

Pengenalan Citra Bunga Menggunakan Segmentasi Otsu Threshold dan Naïve Bayes

Perani Rosyani¹, Oke Hariansyah²

Universitas Pamulang

e-mail: ¹dosen00837@unpam.ac.id, ²dosen00840@unpam.ac.id

Diajukan: 10 Mei 2020; Direvisi: 9 Juni 2020; Diterima: 30 Juni 2020

Abstrak

Bunga merupakan salah satu bagian tumbuhan yang mempunyai warna mencolok dibanding batang kayu dan daun. Bunga merupakan salah satu fitur penting dalam proses pengenalan objek. Proses pengenalan objek pada segmentasi digital sangat penting untuk memisahkan antara background dan foreground dari sebuah citra. Hal ini bertujuan untuk mendapatkan fitur-fitur yang dibutuhkan. Citra bunga pada penelitian ini memiliki kompleksitas gambar cukup sulit karena ada daun dan batang pohon di sekeliling gambar bunga tersebut. Oleh karena itu, dalam penelitian ini segmentasi yang diusulkan menggunakan Otsu Threshold sebagai metode untuk memisahkan foreground dan background. Proses segmentasi sangat menentukan untuk mendapatkan fitur bentuk berupa area, eccentricity, dan perimeter. Hasil kalkulasi dari fitur-fitur tersebut akan diklasifikasi menggunakan algoritma Naïve Bayes dengan menggunakan 120 citra bunga dari dataset 17 flower. Dataset tersebut akan dibagi menjadi data tes dan data training, dan menggunakan cross validasi ($k=10$). Hasil dari klasifikasi menggunakan Naïve Bayes mendapatkan tingkat akurasi yang lebih tinggi sebesar 99.168% dengan relative absolute error sebesar 8.0937% dibandingkan dengan penelitian sebelumnya sebesar 83.83%.

Kata kunci: Pengenalan bunga, Otsu Threshold, Naïve Bayes, Klasifikasi.

Abstract

Flower are one part of the plant that has striking color compared to tree trunks and leaves. Flower is one of the important features in the object recognition process. In the process of introducing digital segmentation objects, it is very important to separate the background and foreground from an image. It aims to get the features needed. Image of flowers in this study has a complexity of images is quite difficult because there are leaves and tree trunks around the flower image. Therefore, in this study the proposed segmentation uses the Otsu Threshold as a method for separating the foreground and background. The segmentation process is crucial to get the shape features in the from area, eccentricity, and perimeters. The results of the calculations of these features will be classified using the naïve Bayes algorithm by using 120 flower images of 17 flower datasets. The dataset will be divided into test data and training data and use cross-validation ($k = 10$). The results of the classification using Naïve Bayes get a higher accuracy rate of 99.168% with a relative absolute error of 8.094% compared with previous studies of 83.83%.

Keywords: Recognition, Flower, Otsu Threshold, Naïve Bayes, Classification.

