

Analysis of Recognition Pattern Leaves uses the Method Linear Discriminant Analysis (LDA) and the Distance Minkowski

Dian Ami Widyati¹, R. Rizal Isnanto², Munawar Agus Riyadi³

¹Program Studi Magister Teknik Elektro, Fakultas Teknik, Universitas Diponegoro
Jl. Prof. Sudharto,SH, Tembalang, Semarang,Jawa Tengah, Indonesia, e-mail:
dianamiwidyati@gmail.com

²Program Studi Magister Teknik Elektro, Fakultas Teknik, Universitas Diponegoro
Jl. Prof. Sudharto,SH, Tembalang, Semarang,Jawa Tengah, Indonesia, e-mail:
rizal@ce.undip.ac.id

³Program Studi Magister Teknik Elektro, Fakultas Teknik, Universitas Diponegoro
Jl. Prof. Sudharto,SH, Tembalang, Semarang,Jawa Tengah, Indonesia, e-mail:
munawar@elektro.undip.ac.id

ARTICLE INFO

Article history:

Received 03 January 2021
Received in revised form 06 January
2021
Accepted 15 January 2021
Available online 31 January 2021

ABSTRACT

Indonesia is a tropical country that has a diversity of plant species. Plants consist of three basic organs namely roots, stems and leaves. The leaves are one of the parts that are often used to classify plants, since each type of plant has different characteristics. The shape of leaf edges can be used for reference in leaf classification. The human brain has limitations in processing or extracting information on the types of plants based on leaves. Therefore, it requires the transfer of manual knowledge to a digital system. So in this study made a system that is able to do the introduction of leaves using the extraction of features on the leaves using linear metide Discriminant Analysis (LDA) and minkowski distance.

The process of controlling leaf image patterns begins with the taking of leaf imagery, then goes to the pre-management stage to distinguish objects from the background. After that it goes into the feature extraction stage using Linear Discriminant Analysis (LDA) to obtain characteristic features of the imagery and Minkowski Distance to perform an introduction of the leaf pattern. Based on the results of the study with the amount of data as many as 40 classes with each class as many as 6 images, with a trained image of 160 leaf imagery and test imagery as many as 80 leaf imagery. At the time of introduction using distance Minkowski used 3 coefficients namely minkowski coefficients 1, 2, and 3. From each coefficient minkowski obtained a percentage of accuracy. Minkowski 1 coefficient of 41.25%, minkowski 2 coefficient of 33.75%, and minkowski 3 coefficient of 30%. The percentage of accuracy in this study can not produce an estimated 80% because the amount of data greatly affects the percentage result, the more data there is then the percentage value will also be smaller.

