix  $S_{WK}$  not to be positive. Hence, the  $S_{WK}$  must to be formalised before computing the inverse. This is attained by adding a simply multiple  $\beta$  of the identity matrix  $I$  [23]. Thus,  $\omega$  is constituted by the eigenvector matching to the biggest eigenvalue constructions the optimal decision hyper-plane of  $(S_{WK} + \beta I)^{-1} S_{BK}$ .

However, BKNDA shows some difficulties and serious performance degradation in real world application, where data are sequentially acquired or are not present from the outset. Moreover, it requires large storage space and leads to increased training time, chiefly, for large scale datasets. Thus, incremental learning strategy is required.



### 2.2.2 The Novel Incremental version of the BKND

In this subsection, we describe the main principle of our novel Incremental KNDA (IKNDA) method that can process sequentially as new samples are added at any time, without the necessity for recalculating the scatter matrices  $S_{WK}$  and  $S_{BK}$  again. More precisely, IKNDA builds discriminant eigenspace representation without the calculating of the full scale scatter matrices  $S_{WK}$  and  $S_{BK}$ , but rather utilising incremental updates of  $S_{WK}$  and  $S_{BK}$ , as new sample is added.

Besides, the IKNDA realises dimension diminution and classification of data incrementally by drawing upon the nearest neighbour rule to calculate the local means, and it is capable to treat jointly the heteroscedastic and non normal data, while initially employing a restricted sum of items instances for training.

Furthermore, the proposed IKNDA focuses mainly on mapping the data non-linearly into some feature space and then calculating incremental nonparametric discriminant analysis there, thereby implicitly performing a non-linear separation based on local nearest-neighbours in input space. Thus, improving classification accuracy as data classes are nonlinearly separable.

In the implementation of the proposed novel IKNDA method we suppose that we have calculated the  $S_{WK}$  and  $S_{BK}$  scatter matrix from at the minimum two classes.

For now, let's assume that a fresh training pattern  $y$  is arriving to the algorithm, we differentiate between the two cases:

Case1: The fresh training pattern  $y^{C_E}$  belongs to an existing class.

This case is illustrated in **Fig. 1**. Let us assume that the new pattern  $y^{C_E}$  pertains to one of the current classes  $C_E$ , where  $1 < E < L$ .

The liaisons between the different classes in **Fig. 1** represent the old and novice state of the classes, earlier and after, append the newest pattern. So, posterior to introducing of the new element, the expression employed to recursively sum up the  $S_{BK}$  matrix, is disposed by:

$$S'_{BK} = S_{BK} - S_{BK}^{in}(C_E) + S_{BK}^{in}(C_{E'}) + S_{BK}^{out}(y^{C_E}) \quad (12)$$

Where  $C_{E'} = C_E \cup \{y^{C_E}\}$ ,  $y^{C_E}$  is the last pattern and  $E$  is the class of  $y^{C_E}$ .  $S_{BK}^{in}(C_E)$  explicate the covariance matrix between classes, the present and the latent that is likely to be modified,  $S_{BK}^{in}(C_{E'})$  represents the covariance matrix between current classes and the refreshed class  $C_{E'}$ , and by  $S_{BK}^{out}(y^{C_E})$  we point out the covariance matrix amidst the vector  $y^{C_E}$  and the current classes. The expression to compute these matrices are given as follows:

$$S_{BK}^{in}(C_E) = \sum_{j=1, j \neq E}^L \sum_{i=1}^{nc_j} \Psi_i(C_j, C_E) L_j(\varphi(x_i^j)) L_j(\varphi(x_i^j))^T \quad (13)$$

$$S_{BK}^{out}(y^{C_E}) = \sum_{j=1, j \neq E}^L L_j(\varphi(y^{C_E})) L_j(\varphi(y^{C_E}))^T \quad (14)$$

Hence, in this situation the within scatter matrix  $S'_{WK}$  is refreshed applying the following formula:

$$S'_{WK} = \sum_{j=1, j \neq E}^L S_{WK}(C_j) + S_{WK}(C_{E'}). \tag{15}$$

Where,

$$S_{WK}(C_{E'}) = S_{WK}(C_E) + \frac{n_{C_E}}{n_{C_E} + 1} (y^{C_E} - \overline{\varphi(x^{C_E})})(y^{C_E} - \overline{\varphi(x^{C_E})})^T \tag{16}$$

