

Ordinal matrix encoding based facial recognition

Guillène Martiale Wandja · J. S. Armand Eyebe Fouda · Hermann
Djeugoue Nzeuga · Samrat L. Sabat · Wolfram Koepf

Received: date / Accepted: date

Abstract Now-a-days, ordinal patterns are shown to be effective in extracting discriminant image features. In this paper, we present the ordinal matrix encoding (OME) as a method that transforms an 8-bit encoded image into another image with s gray levels. Such an encoding acts as a highpass filter and allows us to enhance the image contours that are useful for feature extraction. In this work, we hybridized the OME technique with the linear discriminant analysis (LDA) approach to define the modified LDA (MLDA) to extract image features. The MLDA considers only interclass matrices of encoded images to highlight their singularities. Subsequently, a support vector machine (SVM) is applied to the MLDA output to perform facial image classification. We validated the proposed classification method using images from the ORL, FERET and FEI standard databases. The results indicate an overall accuracy of 99.07%, 73.61% and 98.78% for the ORL, FERET and FEI databases, respectively. Further, we evaluated the impact of OME by analyzing the classification accuracy

of the SVM-LDA combination on raw images from the ORL database. The accuracy was 95.25% with intra-class matrices and 94.50% without, both lower than the 99.07% achieved with encoded images. This improvement occurs because OME preserves only the essential details of the raw images for feature extraction, enhancing their discriminative ability.

Keywords Feature extraction · Facial recognition · Ordinal pattern · Image classification

1 Introduction

Facial recognition has been an active research field for several decades and has recently gained significant attention in domains like computer vision, machine learning, artificial intelligence, and video surveillance. This growing interest is driven by major advances and the wide range of applications of technology in society. The primary function of facial recognition systems is to identify individuals from static images, video footage, or data streams [1]. The effectiveness of the system depends largely on its ability to extract key features that ensure accurate identification. Linear methods, for example, have been extensively used in computer vision and machine learning over the past two decades. These methods include Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and their variants, which use eigenfaces and fisherfaces for feature extraction [2, 3]. PCA is a dimensional reduction technique and achieved a 95% accuracy rate in the ORL image database for facial recognition [4]. Similarly, LDA has become one of the most popular linear projection techniques for feature extraction and facial recognition. Studies such as those in [5] have used LDA to improve

Guillène Martiale Wandja
Faculty of Science, University of Yaounde 1, Cameroon
E-mail: guillenewandja@gmail.com

J. S. Armand Eyebe Fouda
Faculty of Science, University of Yaounde 1, Cameroon
E-mail: efoudajsa@gmail.com

Hermann Djeugoue Nzeuga
Faculty of Science, University of Yaounde 1, Cameroon
E-mail: mahernzeuga@yahoo.fr

Samrat L. Sabat
Centre for Advanced Studies in Electronics Science and Technology, University of Hyderabad, India
E-mail: slssp@uohyd.ac.in

Wolfram Koepf
Institute of Mathematics, University of Kassel, Germany
E-mail: koepf@mathematik.uni-kassel.de

class separability in image samples. However, it is often recommended to combine LDA with other methods, as its accuracy usually stabilizes around 95%. Gui-Fu et al. introduced an incremental approach to Complete Linear Discriminant Analysis (CLDA), which updates discriminant vectors as new samples are added to the training set thus the efficiency of the model [6]. In [7], a novel LDA-based method achieved an impressive facial recognition accuracy of 97.54%. Despite their success, facial recognition remains challenging in unconstrained environments, where factors such as changing lighting conditions and varied poses affect performance [8,9]. Innovative solutions to improve face recognition against adversarial attacks and mask-induced occlusions were proposed in [10] and [11]. However, these approaches remain complex and resource-intensive. The InceptBlock Enhanced Attention Fusion Network (IBEAFNet) developed in [11] was used to suppress the less relevant mask regions of the face, while focusing on significant fine- and coarse-level features. Its application on the CASIA, Yale, and HIF databases showed an accuracy of 91.00%, 89.5%, and 93.00%, respectively.

Recently, entropy-based methods have been introduced for feature extraction [12]. Building on the concept of permutation entropy (PE), Manas Ghosh et al. developed an algorithm that combines PE with fuzzy G-2DLDA for facial recognition. Their approach resulted in an average accuracy of 99.03% and 66.04% on the ORL and FERET dataset respectively [13]. These findings highlight the effectiveness of ordinal patterns (OP) in image classification. However, the accuracy of these methods varies unpredictably with changes in the size of the training set. Ideally, the accuracy should have shown a consistent improvement with the increase in the size of the training set. Additionally, the relatively low accuracy with the FERET database indicates that challenges remain to be addressed.

This paper proposes a technique to improve facial recognition accuracy by combining the OP transform with LDA and SVM. The process involves the use of OP transform for image pre-processing, the Fisher approach for dimensionality reduction, LDA to extract features that optimize class separability, and finally, SVM as the classification method for testing. Ordinal patterns are typically defined for one-dimensional (1D) data series. When applied to images, 2D matrices are first converted to 1D signals before determining the OPs [14]. In our approach, we introduce a two-dimensional transform called ordinal matrix encoding (OME), which outputs ordinal matrices. OME builds on the 1D OP encoding described in [15], transforming each pixel in the original image into a new pixel value to generate the ordinal image. OME enhances image details,

thus improves their discrimination characteristics. Applying LDA on encoded images, we expect to highlight discriminant features between different image classes, thereby strengthening similarities between images of the same class. This helps to ignore the intra-class scatter matrices in the LDA algorithm, while improving the classification accuracy. We therefore define the modified LDA (MLDA) by replacing original test images by their ordinal representation in the LDA approach where only inter-class scatter matrices are considered. The main contributions of this paper are the following.

1. The applications of ordinal matrix encoding (OME) to highlight image details.
2. The use of MLDA including OME to enhance feature extraction by suppressing the computation of intraclass scatter matrices.
3. The combination of MLDA with SVM to enhance the recognition accuracy.

