

Klasifikasi Jenis Bunga Berdasarkan Ekstraksi Tekstur GLCM Menggunakan Metode K-Nearest Neighbors

Christian Richie Wijaya¹, Yosefa Camilia Moniung²

Universitas Multi Data Palembang; Jl. Rajawali No. 14, Kota Palembang, Sumatera Selatan
30113, telp: (0711) 376 400

Jurusan Informatika, FIKR UMDP

e-mail: richiewijaya@mhs.mdp.ac.id, camiliayosefa@mhs.mdp.ac.id

Abstrak

Penelitian ini bertujuan untuk mengembangkan sebuah sistem klasifikasi jenis bunga berdasarkan ekstraksi fitur GLCM (Gray Level Co-occurrence Matrix) menggunakan algoritma K-Nearest Neighbors (K-NN). GLCM digunakan untuk mengekstraksi fitur dari citra bunga, yang kemudian digunakan sebagai input untuk algoritma K-NN dalam proses klasifikasi. Metodologi penelitian melibatkan pengumpulan dataset citra bunga yang mencakup bunga daisy dan bunga matahari. Ekstraksi fitur dilakukan dengan menghitung matriks GLCM dari citra, yang selanjutnya diolah untuk mendapatkan representasi fitur tekstur. Selanjutnya, algoritma K-NN diterapkan untuk melakukan klasifikasi berdasarkan fitur tekstur yang diekstraksi. Performa sistem dievaluasi menggunakan metrik-metrik seperti akurasi, presisi, dan recall. Hasil eksperimen menunjukkan bahwa pendekatan ini mampu memberikan hasil klasifikasi yang baik untuk jenis-jenis bunga yang berbeda berdasarkan karakteristik tekstur GLCM dengan akurasi 72.5%.

Kata kunci—bunga, Grey Level Co-Occurrence Matrix, K-Nearest Neighbors, klasifikasi

Abstract

This research aims to develop a flower classification system based on the extraction of GLCM (Gray Level Co-occurrence Matrix) features using the K-Nearest Neighbors (K-NN) algorithm. GLCM is employed to extract features from flower images, which are then used as input for the K-NN algorithm in the classification process. The research methodology involves collecting a dataset of flower images that includes daisy and sunflowers. Feature extraction is performed by calculating the GLCM matrix from the images, which is further processed to obtain a representation of texture features. Subsequently, the K-NN algorithm is applied to classify based on the extracted texture features. The system's performance is evaluated using metrics such as accuracy, precision, and recall. Experimental results indicate that this approach is capable of providing good classification outcomes for different types of flowers based on GLCM texture characteristics with accuracy 76.6667%.

Keywords—keywords flower, Grey Level Co-Occurrence Matrix, K-Nearest Neighbors, Classification



This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

1. PENDAHULUAN

Bunga merupakan struktur reproduksi yang ditemukan pada tanaman berbunga (*angiospermae* atau *magnoliophyta*). Dalam ilmu botani (ilmu yang mempelajari tentang tumbuh-tumbuhan), bunga merupakan bagian dari tanaman untuk menghasilkan biji. Pembuahan serta penyerbukan berlangsung pada bunga. Setelah dilakukan pembuahan, selanjutnya bunga akan berkembang lebih lanjut membentuk buah. Terdapat ratusan ribu spesies atau jenis bunga di dunia ini, dengan beragam bentuk, warna, dan struktur[1]. Dari ratusan ribu spesies bunga, masyarakat akan kesulitan untuk membedakannya.

Maka dari itu, teknologi komputer yang berkembang pesat bisa memudahkan dalam memperoleh informasi untuk mengetahui jenis-jenis bunga. Pengolahan citra *dan machine learning* merupakan salah satu jenis teknologi untuk menyelesaikan masalah itu. Pengolahan citra akan membantu dalam *pre-processing* yang selanjutnya akan dilakukan oleh *machine learning*. *Machine learning* dibagi 3, yakni *supervised learning*, *unsupervised learning*, dan *reinforcement learning*. Dalam menyelesaikan masalah ini, akan digunakan *supervised learning*, yakni proses klasifikasi. Klasifikasi data adalah proses pengurutan dan pengelompokan data ke dalam berbagai jenis, bentuk atau kelas berbeda lainnya[2]. Pada hal ini, klasifikasi bunga diperlukan untuk mengatasi masalah klasifikasi bunga secara manual serta mempersingkat waktu dalam identifikasi bunga. Dalam kasus klasifikasi bunga, pemrosesan gambar adalah langkah penting untuk identifikasi spesies tanaman yang dibantu komputer [3].

