handwritten

by Handwritten Handwritten

Handwritten Capital Letter Recognition Using Symmetric Two-Dimensional Linear Discriminant Analysis

Ismail Husein[1], **Rina Widyasari**[1]
[1]Universitas Islam Negeri Sumatera Utara Medan

Article Info

Article history:

Received mm dd, yyyy
Revised mm dd, yyyy
Accepted mm dd, yyyy

Keywords:

Capital Letter
Recognition Using
DLDA

ABSTRACT (10 PT)

Statistical pattern recognition is the process of using statistical techniques to obtain information and make informed decisions based on data measurements. It is possible to solve the doubt inherent in the objective function of the 2-Dimension Linear Discriminant Analysis by employing the symmetrical 2-Dimension Linear Discriminant Analysis approach. Symmetrical 2-dimensional linear discriminant analysis has found widespread use as a method of introducing handwritten capital letters. Symmetric 2-DLDA, according to Symmetric 2-DLDA, produces better and more accurate results than Symmetric 2-DLDA.

Corresponding Author:

Ismail Husein
Department of Mathematics
Universitas Islam Negeri Sumatera Utara Medan
Email: husein_ismail@uinsu.ac.id

1. INTRODUCTION (10 PT)

Pattern recognition is the process of recognizing patterns regularities in data by using automated algorithm. The process in pattern recognition systems starts from the selection of patterns as sensors, then the patterns are entered into processing techniques, representation charts, and finally, the process of modelling decision making (Theodoris *et al*., 2003).

The best approach used for pattern recognition in this study is statistical classification (Jain *et al*., 2002). Various algorithms can be applied for pattern recognition. One of them is Linear Discriminant Analysis, hereinafter abbreviated as writer with LDA, has been successfully applied in computer visualization. As a subspace investigation way to deal with study the low-dimension structure of high-dimension data, LDA is probing a set of vectors that maximize the Fisher Discriminant Criterion. This method concurrently minimizes the distribution in class (S_w) and maximizes the distribution between classes (S_b) in the vector space character projection (Sharma and Paliwal, 2015).

In 2-Dimensional matrices such as images, and in general the image is not symmetrical $X_i \neq X_i^T$, then the distribution matrix between classes and distribution matrices in a class is defined not single:

$$S_b(XX^T) \neq S_b(X^T X), \quad S_w(XX^T) \neq S_w(X^T X),$$

so there are a number of possible choices for determining the appropriate objective function of LDA (Tharwart, *et al*., 2016). The introduction of handwritten capital letters is one of the applications of symmetrical 2-dimension linear discriminant analysis.

 ISSN: 2476-9843

2. RESEARCH METHOD

Linear Discriminant Analysis

In the classification process, first measure the observational characteristics of the sample. Extract all the information contained in the sample to calculate the sample-time value for a curve-shaped pattern, and the level of blackness of the pixels for a figure, as shown in Figure 1 (Li and Yuan, 2005).

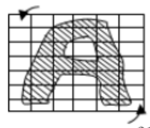

Figure 1. Example measurement of letter patterns

If a data matrix is given $H \in R^{N \times n}$, LDA method intents to encounter a transformation $K \in R^{N \times l}$ that assigns each a_i column from matrix H, for $1 \le i \le p$, in the P dimension space to the bi vector in the dimension l space. Namely K : $h_i \in R^{N \times n} \rightarrow \mathbf{b}_i = K^T h_i \in R^l (l < P)$. In other words, LDA intents to encounter a vector space **K** spanned by $\{k_i\}_{i=1}^{l}$ where $K = [k_1, k_2, ..., k_l]$, so that each h_i is projected to **K** by $(k_1^T.h_i, ..., k_l^T.h_i)^T \in R^l$ (Duda, *et al.*, 2012).

