PENERAPAN METODE ANALISA DISKRIMINAN MAJEMUK DENGAN PENDEKATAN TRANSFORMASI FUKUNAGA KOONTZ

Rully Soelaiman[1] **Wiwik Anggraini**[2] **M.Mujahidillah**[3]
[1,2,3]Fakultas Teknologi Informasi
Institut Teknologi Sepuluh Nopember (ITS) Surabaya, 60111, Indonesia
Email: rully@is.its.ac.id

ABSTRACT
Linear discriminant analysis is one of method frequently used and developed in the field of pattern recognition. This method tries to find the optimal subspace by maximizing the Fisher Criterion. Application of pattern recognition in high-dimensional data and the less number of training samples cause singular within-class distribution matrix.
In this paper, we developed Linear Discriminant Analysis method using Fukunaga Koontz Transformation approach to meet the needs of the nonsingular within-class distribution matrix. Based on Fukunaga Koontz Transformation, the entire space of data is decomposed into four subspaces with different discriminant ability (measured by the ratio of eigenvalue). Maximum Fisher Criterion can be identified by linking the ratio of eigenvalue and generalized eigenvalue.
Next, this paper will introduce a new method called complex discriminant analysis by transforming the data into intraclass and extraclass then maximize their Bhattacharyya distance. This method is more efficient because it can work even though within-class distribution matrix is singular and between-class distribution matrix is zero.

Keywords: *pattern classification, linier discriminant analysis, complex discriminant analysis, Fukunaga Koontz Transformation*

ABSTRAK
Analisa Diskriminan Linier merupakan salah satu metode yang sering digunakan dan dikembangkan pada bidang pengenalan pola. Metode ini mencoba menemukan subspace optimal dengan memaksimalkan Fisher Criterion. Penerapan pengenalan pola pada data berdimensi tinggi dan jumlah sample training yang sedikit menyebabkan matriks sebaran within-class bersifat singular.
Pada makalah ini, dikembangkan metode Analisa Diskriminan Linier dengan pendekatan Transformasi Fukunaga Koontz untuk memenuhi kebutuhan matriks sebaran within-class yang bersifat nonsingular. Berdasarkan Transformasi Fukunaga Koontz, seluruh space data didekomposisi ke dalam empat subspace dengan kemampuan diskriminan yang berbeda-beda (diukur dari rasio eigenvalue). Fisher Criterion maksimal dapat diketahui dengan menghubungkan rasio eigenvalue dan generalized eigenvalue.
Selanjutnya akan diperkenalkan metode baru bernama Analisa Diskriminan Majemuk dengan mentransformasikan data menjadi intraclass dan extraclass dan memaksimalkan jarak Bhattacharyya. Metode ini lebih efisien karena dapat bekerja meskipun matriks sebaran within-class singular dan matriks sebaran between-class bernilai nol.

Kata kunci: *klasifikasi pola, analisa diskriminan linier, analisa diskriminan majemuk, transformasi Fukunaga Koontz*

Saat ini, analisa subspace diskriminan telah dipelajari secara luas pada bidang visi komputer dan pengenalan pola. Hal ini telah umum digunakan untuk ekstraksi fitur dan reduksi dimensi pada pengenalan wajah, dan klasifikasi teks. Salah satu metode yang populer adalah Diskriminan Linier Fisher, disebut juga Analisa Diskriminan Linier. Metode ini mencoba menemukan subspace optimal dengan cara memaksimalkan perpisahan dua kelas. Hal ini dapat diperoleh dengan meminimalkan jarak matriks sebaran *within-class* S_w dan memaksimalkan jarak matriks sebaran *between-class* S_b secara simultan sehingga menghasilkan Fisher Criterion J_F yang maksimal. Dengan memaksimalkan *Fisher Criterion* tersebut, Diskriminan Linier Fisher menemukan subspace dimana kelas-kelas paling terpisah secara linier [1] [2]. Namun pada banyak aplikasi dengan data berdimensi tinggi, contohnya pengenalan wajah, matriks sebaran S_w bersifat singular karena pada umumnya dimensi data lebih besar dari jumlah sample. Hal ini disebut *undersampled* atau *small sample size problem*.

Sampai sekarang, banyak metode ditawarkan untuk menangani masalah kebutuhan *nonsingularity* dari S_w, seperti *Fisherface, Discriminant Common Vectors, Dual Space*, LDA/GSVD dan LDA/QR. Tetapi metode-metode diatas tidak secara langsung bersesuaian dengan generalized eigenvalue λ, alat pengukur essensial kemampuan mendiskriminasi. Kenyataannya, metode yang sudah ada menghasilkan suboptimum pada *Fisher Criterion* karena informasi diskriminan penting diabaikan untuk membuat S_w dapat diinverskan.

Berdasarkan Transformasi Fukunaga Koontz, seluruh space data didekomposisi ke dalam empat subspace dan mengoreksi maksimasi J_F meskipun S_w bersifat singular. Analisa Diskriminan Linier hanya optimal jika diterapkan pada dua distribusi Normal dengan matriks kovarian yang sama. Kasus terburuk terjadi ketika semua kelas mempunyai rata-rata yang sama sehingga matriks sebaran *between-class* $S_b = 0$.

Makalah ini disusun untuk membahas penerapan metode Analisa Diskriminan Majemuk dengan membuat dua kelas dan memaksimalkan jarak Bhattacharyya (batas error dari Bayes Classifier) menggunakan Transformasi Fukunaga Koontz [3] [4]. Subspace diskriminan diperoleh dari Transformasi Fukunaga Koontz sehingga dapat ditemukan global optimum secara langsung (tanpa iterasi). Metode Analisa Diskriminan Majemuk dan Transformasi Fukunaga Koontz ini lebih efisien sehingga dapat bekerja meskipun matriks sebaran *within-class* S_w bersifat singular dan matriks sebaran between-class bernilai nol ($S_b = 0$).