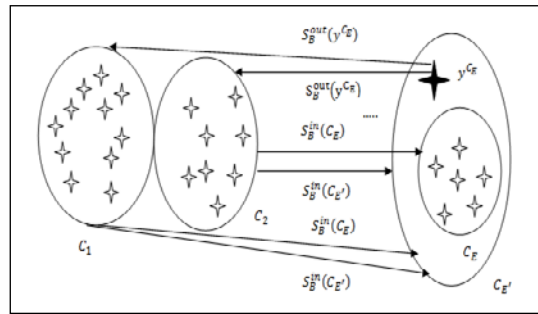


Fig. 1. Case1. The novel appended element pertains to an existing class.

Case2: The novel training element  $y^{C_{L+1}}$  pertains to a novel class.

This situation is described in Fig. 2. A novel element  $y^{C_{L+1}}$  pertains to a novel class  $C_{L+1}$ . The class links present only global  $S_{BK}$  matrix classes that are up to change by the update expression. After the inclusion of the novel element, the equations employed to recursively determinate the  $S_{BK}$  matrix, are expressed by:

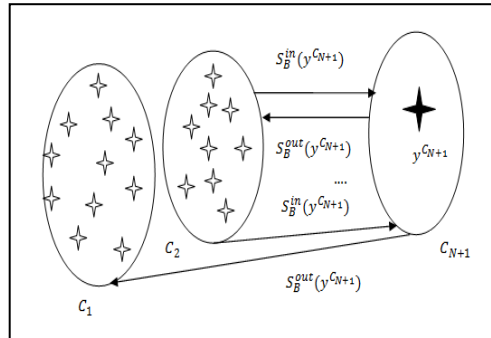
$$S'_{BK} = S_{BK} + S_{BK}^{out}(C_{L+1}) + S_{BK}^{in}(C_{L+1}) \tag{17}$$

Where  $S_{BK}^{out}(C_{L+1})$  and  $S_{BK}^{in}(C_{L+1})$  are defined as follows:

$$S_{BK}^{out}(C_{L+1}) = \sum_{j=1}^L L_j(\varphi(y^{C_{L+1}}))L_j(\varphi(y^{C_{L+1}}))^T \tag{18}$$

$$S_{BK}^{in}(C_{L+1}) = \sum_{j=1}^L \sum_{i=1}^{n_{C_j}} \Psi_i(C_j, C_{L+1})L_j(\varphi(x_i^j))L_j(\varphi(x_i^j))^T \tag{19}$$

Concerning the novel  $S'_{WK}$  matrix, this last remains unchanged:  $S'_{WK} = S_{WK}$ .



**Fig. 2.** Case 2. The new append sample belong to a new class.

The proposed IKNDA could be described by the following algorithm:

---

**Algorithm 2.** Incremental Kernel Nonparametric Discriminant Analysis Algorithm.

---

- (1) Kernelization of feature vectors using RBF Kernel according to Eqs. (5) and (6).
- (2) Case1: If  $\exists y^{C_E}$  and  $1 \leq E \leq L$ . add  $y^{C_E}$  to  $C_E$ .
- (3) Compute  $S_{BK}^{in}(C_E), S_{BK}^{in}(C_{E'})$  and  $S_{BK}^{out}(y^{C_E})$  scatter matrices and use it to update  $S'_{BK}$  using Eqs. (12), (13) and (14).
- (4) Compute  $S_{WK}(C_{E'})$  and use it to update  $S'_{WK}$  according to Eqs. (15) and (16). end Case1
- (5) Case 2: If  $\exists y^{C_E}$  and  $E > L$ . add  $y^{C_E}$  to  $C_E$
- (6) Compute the  $S_{BK}^{out}(C_{L+1})$  and  $S_{BK}^{in}(C_{L+1})$  scatter matrices and update  $S'_{BK}$  using Eqs. (17), (18) and (19). end Case 2
- (7) The solution to the problem (7) consists of the eigenvector associated to the greatest eigenvalue of the matrix  $(S'_{WK} + \beta I)^{-1} S'_{BK}$ .

---

In this part, we will present in details the proposed face recognition system.

### 2.2.3 Face Recognition Method Algorithm

The process of the face recognition system is presented along these lines.