Consider that the initial data in H is subdivided into k classes so that $H = \{\prod_1, \prod_2, ..., \prod_k\}$, where $\prod_i$ loads n_i the data points of class $-i$ and $\sum_{i=1}^{k} p_i = p$. The classic LDA intents to encounter the optimal transformation of G so that the class structure of the initial high-dimensional space data is converted into low-dimensional space (Xiong *et al.*, 2004).
In LDA, the transformation to subspace with lower dimension is

$$y_i = K^T x_i \quad \text{(Luo et al., 2007)} \qquad [1]$$

where K is the transformation to a subspace. Usually also written with $(y_1, ..., y_p) = K^T (x_1, ..., x_p)$ atau $Y = K^T X$.. The main purpose of LDA is to find the value of K so that classes can be more detached in the transformation space and can conveniently be identified from the others.
In the Linear Discriminant Analysis method, there are two distribution matrices, namely the in-class distribution matrix symbolized by S_w, and the inter-class distribution matrix symbolized by S_b, each construed as follows:

$$S_w = \sum_{i=1}^{c} \sum_{x_k \in \prod_i} [x_k - m_i][x_k - m_i]^T \qquad [2]$$

$$S_b = \sum_{i=1}^{c} n_i [m_i - m][m_i - m]^T \qquad [3]$$

Where p_i is the number of fragment in the class x_i, and m_i is the average image of $-i$ and m is the overall average (Viszlay, *et al.*, 2014). The class average formula and the overall average is as follows:

$m_i = \frac{1}{n_i} \sum_{x \in \prod_i} x$ is the average of the i-class, and $m = \frac{1}{n} \sum_{i=1}^{k} \sum_{x \in \prod_i} x$ is the overall average (Fukunaga, 1990). General optimizations in Linear Discriminant Analysis include (Fukunaga, 1990):

$$\max_K J(K) = \text{tr} \frac{S_b(Y)}{S_w(Y)} = \text{tr} \frac{K^T S_b(X) K}{K^T S_w(X) K} \qquad [4]$$

3. RESULTS AND ANALYSIS

3.1 In 2-Dimensional Linear Discriminant Analysis

The preeminent discrepancy between the classic LDA and 2-DLDA that the researchers propose in this study is about data representation. Classical LDA uses vector representations, whereas 2-DLDA works with data in matrix representations. In using the 2-DLDA method, it will be seen that the representation leads to eigen-decomposition of the matrix with a smaller size. More specifically, 2-DLDA involves eigen-decomposition matrices of r × r and c × c sizes, which are much smaller than the classical LDA matrices (Yang and Yu, 2014).

In 2-DLDA it has been agreed that a set of images is symbolized by $X = (X_1, X_2, ..., X_n)$, $X_i \in \Re^{r \times c}$. With the same clairvoyance as the classical LDA, 2-DLDA tries to find a linear transformation

$$\mathbf{Y_i = L^T X_i R} \qquad [5]$$

so the different classes are detached.

Suppose that $M_i = \frac{1}{n_i}\sum_{x \in \Pi_i} X$ is the average of the i-th class, $1 \le i \le k$, and $M = \frac{1}{n}\sum_{i=1}^{k}\sum_{x \in \Pi_i} X$ means the overall average (Stang, 2003). In 2-DLDA, researchers regard images as two-dimensional signals and intend to detect two transformation matrices $L \in \Re^{r \times l_1}$ and $R \in \Re^{c \times l_2}$ and assigns each H_i member for $1 \le x \le n$, to a B_i matrix so that B_i = LT H_i R.

Similarly to classic LDA, 2-DLDA intends to encounter optimal L and R transformations (projections) so the class structure of the initial high-dimensional space is converted to a low-dimensional space. An innate metric affinity between matrices is the Frobenius norm (Viszlay, *et al.*, 2014). The square of the distance from within-class and between classes can be calculated as below:

$$D_w = \sum_{i=1}^{k}\sum_{x \in \Pi_i} \|X - M_i\|_F^2, \qquad D_b = \sum_{i=1}^{k} n_i \|M_i - M\|_F^2 \qquad [6]$$

trace $(M M^T)$= $\|M\|_F^2$, for a matrix *M*, so obtained :

$$D_w = trace\left(\sum_{i=1}^{k}\sum_{x \in \Pi_i} \|X - M_i\|_F^2\right) \qquad [7]$$

$$D_b = trace\left(\sum_{i=1}^{k}\sum_{x \in \Pi_i} \|X - M_i\|_F^2\right) \qquad [8]$$

