

Klasifikasi Jenis Biji Kopi menggunakan Algoritma LDA dan NN

Lutfi Agung Swarga^{1*}, Kukuh Setyadjit², Ahmad Ridhoi³

^{1,2,3} Teknik Elektro, Universitas 17 Agustus 1945 Surabaya, Surabaya

^{1*}lutfiagung@untag-sby.ac.id, ²kukuh@untag-sby.ac.id, ³ridhoi@untag-sby.ac.id

Abstract – Coffee is an agricultural crop that is processed to make a brewed drink that has been ground into powder. Coffee is also a commodity that is often consumed by Indonesian people. It was recorded that in 2001, coffee commodities in Indonesia were able to generate foreign exchange of US\$595.7 million and were ranked first among export commodities in the plantation subsector. However, to find out the type of coffee, you need an expert who knows the type from its shape, because there are similarities in the shape of one type to another. This takes a long time and is inefficient. Therefore, a system was created that determines the types of tobacco from the seeds using two methods: Linear Discriminant Analysis (LDA) algorithms and Artificial Neural Networks (ANN) with data on two types of Arabica and Robusta coffee, which have an accuracy of up to 93%.

Keywords — Agriculture, Coffee, LDA, ANN

Abstrak—Kopi merupakan tanaman agrikultur yang diproses untuk menjadikan minuman hasil seduhan yang telah dihaluskan menjadi bubuk. Kopi juga salah satu komodi yang sering dikonsumsi oleh masyarakat Indonesia. Tercatat pada tahun 2001 komoditas kopi di Indonesia mampu menghasilkan devisa sebesar US\$ 595,7 juta dan menduduki peringkat pertama diantara komoditas ekspor subsektor perkebunan. Kemudian untuk mengetahui jenis kopi perlu adanya pakar yang mengetahui jenis tersebut dari bentuk, dikarenakan adanya kesamaan bentuk dari jenis satu dengan yang lainnya. Hal tersebut memakan waktu lama dan tidak efisien. Maka dari itu dibuatlah system yang mengetahui jenis-jenis tembakau dari bijinya menggunakan 2 metode Algoritma *Linier Discriminant Analysis* (LDA) dan *Artificial Neural Network* (ANN) dengan data 2 jenis kopi *Arabica* dan *Robusta* yang memiliki akurasi rata-rata mencapai diatas 90%.

Kata Kunci—Agrikultur; Coffe; LDA; ANN

I. PENDAHULUAN

Indonesia merupakan salah satu produsen kopi terbesar di dunia. Menurut Indonesia baik.id, Indonesia adalah negara penghasil biji kopi terbesar keempat di dunia setelah Brazil, Vietnam dan Kolombia dengan produksi rata-rata sebesar 639 ribu ton per tahun atau sekitar 8 % dari produksi kopi dunia. Saat ini setelah berkembangnya industry makanan dan minuman, kopi bukan hanya sekadar kebutuhan konsumsi melainkan digunakan sbagai ladang bisnis caffe di Indonesia [1][2][3].

Linier Discriminant Analysis (LDA) adalah salah satu metode algoritma klasifikasi yang digunakan untuk menyelesaikan pengurangan dimensi dan klasifikasi pola. Metode LDA

dengan mudah menangani fitur dan class yang berbeda pada suatu data [4], serta klasifikasi multilabel [5]. Metode ini digunakan klasifikasi jenis tembakau[6], klasifikasi jenis dan bentuk dari wortel sebelum dan sesudah panen[7], dan diagnosis stadium penyakit Alzheimer [8].

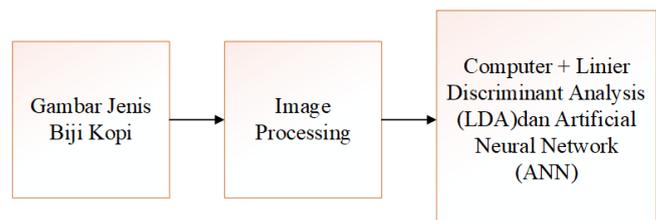
Artificial Neural Networks (ANN) Merupakan metode algoritma klasifikasi yang digunakan untuk menyelesaikan jenis sebuah data baik berupa ciri gambar maupun ciri data numerik. ANN memiliki kelebihan dalam mengatasi klasifikasi dengan data jenis dan data yang besar dengan akurasi yang signifikan[9][10]. Metode ini digunakan untuk klasifikasi hama pada daun tanaman kopi[11], klasifikasi kopi bubuk bermerek menggunakan sensor gas [12] dan klasifikasi penyakit daun kopi menggunakan yolo7[13].