---

**Algorithm 3.** Face Recognition method.

---

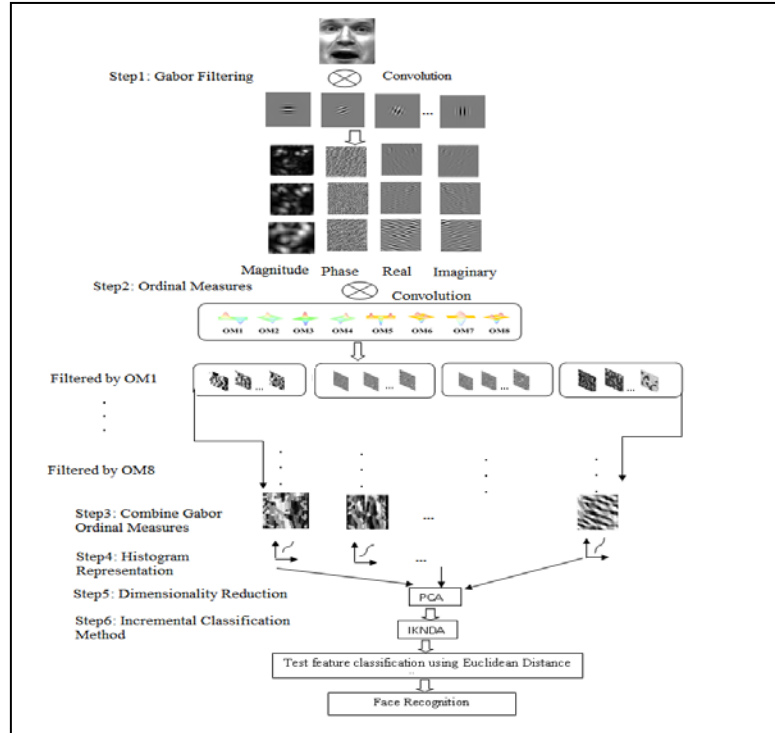
- (1) Features derivation (Gabor Ordinal Measures) for effective feature face analysis and discrimination.
- (2) Mapping each feature vector into Kernel feature space, using RBF Kernel  $K(x, x_i) = e^{-\|x-x_i\|^2/\sigma}$ .
- (3) Training face: Train the IKNDA on the various classes of feature vectors in the Kernel space using Eqs. (12) and (15) so as to obtain nonparametric discriminant features.
- (4) Testing phase: the trained IKNDA is used to project the test pattern into different lower dimension classes.

(5) The Euclidean distance is employed to assort the testing examples into derived different lower dimension classes.

(6) Face Recognition.

## 2.2.4 Block Diagram of the Face Recognition Method

In this part, we will represent in details the process of our face recognition system:



**Fig. 3.** Block Diagram of the Face Recognition Method.

## 3. Experimental Results

In this part, we are going to delineate a decontextualized comparison and assessment of the IKNDA to the BKNDA in the matter of classification accuracy, to highlight the prominence of our incremental model over batch model, as well as, other relevant state-of-the-art incremental classifiers. Then, we will perform an evaluative comparison of the IKNDA in the face recognition context, to show the advantage of incrementally updating eigenspace after each sample is presented.

### 3.1 Decontextualized Evaluation and Comparison

In this subsection, we employ real world datasets to evaluate the IKNDA and compare it to the BKNDA, the Incremental Kernel Fisher's Discriminant Analysis (IKFD) [14], the Incremental Linear Discriminant Analysis (ILDA), and the Incremental Nonparametric Discriminant Analysis (INDA) [12], in terms of classification performance. For data kernelization, we utilize the Gaussian RBF kernel  $K(x, x_i) = e^{-\|x-x_i\|^2/\sigma}$ . In fact, it was shown to be flexible and robust [24]. Here  $\sigma$  denotes positive "width" parameter. Given that the purpose of the decontextualized evaluation is to show case the performance of the

IKNDA in a general context, the selected databases have no connection to Gabor and Ordinal analysis. These datasets are extracted from the UCI machine learning repository and selected minutely, so much so we have heterogeneity of dimensions and sizes, and then we may put to examination the robustness of our IKNDA against different feature sizes. More details regarding the original provenance of these databases are to hand in a highly comprehensive online repository [25]. A thorough depiction of the utilized datasets is represented in **Table 1**.