ANALISA DISKRIMINAN LINEAR, ADL

Jika $A = \{a_1, ..., a_N\}, a_i \in R^D$ merupakan set data dari $D - vektor$ dimensional. Setiap titik data objek bersesuaian tepat satu dengan kelas objek C $\{L_1, ..., L_C\}$. Jumlah vektor data pada kelas L_i dinotasikan dengan N_i; sehingga, $N = \Sigma N_i$. Untuk data berdimensi tinggi seperti gambar wajah, umumnya, $C \leq N << D$ [5] [6]. Matriks sebaran *between-class* S_b, matriks sebaran *within-class* S_w, dan matriks sebaran total S_t didefinisikan sebagai berikut:

$$S_b = \sum_{i=1}^{C} N_i(m_i - m)(m_i - m)^T \qquad = H_b H_b^T \quad (1)$$

$$S_w = \sum_{i=1}^{C} \sum_{a_j \in L_i} (a_j - m_i)(a_j - m_i)^T \quad = H_w H_w^T \quad (2)$$

$$S_t = \sum_{i=1}^{N} (a_i - m)(a_i - m)^T \qquad = H_t H_t^T \quad (3)$$

$$S_t = S_b + S_w \quad (4)$$

Dimana, m_i adalah rata-rata kelas (rata-rata objek yang termasuk anggota sebuah kelas) dan m adalah rata-rata global (rata-rata dari semua kelas) dari A. Matriks $H_b \in R^{DxC}$, $H_w \in R^{DxN}$, dan $H_t \in R^{DxN}$ berturut-turut adalah matriks *precursor* dari matriks sebaran *between-class*, matriks sebaran *within-class*, dan matriks sebaran total,

$$H_b = [\sqrt{N_1}(m_1 - m), ..., \sqrt{N_C}(m_C - m)] \quad (5)$$

$$H_w = [A_1 - m_1.l_1^T, ..., A_C - m_C.l_C^T] \quad (6)$$

$$H_t = [a_1 - m, ..., a_N - m] \quad (7)$$

Dimana, $l_i = (1, ..., l)^T \in R^{N_i}$ dan A_i adalah matriks data untuk kelas L_i. Rank tiap matriks sebaran adalah $r_w = rank(S_w)$, $r_b = rank(S_b)$, dan $r_t = rank(S_t)$. Untuk data berdimensi tinggi $(N << D)$, $r_b \leq C - 1$, $r_w \leq N - C$ dan $r_t \leq N - 1$.

Metode Analisa Diskriminan Linier mencoba menemukan subspace optimal dengan cara memaksimalkan perpisahan dua kelas. Hal ini dapat diperoleh dengan meminimalkan jarak matriks sebaran *within-class* S_b dan memaksimalkan jarak matriks sebaran *between-class* S_w secara simultan. *Fisher Criterion* dapat dituliskan sebagai

$$J_F(\Phi) = trace\{(\Phi^T S_w \Phi)^{-1}(\Phi^T S_b \Phi)\} \quad (8)$$

Dimana Φ adalah matriks transformasi linier. Solusi yang memaksimalkan *Fisher criterion* (J_F) adalah set dari eigenvector $\{\phi_i | i = 1, 2, ..., k\}$ yang bersesuaian dengan k eigenvalue terbesar $\{\lambda_i | i = 1, 2, ..., k\}$. Batas atas dari k adalah $C - 1$, dimana C adalah jumlah kelas.

$$S_b \phi = \lambda S_w \phi \quad (9)$$

TRANSFORMASI FUKUNAGA KOONTZ, TFK

Transformasi Fukunaga Koontz telah didesain untuk masa-lah pengenalan dua kelas [3][7][8]. Jika terdapat matriks data A_1 dan A_2 dari dua kelas, dengan matriks autokorelasi $S_1 = A_1 A_1 T$ dan $S_2 = A_2 A_2 T$ adalah semidefinit positif dan simetris. Jumlah dari kedua matriks ini tetap semi-definit positif dan simetris dan dapat difaktorisasikan dalam bentuk:

$$S = S_1 + S_2 = [U, U_\perp] \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U^T \\ U_\perp^T \end{bmatrix} \quad (10)$$

Tanpa kehilangan sifat umum, S dapat bersifat singular dan $r = rank(S) < D$; sehingga, $D = diag\{\lambda_1, ..., \lambda_r\}$, $\lambda_1 \geq ... \geq \lambda_r \geq 0$. Set eigenvector $U \in R^{DxR}$ bersesuaian dengan eigenvalue selain nol dan set $U_\perp \in R^{Dx(D-r)}$ adalah komplemen orthogonal dari U. S dapat diperjelas dengan operator transformasi $P = UD^{-\frac{1}{2}}$. Jumlah dua matrix S_1 dan S_2 menjadi:

$$P^T S P = P^T(S_1 + S_2)P = \widetilde{S}_1 + \widetilde{S}_2 = I \quad (11)$$

Dimana $\widetilde{S}_1 = P^T S_1 P, \widetilde{S}_2 = P^T S_2 P$, dan $I \in R^{rxr}$ adalah matriks identitas. Anggap eigenvector dari S_1 adalah v dengan eigenvalue λ_1 sehingga $S_1 v = \lambda_1 v$. Karena $\widetilde{S}_1 = I - \widetilde{S}_2$, sehingga dapat dituliskan sebagai berikut:

$$\left(I - \widetilde{S}_2\right) v = \lambda_1 v \quad (12)$$

$$\widetilde{S}_2 v = (1 - \lambda_1)\, v \quad (13)$$

Ini artinya $\widetilde{S}_2$ mempunyai eigenvector yang sama dengan $\widetilde{S}_1$ tetapi eigenvalue yang bersesuaian adalah $\lambda_2 = 1 - \lambda_1$. Konsekuensinya, eigenvector yang dominan pada S_1 adalah eigenvector terendah pada S_2 dan sebaliknya. Akibatnya, pola yang merupakan anggota kelas 1 mendapatkan nilai koefisien besar ketika diproyeksikan pada eigenvector dominan dari S_1 dan sebaliknya. Eigenvector akan membentuk subspace dimana 2 kelas terpisah. Klasifikasi dapat dilakukan dengan mengambil nearest neighbor pada subspace ini.