The remainder of the paper is structured as follows. Section 2 presents the ordinal matrix encoding, Section 3 is devoted to the application of OME to image classification, and simulation results are given in Section 4, followed by conclusions in Section 5.

2 Ordinal matrix encoding

2.1 Brief recall on ordinal patterns

Consider a time series $\{x_t\}_{t=0,1,\dots,L-1}$, where t is the time index and P defines an embedding dimension. Comparing the neighboring values sorted into ascending order in a given embedding vector $\mathbf{x}_i = (x_i, x_{i+\tau}, \dots, x_{i+(P-1)\tau})$, $i = 0, \dots, L - (P - 1)\tau - 1$ and $P \in \mathbb{N}_{\geq 3}$, permutations π_i of length P are obtained. $\tau \in \mathbb{N}_{\geq 1}$ is the delay time of the samples or the time lag [16]. In fact, each embedding vector \mathbf{x}_i is associated with the ordered vector of position $\pi^0 = (0, 1, \dots, P - 1)$. Sorting \mathbf{x}_i in ascending order yields the sorted vector \mathbf{x}_i^0 , and rearranging the values of π^0 according to their previous position in \mathbf{x}_i leads to the ordinal pattern π_i [17]. An example is given for $\{x_t\} = \{1, 9, 8, 6, 1, 7, 3, 5, 2\}$, $P = 3$ and $\tau = 1$. In this case, the first embedding vector $\mathbf{x}_0 = (1, 9, 8)$ in ascending order leads to $\mathbf{x}_0^0 = (1, 8, 9)$, while the corresponding vector of position $(0, 1, 2)$ leads to the ordinal pattern $\pi_0 = (0, 2, 1)$. The complete list of ordinal patterns for this example is shown in Fig.1.

2.2 Ordinal matrix encoding

We define the ordinal matrix encoding as 2D nonlinear transform f . It consists of transforming a matrix

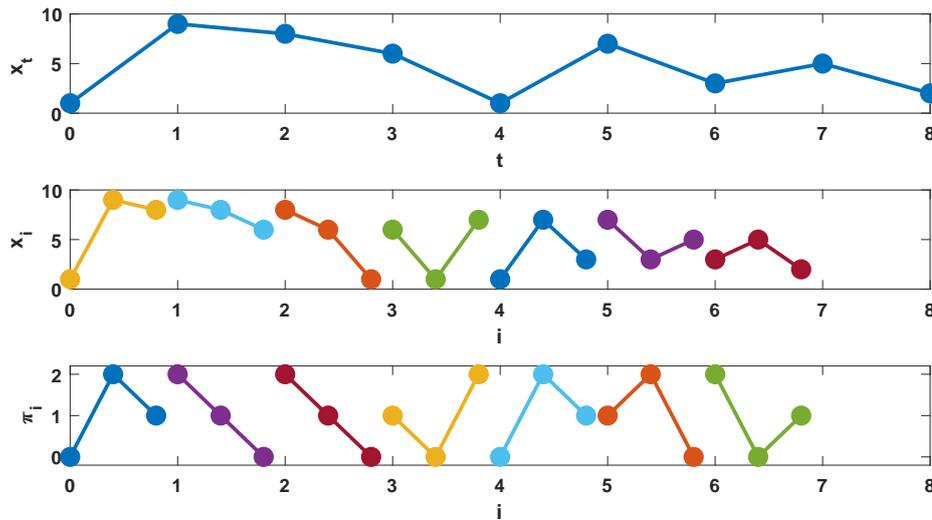


Fig. 1: Example of ordinal patterns of length 3 derived from the time series $\{x_t\} = \{1, 9, 8, 6, 1, 7, 3, 5, 2\}$, $P = 3$. From top to bottom are shown the time series $\{x_t\}$, the embedding vectors \mathbf{x}_i and the corresponding ordinal patterns π_i , respectively.

$W = (\omega_{i,j})_{0 \leq i < P, 0 \leq j < Q}$, $(P, Q) \in \mathbb{N}_{\geq 1}$ into an encoded scalar value $\lambda = f(W)$. Rows or columns of W are first individually sorted into ascending order to obtain an ordinal matrix $\Pi^{(x)}$ or $\Pi^{(y)}$ that is of the size $P \times Q$ as W and whose rows or columns are ordinal patterns $\pi_i^{(x)}$ or $\pi_j^{(y)}$, respectively. Thereafter, the rows or columns of the ordinal matrix are encoded to derive λ . In the case of row sorting, $\Pi^{(x)}$ is a vector of P rows of Q -length ordinal patterns $\pi_i^{(x)}$, the number of distinct patterns $\pi_i^{(x)}$ being $Q!$. According to [15], the set of possible ordinal patterns $\{\pi_i^{(x)}\}$ can be encoded, so that each pattern $\pi_i^{(x)}$ is assigned a unique code $0 \leq \epsilon_i^{(x)} \leq Q! - 1$, $\epsilon_i^{(x)} \in \mathbb{N}$, as shown in Table 1. By assigning each $\pi_i^{(x)}$ in $\Pi^{(x)}$, its corresponding code

Table 1: Example of encoding values of ordinal patterns of length 3.

π_i	(0,1,2)	(0,2,1)	(1,0,2)	(1,2,0)	(2,0,1)	(2,1,0)
ϵ_i	0	1	2	3	4	5

$\epsilon_i^{(x)}$, the ordinal matrix $\Pi^{(x)}$ is transformed into an encoded vector $\Gamma^{(x)} = (\epsilon_0^{(x)}, \epsilon_1^{(x)}, \dots, \epsilon_{P-1}^{(x)})^T$ or simply $(\epsilon_0^{(x)} \epsilon_1^{(x)} \dots \epsilon_{P-1}^{(x)})_{Q!}$. Given that $0 \leq \epsilon_i^{(x)} < Q!$, $\epsilon_i^{(x)}$ can be seen as symbols of the $Q!$ basis and $\Gamma^{(x)}$ represents the number representation in the same basis. Thus, the decimal value $\lambda^{(x)}$ of $\Gamma^{(x)}$ is obtained as

$$\lambda^{(x)} = \sum_{i=0}^{P-1} \epsilon_i^{(x)} \cdot (Q!)^i, \quad (1)$$

where $0 \leq \lambda^{(x)} < (Q!)^P$, $\lambda^{(x)} \in \mathbb{N}$. By applying the same principle in the y -dimension, $\lambda^{(y)}$ is obtained as

$$\lambda^{(y)} = \sum_{j=0}^{Q-1} \epsilon_j^{(y)} \cdot (P!)^j, \quad (2)$$

where $0 \leq \lambda^{(y)} < (P!)^Q$. $\lambda^{(x)}$ and $\lambda^{(y)}$ are directional values. We can also define the non-directional value $\bar{\lambda}$ by considering the mean value of the two directional values as

$$\bar{\lambda} = \left\lfloor \frac{\lambda^{(x)} + \lambda^{(y)}}{2} \right\rfloor. \quad (3)$$