To make certain that the attained results are premeditated and unbiased, we utilize a 10 fold cross validation process [26]. More precisely, each used dataset was arbitrarily divided into 10 subsets of identical size. To establish a model, one of the 10 subsets was eliminated to represent test samples, and the remainder was applied as the training data. Finally, the accuracy over all models is calculated by averaging the 10 obtained accuracy estimates.

For the BKNDA and IKNDA, cross-validation and grid search are used in order to find the combination of hyperparameters, namely,  $\sigma$  and  $\kappa$  that yields the best classification performance. As far as the IKFD is concerned, the tuning of  $\sigma$  is achieved using cross-validation. Concerning INDA, we achieved 10 separate runs with various nearest neighbour numbers  $\kappa \in \{1, 2, \dots, 10\}$ , respectively, and then we choose the value of  $\kappa$  that affords the finest classification efficiency. In order to evaluate and compare two class classifiers in terms of performance, the Receiver Operating Characteristic (ROC) curves [27] are generally used. The ROC curve represents a powerful computation of the efficiency of studied classifiers. It does not call for the number of testing or training data points, but instead it requires only the rates of appropriate and inappropriate samples classification. The ROC curve is determined by plotting the True Positive Rate (TPR) vs the False Positive Rate (FPR). To evaluate the different methods, the Area Under Curve is made use of (AUC) [28], produced by the ROC curves, and we presented them in **Table 2**.

**Table 1.** Details of the real databases

Database name	Number of class1 samples	Number of class2 samples	Number of classes	Number of features
Breast Cancer	233	430	2	9
diabetes	268	500	2	8
diabetes1	268	500	2	7
German	700	300	2	24
Heart Disease	134	159	2	13
liver	145	200	2	6
Thyroid	150	35	2	5
Ionosphere	225	126	2	34
Haberman's survival	225	81	2	3
Twonorm	3697	3703	2	20
Waveform	3304	1696	2	21
SPECT	212	55	2	44
Ringnorm	3736	3664	2	20

We can clearly notice from **Table 2**, that by averaging over 10 different models, IKNDA outperforms the ILDA, IKFD, INDA, IKSVM and BKNDA for all used datasets, in terms of computed AUCs. Unsurprisingly, ILDA gives almost the worst accuracy values. In fact, the ILDA utilizes non-flexible linear boundary so as to discriminate between face classes. Its performance is lacking when exercised on issues that call for nonlinear decision boundaries. Still, the efficiency of ILDA dwindles as the underlying class distribution is not normal. The

IKFD provides better results than the ILDA, since it utilizes the Kernel trick to create a flexible non-linear decision boundary between different classes. Generally speaking, in terms of AUC, there is a greater difference between the ILDA and the INDA, BKNDA and IKNDA than there is between the ILDA and the IKFD. This was expected as the INDA, BKNDA and IKNDA, are formulated on the principle that the normal vectors on the decision boundary are more relevant for distinction making. Thus, in feature space, the normal vectors are used to compute the between-class scatter matrix on a local basis in the neighbourhood of the decision boundary. As pleasing consequence, they allow the relaxation of normality assumption [39]. The IKNDA outperforms the IKSVM since this latter can be misled by data spread, as its resolution boils down to the limited number of support vectors. Obviously, the BKNDA and IKNDA outperform the INDA thanks to the kernelization. Finally, according to **Table 2**, the IKNDA and BKNDA perform almost the same on different datasets, in terms of classification performance. More precisely, the IKNDA slightly outperforms the BKNDA on the majority of datasets. In fact, this is interpreted by the samples introduction order, which has an effect on the evaluation of the nearest neighbours and then the computation of the between scatter matrix  $S_{BK}$ . Also, the estimation of the nearest neighbours in incremental could affords lightly cleaner solution.

To further show case the progress in efficiency by IKNDA, we introduce a number of statistical measurements. We put to use a paired t-tests on the AUC values by pairing up the IKNDA technique with every technique at its definite time. The paired t-test discloses if the two sets of measured values are significantly different. The last row of **Table 2** provides the confidence intervals (in %) attained from the performed t tests. This confidence interval evaluates the probability of the paired distributions being the same. As to when the confidence interval is high, it is unlikely that the underlying distributions are statistically similar. Noticeably, confidence intervals in their entirety are high (very close to 100%), which makes plain that IKNDA unveils many statistically significant accuracy improvements.