In low-dimensional space, the result of linear transformations L and R, the distance between classes and *between-classes* becomes:

$$\bar{D}_w = trace\left(\sum_{i=1}^{k}\sum_{x \in \Pi_i} L^T (X - M_i) RR^T (X - M_i)^T L\right) \qquad [9]$$

$$\bar{D}_b = trace\left(\sum_{i=1}^{k} n_i L^T (X - M_i) RR^T (X - M_i)^T L\right) \qquad [10]$$

The optimum of transformation of L and R will maximize $\bar{D}_b$ and minimize $\bar{D}_w$, because of the difficulty of calculating the optimum of L and R concurrently, the following is the algorithm for 2-DLDA. More particularly, for a settled R, we can calculate the optimum of L by determining the same optimization problem with equation (12). By calculating L, we can then amend R by determining another optimization problem as the only solution in equation (6).

 ISSN: 2476-9843

L calculation

For a settled R, $\bar{D}_w$ and $\bar{D}_b$ can be rephrased as

$$\bar{D}_w = trace\left(L^T S_w^R L\right), \bar{D}_b = trace\left(L^T S_b^R L\right) \quad [12]$$

where

$$S_w^R = \sum_{i=1}^{k} \sum_{X \in \Pi_i} (\mathbf{X} - M_i) R R^T (\mathbf{X} - M_i)^T, \quad S_b^R = \sum_{i=1}^{k} n_i (M_i - M) R R^T (M_i - M)^T \quad [13]$$

optimal L can be calculated by figuring out the succeeding optimization problem: max_L *trace* $((L^T S_w^R L)^{-1} (L^T S_b^R L))$. The solution can be procured by clarifying the problem of generalizing the following eigenvalues: $S_w^R x = \lambda S_b^R x$. Because S_w^R in general it is nonsingular, the optimum L can be obtained by calculating an eigen-decomposition on $(S_w^R)^{-1} S_b^R$. Remark that the size of the matrices S_w^R and S_b^R are $r \times r$ (square matrix), which is smaller than the size of the matrices S_w and S_b in the classic LDA (Zhao *et al*., 2018).

R calculation

Then calculate R for a settled L. $\bar{D}_w$ and $\bar{D}_w$ can be written back as

$$\bar{D}_w = trace\left(R^T S_w^L R\right), \bar{D}_b = trace\left(R^T S_b^L R\right) \quad [15]$$

where

$$S_w^L = \sum_{i=1}^{k} \sum_{X \in \Pi_i} (X - M_i)^T L L^T (X - M_i), \quad S_b^L = \sum_{i=1}^{k} n_i (M_i - M)^T L L^T (M_i - M) \quad [16]$$

Optimal L can be calculated by clarifying the succeeding optimization problem: max_R *trace* $((R^T S_w^L R)^{-1} (R^T S_b^L R))$. The solution can be obtained by solving the problem of generalizing the following eigenvalues: $S_w^L x = \lambda S_b^L x$. Because in general it is nonsingular, the optimum R can be obtained by calculating an eigen-decomposition on $(S_w^R)^{-1} S_b^R$. Remark that the size of the matrices S_w^L and S_b^L are $r \times r$ (square matrix) (Zhao *et al*., 2018).

3.2 Symmetrical 2-Dimension Linear Discriminant Analysis (Symmetrical 2-DLDA)

It has been stated in the previous chapter that the classification approach with 2-Dimension Linear Discriminant Analysis (2 DLDA) raises a fundamental problem of doubt: There are two ways to delineate in-class distribution matrices S_w

$$S_w(\mathbf{X}\mathbf{X}^\mathbf{T}) = \sum_{j=1}^{k} \sum_{X_i \in \pi_j} (X_i - M_j)(X_i - M_j)^T$$

$$S_w(\mathbf{X}^\mathbf{T}\mathbf{X}) = \sum_{j=1}^{k} \sum_{x_i \in \pi_j} (X_i - M_j)^T (X_i - M_j)$$

