

**BUKTI KORESPONDENSI
ARTIKEL JURNAL TERAKREDITASI SINTA 2**

Judul : Algorithm Symmetric 2-DLDA for Recognizing Handwritten Capital Letters

Jurnal : Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 21, No. 2, Maret 2022, hal. 389-402

E-ISSN : 2476-9843

Status : Penulis pertama

Penulis : Ismail Husein, et al.

Pengindeks : Peringkat 2 (SINTA)

Link : <https://journal.universitاسbumigora.ac.id/index.php/matrik/article/view/1254>



Submission Library

View Metadata

Submissions

Algorithm Symmetric 2-DLDA for Recognizing Handwritten Capital Letters

Ismail Husein, Rina Widyasari

Submission

Review

Copyediting

Production

Submission Files

Search



- ▶  5986-1 [ismailhusein22, Author, Handwritten Capital Letter Recognition Using Symmetric Two-Dimensional Linear Discriminant Analysis.docx](#) Research Results
- ▶  6897-1 [khairanmarzuki, Journal editor, 1254-kirim ke reviewer.docx](#) Article Text

Download All Files

Pre-Review Discussions

Add discussion

Name	From	Last Reply	Replies	Closed
Pre Review	khairanmarzuki Jul/21	ismailhusein22 Sep/17	8	<input type="checkbox"/>

 	ismailhusein22	-	0	<input type="checkbox"/>
	Mar/04			

Platform &
workflow by
OJS / PKP



Submissions

Algorithm Symmetric 2-DLDA for Recognizing Handwritten Capital Letters

Ismail Husein, Rina Widyasari

Submission

Review

Copyediting

Production

Round 1

Round 1 Status

Submission accepted.

Notifications

[\[matrik\] Editor Decision](#)

2022-03-29 02:27 PM

Reviewer's Attachments

[Search](#)

No Files

No Files

Review Discussions

[Add discussion](#)

Name	From	Last Reply	Replies	Closed
info revisi	khairanmarzuki Sep/30	ismailhusein22 Oct/05	2	<input type="checkbox"/>
informasi revisi	khairanmarzuki Jan/12	ismailhusein22 Mar/04	9	<input type="checkbox"/>

Platform &
workflow by
OJS / PKP



Submissions

Algorithm Symmetric 2-DLDA for Recognizing Handwritten Capital Letters

Ismail Husein, Rina Widyasari

Submission

Review

Copyediting

Production

Copyediting Discussions

[Add discussion](#)

Name	From	Last Reply	Replies	Closed
informasi copy editing	khairanmarzuki	-	0	<input type="checkbox"/>
	Mar/05			

Copyedited

[Search](#)

 9230-1	khairanmarzuki, Journal editor, 1254-Article Text-9212-1-18-20220302.docx	Article Text
--	---	--------------

Platform &
workflow by
OJS / PKP



Algorithm Symmetric 2-DLDA for Recognizing Handwritten Capital Letters

Ismail Husein, Rina Widyasari

Submissions

Submission

Review

Copyediting

Production

Production Discussions

[Add discussion](#)

Name	From	Last Reply	Replies	Closed
------	------	------------	---------	--------

No Items

Galleys

[▶ pdf](#)

Platform &
workflow by
OJS / PKP

Algorithm Symmetric 2-DLDA for Recognizing Handwritten Capital Letters

Husein¹, Rina Widyasari²
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.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Ismail Husein
Department of Mathematics
UIN 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.[1]

When it comes to pattern identification in this study, statistical classification is the most effective method[2]. Pattern recognition can be done using a variety of algorithms. Writing with LDA is a successful application of Linear Discriminant Analysis (LDA). Latent Dirichlet Allocation (LDA) is a subspace investigation method for studying the low-dimension structure of high-dimension data. In the vector space character projection, this method simultaneously minimizes the distribution within class (S_w) and maximizes the distribution between classes (S_b).[3]

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:

$$S_b (XX^T) \neq S_b (X^T X), \quad S_w (XX^T) \neq S_w (X^T X), \quad (1)$$

so there are a number of possible choices for determining the appropriate objective function of LDA[4]. The introduction of handwritten capital letters is one of the applications of symmetrical 2-dimension linear discriminant analysis.