**Table 2.** Average AUC of every single formula for the 13 real-world datasets (best formula in bold, second best emphasized)

Databases name	ILDA	IKFD	INDA	IKSVM	BKNDA	IKNDA
Breast Cancer	64.21	82.63	89.99	93.92	95.81	<b>95.85</b>
German	52.23	67.33	68.78	79	<b>80.5</b>	80.2
HeartDisease	60.51	68.91	72.31	81.26	82.59	<b>82.61</b>
Ionosphere	55.30	67.00	70.83	94.4	94.57	<b>94.9</b>
Ringnorm	68.44	72.76	79.89	97.5	98.57	<b>98.66</b>
SPECT	69.52	79.87	84.76	96.4	<b>96.52</b>	96.49
Thyroid	68.89	88.74	91.11	96	96.67	<b>96.69</b>
Twonorm	59.94	82.23	86.78	97.5	97.54	<b>97.58</b>
Waveform	68.78	89.92	91.85	90.55	<b>94.3</b>	<b>94.3</b>
Diabetes	67.23	70.61	77.76	80.75	82.88	<b>82.89</b>
Diabetes1	59.50	64.57	70.25	75.45	75.4	<b>75.7</b>
Haberman's survival	67.90	72.56	82.95	83.9	84	<b>84.09</b>
Liver	53.93	67.39	72.50	74.8	<b>78.7</b>	78.4
Confidence	95.7	96.7	97.9	98.4	99.5	-

### 3.2 Contextualized Evaluation and Comparison

In this subsection, we evaluate our novel adaptive face recognition method. Moreover, we make a comparative evaluation of the IKNDA to the BKNDA, ILDA, IKFD, IKSVM and

INDA to showcase the efficiency of the formula we advance, in terms of classification performance and training runtime, in the face recognition context. For this reason, we have used the Yale [32], ORL [33], UMIST [34], AR [35] and Indian [36] face databases. Moreover, for the same datasets, we have compared the different face recognition methods, using Local Normalization- Local Gabor Binary Pattern (LN-LGBP) descriptor [30] and [31] instead of GOM descriptor for face image representation, in order to show the advantage of the GOM descriptor. A detailed description of the used face databases is represented in **Table 4**.

**Table 3.** Description of the face databases

Face Database name	Number of faces	Number of individuals or classes	Number of faces per class
Yale	165	15	11
ORL	400	40	10
UMIST	564	20	20
AR	1190	126	30
Indian	440	40	11

For each Yale, ORL, UMIST, AR and Indian face image, we employed a set of 2D Gabor wavelets made up of 5 frequencies and 8 orientations on every facial image. Ordinal measures are extracted from diverse components of Gabor images, applying di-lobe and four-lobe Multi-Lobe Differential Filters (MLDF), with orientation values equal to  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ , 4 pixel-inter-lobe distance and  $5 \times 5$  lobe size. Subsequently, each acquired image is split into blocks that correspond to feature vectors. Hence forward, we have calculated the histograms of feature vector, which can be shrunk to encapsulate only 64 bins.

Finally, we have fallen back on PCA to reduce the dimension of feature vector.

To evaluate and compare every single method's performance, we have employed classification accuracy. This is determined by  $100 \frac{N_{CC}}{N} \%$ , where  $N$  is the sum of data points (testing data) and  $N_{CC}$  is the sum of points classified correctly. Also, for kernelization we have used the Gaussian RBF. Concerning the computation of the hyperparameters  $\sigma$  and  $\kappa$ , we performed the same setting as the decontextualized comparative evaluation.

**Table 5**, **Table 6** and **Table 6** illustrate, respectively, the training times and classification accuracies of the compared classifiers, using LN-LGBP and GOM descriptors, on the face databases.

According to **Table 6** and **Table 6**, we can see that the IKNDA outperforms the other classifiers when it comes to recognition precision. This was expected, as the IKNDA is an incremental model based on flexible non-linear boundary, and it relaxes normality assumption. Also, we can notice that the face recognition methods based on GOM descriptor outperform the ones based LN-LGBP descriptor. In fact, unlike the GOM descriptor, the LN-LGBP descriptors not robust to noise and face occlusions caused by the image acquisition environment. Moreover, LN-LGBP descriptor considers only the qualitative relationship between two pixels, whereas GOM descriptor deploys ordinal measures among multiple image regions in intensity and feature levels.