and there are two ways to delineate the distribution matrix between classes S_b

$$S_b(\mathbf{XX^T}) = \sum_{j=1}^{k} n_j (M_j - M)(M_j - M)^T$$

$$S_b(\mathbf{XX^T}) = \sum_{j=1}^{k} n_j (M_j - M)^T (M_j - M)$$

Consequently, in the space of transformation can be written

$$S_b(\mathbf{YY^T}),\ S_b(\mathbf{Y^TY}),$$

$$S_w(\mathbf{YY^T}),\ S_w(\mathbf{Y^TY}),$$

In general, images are not symmetrical $X_i \neq X_i^T$, then

$$S_b(\mathbf{YY^T}) \neq S_b(\mathbf{Y^TY}),$$

$$S_w(\mathbf{YY^T}) \neq S_w(\mathbf{Y^TY}),$$

For this argument, the objective function of LDA is dubious, which raises a number of choices:

$$J_1 = \mathrm{tr}\frac{S_b(\mathbf{YY^T})}{S_w(\mathbf{YY^T})}$$

$$J_2 = \mathrm{tr}\frac{S_b(\mathbf{Y^TY})}{S_w(\mathbf{Y^TY})}$$

$$J_3 = \mathrm{tr}\left[\frac{S_b(\mathbf{YY^T})}{S_w(\mathbf{YY^T})} + \frac{S_b(\mathbf{Y^TY})}{S_w(\mathbf{Y^TY})}\right] \qquad J_4 = \mathrm{tr}\left[\frac{S_b(\mathbf{Y^TY})}{S_w(\mathbf{Y^TY})}\frac{S_b(\mathbf{YY^T})}{S_w(\mathbf{YY^T})}\right],$$

$$J_5 = \mathrm{tr}\frac{S_b(\mathbf{YY^T}) + S_b(\mathbf{Y^TY})}{S_w(\mathbf{YY^T}) + S_w(\mathbf{Y^TY})},$$

(Luo *et al*., 2007)

Symmetrical 2-Dimensional Linear Discriminant Analysis in solving the ambiguous problem above inspired by a key observation: if the picture is symmetrical, namely $\mathbf{X_i} = \mathbf{X_i^T}$, then

$$S_w(\mathbf{XX^T}) = S_w(\mathbf{X^TX}),$$

$$S_b(\mathbf{XX^T}) = S_b(\mathbf{X^TX}).$$

The solution of this problem uses a new data representation that is symmetric linear transformation.

$$\begin{pmatrix} 0 & \mathbf{Y_i^T} \\ \mathbf{Y_i} & 0 \end{pmatrix} = \mathbf{\Gamma^T}\begin{pmatrix} 0 & \mathbf{X_i} \\ \mathbf{X_i^T} & 0 \end{pmatrix}\mathbf{\Gamma},\ \ \mathbf{\Gamma} = \begin{pmatrix} 0 & \mathbf{L} \\ \mathbf{R} & 0 \end{pmatrix}$$

Pada Fukunaga(1990), matriks didefinisikan sebagai:

In Fukunaga (1990), Γ matrix is defined as:

$$\begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & \vdots \\ 0 & \cdots & \sigma_n \end{pmatrix}$$

that is, the diagonal matrix is mainly the value of the variance of the data and other elements 0. The linear transformation above is equivalent to the linear transformation $\mathbf{Y_i} = \mathbf{L^T X_i R}$.. The explanation is as follows:

$$\begin{pmatrix} 0 & \mathbf{Y_i^T} \\ \mathbf{Y_i} & 0 \end{pmatrix} = \mathbf{\Gamma^T}\begin{pmatrix} 0 & \mathbf{X_i} \\ \mathbf{X_i^T} & 0 \end{pmatrix}\mathbf{\Gamma},\ \ \mathbf{\Gamma} = \begin{pmatrix} 0 & \mathbf{L} \\ \mathbf{R} & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & \mathbf{Y_i^T} \\ \mathbf{Y_i} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{L} \\ \mathbf{R} & 0 \end{pmatrix}^{\mathbf{T}}\begin{pmatrix} 0 & \mathbf{X_i} \\ \mathbf{X_i^T} & 0 \end{pmatrix}\begin{pmatrix} 0 & \mathbf{L} \\ \mathbf{R} & 0 \end{pmatrix}$$