1. Pendahuluan

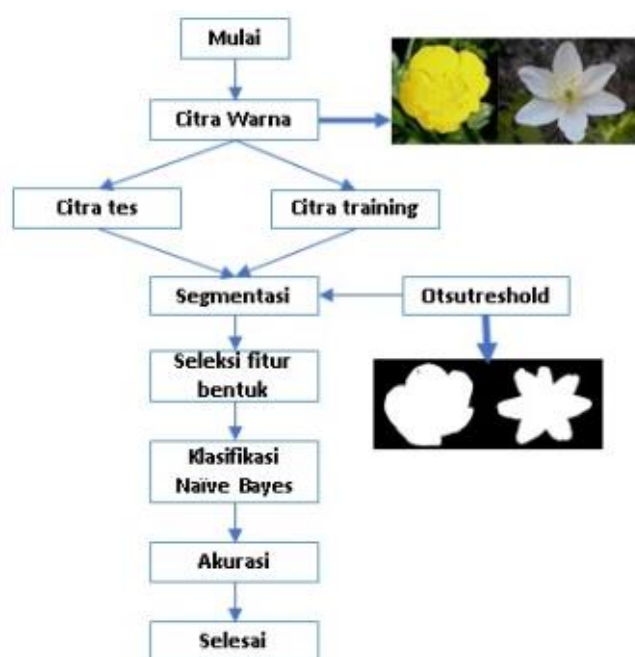
Pengenalan citra digital banyak mendapatkan perhatian dalam beberapa tahun terakhir. Beberapa bidang telah menerapkan pengenalan citra digital di berbagai bidang dan objek, seperti untuk pengenalan citra wajah [1], pengenalan bunga anggrek [2], mengenali pelat kendaraan [3], bahkan di bidang pertanian untuk menentukan kualitas cabai yang bagus [4]. Proses pengenalan objek ini memiliki tahapan-tahapan seperti *preprocessing* yang harus dilalui, yaitu dengan melakukan segmentasi untuk mendapatkan fitur dari objek tersebut. Objek di dalam penelitian ini adalah bunga dengan kondisi *background* yang cukup kompleks. Meskipun daun sangat penting di dalam proses pengenalan tanaman [5], bunga juga termasuk bagian yang mendukung untuk proses tersebut. Pengenalan bunga merupakan proses yang menantang karena beberapa fitur harus dikombinasikan [6], termasuk tekstur warna [7], matriks *co-accurance* tingkat abu-abu, dan filter gabor. Kombinasi beberapa fitur ini untuk menjaga tingkat akurasi tetap tinggi dalam pengenalan objek. Objek yang kami ambil dalam proses pengenalan objek adalah bunga karena bunga merupakan salah satu bagian tumbuhan yang mempunyai warna mencolok dibanding batang dan daun [8].

Dalam penelitian ini, kami menggunakan segmentasi dari Otsu Threshold yang dinilai mudah di dalam proses mendapatkan ciri dari suatu gambar [9]. Segmentasi citra ini bertujuan untuk mendapatkan ekstraksi ciri bentuk [10] dari bunga yaitu area, perimeter, dan *eccentricity*. Ciri bentuk ini didapat setelah melewati penambahan beberapa operasi untuk menghilangkan *noise* di dalam citra seperti operasi *filling holes*, *filtering*, serta *area open*. Hal ini merupakan bentuk optimalisasi setelah proses segmentasi menggunakan Otsu Threshold. Dalam penelitian ini menggunakan Naïve Bayes untuk mengklasifikasi hasil ekstraksi fitur bentuk untuk mengenali objek bunga. Pada penelitian yang dilakukan oleh Bima, dkk menggunakan Naïve Bayes Classifier dengan Eigenface untuk memprediksi tingkat akurasi pengenalan wajah [11], di dalam penelitiannya berhasil meningkatkan tingkat akurasi menjadi 89.5%. sedangkan untuk sistem prediksi menggunakan Naïve Bayes pada penelitian Azman didapat tingkat akurasi sebesar 83.83% [12]

Naïve Bayes Classifier merupakan salah satu metode klasifikasi yang dapat mengenali objek dengan *dataset* minimum [13]. Hal ini sangat menguntungkan karena pada penelitian ini kami menggunakan 120 *dataset* dari 17 *flower* dengan 2 *class* yang berbeda. *Dataset* tersebut akan dibagi menjadi data tes dan data *training* dan menggunakan *cross validasi* (k=10). Tujuan dari penelitian ini adalah untuk mendapatkan tingkat akurasi yang lebih baik daripada penelitian sebelumnya pada proses pengenalan objek, hal ini didukung dengan proses segmentasi citra menggunakan Otsu Threshold dan penambahan operasi-operasi citra lainnya untuk membersihkan *noise*, sehingga objek yang dihasilkan bersih tanpa *noise*.

2. Metode Penelitian

Metode yang diusulkan untuk klasifikasi citra bunga memiliki 2 fase yaitu fase pelatihan dan fase pengujian. Fase pelatihan bertujuan untuk membangun model berdasarkan *subset* gambar yang disebut citra latih [14]. Pertama, gambar tersebut tersegmentasi dengan metode Otsu Threshold dan fitur diekstraksi. Kemudian, fitur akan dihitung dan digunakan untuk pengklasifikasian menggunakan Naïve Bayes.



Gambar 1. Metode penelitian.