Implementasi klasifikasi bunga menggunakan KNN telah dilakukan sebelumnya oleh Lia Farokhah (2020), yaitu Implementasi K-Nearest Neighbor Untuk Klasifikasi Bunga Dengan Ekstraksi Fitur Warna RGB. Penelitian tersebut bertujuan untuk membuktikan kekurangan metode KNN dan ekstraksi fitur menggunakan fitur RGB terhadap karakteristik objek tersebut. Adapun simpulan dalam penelitian itu adalah metode klasifikasi KNN dengan metode ekstraksi fitur warna RGB kurang cocok untuk mengklasifikasi label objek dengan kemiripan warna walaupun bentuknya berbeda, namun masih bisa dipakai dalam mengklasifikasikan objek dengan perbedaan warna yang mencolok di salah satu sisi walaupun bentuk objeknya memiliki struktur bentuk yang mirip.

Penelitian selanjutnya pernah dilakukan oleh Sandi Prayogo, dkk (2022), yaitu Perancangan Sistem Klasifikasi Jenis Bunga Mawar Menggunakan Metode K-Nearest Neighbor (KNN). Penelitian tersebut bertujuan untuk membuat sistem yang dapat mempermudah klasifikasi jenis bunga mawar menggunakan metode KNN karena mempunyai performa yang tinggi terhadap akurasi dalam klasifikasi citra terhadap objek data pembelajaran berdasarkan jarak terdekat.

Kemudian, ada juga penelitian yang dilakukan oleh Rafie (2021), yaitu Klasifikasi Bunga Menggunakan Naïve Bayes Berdasarkan Fitur Warna dan Texture. Penelitian ini bertujuan untuk mengklasifikasi bunga berdasarkan jenis bunga menggunakan teknik pemrosesan citra. Ekstraksi fitur yang digunakan awalnya adalah Hue, Saturation, Value untuk mendapatkan citra warna. Ekstraksi fitur warna dan fitur texture GLCM dapat digunakan untuk ekstraksi ciri pada citra. Hasil percobaan pertama klasifikasi Naive Bayes 7 fitur tingkat akurasinya sebesar 66 %, percobaan kedua dengan distribusi kernel akurasinya 73%. Sedangkan dengan menerapkan 10 fitur dengan distribusi kernel menunjukkan tingkat akurasi lebih besar yaitu 77 %.

Berdasarkan ketiga penelitian tersebut, peneliti memperoleh ide untuk melakukan klasifikasi sederhana terhadap bunga ke dalam 2 kelas, yakni bunga Daisy dan Sunflower menggunakan metode KNN namun dengan ekstraksi fitur tekstur GLCM.

2. METODE PENELITIAN

2.1 Bunga daisy dan matahari



Gambar 1
Bunga Daisy



Gambar 2
Bunga Matahari

Bunga mempunyai banyak jenis yang beraneka ragam, dari warna, bentuk, ukuran, makna, dan lain – lain. Adapun bunga yang terkadang terlihat mirip sehingga sulit dibedakan, contohnya saja bunga matahari dan bunga daisy. Keduanya memiliki inti bunga yang mirip, memiliki kelopak bunga yang banyak. berbentuk mekar, dan batang tumbuhan yang berwarna hijau[4].

2.2 Pengolahan Citra Digital

Citra atau gambar dalam bahasa latin *imago* adalah suatu representasi, kemiripan, atau imitasi dari suatu obyek atau benda. Pengolahan citra merupakan sebuah bentuk pemrosesan sebuah citra atau gambar dengan proses numerik dari gambar tersebut, dalam hal ini yang diproses adalah masing-masing pixel atau titik dari gambar tersebut. Salah satu teknik pemrosesan citra memanfaatkan komputer sebagai peranti lunak memproses masing-masing pixel dari sebuah gambar. Oleh karena itu muncul istilah pemrosesan citra secara digital atau digital image processing[5].

2.3 Grayscale

Pada penelitian ini, citra input yang digunakan merupakan citra berwarna yang terdiri dari 3 layer, yaitu Red, Green, dan Blue (RGB). Untuk mengubah RGB ke grayscale, dapat melalui persamaan:

$$I_y = 0.333Fr + 0.5Fg + 0.1666Fb \quad (1)$$

I_y = image grayscale

Fr = function red

Fg = function green

Fb = function blue

2.4 Metode Otsu

Metode Otsu merupakan salah satu metode untuk segmentasi citra digital dengan menggunakan nilai ambang secara otomatis, yakni mengubah citra digital warna abu-abu menjadi hitam putih berdasarkan perbandingan nilai ambang dengan nilai warna piksel citra digital [6]. Untuk mendapatkan nilai threshold ada perhitungan yang harus dilakukan. Langkah awal yang harus dilakukan adalah membuat histogram. Histogram digunakan untuk mengetahui jumlah piksel untuk setiap tingkat keabuan [7].