ALGORITMA ADL/TFK

Transformasi Fukunaga Koontz diterapkan untuk mengatasi masalah singularity S_w. Umumnya, pada permasalahan Analisa Diskriminan Linier terdapat lebih dari dua kelas. Untuk menangani masalah banyak kelas, matriks autokorelasi S_1 dan S_2 disubstitusi dengan matriks sebaran S_b dan S_w. Transformasi Fukunaga Koontz dapat diaplikasikan pada S_b, S_w, dan S_t karena S_b, S_w, dan S_t bersifat semi-definit positif dan simetris dimana $S_t = S_b + S_w$. Metode tersebut dinamakan Analisa Diskriminan Linier/Transformasi Fukunaga Koontz (ADL/TFK). Seluruh space data didekomposisikan menjadi U dan $U_\perp$.

$U_\perp$ adalah suatu set eigenvector yang berhubungan dengan eigenvalue nol dari S_t. Set eigenvector ini adalah titik temu null space S_b dan Sw dan tidak memiliki informasi diskriminan. Disisi lain, U adalah set eigenvector yang berhubungan dengan eigenvalue bukan nol dari S_t dan berisi informasi diskriminan.

Berdasarkan rumus Transformasi Fukunaga Koontz yaitu $\widetilde{S}_b = P^T S_b P$ dan $\widetilde{S}_b = P^T S_w P$ membagi eigenspace yang sama dan jumlah dari dua eigenvalue yang berhubungan dengan eigenvector yang sama adalah sama dengan 1.

Berdasarkan rasio eigenvalue $\frac{\lambda_b}{\lambda_w}$, U dapat didekomposisi menjadi tiga subspace. $U_\perp$ disertakan sebagai subspace keempat untuk menjaga integritas keseluruhan space data [3][1].

Subspace 1. span (S_b) $\cap$ null (S_w), set eigenvector $\{v_i\}$ berhubungan dengan λ_w dan λ_b = 1. Karena $\frac{\lambda_b}{\lambda_w}$ = ∞, pada subspace ini rasio eigenvalue maksimal.

Subspace 2. span (S_b) $\cap$ span (S_w), set eigenvector $\{v_i\}$ berhubungan dengan dan 0 < λ_w <1 dan 0 < $lambda_b$ <1. Karena 0 < $\frac{\lambda_b}{\lambda_w}$ < ∞, rasio eigenvalue terbatas dan lebih rendah dari Subspace 1.

Subspace 3. null (S_b) $\cap$ span (S_w), set eigenvector $\{v_i\}$ berhubungan dengan λ_w = 1 dan λ_b = 0. Karena $\frac{\lambda_b}{\lambda_w}$ = 0, rasio eigenvalue minimal.

Subspace 4. null (S_b) $\cap$ null (S_w), set eigenvector berhubungan dengan eigenvalue nol dari S_t.

Pada prakteknya, beberapa subspace dari keempat subspace ini mungkin tidak ada, tergantung rank dari S_b, S_w, dan S_t.

ALGORITMA ADM/TFK

Dari perspektif Bayes Classifier, Analisa Diskriminan Li-nier hanya optimal jika diterapkan pada dua distribusi Normal dengan matriks kovarian yang sama. Metode berbasis Analisa Diskriminan Linier akan tidak optimal jika dibanding dengan Bayes Classifier [6][7]. Kasus terburuk terjadi ketika semua kelas mempunyai rata-rata yang sama. Pada kasus ini, S_b = 0 dan semua metode berbasis Analisa Diskriminan Linier akan gagal. Untuk menangani permasalahan ini, masalah banyak kelas diubah menjadi masalah klasifikasi pola *binary* dengan memperkenalkan $\triangle = a_i - a_j$ dan mendefinisikan space intraclass $\Omega_I = \{(a_i - a_j)|L(a_i)\} = L(a_j)$, begitu juga extraclass $\Omega_E = \{(a_i - a_j)|L(a_i)\} \neq L(a_j)$ dimana $L(a_i)$ adalah label kelas dari a_i.

$$m_I = m_E = 0 \tag{14}$$

$$\sum I = \frac{1}{N_I} \sum_{L(a_i)=L(a_j)} (a_i - a_j)(a_i - a_j)^T \tag{15}$$

$$\sum E = \frac{1}{N_E} \sum_{L(a_i)\neq L(a_j)} (a_i - a_j)(a_i - a_j)^T \tag{16}$$

$$N_I = \frac{1}{2} \sum n_i(n_i - 1) \tag{17}$$

$$N_E = \sum_{L_i \neq L_j} n_i n_j \tag{18}$$

Dimana, N_I adalah jumlah sample pada Ω_I dan N_E adalah sejumlah sample pada Ω_E.

