

PERANCANGAN SISTEM KLASIFIKASI JENIS BUNGA MAWAR MENGUNAKAN METODE *K-NEAREST NEIGHBOR* (KNN)

Sandi Prayogo¹, Ahmad Abdul Chamid², Alif Catur Murti³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muria Kudus
Email: ¹201851013@std.umk.ac.id, ²abdul.chamid@umk.ac.id, ³alif.catur@umk.ac.id

(Naskah masuk: 13 Mei 2022, diterima untuk diterbitkan: 30 Juni 2022)

Abstrak

Bunga mawar adalah bunga yang sangatlah indah dan memiliki bau yang harum dan juga terkenal di seluruh penjuru dunia. Meskipun bunga mawar menjadi primadona di setiap acara penting, tetapi masih banyak orang yang belum memahami tentang bunga berduri satu ini. Ada banyak varietas mawar, dan setiap varietas memiliki kesamaan sehingga mempersulit dalam membedakan untuk menentukan jenis bunga mawar, diperlukan waktu yang cukup lama dan ketelitian jika membedakan dengan hanya melihat dengan panca indera untuk dapat mengetahui jenis bunga mawar tersebut. Untuk memecahkan permasalahan tersebut, penulis dalam penelitian ini membuat sistem yang bisa mengklasifikasi dan pengelompokan terhadap klasifikasi jenis bunga mawar. Algoritma *K-Nearest Neighbor* (KNN) digunakan saat merancang program aplikasi berbasis pengolahan citra digital. Sehingga memudahkan seseorang untuk mengenali bunga mawar berdasarkan citra pelatihan dengan jarak yang lebih dekat. Dalam penelitian ini menggunakan metode uji *confusion matrix*, lalu diperoleh hasil yaitu akurasi sebesar 92%, presisi sebesar 93% dan *recall* sebesar 92%.

Kata kunci: Bunga Mawar, Pengolahan Citra Digital, *K-Nearest Neighbor* (KNN)

DESIGN OF ROSE TYPE CLASSIFICATION SYSTEM USING *K-NEAREST NEIGHBOR* (KNN) METHOD

Abstract

*The rose is a beautiful flower and has a fragrant smell which is also famous all over the world. Although roses are the prima donna at every important event, many people still don't understand this thorny flower. There are many varieties of roses, and each variety has similarities, making it difficult to distinguish to determine the type of rose, it takes a considerable amount of time and accuracy to distinguish the types of roses with the five senses. To solve the rose problem, the authors of this study create a system that can classify and group flower type classifications. The *K-Nearest Neighbor* (KNN) algorithm is used when designing application programs based on digital image processing. Making it easier for someone to recognize roses based on training image at a closer distance. In this study using the confusion matrix test method, then the results obtained are 92% accuracy, 93% precision and 92% recall.*

Keywords: Rose, Digital Image Processing, *K-Nearest Neighbor* (KNN)

1. PENDAHULUAN

Bunga mawar (*Rosa sp.*) merupakan bunga yang memiliki bau harum dan wangi sehingga banyak disukai oleh masyarakat. Tanaman yang mulai tersebar dari wilayah Eropa Timur, Timur Tengah dan dataran Cina. Perkembangannya sangat pesat, tersebar dari daerah iklim dingin (*subtropis*) sampai daerah panas (*tropis*), dengan perantara orang-orang Belanda mawar mulai masuk ke Indonesia dari Eropa (Rohma dan Jazilah 2020).

Sekarang ini mawar merupakan salah satu bunga yang banyak digemari masyarakat. Hal tersebut tidak diragukan lagi karena hampir seluruh

dunia mengetahui dan mengenal bunga mawar. Disamping warna bunga yang cantik menawan dengan aneka ragam warna-warni. Seakan menghidupkan suasana taman menjadi semarak, ditambah lagi pesona harumnya yang semerbak wangi, bunga mawar juga mempunyai banyak manfaat bagi masyarakat luas. Selain itu, ada berbagai jenis bunga mawar, dan membutuhkan waktu yang cukup lama untuk memahami jenis bunga mawar hanya dengan panca indera, dan akurat. Untuk mengetahuinya, Anda perlu mengidentifikasi jenis mawar.

Berdasarkan permasalahan yang terjadi, dilakukan klasifikasi terhadap jenis bunga mawar menggunakan metode *K-Nearest Neighbor* (KNN) karena mempunyai performa yang tinggi terhadap akurasi dalam klasifikasi citra terhadap objek data pembelajaran berdasarkan objek terdekat. Ekstraksi ciri digunakan untuk mendapatkan nilai, setelah itu objek dikategorikan berdasarkan jarak terdekat.