As for computational complexity, the most time consuming phase consists of determining the  $S_{BK}$  matrix (consisting finding the most proximate neighbors). Envisaging to show the performance of IKNDA, we calculated the computational time required to update  $S_{BK}$  (using Eqs (12) and (17)) and compared it with time required by the Batch KNDA to recalculate  $S_{BK}$  (using Eq (9)) each time from the beginning. More precisely, for each used face database,

we successively updated the  $S_{BK}$  matrix with all face images. According to **Table 5**, the time needed for IKNDA to update the  $S_{BK}$  matrix employing the incremental technique, as new sample is added, is significantly inferior than recalculating the  $S_{BK}$  from the beginning, using BKNDA. IKNDA approach more relevant for real-time applications.

**Table 4.** Training times in seconds for BKNDA and IKNDA on the face databases (using GOM descriptor).

Databases name	BKNDA	IKNDA
ORL	0.85	0.06
Yale	0.48	0.04
UMIST	1.23	0.05
AR	1.75	0.56
Indian	0.67	0.05

**Table 5.** Classification accuracy of each classifier on the face databases, using GOM descriptor (best method in bold, second best emphasized).

Databases name	ILDA	IKFD	INDA	IKSVM	BKNDA	IKNDA
ORL	30.67	77.9	80.19	83.15	98.57	<b>98.60</b>
Yale	33.3	88.8	89.14	90	90	<b>91.12</b>
UMIST	50.8	72.34	83.1	96.2	<b>99.68</b>	99.63
AR	53.6	70.1	82.2	84.54	90.31	<b>90.36</b>
Indian	60.6	85.5	90.7	93.65	95.80	<b>95.89</b>

**Table 6.** Classification accuracy of each classifier on the face databases, using LN-LGBP descriptor (best method in bold, second best emphasized).

Databases name	ILDA	IKFD	INDA	IKSVM	BKNDA	IKNDA
ORL	30.2	70.3	77.9	78.32	87.3	<b>87.5</b>
Yale	32	78.1	81.5	86.56	88.76	<b>88.78</b>
UMIST	45.34	65.98	77.43	95.1	90.2	<b>90.6</b>
AR	48.9	66.3	75.1	80.2	80.33	<b>80.46</b>
Indian	57.36	79.2	81.52	92.7	<b>85.10</b>	85.09

#### 4. Conclusion

In the present, a novel adaptive face recognition system formulated on Incremental Kernel Nonparametric Discriminant Analysis (IKNDA) is what we advance. This latter is based on incremental learning process, where new items may be adjoined asynchronously, over distinct time stamps, the moment they become accessible. The IKNDA draws upon the most proximal neighbor rule. We fell back on a decontextualized comparison of the IKNDA to relevant state-of-the-art classification algorithms ILDA, IKFD, IKSVM and INDA on real world datasets. Experiments proved that the IKNDA outdid these algorithms, on classification efficiency. Moreover, a contextualized comparison and evaluation has shown the robustness of the presented adaptive face recognition method and the superiority of the IKNDA in terms of face recognition performance and computational running time. In future work, since IKNDA was shown to be more performant than the incremental SVM, we can investigate its effectiveness to classify deep face features and to improve face recognition performance.