An example is given for $W = \begin{pmatrix} 120 & 191 & 47 \\ 233 & 188 & 153 \\ 27 & 14 & 77 \end{pmatrix}$, $P = Q =$

3. The ordinal matrices in the x and y direction are, respectively,

$\Pi^{(x)} = \begin{pmatrix} 2 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}$ and $\Pi^{(y)} = \begin{pmatrix} 2 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$. Ac-

cording to Table 1, $\Pi^{(x)}$ is encoded as $\Gamma^{(x)} = (4, 5, 2)^T$ or simply 452_6 ; and $\Pi^{(y)}$ as $\Gamma^{(y)} = (4, 5, 1)$ or simply 451_6 . Taking into account Eqs. (1)-(3), the corresponding scalar values on basis 10 are, respectively, $\lambda^{(x)} = 106$, $\lambda^{(y)} = 70$ and $\bar{\lambda} = 88$.

2.3 Ordinal image encoding

Let an original image $A = (a_{m,n})_{0 \leq m < M, 0 \leq n < N}$, $(N, M) \in \mathbb{N}_{\geq 1}$, the corresponding OMEI $\Lambda = (\lambda_{m,n})$ is obtained by transforming each sub-matrix $W^{(m,n)} =$

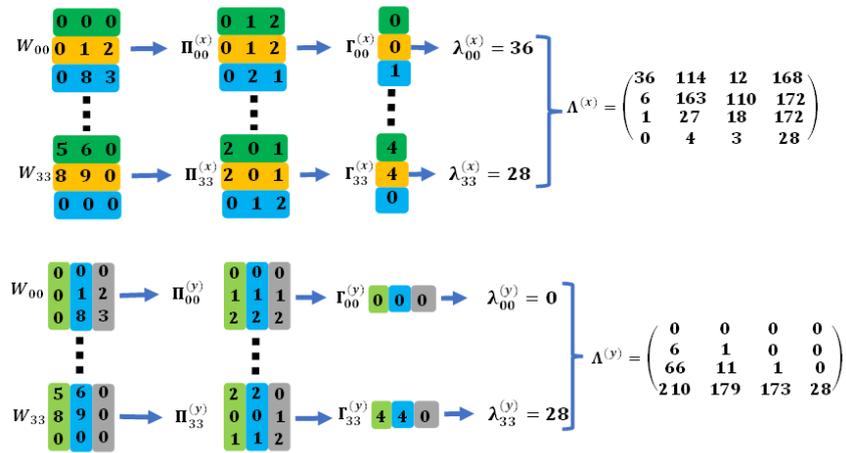


Fig. 2: Example of OME of matrix A. The OME parameters are set as $P = Q = 3$ respectively along x and y direction

$(\omega_{i,j})_{P \times Q}$, $(P, Q) \in \mathbb{N}_{\geq 1}^2$ centered at $a_{m,n}$ with $P \leq M$, $Q \leq N$. Thus, each pixel $\lambda_{m,n}$ of Λ is obtained by applying the OME to the local region covered by $W^{(m,n)}$. This process involves the sub-matrix moving pixel by pixel across the image, starting from the top left corner and moving to the right, then moving down line by line until the bottom right corner. A zero padding is considered for the processing of the image borders. As an example, let us consider the image A defined as

$$A = \begin{pmatrix} 1 & 2 & 1 & 4 \\ 8 & 3 & 4 & 5 \\ 6 & 7 & 5 & 6 \\ 5 & 6 & 8 & 9 \end{pmatrix}, \quad (4)$$

with $M = N = 4$. By considering sub-matrices $W^{(m,n)}$ of size 3×3 and performing the OME in the x or y direction, we obtain the OMEs $\Lambda^{(x)}$ or $\Lambda^{(y)}$, respectively. The corresponding results are shown in Fig.2. The results obtained by applying the proposed image encoding method to an image of the ORL database are depicted in Fig.3. The corresponding spectra are shown in Fig. 4 to confirm that the bandwidth of low frequencies depicted by the relatively large white spot in the center of the original image spectrum is narrowed in the transformed image spectra, whereas the high frequencies are accentuated. The spectrum of the mean encoded image is brightened beyond low frequencies and in all directions. The low frequencies in the transformed image preserve structural information (general shapes and contours of objects within an image) that is crucial for object recognition and classification [18]. The accentuation of high frequencies allows us to highlight fine details, such as image texture and edges [19]. Given that the application of the OME preserves es-

sential structural information and accentuates details of the image, it could constitute an important tool for the discrimination and classification of images. Therefore, we suggest combining OME with LDA to perform facial image feature extraction. .

3 Application of OME to image classification

Our model for feature extraction combines OME and LDA. In this section, we provide details on the algorithmic steps of the proposed image classification approach.

3.1 Brief recall of LDA

LDA is a commonly used technique in supervised machine learning to solve multi-class classification problems. It is a classical approach for dimensionality reduction and supervised classification. Consider that there are C classes and D training face images for each class. Assuming that \mathcal{I}_{ij} is the j -th face image of the i -th class in the \mathcal{X} database, then \mathcal{X} can be represented as $\mathcal{X} = \{\mathcal{I}_{ij}\}$ with $1 \leq i \leq C$ and $1 \leq j \leq D$.