Commented [km1]:

Perlu revisi
Abstrak : Apa masalahnya?. Tujuan?, Metodologi (Oke), Hasil (Ukuran?, perbandingan?).
Pendahuluan : Belum menunjukkan masalah & pentingnya penelitian ini ?.
Metodologi : Oke

Kesimpulan : 2-DLDA memberikan hasil yang lebih baik dan akurat dibandingkan 2-DLDA bila diterapkan pada data berdimensi tinggi??.
Apa perbandingannya?. Apa yang diukur: waktu?, akurasi?
Saran untuk penelitian selanjutnya..

Artikel ini terlalu banyak rumus sehingga membingungkan tolong dikasih penjelasan pada rumus tersebut tidak hanya ditulis rumusnya saja..

2. RESEARCH METHOD

2.1. 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 [5].

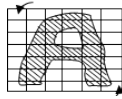


Figure 1. Example measurement of letter patterns

If a data matrix is given $H \in R^{N \times n}$, LDA method intends to encounter a transformation $K \in R^{N \times l}$ that assigns each a_i column from matrix H , for $1 \leq i \leq p$, in the P dimension space to the l vector in the dimension l space. Namely K :

$$h_i \in R^{N \times n} \rightarrow b_i = K^T h_i \in R^l \quad (l < P)$$

In other words, LDA intends to encounter a vector space \mathbf{K} spanned by $\{k_i\}_{i=1}^l$ where $K = [k_1, k_2, \dots, k_l]$, so that each h_i is projected to \mathbf{K} by $(k_1^T h_i, \dots, k_l^T h_i)^T \in R^l$ [6].

Consider that the initial data in H is subdivided into k classes so that $H = \{\Pi_1, \Pi_2, \dots, \Pi_k\}$, where Π_i loads n_i the data points of class $-i$ and $\sum_{i=1}^k P_i = P$. The classic LDA intends to encounter the optimal transformation of G so that the class structure of the initial high-dimensional space data is converted into a low-dimensional space.[7]

The transformation to a lower dimension subspace is used in LDA is

$$y_i = K^T x_i \quad (2)$$

where K is the transformation to a subspace. Usually also written with $(y_1, \dots, y_p) = K^T (x_1, \dots, x_p)$ 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:

$$S_w = \sum_{i=1}^c \sum_{x_i \in \Pi_i} [x_i - m_i][x_i - m_i]^T \quad (3)$$

$$S_b = \sum_{i=1}^c n_i [m_i - m][m_i - m]^T \quad (4)$$

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 [9]. The class average formula and the overall average is as follows:

$$m_i = \frac{1}{n_i} \sum_{x \in \Pi_i} x \quad \text{is the average of the } i\text{-class, and} \quad m = \frac{1}{n} \sum_{i=1}^k \sum_{x \in \Pi_i} x \quad \text{is the overall average. General optimizations in Linear Discriminant Analysis[10]:}$$

$$\max_K J(K) = \text{tr} \frac{S_b(Y)}{S_w(Y)} = \text{tr} \frac{K^T S_b(X) K}{K^T S_w(X) K} \quad (5)$$

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 $r \times r$ and $c \times c$ sizes, which are much smaller than the classical LDA matrices [11].

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

$$Y_i = L^T X_i R \quad (6)$$

so the different classes are detached.

$$M_i = \frac{1}{n_i} \sum_{x \in \Pi_i} X \quad M = \frac{1}{n} \sum_{i=1}^k \sum_{x \in \Pi_i} X$$

Suppose that is the average of the i -th class, $1 \leq i \leq k$, and M means the overall average [12]. In 2-DLDA, researchers regard images as two-dimensional signals and intend to detect two transformation matrices $L \in \mathfrak{R}^{r \times l_1}$ and $R \in \mathfrak{R}^{c \times l_2}$ and assigns each H_i member for $1 \leq x \leq n$, to a B_i matrix so that $B_i = L^T H_i R$.

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 [9]. The square of the distance from within-class and between classes can be calculated as below:

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

$\text{trace}(MM^T) = \|M\|_F^2$, for a matrix M , so obtained :

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

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

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

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

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

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).

Commented [km2]: Apakah ini rumus tolong di kasih keterangan rumus keberapa..