Penelitian ini bertujuan untuk mengklasifikasi biji kopi berdasarkan bentuk gambar. *Image processing* sebagai proses pengambilan gambar menjadi data ciri pada jenis kopi tersebut. Metode LDA dan ANN dilibatkan sebagai algoritma untuk klasifikasi jenis kopi.

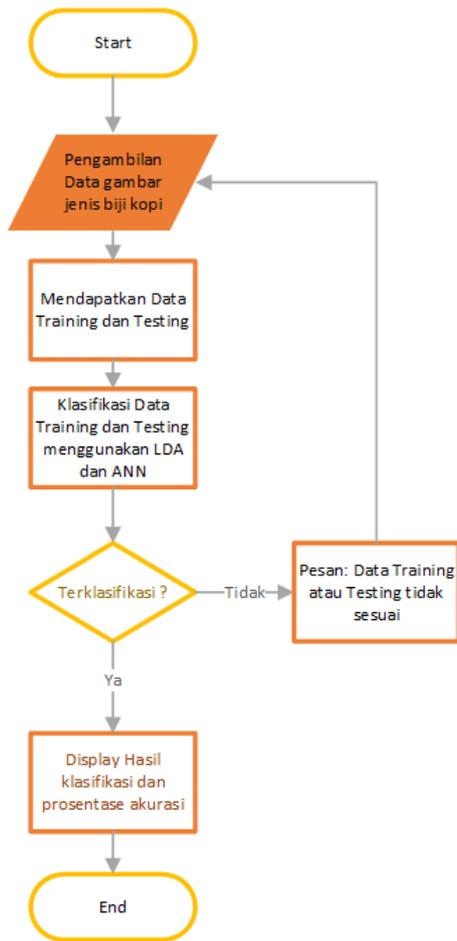
II. METODE PENELITIAN

A. Metode

Sistem klasifikasi jenis biji kopi ini terdiri dari *Image processing* beserta algoritma LDA dan ANN. Bahasa pemrograman pada 7 able 7 ini menggunakan MATLAB R2022a. Jenis biji kopi terdiri dari Arabica dan Robusta. Gambar jenis kopi yang sudah diambil akan di proses pengambilan data pixel yang terdiri dari RGB, HSV, Biner dan GLCM. Kemudian hasil dari data tersebut kemudian digunakan sebagai data latih dan testing bagi metode LDA dan ANN. Metode ini berfungsi untuk mengklasifikasi jenis biji kopi dari bentuk dan warna yang diproses. Diagram blok 7able7 klasifikasi biji kopi ditunjukkan pada Gambar 1.



Gambar 1. Blok diagram 7able7 klasifikasi jenis kopi.

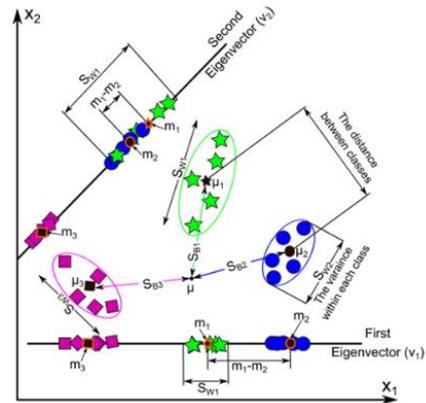


Gambar 2. Flowchart klasifikasi jenis tembakau.

Algoritma proses klasifikasi jenis biji kopi ditunjukkan pada Gambar 2, yang mana dimulai dari pengambilan data gambar biji kopi menggunakan kamera. Setelah pengambilan data training dan data testing, maka dilakukan klasifikasi menggunakan algoritma LDA dan ANN. Jika proses klasifikasi sesuai, maka akan menampilkan hasil berupa classification result dan prosentase akurasi. Jika terjadi error, maka akan memberikan pesan dan melakukan pengambilan data gambar.

A. Linier Discriminant Analysis

Analisis Diskriminan Linier (LDA) adalah metode yang umum untuk masalah pengurangan dimensi sebagai preprocessing untuk pembelajaran mesin dan aplikasi klasifikasi pola [14]. Metode ini memaksimalkan rasio varians between-class, serta within-class dalam kumpulan data tertentu sehingga menjamin keterpisahan maksimal. Gambar 3 adalah contoh visual untuk mengilustrasikan teori LDA.



Gambar 3. Visualisasi dua dimensi untuk 3 kelas yang berbeda [12].