 ISSN: 2476-9843

$$\begin{pmatrix} 0 & \mathbf{Y}_i^T \\ \mathbf{Y}_i & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{R}^T \\ \mathbf{L}^T & 0 \end{pmatrix} \begin{pmatrix} 0 & \mathbf{X}_i \\ \mathbf{X}_i^T & 0 \end{pmatrix} \begin{pmatrix} 0 & \mathbf{L} \\ \mathbf{R} & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & \mathbf{Y}_i^T \\ \mathbf{Y}_i & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{R}^T\mathbf{X}_i^T & 0 \\ 0 & \mathbf{L}^T\mathbf{X}_i \end{pmatrix} \begin{pmatrix} 0 & \mathbf{L} \\ \mathbf{R} & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & \mathbf{Y}_i^T \\ \mathbf{Y}_i & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{R}^T\mathbf{X}_i^T\mathbf{L} \\ \mathbf{L}^T\mathbf{X}_i\mathbf{R} & 0 \end{pmatrix}$$

obtained $\mathbf{Y}_i = \mathbf{L}^T \mathbf{X}_i \mathbf{R}$ and $\mathbf{Y}_i^T = \mathbf{R}^T \mathbf{X}_i^T \mathbf{L}$. We also have

$$\left\| \begin{pmatrix} 0 & \mathbf{X}_i^T \\ \mathbf{X}_i & 0 \end{pmatrix} - \mathbf{\Gamma} \begin{pmatrix} 0 & \mathbf{Y}_i \\ \mathbf{Y}_i^T & 0 \end{pmatrix} \mathbf{\Gamma}^T \right\|^2 = 2\left\|\mathbf{X}_i - \mathbf{L}\mathbf{Y}_i\mathbf{R}^T\right\|^2$$

Therefore, optimization using (**L**, **R**) is equivalent to optimization using $\mathbf{\Gamma}$.
Other than that, by using symmetric linear transformations produced a theorem:

Theorem 1: The single objective function of LDA for2-DLDA is

$$J_{ADL2-D} = \mathrm{tr}\frac{\mathbf{S_b(YY^T)}}{\mathbf{S_w(YY^T)}} = \mathrm{tr}\frac{\mathbf{S_b(Y^TY)}}{\mathbf{S_w(Y^TY)}} \quad [17]$$

$$J_{ADL2-D} = \mathrm{tr}\left(\frac{\mathbf{R^TS_b^LR}}{\mathbf{R^TS_w^LR}} + \frac{\mathbf{L^TS_b^RL}}{\mathbf{L^TS_w^RL}}\right) \quad [18]$$

Using theorem 1, in the case non symmetrics matrices which cause S_w *and* S_b in X space to be doubly defined it also causes S_w and S_b in Y space to be doubly defined. So

$$J_1^{'} = \mathbf{tr}\frac{S_b(YY^T)}{S_w(YY^T)} = \mathbf{tr}\frac{R^T\mathbf{S}_\mathbf{b}^L R}{R^T\mathbf{S}_\mathbf{w}^L R}$$

$$J_2^{'} = \mathbf{tr}\frac{S_b(Y^TY)}{S_w(Y^TY)} = \mathbf{tr}\frac{\mathbf{L^TS_b^RL}}{\mathbf{L^TS_w^RL}}$$

In an independent optimization approach, to get R can be done by maximizing J_1 '(reject J_2') and then obtaining L by maximizing J_2 '(rejecting J_1'). This is not consistent in optimizing the objective function, which is when maximizing J_1', J_2' has decreased and vice versa. This problem can be solved by two techniques namely first, when maximizing J_1', must calculate J_2'. But, on the other hand also need to know how to combine J_1'and J_2'. The simple combination that can be done is $J = J_1' + J_2'$, i.e. :

$$J = \mathbf{tr}\frac{\mathbf{R^TS_b^RR + L^TS_b^LL}}{\mathbf{R^TS_w^RR + L^TS_w^LL}} \quad [19]$$

(Luo et *al*., 2005).
Secondly is how to optimize the objective function. The result to maximizing $\max_R J$ can be simply done by calculating eigenvectors from $S_w^{-1}S_b$, the same calculations as the Linear Discriminant Analysis method.