Sebagai catatan, biasanya, rank ($\sum_E$) dan rank ($\sum_I$) keduanya lebih besar dari $C - 1$, dimana C adalah jumlah kelas. H_I dan H_E adalah matriks precursor dari $\sum_I$ dan $\sum_E$ dinyatakan dengan

$$H_I = \frac{1}{\sqrt{N_I}} [..., (a_i - a_j), ...] \tag{19}$$

Dimana $L(a_i) = L(a_j)$

$$H_E = \frac{1}{\sqrt{N_E}} [..., (a_i - a_j), ...] \tag{20}$$

Dimana $L(a_i) \neq L(a_j)$

Tujuan utama adalah menemukan subspace Φ dimana Ω_I dan Ω_E dapat dipisahkan sejauh mungkin. Jarak Bhattacharyya dapat digunakan untuk menghitung overlap dari berbagai fungsi padat peluang (*probability density function*). Berdasarkan referensi [6], subspace optimal Φ dapat dihitung dengan memaksimalkan criterion baru dimana jarak Bhattacharyya:

$$\begin{aligned} J_{MDA} = ln|(\Phi^T \sum_E \Phi)^{-1} (\Phi^T \sum_I \Phi) + \\ (\Phi^T \sum_I \Phi)^{-1} (\Phi^T \sum_E \Phi) + 2I_d| \end{aligned} \tag{21}$$

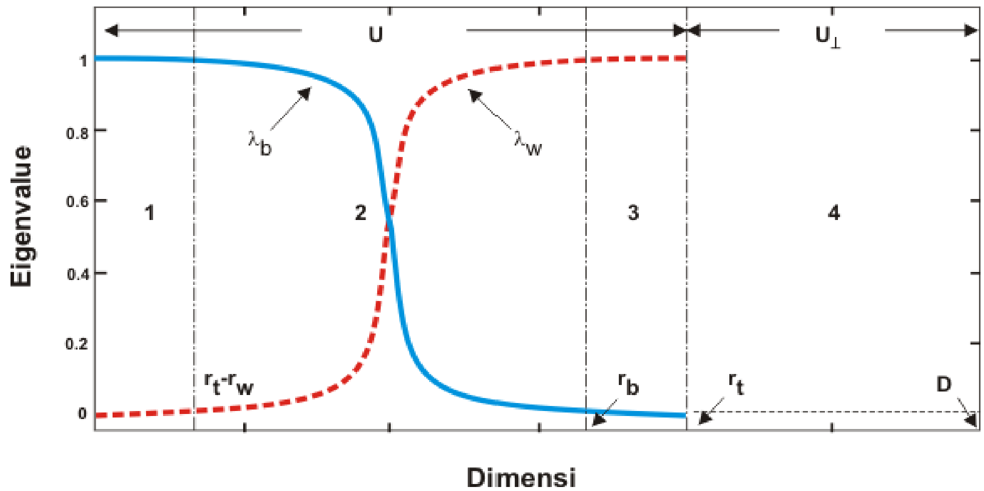
Dimana I_d adalah matriks identitas pada space dimensi rendah.

Jika $\sum_I$ atau $\sum_E$ singular, subspace tidak dapat diperoleh secara langsung dengan inversi matriks. Namun, dengan mengaplikasikan Transformasi Fukunaga Koontz pada $\sum_I$ dan $\sum_E$ dapat diperoleh empat subspace dengan dua kurva eigenvalue seperti pada Gambar 1. Anggap λ_I dan λ_E adalah eigenvalue yang berasosiasi dengan eigenvector yang sama dan λ adalah generalized eigenvalue dari pasangan matriks ($\sum_I, \sum_E$), maka generalized eigenvalue $\lambda = \frac{\lambda_I}{\lambda_E}$ dan $\lambda_I + \lambda_E = 1$.

Berdasarkan referensi [6], J_{MDA} dapat dimaksimalkan dengan menghitung eigenvector yang bersesuaian dengan $\lambda + \frac{1}{\lambda}$. Teori ini dinamakan Analisa Diskriminan Majemuk / Transformasi Fukunaga Koontz, ADM/TFK. ADM memiliki beberapa fitur dengan fitur pertama mengenai subspace diskriminan yang dihasilkan optimal pada jarak Bhattacharyya. Fitur kedua menyatakan bahwa ADM menemukan subspace optimal global dengan memaksimalkan jarak Bhattacharyya yang merupakan batas error Bayes Classifier. Fitur terakhir tentang eigenvector diskriminan yang dihasilkan bisa lebih dari C–1 karena umumnya rank $\sum_I$ dan $\sum_E$ lebih besar dari C–1, batas atas rank S_b.

Berdasarkan *criterion* baru pada persamaan (20), analisa pada Transformasi Fukunaga Koontz menunjukkan bahwa Subspace 1 dan Subspace 3 adalah subspace paling diskriminan. Tetapi, akan beresiko jika menggunakan $\sum_I \in R^{D\times D}$ dan $\sum_E \in R^{D\times D}$ secara langsung. Karena $\sum_I$ dan $\sum_E$ mungkin bersifat singular atau terlalu besar untuk dibentuk. Matriks precursor $H_I \in R^{D\times N_I}$ dan $H_E \in R^{D\times N_E}$ dapat digunakan sebagai alternatif. Namun menggunakan H_E juga tidak efisien karena N_E terlalu besar. Hubungan antara S_t, $\sum_I$, dan $\sum_E$ adalah

$$NS_t = N_I \sum_I + N_E \sum_E \tag{22}$$



Gambar 1: Dekomposisi seluruh data menjadi 4 subspace

Jika didefinisikan $\sum'_I = \frac{N_I}{N}\sum_I$ dan $\sum'_E = \frac{N_E}{N}\sum_E$, maka

$$S_t = \sum{}' I + \sum{}' E \tag{23}$$

Persamaan generalized eigenvalue untuk pasangan matriks ($\sum_I$, $\sum_E$) adalah $\sum_E v = \lambda \sum_I v$.
Karena $\sum_I = \frac{N}{N_I}\sum'_I$ dan $\sum_E = \frac{N}{N_E}\sum'_E$
maka $\frac{N}{N_E}\sum'_E v = \lambda \frac{N}{N_I}\sum'_I v$, sehingga