Keywords: leaf, Linear Discriminant Analysis (LDA), Minkowski Distance

1. Pendahuluan

Tumbuhan terdiri dari tiga organ dasar yaitu akar, batang dan daun[1]. Daun merupakan salah satu bagian tanaman yang sering digunakan untuk mengklasifikasi jenis tanaman, karena setiap jenis tanaman memiliki fitur yang berbeda[2]. Ilmu tentang tumbuhan mengalami kemajuan yang pesat salah satunya adalah morfologi tumbuhan yang mempelajari bentuk daun dan susunan tubuh tumbuhan. Bentuk tepian daun bisa digunakan untuk acuan klasifikasi daun. Proses klasifikasi tumbuhan dapat dilakukan dengan cara mengidentifikasi citra bentuk daun dari tumbuhan[3]. Sistem pengenalan dan pengklasifikasian daun yang otomatis sangat berguna karena dapat mendukung pengklasifikasian tumbuhan dengan cepat. Daun merupakan salah satu ciri tumbuhan yang unik dan mudah diamati serta cukup representatif sehingga bisa dijadikan objek untuk ekstraksi fitur tumbuhan[3]. Dalam penelitian ini banyak metode yang dapat digunakan salah satunya adalah metode *Linear Discriminant Analysis* (LDA). Metode *Linear Discriminant Analysis* (LDA) dapat digunakan untuk ekstraksi fitur dalam kelas-kelas yang berbeda[4]. Metode *Linear Discriminant Analysis* (LDA) memiliki kemampuan untuk memisahkan data antar kelas menjadi lebih terpisah dengan cara memaksimalkan nilai *between-class scatter* dan meminimalkan *within-class scatter*. Metode *Linear Discriminant Analysis* (LDA) digunakan untuk mendapatkan karakteristik ciri dari citra dan memperbaiki jarak antar kelas[5]. Metode *Linear Discriminant Analysis* (LDA) digunakan sebagai ekstraksi ciri, dan Jarak Minkowski digunakan untuk mengenali pola dari daun. Sebelum memasuki proses ekstraksi ciri akan melalui proses segmentasi citra dengan melalui beberapa tahapan yaitu proses pemisahan antara objek dengan latar belakang, kemudian dilakukan proses pemotongan citra, setelah citra dipotong dilakukan perubahan ke dalam aras keabuan, kemudian dilakukan proses perubahan dimensi, dimana perubahan dimensi merupakan proses perubahan dari citra menjadi nilai vector, nilai vector yang di dapat akan di jadikan sebagai masukkan dalam ekstraksi ciri menggunakan metode *Linear Discriminant Analysis* (LDA). Masukkan untuk metode *Linear Discriminant Analysis* (LDA) berupa nilai vektor diolah dalam metode ini maka akan menghasilkan nilai keluaran berupa nilai nilai eigen. Nilai eigen yang didapatkan akan mejadi masukkan untuk tahap pengenalan pola dari daun menggunakan jarak minkowski. Keluaran yang dihasilkan dari pengenalan pola berupa nilai eigen. Nilai eigen dengan nilai terkecil akan dikenali sebagai citra uji. Proses pengenalan pola ini data yang akan digunakan berupa gambar citra daun.

2. Metode Penelitian

2.1 Pengolahan Citra Digital

Pengolahan citra merupakan bidang yang bersifat multidisiplin yang terdiri dari banyak aspek antara lain: fisika (optik, nuklir, gelombang dan lain-lain), elektronika, matematika, seni, fotografi dan teknologi komputer. Pengolahan Citra (*images processing*) memiliki hubungan yang sangat erat engan disiplin ilmu yang dinyatakan dalam bentuk masukkan berupa citra dan keluaran berupa citra[6]. Pengolahan citra digital adalah ilmu ilmu yang mempelajari hal- hal yang berkaitan dengan kualitas gambar, transformasi gambar, melakukan pemilihan citra yang optimal, melakukan proses penarikan informasi atau deskripsi objek yang terkandung pada citra. Pengolahan citra dikelompokkan menjadi dua yaitu memperbaiki citra yang sesuai dengan kebutuhan dan mengolah informasi yang terdapat di dalam citra[7].

2.2 *Linear Discriminant Analysis* (LDA)

Metode *Linear Discriminant Analysis* (LDA) diunakan untuk ekstraksi fitur kedalam kelas-kelas yang berbeda[8]. Metode *Linear Discriminant Analysis* (LDA) mampu memisahkan antarkelas menjadi lebih terpisah dengan memaksimalkan nilai *between-class scatter* dan meminimalkan *within-class scatter*. Pada ekstraksi ciri menggunakan *Linear Discriminant Analysis* (LDA) dataset lokasinya tetap namun kelas yang dibentuk menjadi lebih terpisah sehingga dapat menyebabkan jarak antar kelas menjadi lebih besar, sedangkan jarak antar data pelatihan dalam satu kelas akan menjadi lebih kecil[5]. Berikut merupakan Langkah- langkah proses ekstraksi ciri menggunakan *Linear Discriminant Analysis* (LDA)[5].

- a. mengubah matriks dua dimensi menjadi satu dimensi atau ke dalam vektor atau vektor kolom.
- b. Mengkelompokkan data latih kedalam matriks sejumlah kelas () x_i
- c. Hitung rerata (*mean*) dari setiap kelas (μ_i). Perhitungan nilai rerata dari tiap kelas dapat dihitung dengan persamaan (1).

$$\mu_i = \frac{1}{N_i} \sum_{x \in \omega_i} x \quad (1)$$

Jika setiap data pelatihan ditransformasikan dalam bentuk vektor baris, maka perhitungan rerata menggunakan model kolom. Jika data pelatihan ditransformasikan ke dalam bentuk kolom maka rerata dimensi dihitung berdasarkan baris, sehingga jumlah dimensi rerata yang dihasilkan sama dengan jumlah dimensi dataset.