However, the objective function described in equation (19) cannot be used to determine the trace of a single ratio of the two distribution matrices. This happens because the objective function cannot be solved in the same direction through eigenvector calculations (same as standard LDA). However, this can be overcome by developing an efficient algorithm using the gradient-up approach. This approach reduces objective functions. The derivative of the matrix function is done by using the basic matrix algebra contained in Fukunaga's book (1990). The results are shown in the following Lemmas:

Lemma 2: *Let* $P_L = \mathbf{L^TS_b^RL}$, $Q_L = \mathbf{L^TS_w^RL}$, $P_R = \mathbf{R^TS_b^LR}$ *. and* $Q_R = \mathbf{R^TS_b^LR}$.
Dervative of objective function $J_{ADL2\text{-}D}$ in *equation* (3.2) *as follow*

For $\frac{\partial J}{\partial R}$ *obtained*

$$\frac{\partial}{\partial R}\operatorname{tr}\frac{R^T S_b^L R}{R^T S_w^L R} = 2S_b^L R Q_R^{-1} - 2S_w^L R Q_R^{-1} P_R Q_R^{-1}$$ *and*

$$\frac{\partial}{\partial R}\operatorname{tr}\frac{L^T S_b^R L}{L^T S_w^R L} = 2\sum_{k=1}^{K}\sum_{A_i\in\pi_k}(H_i - M_k)^T L Q_L^{-1} L^T (H_i - M_k) R - 2\sum_{k=1}^{K}(M_k - M)^T L Q_L^{-1} P_L Q_L^{-1} L^T (M_k - M) R \quad (20)$$

For $\frac{\partial J}{\partial L}$ *obtained*

$$\frac{\partial}{\partial L}\operatorname{tr}\frac{L^T S_b^R L}{L^T S_w^R L} = 2S_b^R L Q_L^{-1} - 2S_w^R L Q_L^{-1} P_L Q_L^{-1}$$ *and*

$$\frac{\partial}{\partial L}\operatorname{tr}\frac{R^T S_b^L R}{R^T S_w^L R} = 2\sum_{k=1}^{K}\sum_{A_i\in\pi_k}(H_i - M_k) R Q_L^{-1} R^T (H_i - M_k)^T L - 2\sum_{k=1}^{K}(M_k - M) R Q_R^{-1} P_R Q_R^{-1} R^T (M_k - M)^T L \quad (21)$$

(Luo *et al*., 2005).

Using the explicit gradient formula above, an algorithm can be developed like algorithm 1 to facilitate the classification applied to computer visualization, following an efficient algorithm using the gradient-up approach.

Algorithm 1 Symmetrical 2-DLDA using Gradient

Input

a) Set of figure $\{\mathbf{X}_i\}_{i=1}^{n}$ and label of each class

b) L_0, R_0 inisialization

c) Frequency c for orthogonalisazion

Inisialization

a) $L \leftarrow L_0, R \leftarrow R_0$

b) Compute $M_k, k = 1. 2. .... K$ and M

M_k is average of each class, and M is overall average

c) $t \leftarrow 0$

Do

Compute $S_w^R,\ S_w^L,\ S_b^R,\ S_b^L$

$R \leftarrow R + \delta\frac{\partial J}{\partial R}$

$L \leftarrow L + \delta\frac{\partial J}{\partial L}$

$t \leftarrow t + 1$

if (t mod c) = 0

$R \leftarrow$ eigenvector from $(S_w^L)^{-1} S_b^L$

$L \leftarrow$ eigenvector from $(S_w^R)^{-1} S_b^R$

endif

Output ***L*, *R***

 ISSN: 2476-9843

(Ye *et al*., 2005)

3.3 Application of the use of Symmetrical 2-Dimension Linear Discriminant Analysis Method in an Example of Character Pattern Recognition

To facilitate understanding of Linear Discriminant Analysis (LDA) and 2-Dimension Linear Discriminant Analysis (2-DLDA), researchers present how to recognize patterns of a character using these methods and how they compare with each other. The following are examples of 2 characters A and B, with each character A and B having two patterns.