Performing LDA on this data set for classification purposes involves five steps as described below.

Step 1: Compute the dataset global mean image μ as

$$\mu = \frac{1}{C \cdot D} \sum_{i=1}^C \sum_{j=1}^D \mathcal{I}_{ij} \quad (5)$$

Step 2: Compute the mean images for each class i as

$$\mu_i = \frac{1}{D} \sum_{j=1}^D \mathcal{I}_{ij} \quad (6)$$

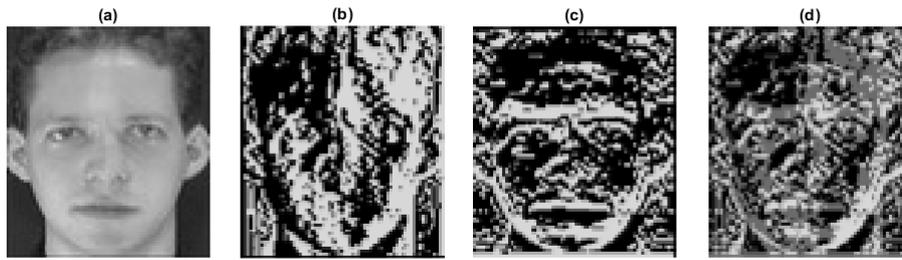


Fig. 3: Example of OME applied to an image of the ORL database: (a) original image, (b) x -direction encoded image $A^{(x)}$, (c) y -direction encoded image $A^{(y)}$ and (d) mean encoded image \bar{A}

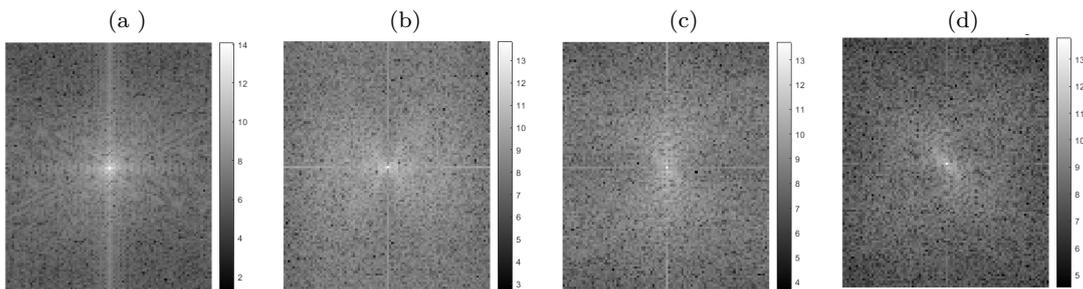


Fig. 4: Evidence of high-pass nonlinear filtering with the OME: (a) spectrum of the original image, (b) spectrum of the x -direction encoded image, (c) spectrum of the y -direction encoded image, (d) spectrum of the mean encoded image \bar{A}

Step 3: Compute the scatter inter-class matrix S_b and the scatter intra-class matrix S_w as

$$S_b = \sum_{i=1}^C D \cdot (\mu_i - \mu)(\mu_i - \mu)^T \quad (7)$$

$$S_w = \sum_{i=1}^C \sum_{j=1}^D (\mathcal{I}_{ij} - \mu_i)(\mathcal{I}_{ij} - \mu_i)^T \quad (8)$$

Step 4: Compute the eigenvalues of $S_w^{-1}S_b$ and derive the projection matrix $\vartheta = (\theta_0, \theta_1, \dots, \theta_{L-1})$, where θ_l are eigenvectors corresponding to the above eigenvalues, with $L \leq M$.

Indeed, the objective of LDA is to find the projection matrix ϑ that maximizes the ratio

$$J(\vartheta) = \frac{|\vartheta^T S_b \vartheta|}{|\vartheta^T S_w \vartheta|} \quad (9)$$

Such an optimal solution is obtained by considering the matrix $\vartheta = (\theta_0, \theta_1, \dots, \theta_{L-1})$ of eigenvectors θ_l corresponding to the L largest eigenvalues of $S_w^{-1}S_b$, with $0 \leq l < L$ and $L \in \mathbb{N}_{\geq 1}$. L is the dimension of the projected subspace. Thus, LDA consists of projecting the image \mathcal{I}_{ij} onto an optimal discriminant subspace

to obtain the projected image ξ_{ij} used for classification such that

$$\xi_{ij} = \vartheta^T (\mathcal{I}_{ij} - \mu) \quad (10)$$

$\xi_{ij} = (\varepsilon_{m,n})$, $0 \leq m < M$ and $0 \leq n < L$, is the projection of \mathcal{I}_{ij} in the space of ϑ .

3.2 Modified LDA for image feature extraction

In this section, we propose the modified LDA (MLDA) as a method to improve the extraction of image features. MLDA is based on the combination of LDA with OME techniques. We suggest increasing the distance between means of the various image classes by enhancing the image contours, and hence discriminating images on the basis of their details. For this purpose, we apply the OME prior to images and then consider only inter-class scatter matrices, instead of both inter-class and intra-class as in the LDA approach. Thus, the features obtained in that case are optimal projections of OMEIs, where the projection matrix ϑ is derived from the inter-class scatter matrix. The corresponding steps of the MLDA are listed below:

Step 1: Apply the OME transform to the database images,

$$\text{OME}(\{\mathcal{I}_{ij}\}) = \{A_{ij}\}; 1 \leq i \leq C, \quad 1 \leq j \leq D,$$

Step 2: Compute the mean of encoded images

$$\mu = \frac{1}{C \cdot D} \sum_{i=1}^C \sum_{j=1}^D A_{ij} \quad (11)$$

Step 3: Compute the mean of each class of encoded face images

$$\mu_i = \frac{1}{D} \sum_{j=1}^D A_{ij} \quad (12)$$

Step 4: Calculate the inter-class dispersion matrix,

$$S_b = \sum_{i=1}^C D \cdot (\mu_i - \mu)(\mu_i - \mu)^T \quad (13)$$

Step 5: Compute the eigenvalues of S_b and derive the projection matrix $\vartheta = (\theta_0, \theta_1, \dots, \theta_{L-1})$.