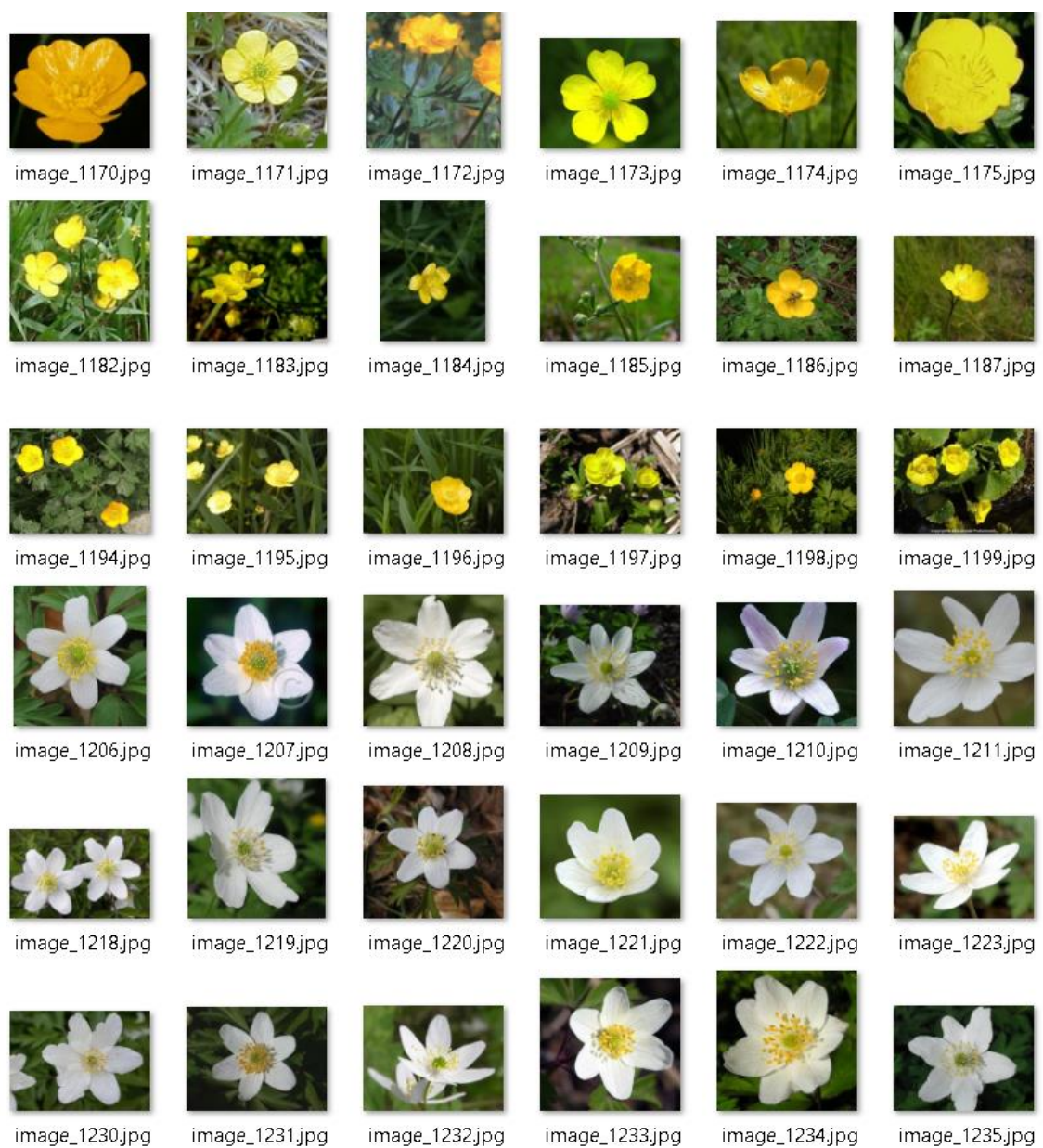
Berikut penjelasan Gambar 1 untuk metode penelitian yang kami usulkan:

1. Pertama adalah pemilihan citra dari data set 17 *flower* di mana *dataset* tersebut mengambil 2 jenis bunga dengan bunga 1 berwarna kuning dan bunga 2 berwarna putih sebanyak 120 citra. Citra tersebut memiliki kompleksitas dari sisi *background*, karena antara *background* dan *foreground* harus dipisahkan untuk mendapatkan objek utuh sehingga dapat dihitung fitur bentuk yang dihasilkan.
2. 120 citra warna yang terpilih akan dibagi menjadi citra latih sebanyak 100 citra dan citra tes sebanyak 20 citra.