.	.	#	#	.	.		#	#	#	#	.	.
.	#	.	.	#	.		#	.	.	.	#	.
#	.	#	#	.	#		#	#	#	#	#	.
#	.	.	.	.	#		#	.	.	.	.	#
#	.	.	.	.	#		#	.	.	#	#	#
#	.	.	.	.	#		#	#	#	.	.	.

H1 H2

.	.	#	#	.	.		#	.	#	#	#	.
.	#	.	.	#	.		#	#	.	.	#	.
#	#	.	.	#	#		#	.	.	#	.	.
#	.	#	#	.	#		#	#	#	#	#	.
#	.	.	.	.	#		#	.	.	.	.	#
#	.	.	.	.	#		#	#	#	#	#	.

H3 H4

Each character pattern above is represented in a $6 \times 6 = 36$ elements matrix, then the matrix is detached into 6 classes, $H_i = \{\prod_1, \prod_2, \ldots, \prod_6\}$. $X_i = 0$ if the element represented is a dot and $X_i = 1$ if the element represented is #, the representation matrix as follow:

$$H_1 = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

In the example of the characters above, it can be seen that the characters A_1 and A_3 characters are symmetrical characters, but the matrix of character representation is not a symmetrical matrix. For the pattern recognition process, then each character representation matrix is partitioned into 6 classes viz.

$$H_{1\_column1} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, H_{1\_column2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, H_{1\_column3} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, H_{1\_column4} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, H_{1\_column5} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$H_{1\_column6} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

And so on for H_2, H_3 and H_4 matrices, then calculate each average.

In high-dimensional data

The average of each class is

M_1 = 2,0553 M_2 = 0,92912 M_3 = 0,0726266 M_4 = -0,18354

M_5 = -0,386472 M_6 = 0,05335 M_7 = -0,09874 M_8 = -0,26913

M_9 = -0,07844

Overall average $M = 0.2331295$.

Linear Discriminant Analysis Method

Eigenvalue matrix in-class distribution (**S_w**) above is

$$eigenvalue(S_w) = \begin{pmatrix} -0.0000 \\ -0.0000 \\ 0.0000 \\ 0.0000 \\ 0.0000 \\ 0.0000 \\ 0.7703 \\ 2.9668 \\ 3.9791 \\ 9.7286 \\ 13.1248 \\ 17.8440 \\ 30.8847 \\ 96.9425 \\ 603.0483 \end{pmatrix}$$

trace (**S_w**) = 779.2891. Trace is the number of eigenvalues in a square matrix of size $n \times n$ which is also the sum of the diagonal elements of the matrix.

The in-class distribution matrix eigenvalue (**S_w**) above is

$$eigenvalue(S_b)=\begin{pmatrix}-0.0000\\-0.0000\\-0.0000\\-0.0000\\0.0000\\0.0000\\0.0000\\0.0000\\73.1690\end{pmatrix}$$

trace ($\mathbf{S_b}$) = number of diagonal elements ($\mathbf{S_b}$) = 73.1690.

$$\text{Optimum objective function LDA} = \frac{\mathbf{tr}(\mathbf{S}_b)}{\mathbf{tr}(\mathbf{S}_w)} = \frac{73.1690}{779.821}$$

$$= 0.093827$$

4. CONCLUSION

The problem of doubt created by the objective function of the 2-Dimensional Linear Discriminant Analysis can be dealt with prior to implementing the symmetrical 2-Dimension Linear Discriminant Analysis approach by resolving the issue before using the symmetrical 2-Dimension Linear Discriminant Analysis approach. As a result, there is a complete objective function. By dividing the traction formulation (Sw) by the inter-class distribution matrix (Sw), the symmetrical 2-DLDA formula was obtained (Sb). The all-encompassing objective function includes everything.