Konsep LDA yang dapat diterapkan pada masalah dua kelas atau lebih dengan rumus sebagai berikut [14],[15].

1. Memberikan set dari N sampel $[x_i]_{i=1}^N$ yang masing-masing merepresentasikan sebagai deret Panjang M .

$$X = \begin{bmatrix} x_{(1,1)} & x_{(1,2)} & \dots & x_{(1,M)} \\ x_{(2,1)} & x_{(2,2)} & \dots & x_{(2,M)} \\ \vdots & \vdots & \vdots & \vdots \\ x_{(N,1)} & x_{(N,2)} & \dots & x_{(N,M)} \end{bmatrix} \quad (4)$$

2. Menghitung mean pada setiap class.

$$\mu_i = \frac{1}{N_j} \sum_{x_i \in \omega_j} x_i \quad (5)$$

3. Menghitung mean total dari semua data.

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i = \sum_{i=1}^c \frac{n_i}{N} \mu_i \quad (6)$$

4. Kemudian menghitung *Between-class* matrix (S_B).

$$S_B = \sum_{i=1}^c n_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (7)$$

5. Kemudian menghitung *Within-class* pada setiap class.

$$S_{W_j} = \sum_{x_i \in \omega_j} (\mu_i - \mu_j)(\mu_i - \mu_j)^T \quad (8)$$

6. Kemudian membuat transformasi matrix pada setiap class (W_i).

$$W_i = S_{W_i}^{-1} S_B \quad (9)$$

- Memproyeksikan sampel pada setiap class menjadi *lower dimensional space*.

$$\Omega_j = x_i V_k^j, x_i \in \omega_j \quad (10)$$

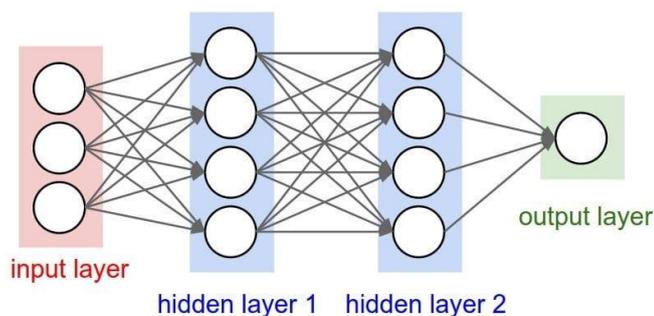
Dimana Ω_j mewakili proyeksi sampel dari class ω_j .

B. Artificial Neural Network

Artificial Neural Network adalah salah satu 9able9 dalam kecerdasan buatan yang terinspirasi oleh cara kerja otak manusia. ANN dirancang untuk mengenali pola, memproses informasi data dengan cara yang mengenali ciri suatu jenis data baik data numerik maupun gambar.

ANN terdiri dari beberapa lapisan neuron (unit pemrosesan) yang saling terhubung. Ada tiga jenis lapisan utama dalam ANN:

- Lapisan Input:** Ini adalah lapisan pertama yang menerima data mentah. Setiap neuron di lapisan ini mewakili fitur atau atribut dari data yang dimasukkan.
- Lapisan Tersembunyi (Hidden Layers):** Setelah data melewati lapisan input, data tersebut diproses melalui satu atau lebih lapisan tersembunyi. Neuron di lapisan ini melakukan perhitungan matematis untuk mengekstrak pola atau fitur dari data. Setiap neuron di lapisan ini terhubung ke neuron di lapisan sebelumnya dan berikutnya, dengan bobot yang mengatur kekuatan koneksi antara neuron.
- Lapisan Output:** Ini adalah lapisan terakhir yang menghasilkan hasil akhir dari jaringan saraf. Neuron di lapisan ini menggabungkan informasi dari lapisan tersembunyi untuk menghasilkan prediksi atau keputusan berdasarkan data yang telah diproses.



Gambar 4. Lapisan Neuron pada algoritma ANN

Konsep ANN juga mempunyai metode lain untuk meningkatkan generalisasi yang diterapkan di Perangkat lunak dengan *regularization & early stopping*. Fungsi kinerja khas yang digunakan untuk pelatihan jaringan saraf umpan maju adalah jumlah rata-rata kuadrat kesalahan jaringan dengan rumus sebagai berikut [].

$$mse = \frac{1}{n} \sum_{i=1}^n (e_i)^2 = \frac{1}{n} \sum_{i=1}^n (X_{real}(k) - X_{predicted}(k))^2 \quad (1)$$