$$\sum{}' Ev = \frac{N_E}{N_I}\lambda \sum{}' Iv \tag{24}$$

Jika dibandingkan dengan persamaan generalized eigenvalue ($\sum'_I$, $\sum'_E$) adalah $\sum'_E v' = \lambda' \sum'_I v'$, dapat diketahui bahwa

$$v = v' \tag{25}$$

dan

$$\lambda = \frac{N_I}{N_E}\lambda' \tag{26}$$

IMPLEMENTASI ADM/TFK

Berikut adalah tahapan-tahapan implementasi menggunakan algoritma ADM/TFK. Pertama-tama input data awal adalah matriks A. Hitung jumlah seluruh data N dan rata-rata seluruh data m, jumlah N_I sesuai persamaan (17), serta H_t dan H_I sesuai persamaan (7) dan (19). Kemudian terapkan dekomposisi QR pada H_t. Langkah berikut adalah menghitung $\widetilde{S}_t = RR^T$ karena

$$\widetilde{S}_t = Q^T S_t Q = Q^T H_t H_t^T Q = RR^T$$

Asumsi bahwa $Z = Q^T * H_I$, untuk menghitung $\widetilde{\sum}'_I = \frac{N_I}{N} ZZ^T$ karena

$$\widetilde{\sum}'_I = Q^T \sum_I{}' Q = \frac{N_I}{N} Q^T H_I H_I^T Q = \frac{N_I}{N} ZZ^T$$

Selanjutnya hitung $\widetilde{\sum}'_E = \widetilde{S}_t - \widetilde{\sum}'_I$. Hitung eigenvector v dan eigenvalue σ_i dari $\widetilde{\sum}_I^{-1} \widetilde{\sum}'_E$ serta jumlah N_E sesuai persamaan (18). Diawali dengan menghitung generalized eigenvalue λ_i menggunakan eigenvalue σ_i sesuai persamaan (26), kemudian urutkan eigenvector v berdasarkan $\lambda + \frac{1}{\lambda}$. Langkah terakhir pada implementasi ADM/TFK adalah menghitung matriks proyeksi akhir dimana $\phi_{MDA} = QV_k$ (k adalah kolom pertama dari V dan V adalah set eigenvector v_i).

UJI COBA

Uji coba metode ADM/TFK yang dilakukan menggunakan data buatan dan akan dilakukan perbandingan kinerja dengan metode lain. Metode pembanding yang dipergunakan adalah Analisa Komponen Utama (Principal Common Analysis), dan Analisa Diskriminan Linier (Linear Discriminant Analysis).

Uji Coba 1

Uji coba 1 dilakukan dengan menggunakan data yang terdiri dari tiga kelas dan berada pada tiga dimensi. Tujuannya adalah memproyeksikan data pada dua dimensi. Tiap kelas memiliki 100 data, rata-rata tiap kelas adalah nol dan matriks kovarian adalah:
$C_1 = [1,1,0]^T * [1,1,0] + 0.1[0,1,1]^T * [0,1,1]$
$C_2 = [0,1,1]^T * [0,1,1] + 0.1[1,0,1]^T * [1,0,1]$
$C_3 = [1,0,1]^T * [1,0,1] + 0.1[1,1,0]^T * [1,1,0]$

Uji Kebenaran 1

Pertama akan dilakukan pengujian kebenaran pada algoritma ADM/TFK akan ditampilkan **perhitungan langsung** generalized eigenvalue dan eigenvector yang bersesuaian dari $\widetilde{\sum}_I^{-1} \widetilde{\sum}_E$.

Untuk data training sample 1, matriks generalized eigenvalue dan eigenvector yang bersesuaian serta hasil dari $\lambda + \frac{1}{\lambda}$ adalah:

$$\lambda = \begin{pmatrix} 1.0091 & 0 & 0 \\ 0 & 0.9950 & 0 \\ 0 & 0 & 0.9900 \end{pmatrix}$$

$$v = \begin{pmatrix} 0.4167 & 0.6535 & 0.3019 \\ 0.7433 & -0.5956 & -0.6510 \\ 0.5233 & -0.4672 & 0.6965 \end{pmatrix}$$

$$\lambda + \frac{1}{\lambda} = \begin{pmatrix} 2.0001 & 2.0000 & 2.0001 \end{pmatrix}$$

Dari hasil $\lambda + \frac{1}{\lambda}$ diatas dapat diketahui bahwa untuk proyeksi 2 dimensi, eigenvector yang terpilih adalah dari kolom 1 dan 3.

Berikut ini adalah data eigenvalue dan eigenvector, generalized eigenvalue λ dan hasil dari $\lambda + \frac{1}{\lambda}$ yang diperoleh dari Implementasi ADM/TFK:

$$\sigma = \begin{pmatrix} 2.0385 & 0 & 0 \\ 0 & 2.0100 & 0 \\ 0 & 0 & 2.000 \end{pmatrix}$$

$$v = \begin{pmatrix} 0.4167 & 0.6535 & 0.3019 \\ 0.7433 & -0.5956 & -0.6510 \\ 0.5233 & -0.4672 & 0.6965 \end{pmatrix}$$

$$\lambda = \begin{pmatrix} 1.0091 & 0.9950 & 0.9900 \end{pmatrix}$$

$$\lambda + \frac{1}{\lambda} = \begin{pmatrix} 2.0001 & 2.0000 & 2.0001 \end{pmatrix}$$

Untuk proyeksi 2 dimensi, eigenvector terpilih juga berasal dari kolom 1 dan 3.