- d. Hitung nilai rerata total dari semua kelas (). Perhitungan nilai rerata total dari keseluruhan kelas dapat dihitung dengan persamaan (2):

$$\mu = \frac{1}{N_1 + \dots + N_c} \quad (2)$$

- e. Hitung matriks *between-class scatter* (S_B) dan matriks *within-class scatter* (S_W). Perhitungan matriks S_B dapat dihitung dengan persamaan (3) sedangkan matriks S_W dapat dihitung dengan menggunakan persamaan (4):

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu)^T (\mu_i - \mu) \quad (3)$$

$$S_W = \sum_{i=1}^c \sum_{j=1}^{N_i} \left((x_j - \mu_i)^T (x_j - \mu_i) \right) \quad (4)$$

Dimana c merupakan jumlah seluruh kelas, N_i adalah jumlah sampel tiap kelas, dimana I adalah mewakili jumlah kelas dari seluruh kelas.

- f. Hitung nilai *covariance* matriks (S). Perhitungan nilai matriks kovarian dapat dihitung dengan menggunakan persamaan (5):

$$ArgMax S = S_B * (S_W)^{-1} \quad (5)$$

Dimana $ArgMax S$ adalah mencari nilai tertinggi dari matriks kovarian yang merupakan matriks hasil reduksi dari ekstraksi *Linear Discriminant Analysis* (LDA) yang memiliki dimensi lebih kecil dibandingkan dengan dimensi matriks citra asli.

- g. Hitung *eigenvalue*(v) dan *eigenvector*(d). Selanjutnya $ArgMax S$ ditentukan maka langkah selanjutnya menentukan matriks *eigenvalue* (λ) matriks *eigenvector*(v) dapat dihitung dengan persamaan (6):

$$\left| (S_W S_B^{-1})^T \lambda I \right| = 0 \quad (6)$$

Setelah dilakukan proses faktorisasi maka akan didapatkan nilai yang nantinya akan digunakan untuk mencari nilai *eigenvector*(v) dapat dihitung dengan persamaan (7)

$$[v, d] = eig(S)$$

$$|S - \lambda I| v = 0 \quad (7)$$

Dimana v merupakan matriks kolom dengan elemen (x_1, x_2, \dots, x_i) di dalamnya matriks inilah yang disebut *eigenvector*. Kedua matriks tersebut dikalikan hingga didapatkan sebuah persamaan (8):

$$(S_{11} - \lambda_1)x_1 + S_{12}x_2 + \dots + S_{cn}x_n = 0 \quad (8)$$

Hasil proyeksi *eigenvector* yang berkorelasi dengan *eigenvalue* lebih mudah dipisahkan

dibandingkan menggunakan *eigenvalue* yang berkolerasi dengan *eigenvalue* lebih kecil.

- h. Hasil matrik proyeksi dan matrik bobot. Perhitungan matrik proyeksi dapat dihitung dengan persamaan (9). Setelah itu, menghitung matriks bobot dari proses ekstraksi LDA dapat dihitung dengan persamaan (10):

$$(9) W_{train} = (x - \mu_i)^T * v$$

$$(10) Eigen_{train} = (x - \mu_i) * v$$

2.3 Jarak Minkowski

Jarak minkowski digunakan untuk mendapatkan jarak terdekat atau terkecil antara citra masukan (citra uji) dengan citra pada basis data sehingga citra masukan dapat didefinisikan. Semakin kecil perbedaan antara dua buah citra, maka jaraknya akan semakin kecil. Begitu juga sebaliknya, semakin besar perbedaan antara dua buah citra maka jaraknya juga akan menjadi semakin besar. Jarak Minkowski mempunyai koefisien $1 \leq \lambda < \infty$ (tak hingga). Contoh nilai vektor masukan adalah $A_i = (A_1, A_2, A_3, \dots, A_n)$ dan nilai vektor citra basis data adalah $B_i = (B_1, B_2, B_3, \dots, B_n)$ maka jarak minkowski dapat dihitung dengan rumus sebagai berikut[9].