Generalisasi dapat ditingkatkan dengan memodifikasi fungsi kinerja dan menambahkan istilah yang terdiri dari rata-rata jumlah kuadrat bobot jaringan & bias.

$$msereg = \lambda mse + (1 - \lambda) msw \quad (2)$$

Dimana λ adalah rasio kinerja,

$$msw = \frac{1}{n} \sum_{j=1}^n w_j^2 \quad (3)$$

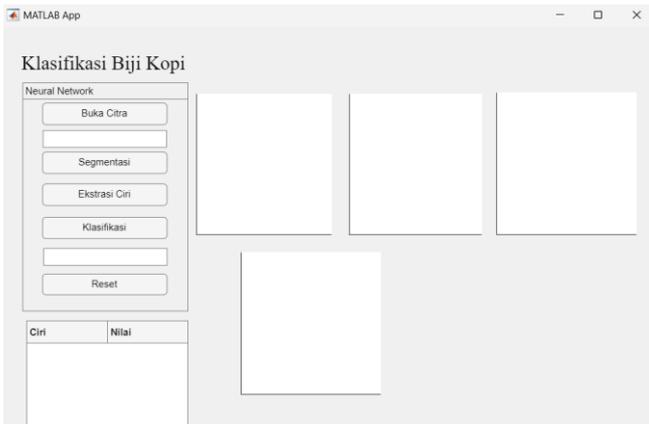
Menggunakan konsep ini memiliki jaringan bobot & bias yang lebih kecil, & ini memaksa respons jaringan menjadi lebih halus & kecil kemungkinannya untuk *over fit* atau terjadi error.

III. HASIL DAN PEMBAHASAN

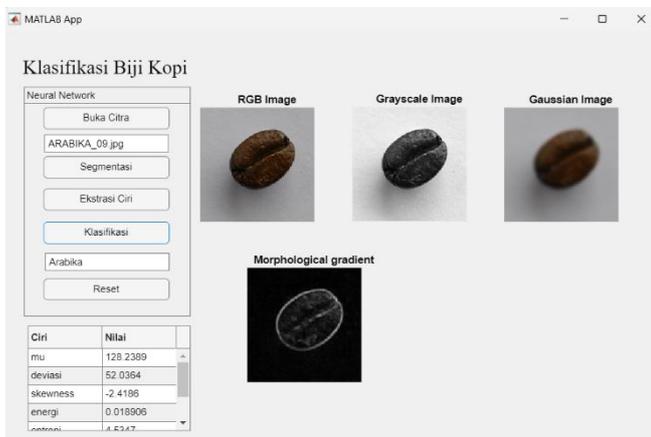
Pada Pada gambar 5 adalah jenis biji kopi yang digunakan dalam penelitian ini. Kemudian 9 able 9 User Interface klasifikasi biji kopi dengan *Image Processing* ditunjukkan pada Gambar 6. Pada tabel 1 menunjukkan data jenis biji kopi dari *image processing* digunakan dalam eksperimen ini. Hal ini menunjukkan bahwa eksperimen pengambilan data gambar diambil sebanyak 70 untuk jenis biji kopi Arabika dan 70 gambar untuk jenis biji kopi robusta yang terbagi menjadi data latih dan testing dengan hasil prosentasi klasifikasi ditunjukkan pada table 2.



Gambar 5. (a) Jenis Kopi Arabika dan (b) Jenis Kopi Robusta.



Gambar 6. User Interface Klasifikasi jenis biji Kopi.

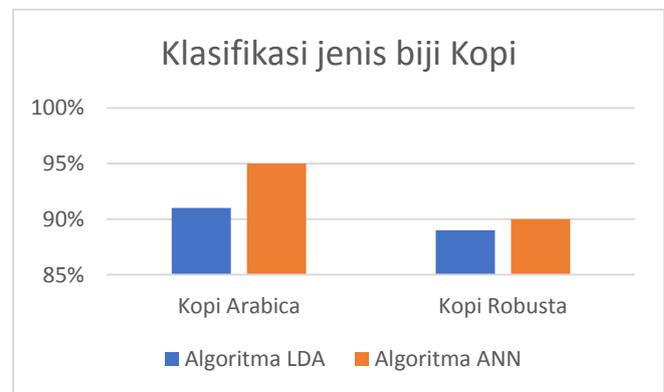


Tabel 1. Data Klasifikasi Jenis Biji Kopi.