## References

- [1] A. K Jain and Li. S. Z, “Handbook of Face Recognition,” *New York, Springer*, 2011.
- [2] Patel Anjali and Verma, Ashok, “IOT based Facial Recognition Door Access Control Home Security System,” *International Journal of Computer Applications*, vol.172, no.7, pp.11-17, 2017. [Article \(CrossRef Link\)](#).
- [3] Zhao Wenyi, Chellappa Rama, Phillips P. Jonathon and Azriel Rosenfeld, “Face recognition: a literature survey,” *ACM Computing Surveys. New York. NY. USA*, vol. 35, no. 4, pp. 399–458, 2003. [Article \(CrossRef Link\)](#).
- [4] Bonnen Kathryn, Klare Brendan F and JAIN Anil K, “Component-based representation in automated face recognition,” *IEEE Transactions on Information Forensics and Security (IFS)*, vol. 8, no.1, pp. 239–253, 2013. [Article \(CrossRef Link\)](#).
- [5] Chai Zhenhua, Sun Zhenan, Méndez-Vázquez Heydi, Ran He and Tiemiu Tan, “Gabor Ordinal Measures for Face Recognition,” *IEEE Transactions on Information Forensics and Security*, vol. 9, no.1, pp-14-26, 2014. [Article \(CrossRef Link\)](#).
- [6] Li Zhifeng, Lin Dahua and Tang Xiaoou, “Nonparametric Discriminant Analysis for Face Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no.4, pp. 755-761, 2009. [Article \(CrossRef Link\)](#).
- [7] Arbia Soula, Salma Ben Said, Riadh Ksantini and Zied Lachiri, “A Novel Kernelized Face Recognition System,” in *Proc. of IEEE conf. on Control Engineering & Information Technology*, pp.1-5, 2016. [Article \(CrossRef Link\)](#).
- [8] Arbia Soula, Salma Ben Said, Riadh Ksantini and Zied Lachiri, “A Novel Face Recognition System Base on Nonparametric Discriminant Analysis,” in *Proc. of IEEE conf. on Control, Automation and Diagnosis*, pp. 1-5, 2017. [Article \(CrossRef Link\)](#).
- [9] Parchami Mostafa, Bashbaghi Saman and Granger Eric, “Video-based face recognition using ensemble of haar-like deep convolutional neural networks,” in *Proc. of International Joint Conference in Neural Networks (IJCNN)*, pp. 4625-4632, 2017. [Article \(CrossRef Link\)](#).
- [10] Dat Tien Nguyen, Tuyen Danh Pham, Na Rae Baek and Kang Ryoung Park, “Combining Deep and Handcrafted Image Features for Presentation Attack Detection in Face Recognition Systems Using Visible-Light Camera Sensors,” *Sensors*, vol.18, no. 3, pp.699, 2018. [Article \(CrossRef Link\)](#).
- [11] Naimul Mefraz Khan, Xiaoming Nan, Azhar Quddus, Edward Rosales and Ling Guan, “On video based face recognition through adaptive sparse dictionary,” in *Proc. Of 11th IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, pp. 1-6, 2015. [Article \(CrossRef Link\)](#).
- [12] Raducanu Bogdan and Vitrià Jordi, “Online nonparametric discriminant analysis for incremental subspace learning and recognition,” *Pattern Analysis and Application*, vol. 11, no. 3-4, pp. 259-268, 2008. [Article \(CrossRef Link\)](#).
- [13] Arbia Soula, Salma Ben Said, Riadh Ksantini and Zied Lachiri, “A Novel Incremental Face Recognition Method Based on Nonparametric Discriminant Model,” in *proc. Of 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp.1-5, 2018. [Article \(CrossRef Link\)](#).
- [14] Hsien-Ting Cheng, Yi-Hsiang Chao, Shih-Liang Yeh, Chu-Song Chen, Hsin-Min Wang and Yi-Ping Hung “An efficient approach to multimodal person identity verification by fusing face and voice information,” in *proc. Of IEEE International Conference on Multimedia and Expo ( ICME)*, pp.542-545, 2005. [Article \(CrossRef Link\)](#).
- [15] Paweł Karczmarek, Adam Kiersztyn, Witold Pedrycz and Michał Dolecki, “An application of chain code-based local descriptor and its extension to face recognition,” *Pattern Recognition*, vol. 65, pp. 26-34, 2017. [Article \(CrossRef Link\)](#).
- [16] Lufan Li, Zhang Jun, Jiawei Fei and Shuohao Li, “An incremental face recognition system based on deep learning,” in *Proc. of 15th IAPR International Conference on Machine Vision Applications (MVA)*, 2017. [Article \(CrossRef Link\)](#).