Uji Efisiensi 1

Biaya komputasional minimal perhitungan langsung
$\widetilde{\sum}_E = ((Q^T *)(H_E * H_E^T) * Q)$ = 270054
$\widetilde{\sum}_I = ((Q^T *)(H_I * H_I^T) * Q)$ = 133704
Biaya komputasional minimal dengan algoritma ADM/TFK
$\widetilde{S}_t$ = $R * R^T$ = 2700, $\widetilde{\sum}_I'$ = (NI/N) * $\widetilde{\sum}_I$ = 133714
Pembentukan matriks $\widetilde{\sum}_E$ dan $\widetilde{\sum}_I$ memerlukan biaya komputasional sebanyak 270054 + 133704 = 403758, sedangkan biaya komputasional ADM/TFK $\widetilde{S}_t$ dan $\widetilde{\sum}_I'$ hanya berjumlah 2700 + 133714 = 136414.

Perbandingan Kinerja 1

Proyeksi menggunakan ADM/TFK dapat dilakukan karena matriks sebaran intraclass $\sum_I$ dan extraclass $\sum_E$ tidak bernilai nol. Hasil proyeksi menunjukkan perpisahan tiga kelas yang cukup jelas meskipun terdapat daerah yang merupakan perpotongan tiga kelas. Selanjutnya hasil proyeksi menggunakan ADM/TFK dibandingkan dengan metode Analisa Komponen Utama dan Analisa Diskriminan Linier.

Hasil proyeksi menggunakan Analisa Komponen Utama menunjukkan penumpukan daerah klasifikasi antara kelas pertama (warna hitam) dan kelas kedua (warna biru). Analisa Komponen Utama gagal melakukan proyeksi pada dua dimensi karena hanya memperhitungkan matriks sebaran total sementara semua kelas memiliki rata–rata nol. Matriks sebaran lain yang menjadi nilai pembeda kelas seperti matriks sebaran within–class, between-class, intra–class, maupun extraclass tidak diperhitungkan.

Proyeksi menggunakan Analisa Diskriminan Linier dan metode apapun yang berbasis Analisa Diskriminan Linier (misalnya: ADL/TFK) tidak dapat dilakukan. Matriks sebaran between–class S_b bernilai nol karena nilai rata–rata pada pada ketiga kelas adalah nol. Akibatnya generalized eigenvalue yang dihasilkan lebih kecil atau sama dengan nol ($\lambda \leq 0$).

Uji Coba 2

Uji coba 2 dilakukan dengan menggunakan data yang hanya memiliki dua kelas dengan tiga dimensi. Kelas pertama memiliki 50 data dengan rata-rata nol dan matriks kovarian $0.5I$. Kelas kedua terdiri dari tiga komponen dengan rata-rata:$[1, 4, 0]$, $\lfloor 2\sqrt{3}, -2, 0 \rfloor$, dan $\lfloor -2\sqrt{3}, 2, 0 \rfloor$. Tiap komponen terdiri dari 25 data dan matriks kovarian $0.5I$. I adalah matriks identitas dengan jumlah baris dan kolom sama dengan jumlah dimensi (dalam hal ini memiliki 3 dimensi).

Uji Kebenaran 2

Hasil perhitungan langsung ($\widetilde{\sum}_I^{-1} \widetilde{\sum}_E$) menghasilkan generalized eigenvalue, eigenvector, serta $\lambda + \frac{1}{\lambda}$:

$$\lambda = \begin{pmatrix} 0.9525 & 0 & 0 \\ 0 & 0.7213 & 0 \\ 0 & 0 & 0.7413 \end{pmatrix}$$

$$v = \begin{pmatrix} 0.6791 & -0.7014 & 0.0059 \\ -0.3702 & -0.5519 & -0.7052 \\ -0.6338 & -0.4510 & 0.7090 \end{pmatrix}$$

$$\lambda + \frac{1}{\lambda} = \begin{pmatrix} 2.0024 & 2.1076 & 2.0902 \end{pmatrix}$$

Dari hasil diatas dapat diketahui bahwa untuk proyeksi 2 dimensi, eigenvector yang terpilih adalah dari kolom 2 dan 3. Berikut ini adalah data eigenvalue dan eigenvector, generalized eigenvalue dan hasil dari yang diperoleh dari Implementasi ADM/TFK:

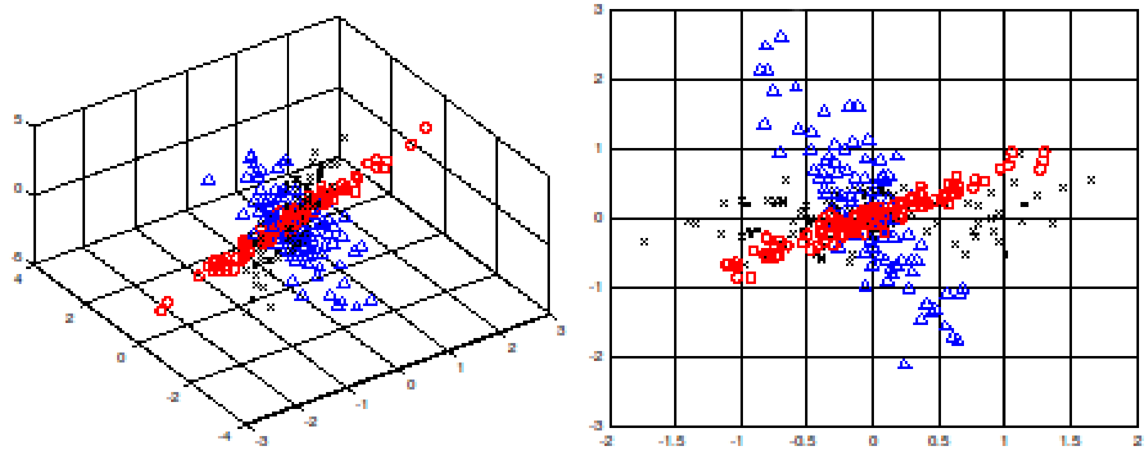
$$\sigma = \begin{pmatrix} 0.8930 & 0 & 0 \\ 0 & 0.6763 & 0 \\ 0 & 0 & 0.6950 \end{pmatrix}$$

$$v = \begin{pmatrix} 0.6791 & -0.7014 & 0.0059 \\ -0.3702 & -0.5519 & -0.7052 \\ -0.6338 & -0.4510 & 0.7090 \end{pmatrix}$$

$$\lambda = \begin{pmatrix} 0.9525 & 0.7213 & 0.7413 \end{pmatrix}$$

$$\lambda + \frac{1}{\lambda} = \begin{pmatrix} 2.0024 & 2.1076 & 2.0902 \end{pmatrix}$$