Dalam pengembangan sistem cerdas untuk pemrosesan gambar digital, pemrosesan gambar digital membantu pecinta mawar mengidentifikasi jenis bunga mawar. Aplikasi citra digital yang dibuat ini bermanfaat dan diharapkan dapat bermanfaat bagi yang membutuhkan informasi tentang jenis bunga mawar.

2. SUMBER PUSTAKA/RUJUKAN

Berikut ini adalah landasan teori berdasarkan penelitian yang saya buat.

2.1. Bunga Mawar

Mawar atau *Rosa hybrida* yang termasuk dalam famili Mawar (*Rosaceae*) merupakan salah satu tanaman hias yang banyak tumbuh dan berkembang di Indonesia. Tidak hanya indah mawar memiliki nilai ekonomi yang tinggi. Bunga mawar banyak digunakan sebagai bunga tangkai, bunga potong, hiasan taman, selain itu bunga mawar juga dimanfaatkan sebagai bunga tabur dan sebagai bahan baku industri kosmetik dan parfum (Muzaki et al. 2021).

2.2. Citra

Citra adalah kombinasi titik, garis, bidang, dan warna untuk menciptakan rupa atau gambar suatu objek, biasanya objek atau orang (Sianturi et al. 2020).

2.3. Pengolahan Citra Digital

Pengolahan citra digital merupakan ilmu yang mempelajari faktor-faktor yang berkaitan dengan peningkatan kualitas citra seperti meningkatkan kontras, mengubah warna, memperbaiki citra, mentransformasi citra (menerjemahkan citra, dan lain-lain), mentransformasikan, mengubah rotasi, skala, geometri, memilih gambar fitur yang optimal untuk analisis, untuk melakukan penyimpanan data yang sebelumnya kecil dan terkompresi, transmisi data, dan waktu pemrosesan, pengolahan data (Munantri et al. 2020).

2.4. *K-Nearest Neighbor*

Algoritma ini tergolong *supervised learning* dengan data eksperimen untuk mengklasifikasikan berdasarkan banyaknya kelas. Algoritma telah dikembangkan oleh Fix and Hodges dan termasuk algoritma klasifikasi sederhana dengan akurasi yang baik. Algoritma *K-Nearest Neighbor* bekerja dengan

mengklasifikasikan data uji dengan mempelajari seberapa dekat data latih. Kelas data uji ini sebagian besar berasal dari kelas data K terbesar, dimana K adalah jumlah data dengan jarak terdekat dari data uji (Abdulbar et al. 2019).

2.5. Confusion Matrix

Confusion matrix adalah ukuran kinerja untuk masalah klasifikasi *machine learning* dimana outputnya dapat berupa dua kelas atau lebih. *Confusion matrix* juga dikenal sebagai matriks kesalahan. Dalam tujuannya, *confusion matrix* tentang mencari perbandingan data hasil pengujian sistem klasifikasi dengan target klasifikasi yang sebenarnya. (Hozairi et al. 2021). *Confusion matrix* ditunjukkan pada tabel 1.

Tabel 1. *Confusion Matrix*

Classification	Predicted Class	
	True	False
Actual: True	True Positif (TP)	False Negatif (FN)
Actual: False	True Negatif (TN)	False Positif (FP)

$$Accuracy = \frac{TP+TN}{Jumlah\ data} \times 100\% \quad (1)$$

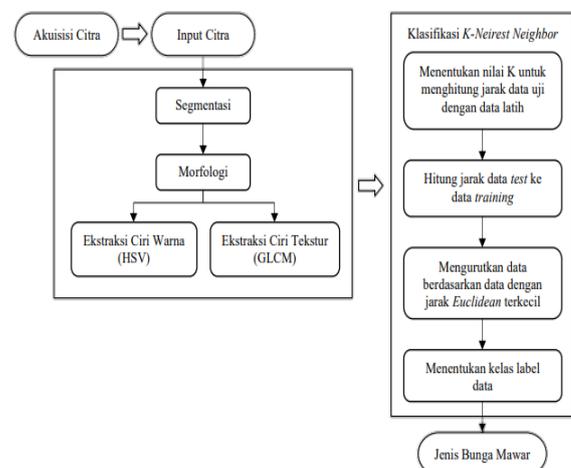
$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

3. METODOLOGI

3.1. Alur Penelitian

Pada tahap ini peneliti mempelajari tentang pengolahan citra digital, dilanjutkan dengan informasi dari jurnal penelitian lain yang dapat digunakan untuk menjelaskan perbedaan yang telah dibuat dan dilakukan sebelumnya. Alur penelitian ditunjukkan pada gambar 1.