$$J = \mathbf{tr}\frac{\mathbf{R}^T\mathbf{S}_b^R\mathbf{R}}{\mathbf{R}^T\mathbf{S}_w^R\mathbf{R}} + \mathbf{tr}\frac{\mathbf{L}^T\mathbf{S}_b^L\mathbf{L}}{\mathbf{L}^T\mathbf{S}_w^L\mathbf{L}}$$

In addition, an efficient computational algorithm for solving objective functions in symmetric 2-DLDA is provided. 2-DLDA gives better and more accurate results than 2-DLDA when applied to high-dimensional data.

REFERENCES

[1] Duda, R.O., Hart, P.E., and Stork, D.G. 2012. *Pattern classification Second Edition*. New Jersey: John Wiley & Sons.

[2] Fukunaga, K. 1990. *Introduction to Statistical Pattern Classification* . San Diego, California, USA : Academic Press.

[3] Jain, Anil, K., Fellow, Duin, Robert, P.W., and Jianchang Mao, Jianchang., Senior Member, IEEE. Statistical Pattern Recognition: A Review. In *IEEE Transactions On Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, January 2000.

[4] Li, M. and Yuan, B. 2005. 2d-lda: A novel statistical linear discriminant analysis for image matrix. *Pattern Recognition Letters*, **26(5)**: page. 527–532.

[5] Luo, D., Ding, C., and Huang, H. 2007. *Symmetric Two Dimensional Linear Discriminant Analysis* (*2DLDA*). University of Texas at Arlington Press.

[6] Sharma, A., Paliwal, K.K. 2015. A deterministic approach to regularized linear discriminant analysis. *Neurocomputing* (**151**) pages 207-214. Sciencedirect 2015

[7] Strang, G. 2003. *Introduction to linear algebra Four Edition*. Massachusetts: Wellesley-Cambridge Press

[8] Tharwat, A., Gaber, T., and Hassanien, A.E. 2016. One-dimensional vs. two-dimensional based features: Plant identification approach. *Journal of Applied Logic* 24 DOI:10.1016/j.jal.2016.11.021

[9] Theodoridis, Sergios. and Koutroumbas, Kontantinos. 2003. *Pattern Recognition*. 2nd *Edition*. New York, USA: Academic Press.

[10] Viszlay, P., Lojka, M. and Juhar, J. 2014. Class-dependent two-dimensional linear discriminant analysis using two-pass recognition strategy. In Proceedings of the 22nd European Signal Processing Conference (EUSIPCO), pages 1796–1800. IEEE.

[11] Xiong, H., Swamy, M., and Ahmad, M. 2005. Two-dimensional fld for face recognition. *Pattern Recognition*, **38(7)**: page. 1121–1124.

[12] Yang, W. and Wu, H. 2014. Regularized complete linear discriminant analysis. *Neurocomputing*, **137**:185–191, Science Direct 2014.

[13] Ye, J., Janardan, R., and Li, Q. 2004. Two-Dimensional Linear Discriminant Analysis.

[14] Zhao, D., Chen, Z., Liu, C., Peng, Y. 2018. Two-dimensional Linear Discriminant Analysis for Low-Resolution Face Recognition. *IEEE*. DOI: 10.1109/CAC.2017.8242858

handwritten

ORIGINALITY REPORT

18%	17%	8%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	issuu.com Internet Source	5%
2	ranger.uta.edu Internet Source	4%
3	repository.universitasbumigora.ac.id Internet Source	2%
4	pdfs.semanticscholar.org Internet Source	2%
5	jurnal.uinsu.ac.id Internet Source	2%
6	www.eurasip.org Internet Source	1%
7	www.jurnal.uinsu.ac.id Internet Source	1%
8	www.cmlab.csie.ntu.edu.tw Internet Source	1%
9	Rong-Hua Li, Shuang Liang, George Baciu, Eddie Chan. "Equivalence Between LDA/QR and Direct LDA", International Journal of	1%

Cognitive Informatics and Natural Intelligence, 2011

Publication

Exclude quotes Off

Exclude bibliography On