3. Setiap citra tes dan citra latih akan disegmentasi menggunakan metode Otsu Threshold, dan melalui operasi-operasi untuk membersihkan *noise* yang ada, seperti *medfilte*, *opening*, *closing*, *filling holes*, dan *openarea*
4. Seleksi fitur bentuk, di dalam tahapan ini hasil segmentasi citra Otsu Threshold akan dihitung nilai perimeter, *eccentricity*, dan areanya.
5. Klasifikasi Naïve Bayes, hasil kalkulasi fitur bentuk akan diklasifikasi menggunakan metode Naïve Bayes dengan tujuan untuk mendapatkan tingkat akurasi dari pengenalan citra bunga ini.
6. Setelah fitur dihitung menggunakan algoritma Naïve Bayes kita akan dapatkan nilai akurasi dari perhitungan tersebut.

2.1. Image Data Colletion

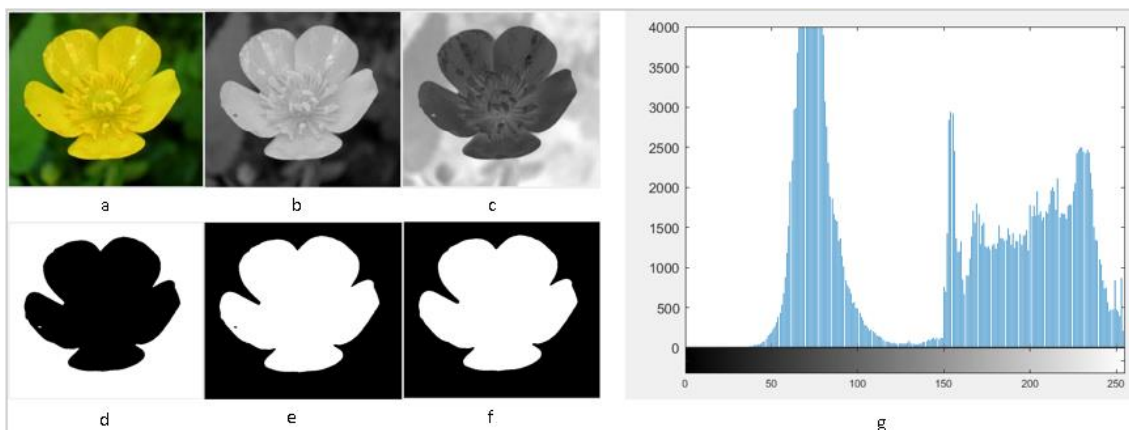
Dataset yang digunakan menggunakan *dataset* dari 17 *flower* sebanyak 120 gambar dengan 2 *class* bunga. Masing-masing dengan 60 gambar yang akan dibagi menjadi 90% citra latih dengan total 50 gambar dan 10% citra uji dengan total 10 gambar.



Gambar 2. Contoh *dataset*.

2.2. Image Data Pre-processing

Dalam tahapan *preprocessing*, langkah yang harus dilewati untuk mendapatkan bentuk yang sempurna, kita memerlukan ekstraksi fitur pada gambar bunga ini. Tahapan tersebut dapat kita lihat pada Gambar 3. Pada keterangan gambar tersebut langkah pertama adalah konversi citra *true color* menjadi citra *grayscale*. Dari citra *grayscale* kita konversi menjadi citra *incomplement*, dan langkah selanjutnya mengubahnya menjadi citra hitam putih. Langkah selanjutnya mengubah nilai matriks yang ada di dalam citra hitam putih menjadi kebalikannya, nilai 0 pada matriks citra tersebut diubah menjadi angka 1, begitu juga sebaliknya, hal ini bertujuan untuk mendapatkan region yang jelas seperti pada poin e. Langkah selanjutnya pada f merupakan citra hasil pembersihan *noise* melalui operasi-operasi citra seperti *filling holes*, *erosi*, *opening*, *closing*, dan *openarea*.