Gambar 1. Alur penelitian

3.2. Pengumpulan Data

Tahap pengumpulan data ini adapun langkah-langkah yang harus dilakukan dimulai dengan persiapan objek yang akan diambil citranya, sampai dengan proses pencitraan. Pada proses ini pengambilan citra dilakukan dengan menggunakan kamera *handphone* Infinix hot 9 dengan total *dataset* sebanyak 250 citra dengan perbandingan 90% data *training* dan 10% data *testing*. Data *training* yaitu 225 citra dibagi 2 yaitu 150 citra merupakan citra bunga mawar dan 75 citra merupakan bukan bunga mawar, lalu 25 citra digunakan sebagai data *testing*.

3.3. Preprocessing

Tahap ini beberapa proses yang akan dilakukan, seperti *cropping*, segmentasi, dan menghilangkan *noise*. Proses yang dilakukan adalah mengubah ukuran gambar, memisahkan *background* dari objek bunga mawar, kemudian menggunakan aplikasi Canva untuk mengganti *background* dengan *background* putih, mengubah gambar RGB menjadi gambar HSV, kemudian gambar RGB mengkonversi ke gambar skala abu-abu dengan menggunakan sistem yang dibuat.

3.4. Klasifikasi KNN

Tahap klasifikasi pengujian data ini menggunakan suatu metode klasifikasi untuk menentukan label (kelas) suatu objek baru dengan nilai *k* yang digunakan adalah berjarak sebesar 5 data pada metode KNN.

4. HASIL DAN BAHASAN

4.1. Antarmuka Aplikasi

Aplikasi terdiri dari 3 tampilan yaitu halaman utama, klasifikasi citra dan tentang aplikasi. Pada halaman utama, terdapat 3 sub menu yaitu klasifikasi citra, tentang aplikasi dan keluar. Pada klasifikasi citra terdapat sub menu halaman utama, tentang aplikasi dan keluar. Lalu, pada tampilan tentang aplikasi terdapat sub menu halaman utama, klasifikasi citra dan keluar.

1. Tampilan Halaman Utama

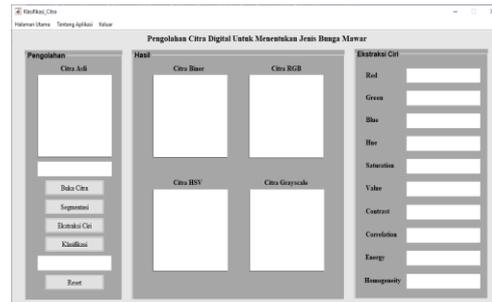
Halaman utama berisi informasi penulis yaitu tema penelitian, logo universitas, program studi, nama universitas dan tahun. Tampilan halaman utama ditunjukkan pada gambar 2.



Gambar 2. Tampilan halaman utama

2. Tampilan Klasifikasi Citra

Pada klasifikasi citra terdapat 3 panel yaitu pengolahan, hasil dan ekstraksi ciri. Panel pengolahan berisi citra asli dan 5 *button* yang digunakan sebagai pengolahan citra yaitu buka citra, segmentasi, ekstraksi ciri, klasifikasi dan *reset*. Panel hasil berisi hasil pengolahan citra asli yaitu citra *biner*, citra RGB, citra HSV dan citra *grayscale*. Tampilan klasifikasi citra ditunjukkan pada gambar 3.



Gambar 3. Tampilan klasifikasi citra

3. Tampilan Tentang Aplikasi

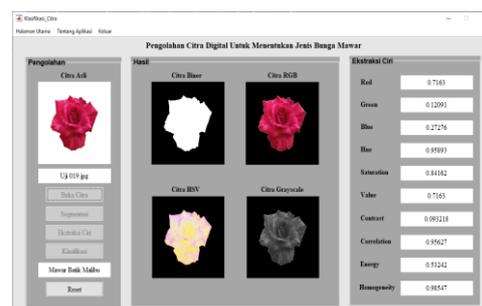
Tampilan ini menampilkan informasi tentang aplikasi yang dibuat. Tampilan tentang aplikasi ditunjukkan pada gambar 4.