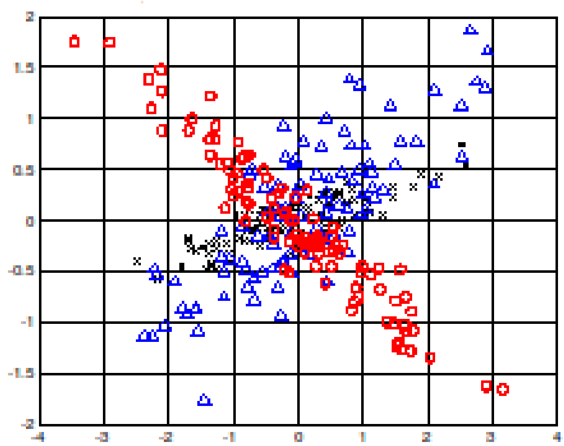
Untuk proyeksi 2 dimensi, eigenvector terpilih juga berasal dari kolom 2 dan 3. Data pada uji coba 1 dan 2 di atas membuktikan bahwa penerapan Transformasi Fukunaga Koontz akan menghasilkan eigenvector yang sama (*shared eigenvector*). Lebih lanjut lagi, metode ADM/TFK yang telah dibuat mampu menghitung eigenvector tanpa menghitung matriks $\sum_E$ secara langsung yang pada umumnya terlalu besar untuk dibentuk.



(a) Representasi data pada 3 dimensi

(b) Proyeksi 2 dimensi dengan ADM/TFK

Gambar 2: Data uji coba 2



Gambar 3: Data uji coba 1: proyeksi 2 dimensi dengan Analisa Komponen Utama

Uji Efisiensi 2

Biaya komputasional minimal perhitungan langsung

$\widetilde{\sum}_E = ((QT * (H_E * H_E^T)) * Q) = 33804$

$\widetilde{\sum}_I = ((QT * (H_I * H_I^T)) * Q) = 36054$

Perhitungan menggunakan algoritma Analisa Diskriminan Majemuk atau Transformasi Fukunaga Koontz. Biaya komputasional minimal dengan algoritma ADM/TFK.

$\widetilde{S}_t = R * R^T = 1125$

$\widetilde{\sum}'_I = (NI/N) * \widetilde{\sum}_I = 36064$

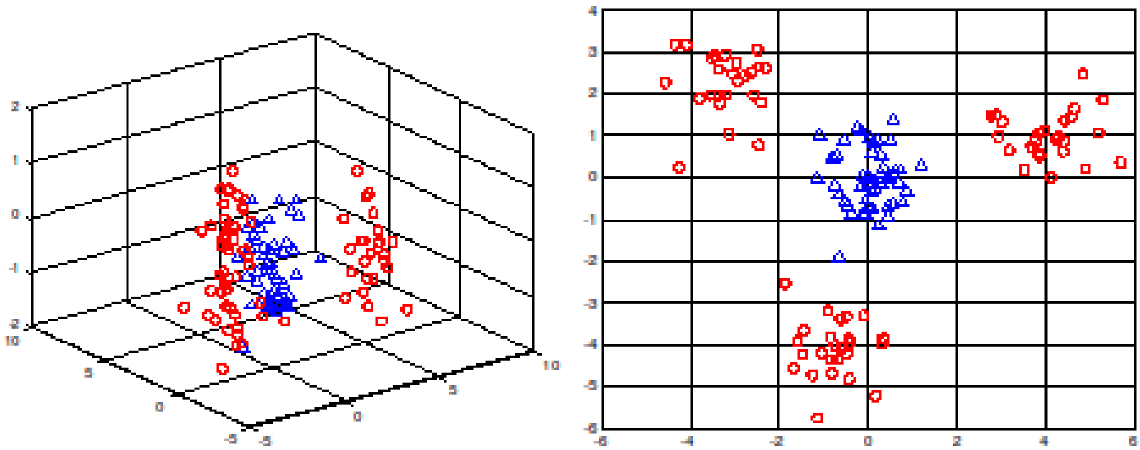
Pembentukan matriks $\widetilde{\sum}_E$ dan $\widetilde{\sum}_I$ memerlukan biaya komputasional sebanyak 33804 + 36054 = 69858, sedangkan biaya komputasional ADM/TFK $\widetilde{S}_t$ dan $\widetilde{\sum}'_I$ hanya berjumlah 1125 + 36064 = 37189.

Dari perbandingan dua metode yang menghasilkan eigenvector yang sama dapat diketahui bahwa perhitungan algoritma ADM/TFK memerlukan biaya komputasional pada pembentukan matriks yang lebih sedikit daripada metode langsung.

Perbandingan Kinerja 2

Hasil proyeksi menunjukkan perpisahan dua kelas yang sangat jelas meskipun kelas kedua terdiri dari tiga komponen yang menyebar disekitar kelas pertama. Hal ini karena analisa yang dilakukan pada matriks sebaran intraclass $\sum_I$ dan extraclass $\sum_E$ mengizinkan untuk melakukan proyeksi mengingat rank($\sum_I$) dan rank($\sum_E$) lebih besar dari C-1 dimana C adalah jumlah kelas. Rank($\sum_I$) dan rank($\sum_E$) diperoleh dari jumlah data sample yang digunakan bukan jumlah kelas. Jika subspace diskriminatif pada ADM/TFK semakin besar maka perpisahan data pada masing-masing kelas dapat dilakukan.