Keterangan:

$$D(A, B) = \sqrt[\lambda]{\sum_{i=1}^n (|A_i - B_i|)^\lambda}$$

$D(A, B)$: Jarak Minkowski antara A dan B

A : vektor ciri citra masukan

B : vektor ciri citra basis data

n : Panjang vektor A dan vektor B

λ : koefisien Minkowski

3. Hasil dan Analisa

Pada penelitian ini menggunakan citra daun yang akan di kenali pola daun dengan menggunakan metode *Linear Discriminant Analysis* (LDA) dan jarak Minkowski.

3.1. Pengenalan Citra Latih

Pengenalan citra latih dilakukan pada 160 jenis daun dari 40 kelas. Citra daun yang digunakan untuk pengenalan citra latih diambil 4 citra sebagai citra latih dari setiap kelas ada 6 citra daun dan 2 citra daun yang tersisa akan dijadikan citra uji.

Tabel 1. pengenalan citra latih

Nama File	Kelas Target	Jarak Minkowski	Kelas Keluaran
Alpukat 1	Alpukat	0	Alpukat
Alpukat 2	Alpukat	0	Alpukat
Belimbing 1	Belimbing	0	Belimbing
Belimbing 2	Belimbing	0	Belimbing
Bougenvile 1	Bougenvile	0	Bougenvile
Bougenvile 2	Bougenvile	0	Bougenvile
Binahong 3	Binahong	0	binahong

Pengujian citra latih ini untuk melihat hasil dari jarak Minkowski adalah 0. Dalam program ini jika hasil dari jarak Minkowski adalah 0 maka program ini berfungsi baik. Jika hasil dari jarak Minkowski tidak 0 maka program dalam pengenalan citra latih ini tidak berfungsi dengan baik atau terjadi kesalahan pada program.

3.2. Pengenalan Citra Uji

Pada pengenalan citra uji ini akan menguji 80 citra daun. Melakukan pengujian ini digunakan untuk mengetahui tingkat persentase dari hasil pengujian. Untuk menghitung persentase dari pengujian ini akan menggunakan persamaan (11)

$$p = \frac{\sum \text{citra uji sesuai}}{\sum \text{data citra uji}} \times 100\% \quad (11)$$

Tabel 2, tabel 3, dan tabel 4 merupakan sampel dari pengujian citra dengan koefisien dari jarak Minkowski 1, 2, dan 3, serta dengan sudut 90° dan 180° . Hasil yang ditunjukkan tidak secara keseluruhan hanya di ambil dari beberapa sampel dalam pengujian citra latih dari 80 data.

Tabel 2. pengenalan citra uji dengan koefisien minkowski 1, dengan citra semua sudut

Nama File	Kelas Target	Jarak Minkowski	Kelas Keluaran
Alpukat 5	Alpukat	28,2542	Rambutan
Cabai 5	cabai	21,5456	Cabai
Binahong 6	Binahong	18,5242	Kopi
Jambu biji 6	Jambu biji	12,6161	Jambu biji
Jati 5	Jati	37,2112	Kamboja
Lavender 6	Lavender	12,0739	Lavender
Lemon 5	Lemon	13,2353	Lemon

Tabel 3. pengenalan citra uji dengan koefisien Minkowski 2, dengan citra semua sudut

Nama File	Kelas Target	Jarak Minkowski	Kelas Keseluruhan
Kaca piring 5	Kaca piring	3,0704	Kelengkeng
Kenanga 6	Kenanga	0	Kenangan
Kopi 5	Kopi	1,6553	Sirsak
Lengkuas 6	Lengkuas	2,1932	Lengkuas
Seri 6	Seri	1,7736	Seri
Srikaya 5	Srikaya	1,344	Srikaya
Singkong 6	Singkong	5,138	Alpukat

Tabel 4. pengenalan citra uji dengan koefisien Minkowski 3, dengan citra semua sudut

Nama File	Kelas Target	Jarak Minkowski	Kelas Keseluruhan
Terong 5	Terong	1,2019	Kunyit
Sirsak 5	Sirsak	1,0543	Sirsak
Nangka 6	Nangka	2,3705	Nangka
Mangkok 6	Mangkok	1,8663	Mangkok
Pecut kuda 5	Pecut kuda	2,7747	Jati
Mahkota dewa 6	Mahkota dewa	1,1419	Terong
Kari 5	Kari	3,1508	Kenangan