Depending on the the OMEI used in the MLDA algorithm, we can define the x-directed MLDA (xMLDA) based on $\Lambda^{(x)}$, the y-directed MLDA (yMLDA) based on $\Lambda^{(y)}$, and the mean MLDA (mMLDA) based on $\bar{\Lambda}$.

3.3 MLDA based image classification

We used the Support Vector Machine (SVM) as a classifier as it offers high adaptability, robustness to noisy data, built-in regularization, some interpretability, and the ability to achieve good performance in many classification and regression tasks. SVM features are projected images converted into row vectors as

$$\mathbf{F}_{ij} = (\varepsilon_{0,0}, \dots, \varepsilon_{0,L-1}, \dots, \varepsilon_{M-1,0}, \dots, \varepsilon_{M-1,L-1}). \quad (14)$$

Each image is thus represented by a vector in the feature space, thereby each dimension corresponds to a specific feature. Let us consider for example the inter-class matrix below defined as

$$S_b = \begin{pmatrix} 36 & 17 & 12 & 168 \\ 6 & 163 & 110 & 172 \\ 1 & 17 & 18 & 172 \\ 0 & 4 & 3 & 28 \end{pmatrix}.$$

The corresponding matrix of eigenvectors ϑ^0 is

$$\vartheta^0 = \begin{pmatrix} -0.5989 & -0.9242 & -0.9242 & 0.7954 \\ -0.7862 & 0.0845 & 0.0845 & -0.3299 \\ -0.1496 & 0.0434 & 0.0434 & 0.5065 \\ -0.0284 & -0.0310 & -0.0310 & -0.0438 \end{pmatrix}.$$

Using, for example, Fisher's criterion that consists of selecting $L < M$ eigenvectors corresponding to the most significant eigenvalues ranged into descending order, we obtain the following projection matrix for $L = 2$

$$\vartheta = \begin{pmatrix} -0.5989 & -0.9242 \\ -0.7862 & 0.0845 \\ -0.1496 & 0.0434 \\ -0.0284 & -0.0310 \end{pmatrix}$$

The corresponding projection of matrix A given in section 2 is $\xi = \vartheta^T \times A$ such that

$$\xi = \begin{pmatrix} -7.9281 & -4.7740 & -4.7189 & -7.4798 \\ -0.1428 & -1.4771 & -0.6172 & -3.2929 \end{pmatrix}$$

This matrix is then transformed into a single-row vector \mathbf{F} to serve as input for SVM such that

$$\mathbf{F} = (\xi_1, \xi_2)$$

where $\xi_1 = (-7.9281, -4.7740, -4.7189, -7.4798)$ and $\xi_2 = (-0.1428, -1.4771, -0.6172, -3.2929)$. To perform the classification, the image database is divided into the training set $\mathcal{X}_a = \{\mathcal{I}_{ij}^a\}$ and test set $\mathcal{X}_b = \{\mathcal{I}_{kl}^b\}$, with $C_a = C_b = C$, $D_a + D_b = D$, ($D_a < D$), $1 \leq i, k \leq C$, $1 \leq j \leq D_a$, and $1 \leq l \leq D_b$. The projection matrix ϑ of the MLDA is obtained by applying the OME exclusively to the training set \mathcal{X}_a . Thereafter, all the images of \mathcal{X}_a are projected to construct the feature matrix

$$\mathcal{M} = (\mathbf{F}_{i,j}). \quad (15)$$

\mathcal{M} is used to build the SVM model. Each line of \mathcal{M} is obtained as described in Eq. (14), i.e. the projected image is transformed into a row vector. The SVM algorithm aims to identify the optimal hyperplane that separates the different classes within the feature space [20]. The constructed SVM model is then used to classify \mathcal{X}_b images based on the learned features. Fig.5 shows the flow chart that outlines the proposed SVM-MLDA image classification method.

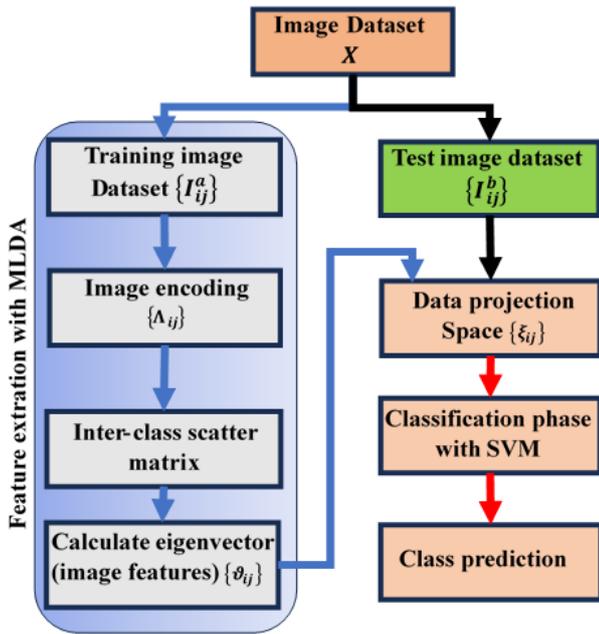


Fig. 5: Flowchart of the proposed SVM-MLDA based classification method

4 Results and Discussion

To evaluate the efficiency of our proposed classification approach, we used the ORL and FERET databases [21, 22]. Metrics such as accuracy (β), receiver operating characteristic (ROC) plots, and confusion matrix are considered.

4.1 Application to ORL database

The ORL database is one of the most used for the evaluation of facial recognition algorithms [21]. This facial database contains 400 images of 40 individuals, with 10 distinct images per subject. The images of the ORL database were captured between April 1992 and April 1994 in a laboratory setting, showcasing a variety of conditions such as lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (with/without glasses). All images were taken against a dark and uniform background, with subjects in a vertical and frontal position, allowing for some lateral movement. The dimension of the database image is 112×92 pixels, with a color depth of 256 gray levels per pixel [21]. As each individual is considered to be a class, there are 10 images per class and 40 distinct classes.