- [17] Wiskott, Laurenz, Fellous, Jean-Marc, Krüger, Norbert and Christoph. vonder Malsburg, "Face recognition by elastic bunch graph matching," *PAMI*, vol. 19, pp.775–779, 1997. [Article \(CrossRef Link\)](#).
- [18] Ángel Serrano, Isaac Martín de Diego, Cristina Conde and Enrique Cabello, "Recent advances in face biometrics with Gabor wavelets: a review," *Pattern Recognition Letters (PRL)*, vol.31, no.5, pp. 372–381, 2010. [Article \(CrossRef Link\)](#).
- [19] V. Vapnik, "Statistical Learning Theory," *Wiley, New York*, 1998.
- [20] R.O. Duda, P.E Hart and D.G Stork, "Pattern Classification," *2nd edition, Wiley, New York*, 2000.
- [21] K. Fukunaga, Introduction to Statistical Pattern Recognition, *2<sup>nd</sup> edition. Academic Press, London*, 2000.
- [22] Keinosuke Fukunaga and Hostetler Larry, "The estimation of the gradient of a density function, with applications in pattern recognition." in *Proc. of IEEE Transactions on Information Theory*, vol.1, no.21, pp-32–40, 1975. [Article \(CrossRef Link\)](#).
- [23] S. Mika, G. Ratsch, J. Weston, B. Scholkopf and K. Mullers, "Fisher discriminant analysis with kernels," in *Proc. of IEEE signal processing society workshop In Neural networks for signal processing*, pp. 41-48, 1999. [Article \(CrossRef Link\)](#).
- [24] M. Cristianini and J. Shawe-Taylor, "An Introduction to Support Vector Machines," *Cambridge University Press, Cambridge*, 2000.
- [25] M. Lichman, "Machine Learning Repository," *University of California, Iverine, School of Information and computer Sciences*, 2013. [Article \(CrossRef Link\)](#).
- [26] Cesare Alippi and Manuel Roveri, "Virtual k-fold cross validation: an effective method for accuracy assessment," in *Proc. of International Joint Conference on Neural Networks*, pp.1-6, 2010. [Article \(CrossRef Link\)](#).
- [27] Tom Fawcett, "An introduction to ROC analysis," *Pattern Recognition letters*, vol. 27, pp. 861–874, 2006. [Article \(CrossRef Link\)](#).
- [28] Hanley James A and Mcneil Barbara J, "A method of comparing the areas under receiver operating characteristic curves derived from the same cases," *Radiology*, vol. 148, no. 3, pp. 839–843, 1983. [Article \(CrossRef Link\)](#).
- [29] K. Fukunaga, "Introduction to Statistical Pattern Recognition," *2<sup>nd</sup> edition. Academic Press, London*, 2000.
- [30] Xie Xudong and Lam Kin-Man, "An efficient illumination normalization method for face recognition," *Pattern Recognition Letters*, vol. 27, no.6, pp. 609-617, 2006. [Article \(CrossRef Link\)](#).
- [31] Jiang Yanxia and Ren Bo, "Face Recognition using Local Gabor Phase Characteristics," in *Proc. of IEEE International Conference on Intelligence and Software Engineering*, pp. 1-4, 10-12 December 2010. [Article \(CrossRef Link\)](#).
- [32] Georghiadis, "A- Yale face database," 1997.
- [33] Yu Hua and Yang Jie, "A direct LDA algorithm for high-dimensional data with application to face recognition," *Pattern Recognition*, vol.34, pp. 2067-2070, 2001. [Article \(CrossRef Link\)](#).
- [34] Face Database. [Article \(CrossRef Link\)](#).
- [35] AM. Martinez, R. Benavente, "The AR face database," *Computer Vision Center*, 1998.
- [36] V. Jain, A. Mukherjee, "The indian face database," 2002.



**Arbia Soula** is currently pursuing the PhD degree at Signal, Image and Information Technology laboratory (LR-SITI) at the National School of Engineering of Tunis (ENIT-Tunisia). Her research interests include signal and image processing, machine learning, Face Recognition.



**Salma Ben Saïd** received the Ph.D. degree from ENIT, in Electrical engineering in 2011. Her research interests include hyperspectral data analysis and remote sensing data applications.



**Riadh Ksantini** received the Ph.D. degree in Computer Science from the University of Sherbrooke, Sherbrooke, QC, Canada, in 2007. He is currently Graduate Faculty in the School of Computer Science, University of Windsor, Canada. His research interest includes: computer vision, machine learning and pattern recognition.



**Zied Lachiri** received the PHD degree in Electrical Engineering from the ENIT, 2002. He is currently a Professor at the Department of Electrical Engineering, ENIT. His research interests include pattern recognition, signal/image processing and applied in biomedical, multimedia and man machine communication.