Jenis Biji Kopi	Ciri	Nilai
Arabika	Mu	130,058
	Deviasi	55,49
	Skewnes	-3
	Energi	0,020
	Entropi	4,34
	Smoothness	0,96
	Gaussian	177
Robusta	Mu	110
	Deviasi	67
	Skewnes	-0,69
	Energi	0,01
	Entropi	4,97
	Smoothness	0,97
	Gaussian	198

Tabel 2. Prosentase Klasifikasi jenis biji Kopi

Jenis Kopi	Algoritma LDA	Algoritma ANN
Kopi Arabika	91%	95%
Kopi Robusta	89%	90%



Gambar 7. Grafik Prosentase Klasifikasi jenis biji Kopi.

IV. KESIMPULAN

Penelitian ini telah mengembangkan sistem untuk mendeteksi jenis biji kopi menggunakan image processing, algoritma LDA dan algoritma ANN. Data gambar yang sudah diambil kemudian dibagi menjadi data training dan testing yang diumpungkan ke algoritma tersebut. Hasil percobaan menunjukkan bahwa system ini dapat membedakan jenis kopi menggunakan 2 algoritma dengan tingkat keberhasilan rata-rata diatas 90 %.

V. DAFTAR PUSTAKA

- [1] R. Windiarti and A. Kusmiati, "Analisis Wilayah Komoditas Kopi Di Indonesia," *J. Sos. Ekon. Pertan.*, pp. 47–58, 2011.
- [2] D. S. T. Manalu and H. Harianto, "Permintaan Kopi Biji Indonesia di Pasar Internasional," *Agriekonomika*, pp. 114–126, 2020.
- [3] B. Foedinatha and D. D. Hartanto, "Perancangan Aplikasi sebagai Wadah Penggemar Kopi di Indonesia," *J. Desain Komun. Vis. Nirmana*, pp. 38–53, 2022.
- [4] F. Nie, Z. Wang, R. Wang, Z. Wang, and X. Li, "Adaptive local linear discriminant analysis," *ACM Trans. Knowl. Discov. Data*, pp. 1–19, 2020.
- [5] C. Ricciardi, A. S. Valente, and K. Edmund, "Linear discriminant analysis and principal component analysis to predict coronary artery disease," *Health Informatics J.*, pp. 1–

12, 2020.

- [6] L. A. Swarga, M. Rivai, and H. Kusuma, "Tobacco Aroma Classification using EHTS, Gas Sensor Array, and LDA Algorithm," *2022 Int. Semin. Intell. Technol. Its Appl. Adv. Innov. Electr. Syst. Humanit. ISITIA 2022 - Proceeding*, pp. 103–107, 2022.
- [7] A. Jahanbakhshi and K. Kheiralipour, "Evaluation of image processing technique and discriminant analysis methods in postharvest processing of carrot fruit," *Food Sci. Nutr.*, pp. 3346–3352, 2020.
- [8] W. Lin, Q. Gao, and M. Du, "Multiclass diagnosis of stages of Alzheimer's disease using linear discriminant analysis scoring for multimodal data," *Comput. Biol. Med.*, pp. 1–8, 2021.
- [9] J. Zupan, "Introduction to Artificial Neural Network (ANN) Methods : What They Are and How to Use Them.," pp. 327–352, 1994.
- [10] A. D. Dongare, R. R. Kharde, and A. D. Kachare, "Introduction to Artificial Neural Network (ANN) Methods," *Int. J. Eng. Innov. Technol.*, pp. 189–194, 2012.
- [11] D. Irfansyah, M. Mustikasari, and A. Suroso, "Arsitektur Convolutional Neural Network (CNN) Alexnet Untuk Klasifikasi Hama Pada Citra Daun Tanaman Kopi," *J. Inform. J. Pengemb. IT*, pp. 87–92, 2021.
- [12] Maria Ulfa, Haryanto, and Kunto Aji Wibisono, "Desain Sistem Pengenalan dan Klasifikasi Kopi Bubuk Bermerek dengan Menggunakan Electronic Nose Berbasis Artificial Neural Network (ANN)," *J-Eltrik*, pp. 51–60, 2021.
- [13] A. Ardiansyah and N. F. Hasan, "Deteksi dan Klasifikasi Penyakit Pada Daun Kopi Menggunakan Yolov7," *J. Sisfokom (Sistem Inf. dan Komputer)*, pp. 30–35, 2023.
- [14] A. Tharwat and T. Gaber, "Linear discriminant analysis: A detailed tutorial," *AI Commun.*, pp. 1–22, 2017.
- [15] M. J. Zaki and W. Meira, Jr, "Linear Discriminant Analysis," in *Data Mining and Machine Learning*, 2020, pp. 1–7.