We divided the database into training sets and test sets by randomly selecting images from each individual. We repeated the experiment 20 times and evaluated the

average accuracy. We first considered the MLDA with encoded images along the x-direction and y-direction, with 80% images per individual for training and 20% for test. The corresponding average accuracy values for 20 experiments are $\beta = 97.88\%$ for the x-directed MLDA, and $\beta = 98.79\%$ for the y-directed MLDA. These results suggest that the MLDA performs better in the y-direction than in the x-direction. This observation is confirmed by the computed ROC area under the curve (AUC) as $\alpha = 0.99$ for the x-directed curve and $\alpha = 1$ for the y-directed curve. In Fig.3, the contours of the mean OMEI were better observed than those of the directed OMEIs. This observation suggests that the nondirected MLDA may perform better than the y-directed MLDA. Therefore, we repeated the experiment above with the mean MLDA. The corresponding average accuracy is $\beta = 99.07\%$, which confirms our assumption. The corresponding ROC curve with $\alpha = 1$ is shown in Fig.6. There are experiments for which the 120 test images were all well classified. We used the mean MLDA for the rest of the experiments in the paper.

We considered different values of P and Q and compared the results obtained. Table 2 shows that $P = Q = 3$ achieves better results than other combinations considered in comparison.

Table 2: Average classification accuracy $\beta(\%)$ obtained with 10 experiments and different combinations of the pair (P, Q) , $s = 8$.

(P, Q)	$(\beta \pm \sigma)\%$
(2, 2)	97.50 ± 1.76
(2, 3)	98.75 ± 2.14
(3, 2)	98.86 ± 1.84
(3, 3)	99.07 ± 1.01

Table 3: Average classification accuracy $\beta(\%)$ obtained with 20 experiments and various sizes s of training sets.

s	DLDA	CLDA	CDLDA/New	mMLDA
3	83.7	89.4	89.4	93.86
4	89.7	94.3	94.3	96.29
5	92.1	97.1	97.1	97.39

We compared our proposed MLDA with other existing classification algorithms such as the DLDA method [27], the CLDA method [23], and CDLDA/New [6]. For this purpose, we randomly selected $s = 3, 4, 5$ images of an individual over 10 for training and the rest for validation, and repeated the experiment 20 times. The average classification accuracy values are given in Table 3.

Table 4: Comparison of the average recognition accuracy $\beta(\%)$ and standard deviation $\sigma(\%)$ with other methods in ORL database.

Method	$(\beta \pm \sigma)\%$					
	$s = 3$	$s = 4$	$s = 5$	$s = 6$	$s = 7$	$s = 8$
EnFFG-2DLDA [13]	93.71 ± 1.92	96.22 ± 1.32	98.25 ± 0.41	99.03 ± 1.32	98.92 ± 0.66	-
G-2DFLD [23]	92.82 ± 2.67	95.94 ± 1.21	97.68 ± 0.91	98.72 ± 0.92	98.42 ± 1.11	-
FG-2DLDA [24]	93.41 ± 1.11	96.07 ± 1.59	98.00 ± 1.07	98.78 ± 0.72	98.50 ± 0.74	-
WFG-2DFLD [25]	93.25 ± 2.30	96.16 ± 1.33	98.00 ± 1.10	98.91 ± 0.64	98.82 ± 1.06	-
FGD-2DIFDA [26]	93.54 ± 2.60	96.03 ± 1.63	98.12 ± 0.91	98.88 ± 0.72	98.74 ± 0.57	-
LDA	84.14 ± 2.56	93.07 ± 1.59	93.26 ± 2.30	94.50 ± 0.91	95.17 ± 1.14	95.25 ± 1.92
mMLDA	93.86 ± 1.94	96.29 ± 1.71	97.39 ± 1.79	98.59 ± 0.78	98.86 ± 1.81	99.07 ± 1.01

We extended the comparison to other methods that used more than $s = 5$ images per class for training. The corresponding results are summarized in Table 4. We observe that the accuracy of the MLDA increases uniformly with the training size s , suggesting better performance. A good classification approach performs better as the training set is large. This condition is not verified for the other methods listed in Table 4. They cannot guarantee a good performance by increasing the size of the training set.

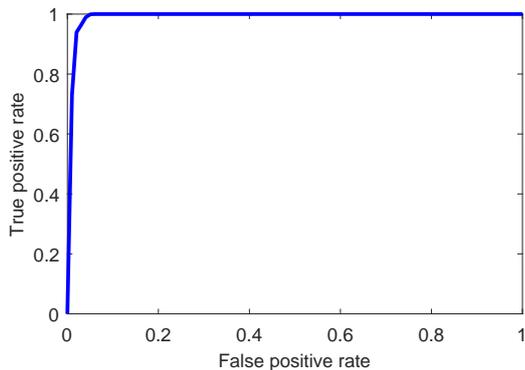


Fig. 6: ROC curve of the [80, 20] mean MLDA classification model

4.2 FERET database

The Facial Recognition Technology (FERET) database contains a total of 14,126 images belonging to 1,199 individuals, as well as 365 sets of duplicate images taken on different days [22]. The images were collected between December 1993 and August 1996. In this study, we will consider 1400 images of 200 individuals, which means 7 images per individual [22]. Depending on the location of the eyes, the original images of the database

Table 5: Evaluation of average recognition accuracy $\beta(\%)$ and standard deviation $\sigma(\%)$ with some classical methods on the FERET face database.

Method	$(\beta \pm \sigma)\%$		
	$s = 2$	$s = 3$	$s = 4$
EnFFG-2DLDA	49.73 ± 0.97	59.69 ± 1.76	66.04 ± 2.18
G-2DFLD	48.51 ± 1.2	56.77 ± 2.11	64.49 ± 2.01
FG-2DLDA	49.05 ± 1.01	58.81 ± 1.13	65.51 ± 2.26
WFG-2DFLD	49.46 ± 1.52	59.41 ± 1.11	65.88 ± 2.62
FGD-2DIFDA	49.65 ± 0.70	58.75 ± 1.29	65.75 ± 2.47
LDA	49.08 ± 0.96	53.85 ± 1.26	70.50 ± 0.47
mMLDA	54.75 ± 1.46	62.29 ± 2.16	73.61 ± 1.76

were cropped and resized to 80×80 pixels as described in [13].