Commented [km3]: Tolong dikasih keterangan rumus keberapa dan kurungnya bukan seperti itu kalau rumus tolong di periksa lagi semua rumus untuk dikasih keterangannya

Commented [km4]: Idem dengan pertanyaan yg diatas

Commented [km5]: Idem denga pertanyaan diatas

1. L calculation

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

$$\bar{D}_w = \text{trace}(L^T S_w^R L), \quad \bar{D}_b = \text{trace}(L^T S_b^R L) \quad [11]$$

where

$$S_w^R = \sum_{i=1}^k \sum_{X \in \Pi_i} (X - M_i) R R^T (X - M_i)^T, \quad S_b^R = \sum_{i=1}^k n_i (M_i - M) R R^T (M_i - M)^T \quad [12]$$

optimal L can be calculated by figuring out the succeeding optimization problem: $\max_L \text{trace}((L^T S_w^R L)^{-1} (L^T S_b^R L))$.

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[13].

2. R calculation

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

$$\bar{D}_w = \text{trace}(R^T S_w^L R), \quad \bar{D}_b = \text{trace}(R^T S_b^L R) \quad [13]$$

where

$$S_w^L = \sum_{i=1}^k \sum_{X \in \Pi_i} (X - M_i)^T L L^T (X - M_i), \quad S_b^L = \sum_{i=1}^k n_i (M_i - M)^T L L^T (M_i - M) \quad [14]$$

Optimal L can be calculated by clarifying the succeeding optimization problem: $\max_R \text{trace}((R^T S_w^L R)^{-1} (R^T S_b^L R))$.

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 $(S_w^L)^{-1} S_b^L$. Remark that the size of the matrices S_w^L and S_b^L are $r \times r$ (square matrix) [13]

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

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

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

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

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

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

Consequently, in the space of transformation can be written

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

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

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

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

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

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

$$J_1 = \text{tr} \frac{S_b(\mathbf{Y}\mathbf{Y}^T)}{S_w(\mathbf{Y}\mathbf{Y}^T)}$$

$$J_2 = \text{tr} \frac{S_b(\mathbf{Y}^T\mathbf{Y})}{S_w(\mathbf{Y}^T\mathbf{Y})}$$

$$J_3 = \text{tr} \left[\frac{S_b(\mathbf{Y}\mathbf{Y}^T)}{S_w(\mathbf{Y}\mathbf{Y}^T)} + \frac{S_b(\mathbf{Y}^T\mathbf{Y})}{S_w(\mathbf{Y}^T\mathbf{Y})} \right]$$

$$J_4 = \text{tr} \left[\frac{S_b(\mathbf{Y}^T\mathbf{Y})}{S_w(\mathbf{Y}^T\mathbf{Y})} \frac{S_b(\mathbf{Y}\mathbf{Y}^T)}{S_w(\mathbf{Y}\mathbf{Y}^T)} \right],$$

$$J_5 = \text{tr} \frac{S_b(\mathbf{Y}\mathbf{Y}^T) + S_b(\mathbf{Y}^T\mathbf{Y})}{S_w(\mathbf{Y}\mathbf{Y}^T) + S_w(\mathbf{Y}^T\mathbf{Y})}, \quad [8]$$

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

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

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

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

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

Pada Fukunaga(1990), matriks didefinisikan sebagai:

In Fukunaga (1990), $\mathbf{\Gamma}$ 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 $\mathbf{Y}_i = \mathbf{L}^T \mathbf{X}_i \mathbf{R}$. The explanation is as follows:

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

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

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

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

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

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

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

Therefore, optimization using (\mathbf{L}, \mathbf{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 for 2-DLDA is

$$J_{ADL2-D} = \text{tr} \frac{\mathbf{S}_b(\mathbf{Y}\mathbf{Y}^T)}{\mathbf{S}_w(\mathbf{Y}\mathbf{Y}^T)} = \text{tr} \frac{\mathbf{S}_b(\mathbf{Y}^T\mathbf{Y})}{\mathbf{S}_w(\mathbf{Y}^T\mathbf{Y})} \quad [15]$$

$$J_{ADL2-D} = \text{tr} \left(\frac{\mathbf{R}^T \mathbf{S}_b^L \mathbf{R}}{\mathbf{R}^T \mathbf{S}_w^L \mathbf{R}} + \frac{\mathbf{L}^T \mathbf{S}_b^R \mathbf{L}}{\mathbf{L}^T \mathbf{S}_w^R \mathbf{L}} \right) \quad [16]$$