Gambar 3. Proses pengambilan bentuk. a. citra asli; b. citra *grayscale*; c. citra *incomplement*; d. citra biner; e. kebalikan citra biner; f. citra hasil penambahan operasi; g. grafik dari citra biner.

2.3. Citra Training

Citra *training* berisi nilai fitur bentuk perimeter, area, serta *eccentricity*. Citra *training* terdiri dari 100 gambar dengan 2 jenis bunga yang berbeda.

2.4. Naïve Bayes

Naïve Bayes adalah metode pengklasifikasi yang diperkenalkan oleh Thomas Bayes. Metode ini mempelajari data dan memprediksi kelas yang memiliki probabilitas [11].

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{1}$$

Di mana $P(A)$ dan $P(B)$ adalah probabilitas dari pengamatan A dan B [15]. $P(B/A)$ adalah untuk mengamati peristiwa B yang diberikan adalah benar.

Persamaan Naïve Bayes diwakili oleh $P(A/B)$, A adalah *a vector input* yang memiliki fitur dan B adalah label kelas [16]. Berdasarkan pada informasi dari data pelatihan, untuk setiap kombinasi dan B , probabilitas akhir $P(B/A)$ model harus dilatih. Dengan model itu, pengujian data A dapat dideklarasikan dengan pencarian nilai B dengan memaksimalkan nilai $P(A^*/B^*)$. Lalu untuk klasifikasi, rumus Naïve Bayes dapat dinyatakan sebagai persamaan.

$$P(B|A) = \frac{P(Y) \prod_{i=1}^q P(A_i|B)}{P(A)} \tag{2}$$

Di mana $P(B/A)$ adalah probabilitas untuk *vector adikelas* Y . $P(Y)$ adalah probabilitas awal kelas Y .

$\prod_{i=1}^q P(A_i|B)$ adalah probabilitas independen kelas B dari semua fitur dalam *vector* A . nilai $P(A)$ selalu merupakan nilai tetap demikian dalam perhitungan selanjutnya kita perlu menghitung $(B) \prod_{i=1}^q P(A_i|B)$ untuk memilih nilai maksimum dari kelas yang dipilih sebagai hasil prediksi dari probabilitas independen. Selain itu probabilitas independen $\prod_{i=1}^q P(A_i|B)$ adalah pengaruh semua fitur dari data untuk setiap kelas B .

$$P(x = v|c) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(v-\mu)^2}{2\sigma^2}} \tag{3}$$

3. Hasil dan Pembahasan

Hasil dari penelitian ini untuk mendapatkan tingkat akurasi terbaik dari penelitian sebelumnya dengan menerapkan algoritma Naïve Bayes sebagai metode klasifikasi pada proses pengenalan citra bunga. Pengolahan citra menggunakan *software* Matlab dan pengklasifikasian menggunakan Weka *tools* dengan menggunakan algoritma Naïve Bayes.

3.1. Variable Naïve Bayes

Variable di dalam perhitungan Naïve Bayes menggunakan fitur bentuk dari perimeter, area, dan *eccentricity*. Bisa kita lihat pada Tabel 1, hasil statistik dari 2 jenis bunga yang berbeda. Bunga 1 mempunyai *mean* dari *variable* perimeter sebesar 964,54 sedangkan Bunga 2 sebesar 1227,31. Hal ini membuktikan perimeter Bunga 2 lebih besar daripada Bunga 1, hal itu dapat terlihat dari gambar Bunga 1 dan Bunga 2. Dari *variable* area pun terlihat jelas, nilai rata-rata *variable* area pada Bunga 2 lebih besar yaitu 122815,29 dibandingkan dengan *variable* area pada Bunga 1 yang memiliki rata-rata *variable* area sebesar 91845,68.

Tabel 1. Hasil *statistic*.