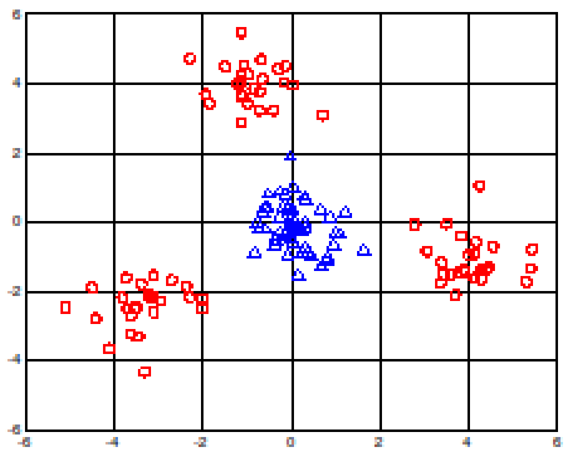
Proyeksi menggunakan Analisa Komponen Utama sangat jelas. Hal ini karena rata-rata yang berbeda antara kelas pertama dengan ketiga komponen pada kelas kedua. Rendahnya nilai matriks kovarian (yaitu 0.5*I) menyebab-



(a) Representasi data pada 3 dimensi

(b) Proyeksi 2 dimensi dengan ADM/TFK

Gambar 4: Data uji coba 2



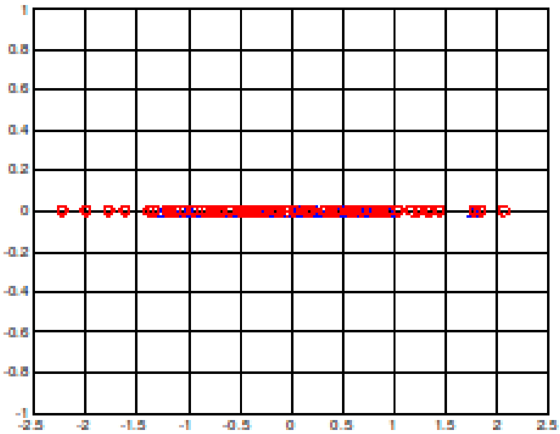
Gambar 5: Data uji coba 1: proyeksi 2 dimensi dengan Analisa Komponen Utama

kan data berkumpul disekitar rata-rata. Sehingga data dapat dibedakan tanpa menggunakan matriks sebaran yang membedakan proyeksi tiap kelas.

Proyeksi menggunakan Analisa Diskriminan Linier berupa garis lurus dan tidak dapat membedakan data pada matriks kelas pertama dan kedua. Proyeksi hanya dapat dilakukan pada dimensi satu karena untuk data dengan dua kelas rank (S_b) = 1. Proyeksi pada dimensi satu tidak dapat digunakan untuk membedakan masing-masing kelas yang ada.

SIMPULAN

Beberapa kesimpulan yang dapat diambil adalah sebagai berikut. Metode Analisa Diskriminan Majemuk atau Transformasi Fukunaga Koontz dibentuk berdasarkan jarak Bhattacharyya yang merupakan batas error dari Bayes Classifier. Sehingga, secara teoritis, metode ini lebih unggul dari Fisher Criterion yang dibentuk berdasarkan matriks sebaran dan tidak berhubungan dengan Bayes Classifier. Metode Analisa Diskriminan Majemuk atau Transformasi Fukunaga Koontz dapat menyediakan subspace diskriminatif yang lebih besar dan tidak dibatasi oleh jumlah kelas. Metode Analisa Diskriminan Majemuk atau Transformasi Fukunaga Koontz dapat bekerja meskipun matriks sebaran between-class $S_b = 0$ dimana semua metode berbasis Analisa Diskriminan Linier gagal diterapkan. Metode Analisa Diskriminan Majemuk atau Transformasi Fukunaga Koontz dapat diterapkan meskipun jumlah data lebih kecil dari jumlah dimensi yang ada. Metode Analisa Diskriminan Majemuk / Transformasi Fukunaga Koontz melakukan proses penghematan space dengan tidak menggunakan $\widetilde{\sum}_E$ secara langsung. Penelitian lebih lanjut dapat dilakukan dengan mengembangkan teori Transformasi Fukunaga Koontz pada analisa subspace diskriminan nonlinier.



Gambar 6: Data uji coba 1: proyeksi 2 dimensi dengan Analisa Diskriminan Linier

DAFTAR PUSTAKA

[1] Anton, H., Rorre, C.: *Elementary Linear Algebra: Application Version*. 7 edn. J. Wiley and Sons (1997)

[2] `http://people.revoledu.com/kardi/tutorial/LDA/index.html`

[3] Zhang, S., Sim, T.: *Discriminant Subspace Analysis: A Fukunaga Koontz Approach*. IEEE Transaction On Pattern Analysis and Machine Intelligence **29**(10) (October 2007) 1732–1745

[4] Duda, R., Hart, P., Stork, D.: *Pattern Recognition*. 2 edn.

[5] Soelaiman, R.: *Sistem Pengenalan Wajah Dengan Penerapan Algoritma Genetika pada Optimasi Basis Eigenface dan Proyeksi Fisherface*. Master's thesis, Universitas Indonesia (2003)

[6] Fukunaga, K.: *Introduction to Statistical Pattern Recognition*. 2 edn. Academic Press (1990)

[7] Jain, A.K., Duin, R.P., Mao, J.: *Statistical Pattern Recognition: A Review*. IEEE Transaction On Pattern Analysis and Machine Intelligence **22**(1) (January 2000) 4–37

[8] Laaksonen, J.: *Subspace Classifiers in Recognition of Handwritten Digits*. Technical report, Helsinki University of Technology (1997)