Using theorem 1, in the case non symmetric 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

$$J'_1 = \text{tr} \frac{S_b(\mathbf{Y}\mathbf{Y}^T)}{S_w(\mathbf{Y}\mathbf{Y}^T)} = \text{tr} \frac{\mathbf{R}^T \mathbf{S}_b^L \mathbf{R}}{\mathbf{R}^T \mathbf{S}_w^L \mathbf{R}}$$

$$J'_2 = \text{tr} \frac{S_b(\mathbf{Y}^T\mathbf{Y})}{S_w(\mathbf{Y}^T\mathbf{Y})} = \text{tr} \frac{\mathbf{L}^T \mathbf{S}_b^R \mathbf{L}}{\mathbf{L}^T \mathbf{S}_w^R \mathbf{L}}$$

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.

$$J = \text{tr} \frac{\mathbf{R}^T \mathbf{S}_b^R \mathbf{R} + \mathbf{L}^T \mathbf{S}_b^L \mathbf{L}}{\mathbf{R}^T \mathbf{S}_w^R \mathbf{R} + \mathbf{L}^T \mathbf{S}_w^L \mathbf{L}} \quad [17]$$

Secondly is how to optimize the objective function. The result to maximizing $\max_R J$ can be simply done by calculating

eigenvectors from $\mathbf{S}_w^{-1} \mathbf{S}_b$, 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 $P_L = \mathbf{L}^T \mathbf{S}_b^R \mathbf{L}$, $Q_L = \mathbf{L}^T \mathbf{S}_w^R \mathbf{L}$, $P_R = \mathbf{R}^T \mathbf{S}_b^L \mathbf{R}$, and $Q_R = \mathbf{R}^T \mathbf{S}_w^L \mathbf{R}$.

Derivative of objective function J_{ADL2-D} in equation (3.2) as follow

$$\frac{\partial J}{\partial R} \text{ obtained}$$

$$\begin{aligned} \frac{\partial}{\partial R} \text{tr} \frac{R^T S_b^L R}{R^T S_w^L R} &= 2S_b^L R Q_R^{-1} - 2S_w^L R Q_R^{-1} P_R Q_R^{-1} \\ &\text{and} \\ \frac{\partial}{\partial R} \text{tr} \frac{L^T S_b^R L}{L^T S_w^R L} &= 2 \sum_{k=1}^K \sum_{A_i \in \pi_k} (H_i - M_k)^T L Q_L^{-1} L^T (H_i - M_k) R \\ &\quad - 2 \sum_{k=1}^K (M_k - M)^T L Q_L^{-1} P_L Q_L^{-1} L^T (M_k - M) R \end{aligned} \quad [18]$$

For $\frac{\partial J}{\partial L}$ obtained [8]

$$\begin{aligned} \frac{\partial}{\partial L} \text{tr} \frac{L^T S_b^R L}{L^T S_w^R L} &= 2S_b^R L Q_L^{-1} - 2S_w^R L Q_L^{-1} P_L Q_L^{-1} \\ &\text{and} \\ \frac{\partial}{\partial L} \text{tr} \frac{R^T S_b^L R}{R^T S_w^L R} &= 2 \sum_{k=1}^K \sum_{A_i \in \pi_k} (H_i - M_k) R Q_L^{-1} R^T (H_i - M_k)^T L \\ &\quad - 2 \sum_{k=1}^K (M_k - M) R Q_L^{-1} P_R Q_L^{-1} R^T (M_k - M)^T L \end{aligned} \quad [19]$$

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

- Set of figure $\{\mathbf{X}_i\}_{i=1}^n$ and label of each class
- L_0, R_0 initialization
- Frequency c for orthogonalization

Initialization

- $L \leftarrow L_0, R \leftarrow R_0$
- Compute $M_k, k = 1, 2, \dots, K$ and M
 M_k is average of each class, and M is overall average
- $t \leftarrow 0$

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 $(t \bmod c) = 0$

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

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

endif

Output L, R [14]

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.