Class	Bunga 1			Bunga 2		
	Atribut	Mean	Std. Dev	Precision	Mean	Std. Dev
Perimeter	964.5444	586.1713	17.8951	1227.3037	478.6421	17.8951
Area	91845.6853	77201.5471	2481.1171	122815.2973	69753.5237	2481.1171
Eccentricity	0.5519	0.2094	0.009	0.5398	0.2019	0.009

3.2. Indikator Evaluasi

Kita menggunakan 3 indikator evaluasi: *recall rate* (R), *precision rate* (P), dan *F value* (F) sebagai bahan analisis untuk menentukan tingkat akurasi yang baik.

$$R = \frac{\text{objek dikenali dengan benar}}{\text{jumlah semua objek didalam gambar}} \times 100\%$$

$$P = \frac{\text{objek dikenali dengan benar}}{\text{jumlah semua objek didalam dataset}} \times 100\%$$

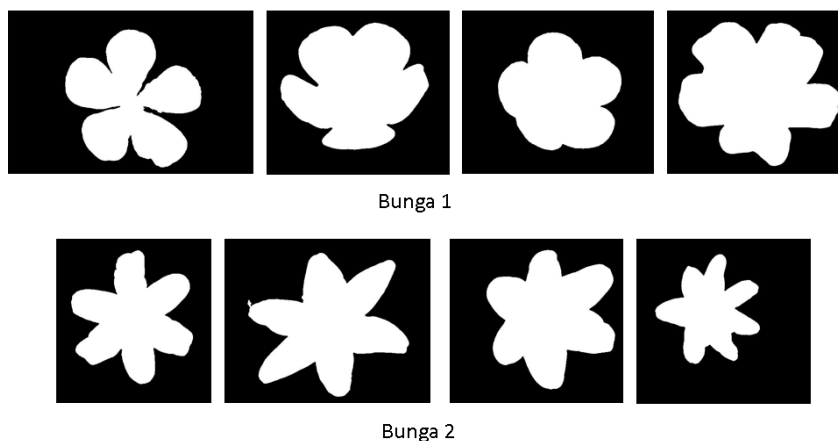
$$F = \frac{2xRxP}{R + P} \times 100\%$$

Dari hasil perhitungan dengan indikator *recall* (R) pada Bunga 1, kita dapatkan hasil akurasi hingga 100% dan 98.3% pada Bunga 2. Untuk indikator *precision rate* (P) pada Bunga 1 sebesar 98.4% dan 100% pada Bunga 2. Untuk indikator *F value* (F) pada Bunga 1 dan Bunga 2 sama-sama mempunyai nilai yang sama yaitu 99.2%. Hal ini dapat dilihat dari Tabel 2 di bawah ini.

Tabel 2. Hasil evaluasi.

	Bunga 1	Bunga 2
TP Rate	100%	98.3%
FP Rate	1.7%	0%
R	100%	98.3%
P	98.4%	100%
F	99.2%	99.2%

Dari hasil uji yang dilakukan dengan menggunakan 120 gambar, dengan menggunakan *cross* validasi (k=10) dapat kita simpulkan bahwa tingkat akurasi yang didapat dengan menggunakan algoritma Naïve Bayes sebesar 99.1667% dengan nilai *relative absolute error* sebesar 8.0937%. Fitur bentuk yang baik dari hasil segmentasi dapat kita lihat dari *sample* pada Gambar 4. Kita bisa lihat bentuknya yang sempurna tanpa *noise*, hal ini sangat mempengaruhi hasil klasifikasi dengan Naïve Bayes, karena di dalam penelitian sebelumnya dengan klasifikasi yang sama hanya mendapatkan tingkat akurasi sebesar 89.5% [11].



Gambar 4. Contoh gambar hasil segmentasi hasil Otsu Threshold.

4. Kesimpulan

Dari hasil penelitian kita bisa ambil kesimpulan sebagai berikut:

1. Berdasarkan penelitian dengan menggunakan 120 citra bunga, maka nilai *False Positive* (FP) pada Bunga 1 sebesar 1,7% sedangkan pada Bunga 2 sebesar 0%.
2. Segmentasi menggunakan Otsu Threshold pada penelitian ini dinilai dapat membersihkan *noise* lebih baik dan meningkatkan tingkat akurasi di dalam proses klasifikasi menggunakan Naïve Bayes sebesar 99,17% dengan tingkat *absolute error* sebesar 8,0937%.