Gambar 4. Tampilan tentang aplikasi

4.2. Implementasi Sistem

Dalam implementasi sistem berikut adalah berhasil dijalankannya sebuah sistem untuk mengklasifikasikan citra bunga mawar kedalam jenis bunga mawar tersebut. Sistem klasifikasi disini digunakan dalam mengidentifikasi jenis bunga mawar menggunakan jarak tetangga terdekat. Implementasi sistem ditunjukkan pada gambar 5.



Gambar 5. Implementasi sistem

4.3. Pengujian Sistem

Pengujian ini bertujuan untuk mengetahui akurasi, presisi dan *recall* dalam sistem ini, maka pengujian klasifikasi sistem sangat diperlukan. Sebagai sampel pengujiannya menggunakan citra yang telah disediakan dengan total pengujian ada 25 citra. Berdasarkan informasi diatas hasil pengujian sistem ditunjukkan pada tabel 2.

Tabel 2. Hasil Pengujian Sistem

Citra	Target Uji	Hasil Uji
Uji 001.jpg	Holland Mohana	Holland Mohana
Uji 002.jpg	Holland Mohana	Holland Mohana
Uji 003.jpg	Holland Mohana	Holland Mohana
Uji 004.jpg	Holland Mohana	Holland Mohana
Uji 005.jpg	Holland Mohana	Holland Mohana
Uji 006.jpg	Gertrude Jekyll	Gertrude Jekyll
Uji 007.jpg	Gertrude Jekyll	Gertrude Jekyll
Uji 008.jpg	Gertrude Jekyll	Gertrude Jekyll
Uji 009.jpg	Gertrude Jekyll	Gertrude Jekyll
Uji 010.jpg	Gertrude Jekyll	Gertrude Jekyll
Uji 011.jpg	Baby Rose	Baby Rose
Uji 012.jpg	Baby Rose	Baby Rose
Uji 013.jpg	Baby Rose	Baby Rose
Uji 014.jpg	Baby Rose	Baby Rose
Uji 015.jpg	Baby Rose	Baby Rose
Uji 016.jpg	Batik Malibu	Batik Malibu
Uji 017.jpg	Batik Malibu	Batik Malibu
Uji 018.jpg	Batik Malibu	Batik Malibu
Uji 019.jpg	Batik Malibu	Batik Malibu
Uji 020.jpg	Batik Malibu	Baby Rose
Uji 021.jpg	Juliet	Juliet
Uji 022.jpg	Juliet	Holland Mohana
Uji 023.jpg	Juliet	Juliet
Uji 024.jpg	Juliet	Juliet
Uji 025.jpg	Juliet	Juliet

Metode uji *confusion matrix* digunakan untuk mengukur kinerja klasifikasi pada penelitian ini..

1. Accuracy

Accuracy merupakan persentase dari total data yang telah diidentifikasi dan ditentukan kebenarannya. Rumus akurasi yang digunakan adalah $(TP / \text{Jumlah Data} \times 100\%)$. Untuk mencari nilai TP maka harus mencari pada masing-masing kelas yang ada pada kondisi kelas aktual mampu diprediksi dengan benar. Untuk perhitungan akurasi sebagai berikut:

$$\text{Accuracy} = \frac{TP}{\text{Jumlah Data}} \times 100\%$$

$$\text{Accuracy} = \frac{(5+5+5+4+4)}{25} \times 100\%$$

$$\text{Accuracy} = \frac{23}{25} \times 100\% = 92\%$$

2. Precision

Precision adalah rasio benar positif berdasarkan data yang dikumpulkan dengan informasi yang tidak akurat atau salah. Rumus presisi adalah $\text{Precision} = TP / (TP + FP)$. Pada fase ini, cari nilai FP untuk setiap kelas.

$$\begin{aligned} \text{FP (Mawar Holland Mohana)} &= 1 \\ \text{FP (Mawar Gertrude Jekyll)} &= 0 \\ \text{FP (Mawar Baby Rose)} &= 1 \\ \text{FP (Mawar Batik Malibu)} &= 0 \\ \text{FP (Mawar Juliet)} &= 0 \end{aligned}$$

Setelah Anda menemukan setiap FP, hitung presisi dari setiap kelas.