.	.	#	#	.	.	#	#	#	#	.	.
.	#	.	.	#	.	#	.	.	.	#	.
#	.	#	#	.	#	#	#	#	#	.	.
#	#	#	#
#	#	#	.	#	#	#	.
#	#	#	#
			H ₁				H ₂				
.	.	#	#	.	.	#	.	#	#	#	.
.	#	.	.	#	.	#	#	.	.	#	.
#	#	.	.	#	#	#	.	#	.	.	.
#	.	#	#	.	#	#	#	#	#	#	.
#	#	#	#
#	#	#	#	#	#	#	.
			H ₃				H ₄				

Each character pattern above is represented in a $6 \times 6 = 36$ elements matrix, then the matrix is detached into 6 classes, $H_i = \{[1], [2], \dots, [6]\}$. $X_i = 0$ 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 A₁ and A₃ 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.

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

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

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

$$\begin{aligned} M_1 &= 2,0553 & M_2 &= 0,92912 & M_3 &= 0,0726266 & M_4 &= -0,18354 \\ M_5 &= -0,386472 & M_6 &= 0,05335 & M_7 &= -0,09874 & M_8 &= -0,26913 \\ M_9 &= -0,07844 \end{aligned}$$

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 (S_b) = number of diagonal elements (S_b) = 73.1690.

$$\begin{aligned} \text{Optimum objective function LDA} &= \frac{\text{tr}(S_b)}{\text{tr}(S_w)} \\ &= \frac{73.1690}{779.821} \\ &= 0.093827 \end{aligned}$$

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 (S_w) by the inter-class distribution matrix (S_w), the symmetrical 2-DLDA formula was obtained (S_b). The all-encompassing objective function includes everything.

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

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] K. Theodoridis, Sergios. and Koutroumbas, *Pattern Recognition. 2nd Edition*. 2003.
- [2] Jain, Anil, K., Fellow, Duin, Robert, P.W., and Jianchang Mao, Jianchang., "Transactions On Pattern Analysis and Machine Intelligence," *IEEE*, vol. 22, no. 1, 2000.
- [3] K. K. Sharma, A., Paliwal, "A deterministic approach to regularized linear discriminant analysis," *Neurocomputing*, vol. 151, pp. 207–214, 2015.
- [4] A. . Tharwat, A., Gaber, T., and Hassanien, "One-dimensional vs. two-dimensional based features: Plant identification approach," *J. Appl. Log.*, vol. 24, 2016.
- [5] B. Li, M. and Yuan, "A novel statistical linear discriminant analysis for image matrix," *Pattern Recognit. Lett.*, vol. 26, no. 5, pp. 527–532, 2005.
- [6] D. . Duda, R.O., Hart, P.E., and Stork, *Pattern classification Second Edition*. New Jersey: John Wiley & Sons. 2012.
- [7] M. Xiong, H., Swamy, M., and Ahmad, "Two-dimensional fld for face recognition," *Pattern Recognit. Lett.*, vol. 38, no. 7, pp. 1121–1124, 2005.
- [8] H. Luo, D., Ding, C., and Huang, "Symmetric Two Dimensional Linear Discriminant Analysis (2DLDA)," in *University of Texas at Arlington Press*, 2007.
- [9] J. Vizslay, P., Lojka, M. and Juhar, "Class-dependent two-dimensional linear discriminant analysis using two-pass recognition strategy," *Proc. 22nd Eur. Signal Process. Conf.*, pp. 1796–1800, 2014.
- [10] K. Fukunaga, "Introduction to Statistical Pattern Classification," in *San Diego, California, USA : Academic Press.*, 1990.
- [11] H. Yang, W. and Wu, "Regularized complete linear discriminant analysis," *Neurocomputing*, vol. 137, pp. 185–191, 2014.
- [12] G. Strang, "Introduction to linear algebra Four Edition," in *Massachusetts: Wellesley-Cambridge Press*, 2003.
- [13] Y. Zhao, D., Chen, Z., Liu, C., Peng, "Two-dimensional Linear Discriminant Analysis for Low-Resolution Face Recognition," *IEEE*, 2018.
- [14] Q. Ye, J., Janardan, R., and Li, *Two-Dimensional Linear Discriminant Analysis*. 2004.

Commented [km6]: Referensi minimal 15.. dan silahkan sesuaikan dengan template referensi jurnal matrik