Hasil pengujian pada penelitian ini dengan menggunakan citra uji sebanyak 80 citra dengan sebanyak 40 kelas. Pada masing masing pengujian terdapat data yang sesuai dan data yang tidak sesuai dengan kelas target yang diinginkan. Dari setiap pengujian dengan koefisien minkowski yang berbeda-beda, akan menghasilkan jumlah citra yang sesuai dan yang tidak sesuai dengan citra uji juga berbeda-beda. Dengan adanya koefisien minkowski yang berbeda-beda juga akan menghasilkan persentase yang berbeda-beda juga. Hasil dari koefisien minkowski 1 menghasilkan citra yang sesuai sebanyak 33 dan yang tidak sesuai sebanyak 47 dengan persentase sebesar 41,25%. Sedangkan hasil dari koefisien minkowski 2 akan menghasilkan citra yang sesuai sebanyak 27 dan yang tidak sesuai sebanyak 53 dengan persentase sebesar 33,75%. Hasil dari

koefisien minkowski 3 menghasilkan citra yang sesuai sebanyak 24 dan yang tidak sesuai sebanyak 56 dengan persentase keberhasilan sebesar 30%. Dari hasil persentase pengujian citra tidak didapatkan persentase keberhasilan diatas angka 80% dikarenakan data yang digunakan begitu banya dengan 40 kelas dan masing-masing kelas terdapat 6 data didalamnya.

4. Kesimpulan

Dari hasil penelitian dan pembahasan dapat disimpulkan yaitu koefisien minkowski mempengaruhi hasil dari jarak yang didapatkan. Penggunaan jumlah citra dalam penelitian ini sangat berpengaruh dengan persentase tingkat keakurasian dari pengenalan. Persentase terendah dalam penelitian ini ada sebesar 30% ketika menggunakan koefisien minkowski 3 dan jumlah data uji yang digunakan sebanyak 80 citra uji. Perlu dilakukan penelitian lebih lanjut dengan menggunakan metode yang berbeda dan dibandingkan dengan penelitian ini.

Referensi

- [1] A. Novitasari, E. P. Purwandari, and F. F. Coastera, "Identifikasi citra daun tanaman jeruk dengan local binary pattern dan moment invariant," *J. Inform. dan Komput.*, vol. 3, no. September, pp. 76–83, 2018.
- [2] S. Y. Riska, L. Cahyani, and M. I. Rosadi, "Klasifikasi Jenis Tanaman Mangga Gadung dan Mangga Madu Berdasarkan Tulang Daun," *J. Buana Inform.*, vol. 6, no. 1, pp. 41–50, 2015, doi: 10.24002/jbi.v6i1.399.
- [3] F. Liantoni, "Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode K-Nearest Neighbor," *J. Ultim.*, vol. 7, no. 2, pp. 98–104, 2016, doi: 10.31937/ti.v7i2.356.
- [4] M. Melinda, "Identifikasi Cacat Amunisi Dengan Menggunakan Pengolahan Citra Digital Dan Linear Discriminat Analysis (LDA)," Institut Teknologi Sepuluh Nopember Surabaya, 2015.
- [5] S. Cahyani, R. Wiryasaputra, and R. Gustriansyah, "Identifikasi Huruf Kapital Tulisan Tangan Menggunakan Linear Discriminant Analysis dan Euclidean Distance," *J. Sist. Inf. Bisnis*, vol. 8, no. 1, p. 57, 2018, doi: 10.21456/vol8iss1pp57-67.
- [6] F. Muwardi and A. Fadlil, "Sistem Pengenalan Bunga Berbasis Pengolahan Citra dan Pengklasifikasi Jarak," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 3, no. 2, p. 124, 2018, doi: 10.26555/jiteki.v3i2.7470.
- [7] L. Ratnawati and D. R. Sulistyningrum, "Penerapan Random Forest untuk Mengukur Tingkat Keparahan Penyakit pada Daun Apel," *J. Sains dan Seni ITS*, vol. 8, no. 2, 2019.
- [8] M. Melinda, "Identifikasi Cacat Amunisi dengan Menggunakan Pengolahan Citra Digital Dan Linear Discriminant Analysis (LDA)," 2015.
- [9] M. Fitriana, "Pengenalan Citra Iris Mata Menggunakan Jarak Minkowski dengan Ekstraksi Ciri Alihragam Gelombang Singkat Daubechies," *Transient*, vol. 4, p. 850, 2015.