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\%$$

$$P (\text{Mawar Holland Mohana}) = \frac{5}{5+1} = 0,83$$

$$P (\text{Mawar Gertrude Jekyll}) = \frac{5}{5+0} = 1$$

$$P (\text{Mawar Baby Rose}) = \frac{5}{5+1} = 0,83$$

$$P (\text{Mawar Batik Malibu}) = \frac{4}{4+0} = 1$$

$$P (\text{Mawar Juliet}) = \frac{4}{4+0} = 1$$

Setelah menghitung presisi dari setiap kelas, bagilah jumlah semua hasil rata-rata dengan jumlah kelas.

$$\text{Precision} = \frac{0,83+1+0,83+1+1}{5} \times 100\%$$

$$\text{Precision} = \frac{4,66}{5} \times 100\% = 93\%$$

3. Recall

Recall adalah rasio benar positif berdasarkan data yang gagal diprediksi dengan benar. Rumus *recall* yang digunakan adalah $\text{Recall} = TP / (TP + FN)$. Hampir sama dengan poin kedua, dengan cari FN terlebih dahulu dari masing-masing kelas.

$$\begin{aligned} \text{FN (Mawar Holland Mohana)} &= 0 \\ \text{FN (Mawar Gertrude Jekyll)} &= 0 \\ \text{FN (Mawar Baby Rose)} &= 0 \\ \text{FN (Mawar Batik Malibu)} &= 1 \\ \text{FN (Mawar Juliet)} &= 1 \end{aligned}$$

Setelah Anda menemukan setiap FN, hitung *recall* dari setiap kelas.

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\%$$

$$R (\text{Mawar Holland Mohana}) = \frac{5}{5+0} = 1$$

$$R (\text{Mawar Gertrude Jekyll}) = \frac{5}{5+0} = 1$$

$$R (\text{Mawar Baby Rose}) = \frac{5}{5+0} = 1$$

$$R (\text{Mawar Batik Malibu}) = \frac{4}{4+1} = 0,8$$

$$R (\text{Mawar Juliet}) = \frac{4}{4+1} = 0,8$$

Setelah menghitung *recall* dari setiap kelas, bagilah jumlah semua hasil rata-rata dengan jumlah kelas.

$$Recall = \frac{1+1+1+0,8+0,8}{5} \times 100\%$$

$$Recall = \frac{4,6}{5} \times 100\% = 92\%$$

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Penelitian ini bertujuan dalam menentukan jenis bunga mawar, dan telah mengumpulkan data-data citra bunga mawar. Dengan total *dataset* sebanyak 250 citra dan dibagi dengan perbandingan 90% data *training* dan 10% data *testing*. Maka, data latih yang digunakan adalah 150 citra bunga mawar dan 75 citra bukan bunga mawar, sisanya yaitu 25 citra digunakan sebagai data uji. Pengujian yang dilakukan mendapatkan hasil adalah akurasi sebesar 92%, presisi sebesar 93% dan *recall* sebesar 92%.

5.2. Saran

Dari penelitian yang sudah dilakukan telah mendapatkan hasil dan juga masih terdapat beberapa saran yang harus dilakukan untuk menunjang pada penelitian selanjutnya. Dari hasil penelitian ini hasil ekstraksi ciri citra jenis bunga mawar disini sistem membuat kesalahan dalam klasifikasi, faktor yang mempengaruhi adalah sudut pengambilan gambar dan kedekatan warna yang sangat mirip satu sama lain. Sehingga diperlukan nilai ekstraksi ciri yang dapat mengklasifikasi nilai selain dari ekstraksi ciri dari segi warna saja.

DAFTAR PUSTAKA

- Abdulbar, H., Adikara, P. P., and Adinugroho, S. (2019). Klasifikasi Genre Lagu dengan Fitur Akustik Menggunakan Metode K-Nearest Neighbor. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(8), 8259–8268.
- Hozairi, H., Anwari, A., and Alim, S. (2021). Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes. *Network Engineering Research Operation*, 6(2), 133.
- Munantri, N. Z., Sofyan, H., & Florestiyanto, M. Y. (2020). Aplikasi Pengolahan Citra Digital Untuk Identifikasi Umur Pohon. *Telematika*, 16(2), 97.
- Muzaki, A., Wahyuni, S., and Hanik, N. R. (2021). Identifikasi Jenis Hama Dan Penyakit Yang Sering Menyerang Tumbuhan Bunga Mawar (*Rosa hybrida* L.) DI DAERAH MANYARAN. *Florea : Jurnal Biologi Dan Pembelajarannya*, 8(1), 52.

Rohma, S. I., and Jazilah, S. (2020). Pengaruh Konsentrasi dan Lama Perendaman Rootone F terhadap Pertumbuhan Stek Mawar (*Rosa* sp.). *Biofarm : Jurnal Ilmiah Pertanian*, 15(1).

Sianturi, M. C., Telaumbanua, F., and ... (2020). Pengamanan Citra Digital Dengan Algoritma Paillier Cryptosystem. ... *Teknologi Komputer & ...*, 405–408.