Daftar Pustaka

- [1] P. Rosyani, "Pengenalan Wajah Menggunakan Metode Principal Component Analysis (PCA) dan Canberra Distance," *J. Inform. Univ. Pamulang*, vol. 2, no. 2, p. 118, 2017.
- [2] D. H. Apriyanti, A. M. Arymurthy, and L. T. Handoko, "Identification of orchid species using content-based flower image retrieval," *Proceeding-2013 Int. Conf. Comput. Control. Informatics Its Appl. "Recent Challenges Comput. Control Informatics", IC3INA 2013*, no. March 2015, pp. 53–57, 2013.
- [3] F. Tanoto, E. Wibowo, F. Lutan, and A. Dharma, "Pengenalan Plat Kendaraan Bermotor Dengan Menggunakan," *J. MATRIK*, vol. 19, no. 1, pp. 27–36, 2019.
- [4] F. Ugm, "Klasifikasi Varietas Cabai Berdasarkan Morfologi Daun Menggunakan Backpropagation Neural Network," vol. 10, no. 2, pp. 161–172, 2016.
- [5] S. Bertrand, R. Ben Ameer, G. Cerutti, D. Coquin, L. Valet, and L. Tougne, "Bark and leaf fusion systems to improve automatic tree species recognition," *Ecol. Inform.*, vol. 46, no. 2017, pp. 57–73, 2018.
- [6] D. S. Guru, Y. H. Sharath Kumar, and S. Manjunath, "Textural features in flower classification," *Math. Comput. Model.*, vol. 54, no. 3–4, pp. 1030–1036, 2011.
- [7] S. Inthiyaz, B. T. P. Madhav, and P. V. V. Kishore, "Flower image segmentation with PCA fused colored covariance and gabor texture features based level sets," *Ain Shams Eng. J.*, vol. 9, no. 4, pp. 3277–3291, 2018.
- [8] P. Rosyani, M. Taufik, A. A. Waskita, and D. H. Apriyanti, "Comparison of color model for flower recognition," *2018 3rd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng.*, pp. 10–14, 2019.
- [9] 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.
- [10] I. R. G. A. Sugiarta, M. Sudarma, and I. M. O. Widyantara, "Ekstraksi Fitur Warna, Tekstur dan Bentuk untuk Clustered-Based Retrieval of Images (CLUE)," *Teknologi Elektro*, vol. 16, no. 1, pp. 85–90, 2017.
- [11] E. B. Putranto, P. A. Situmorang, and A. S. Girsang, "Face recognition using eigenface with naive Bayes," *Proc.-11th 2016 Int. Conf. Knowledge, Inf. Creat. Support Syst. KICSS 2016*, no. 4, pp. 9–12, 2017.
- [12] M. A. Maricar and Dian Pramana, "Perbandingan Akurasi Naïve Bayes dan K-Nearest Neighbor pada Klasifikasi untuk Meramalkan Status Pekerjaan Alumni ITB STIKOM Bali," *J. Sist. dan Inform.*, vol. 14, no. 1, pp. 16–22, 2019.
- [13] S. K. Bhakre and A. Bang, "Emotion recognition on the basis of audio signal using Naive Bayes

-
- classifier,” *2016 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2016*, pp. 2363–2367, 2016.
- [14] A. B. Mabrouk, A. Najjar, and E. Zagrouba, “Image flower recognition based on a new method for color feature extraction,” *VISAPP 2014-Proc. 9th Int. Conf. Comput. Vis. Theory Appl.*, vol. 2, 2014.
- [15] N. Ben Amor, S. Benferhat, and Z. Elouedi, “Naive Bayes vs decision trees in intrusion detection systems,” *Proc. ACM Symp. Appl. Comput.*, vol. 1, pp. 420–424, 2004.
- [16] A. Kelemen, H. Zhou, P. Lawhead, and Y. Liang, “Naive Bayesian Classifier for Microarray Data,” *Proc. Int. Jt. Conf. Neural Networks*, vol. 3, pp. 1769–1773, 2003.