Furthermore, the performance of our method was evaluated by selecting 20 different pairs of training sets ($s = 2, 3, 4$) and test sets ($7-s$) in the FERET database. A comparison between our method and other methods is summarized in Table 5. We observe that mMLDA once more performs better than the other methods in comparison. In addition, the accuracy of the mMLDA increases uniformly with training size s , which confirms its better performance as the training size increases. We computed the AUC of the ROC curve of the [70, 30] mMLDA model as $\alpha = 0.97$, which further confirms its good performance as a classification approach. We also compared with the XCEPTION CNN deep learning architecture in [28], and the combination of neural networks with a genetic algorithm presented in [29], where few image classes of the FERET database are considered, with $s = 4$. Both methods perform, respectively, 96.73% with only 16 classes and 94% with 50 classes, while MLDA performs an average accuracy of 98.48% and 94.08% over 20 epochs in both cases, respectively.

The above results obtained show that our OME based feature extraction method improves image classification. Nonlinear enhancement of the image contours allows one to achieve higher accuracy values. In order to determine the importance of OME, we also considered

raw images of the ORL database, with 80% images per individual for training and 20% for tests. By computing the average accuracy of LDA combined with SVM (SVM-LDA), we found $\beta = 95.25 \pm 1.92\%$. Similarly, the average accuracy of SVM-LDA without considering intraclass matrices S_w has given $\beta = 94.50 \pm 2.19\%$, which is much lower than $\beta = 99.07 \pm 1.01\%$ obtained with encoded images. We did the same experiment with the entire FERET database (200 classes) in case $s = 4$ and achieved a classification accuracy of $\beta = 70.50\%$ with LDA, compared to $\beta = 73.61\%$ for MLDA. Therefore, these results confirm the efficiency of OME in the MLDA approach for feature extraction.

Table 6: Evaluation of average recognition accuracy $\beta(\%)$ and standard deviation $\sigma(\%)$ with some classical methods on the FEI database.

Method	$\beta(\%)$		
	$s = 7$	$s = 8$	$s = 11$
SESRC&LDF [30]	93.67	–	–
DCT-VQ [31]	–	98.00	–
2D-DMWT [32]	–	–	97.447
G-F-LDA [33]	97.84	98.20	98.65
LDA	83.33	94.08	96.67
mMLDA	96.89	98.25	98.78

Additional tests were performed using the FEI face database that has been widely used in the literature to demonstrate the efficiency of MLDA. The database contains 14 face images for each of the 200 distinct individuals, therefore, a total of 2800 images. All images are colorful and taken against a homogeneous white background in an upright frontal position with a profile rotation of up to about 180 degrees [34]. The results obtained are summarized in Table 6 where the high performance of MLDA is observed compared to the results in [30–33].

We also combined MLDA with random forest (RF) and K-nearest neighbors (KNN) as classifier and compared the results obtained with SVM. Table 7 shows that SVM provides high classification rates, compared to RF and KNN. The results are obtained by training and validating simultaneously with SVM, RF, and KNN in the ORL database over 10 epochs, for various training sizes s .

4.3 Speed performance

We designed and executed our classification model on a Windows operating system running MATLAB 2018b. The computer is equipped with an AMD Ryzen Core

Table 7: Comparison of SVM, RF and KNN average classification accuracy $\beta(\%)$, obtained with 10 experiments and various sizes s of training sets

Method	$\beta(\%)$				
	$s = 4$	$s = 5$	$s = 6$	$s = 7$	$s = 8$
SVM	96.25	97.02	98.45	98.78	100
RF	88.75	91.25	93.75	91.25	97.50
KNN	83.75	88.75	86.25	95.00	91.25

i7 processor (2.10 GHz) and 12.00 GB RAM. While running our codes, the average execution time for ORL database is 13.53 minutes per epoch for the MLDA, and 6.8608 minutes for the plain LDA. Although the MLDA computational time is two times greater than the LDA time, it could be further reduced by optimizing the MLDA code, which is not yet the case for the code used. We also evaluated the useful memory for the two methods and found 60.1 MB for plain LDA and 56.71 MB for MLDA.

5 Conclusion

This paper presented the Modified Linear Discriminant Analysis (MLDA) approach, which enhances image feature extraction by combining Ordinal Matix Encoding with Linear Discriminant Analysis (LDA). We also developed a classification method that integrates MLDA with Support Vector Machines (SVM). Our results, tested on the ORL and FERET databases, demonstrate that the proposed approach performs well compared to existing classification methods. This strong performance confirms MLDA’s efficiency in feature extraction. However, a key limitation of MLDA is its computational time, which is currently twice that of LDA. This aspect needs further improvement. We intend to apply MLDA to real-time facial image recognition in video surveillance applications. We will also further investigate the impact of the window size on the performance of MLDA.

Declarations

- Conflict of interest/Competing interests: The authors have no competing interests to declare that are relevant to the content of this article.
- Code availability: The code used in the paper could be available upon reasonable request.
- Author contribution: All authors contributed to the conception and design of the study, material preparation, data collection and analysis. They all read and approved the manuscript.

References

1. Ouyang, D., Zhang, Y. & Shao, J. Video-based person re-identification via spatio-temporal attentional and two-stream fusion convolutional networks. *Pattern Recogn Lett* **117**, 153–160 (2019).
2. Ali, W., Tian, W., Din, S., Iradukunda, D. & Khan, A. A. Classical and modern face recognition approaches: a complete review. *Multimed Tools Appl* **80**, 4825–4880 (2021).
3. Modhej, N., Bastanfard, A., Teshnehlav, M. & Raiesdana, S. Pattern separation network based on the hippocampus activity for handwritten recognition. *IEEE Access* **8**, 212803–212817 (2020).
4. Mahmood, A., Uzair, M. & Al-maadeed, S. Multi-order statistical descriptors for real-time face recognition and object classification. *IEEE Access* **6**, 12993–13004 (2018).
5. Liang, X. & Lin, Y. Enhanced parameter-free diversity discriminant preserving projections for face recognition. *International Journal of Pattern Recognition and Artificial Intelligence* **32**, 1856013 (2018).
6. Lu, G.-F., Zou, J. & Wang, Y. Incremental complete LDA for face recognition. *Pattern Recognition* **45**, 2510–2521 (2012).
7. Chen, L.-F., Liao, H.-Y. M. & Ko, a., Ming-Tat. A new LDA-based face recognition system which can solve the small sample size problem. *Pattern recognition* **33**, 1713–1726 (2000).
8. Peng, P., Portugal, I., Alencar, P. & Donald, C. A face recognition software framework based on principal component analysis. *Plos one* **16**, e0254965 (2021).
9. Garain, J., Kumar, R. K., Kisku, D. R. & Sanyal, G. Addressing facial dynamics using k-medoids cohort selection algorithm for face recognition. *Multimed Tools Appl* **78**, 18443–18474 (2019).
10. Pacheco, S. A. B., Estrada, J. E. & Goyani, M. M. Arai, K. (ed.) *Robust face recognition under adversarial attack using sargan model and improved cross triple mobilenetv1*. (ed. Arai, K.) *Advances in Information and Communication*, 491–510 (Springer Nature Switzerland, Cham, 2025).
11. Trivedi, H. & and, M. G. Adapting face recognition to the masked world: leveraging deep attention networks with a custom dataset. *International Journal of Computers and Applications* **46**, 427–441 (2024).
12. Mu, Z., Hu, J. & Min, J. EEG-based person authentication using a fuzzy entropy-related approach with two electrodes. *Entropy* **18**, 432 (2016).
13. Ghosh, M. & Dey, A. Fractional-weighted entropy-based fuzzy G-2DLDA algorithm: a new facial feature extraction method. *Multimedia Tools and Applications* **82**, 2689–2707 (2023).
14. Dey, A., Sing, J. & Chowdhury, S. Weighted fuzzy generalized 2DFLD: a fuzzy-based feature extraction technique for face recognition. *International Journal of Machine Learning and Computing* **7**, 223–231 (2017).
15. Keller, K., Unakafov, A. M. & Unakafova, V. A. Ordinal patterns, entropy, and EEG. *Entropy* **16**, 6212–6239 (2014).
16. Eyebe Fouda, J. S. A., Koepf, W., Marwan, N., Kurths, J. & Penzel, T. Complexity from ordinal pattern positioned slopes (COPPS). *Chaos, Solitons & Fractals* **181**, 114708 (2024).
17. Pessa, A. A. B. & Ribeiro, H. V. ordpy: A Python package for data analysis with permutation entropy and ordinal network methods. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **31**, 063110 (2021).
18. Gu, S. & Timofte, R. A brief review of image denoising algorithms and beyond. *Inpainting and Denoising Challenges* 1–21 (2019).
19. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **86**, 2278–2324 (1998).
20. Hsu, C.-W. & Lin, C.-J. A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks* **13**, 415 – 425 (2002).
21. Samaria, F. & Harter, A. *Parameterisation of a stochastic model for human face identification*, 138–142 (1994).
22. Phillips, P. J., Moon, H., Rizvi, S. A. & Rauss, P. J. The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**, 1090–1104 (2000).
23. Chowdhury, S., Sing, J., Basu, D. & M., N. Face recognition by generalized two-dimensional fld method and multi-class support vector machines. *Appl. Soft Comput* **11**, 4282–429 (2011).
24. Dey, A. & Ghosh, M. A novel approach to fuzzy-based facial feature extraction and face recognition. *Informatica Si* **43**, 535–543 (2019).
25. Dey, A., Sing, J. K. & Chowdhury, S. Weighted fuzzy generalized 2DFLD: a fuzzy-based feature extraction technique for face recognition. *International Journal of Machine Learning and Computing* **7**, 223–231 (2017).
26. Dey, A., Chowdhury, S. & Sing, J. K. Bhattacharyya, S., Chaki, N., Konar, D., Chakraborty, U. K. & Singh, C. T. (eds) *Feature extraction using fuzzy generalized two-dimensional inverse LDA with gaussian probabilistic distribution and face recognition*. (eds Bhattacharyya, S., Chaki, N., Konar, D., Chakraborty, U. K. & Singh, C. T.) *Advanced Computational and Communication Paradigms*, 553–561 (2018).
27. Yang, J. & Yang, J. Why can LDA be performed in PCA transformed space? *Pattern Recognition* **36**, 563–566 (2003).
28. Lal, P. V., Srilakshmi, U. & Venkateswarlu, D. Face recognition using deep learning xception cnn method. *Journal of Theoretical and Applied Information Technology* **100**, 531–542 (2022).
29. Kamyab, T., Daealhaq, H., Ghahfarokhi, A. M., Behshetinejad, F. & Salajegheh, E. Combination of genetic algorithm and neural network to select facial features in face recognition technique. *International Journal of Robotics and Control Systems* **3**, 50–58 (2023).
30. Liao, M. & Gu, X. Face recognition approach by subspace extended sparse representation and discriminative feature learning. *Neurocomputing* **373**, 35–49 (2020).
31. Aldhahab, A., Al Obaidi, T. & Mikhael, W. B. *Employing vector quantization algorithm in a transform domain for facial recognition*, 1–4 (IEEE, 2016).
32. Aldhahab, A. & Mikhael, W. B. *A facial recognition method based on dmw transformed partitioned images*, 1352–1355 (IEEE, 2017).
33. Hassan, M. M., Hussein, H. I., Eesa, A. S. & Mstafa, R. J. Face recognition based on gabor feature extraction followed by fastica and LDA. *Computers, Materials and Continua* **68**, 1637–1659 (2021).
34. Thomaz, C. E. & Giraldo, G. A. A new ranking method for principal components analysis and its application to face image analysis. *Image and Vision Computing* **28**, 902–913 (2010).