2.5 Metode Ekstraksi Tekstur GLCM

Gray Level Co-occurrence Matrix (GLCM) adalah suatu metode yang digunakan untuk analisis tekstur/ekstraksi ciri [8]. Klasifikasi merupakan cara pengelompokkan benda berdasarkan ciri – ciri yang dimiliki oleh objek klasifikasi. Dalam prosesnya, klasifikasi dapat dilakukan dengan banyak cara baik secara manual ataupun dengan bantuan teknologi. Klasifikasi yang dilakukan secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa adanya bantuan dari algoritma cerdas komputer. Sedangkan klasifikasi yang dilakukan dengan bantuan teknologi, memiliki beberapa algoritma, diantaranya Naïve Bayes, Support Vector Machine, Decision Tree, Fuzzy dan Jaringan Saraf Tiruan[9]. Dengan menggunakan persamaan berikut, kita dapat menghitung berbagai fitur yang dapat digunakan untuk melatih pengklasifikasi [10]:

2.4.1 Contrast

$$Contrast = \sum_i \sum_j (i-j)^2 p(i,j) \quad (2)$$

where $p(i,j)$ is the GLCM matrix

2.4.2 Correlation

$$Correlation = \sum_i \sum_j \frac{ijP_d(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y} \quad (3)$$

where μ_x, μ_y and σ_x, σ_y are the mean and standard deviations of probability matrix *GLCM* along row wise *x* and column wise *y*.

2.4.3 Energy

$$Energy = \sum_i \sum_j p^2(i,j) \quad (4)$$

2.4.4 Homogeneity

$$Homogeneity = \sum_i \sum_j \frac{p(i,j)}{1+|i-j|} \quad (5)$$

2.6 Klasifikasi

Klasifikasi merupakan cara pengelompokkan benda berdasarkan ciri – ciri yang dimiliki oleh objek klasifikasi. Dalam prosesnya, klasifikasi dapat dilakukan dengan banyak cara baik secara manual ataupun dengan bantuan teknologi. Klasifikasi yang dilakukan secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa adanya bantuan dari algoritma cerdas komputer. Sedangkan klasifikasi yang dilakukan dengan bantuan teknologi, memiliki beberapa algoritma, diantaranya Naïve Bayes, Support Vector Machine, Decision Tree, Fuzzy dan Jaringan Saraf Tiruan[11].

2.7 Metode K-Nearest Neighbors

Metode K-Nearest Neighbors (KNN) adalah metode klasifikasi yang populer dalam data mining dan statistik karena implementasinya yang sederhana dan kinerja klasifikasi

yang signifikan [12]. Algoritma K-Nearest Neighbour (KNN) akan mengklasifikasikan citra uji ke dalam kelas dengan jumlah anggota terbanyak. Prinsip kerja KNN adalah mencari jarak terdekat antara data yang akan dievaluasi dengan k tetangga (neighbor) terdekatnya dalam data pelatihan[13]. Persamaan 6 akan digunakan untuk menghitung jarak tetangga terdekat menggunakan metode *Eculidean*.

$$D = \sqrt{(x_1 - x_2)^2 - (x_2 - y_2)^2} \quad (6)$$

Keterangan:

x = sample data

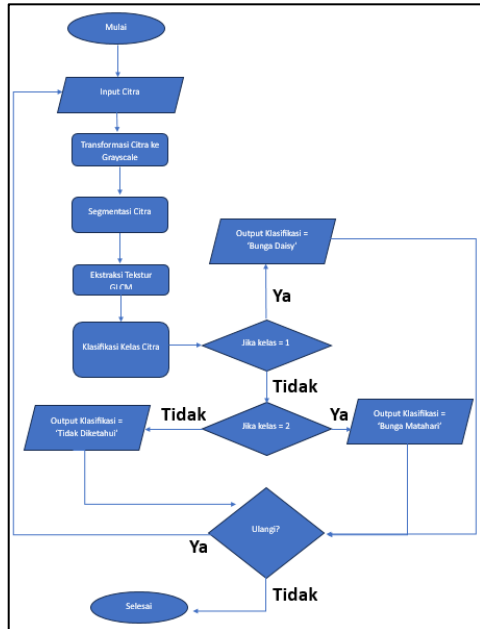
y = data uji

D = jarak euclidean

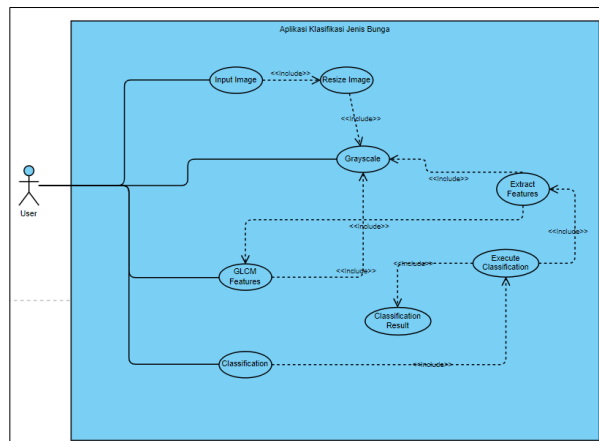
3. HASIL DAN PEMBAHASAN

3.1 Alur Perancangan Sistem

Pada tahap ini peneliti akan menggunakan metode KNN dengan ekstraksi fitur GLCM, yakni *contrast*, *correlation*, *energy*, dan *homogeneity* kemudian mengumpulkan informasi dari jurnal penelitian orang lain yang pernah dibuat sebelumnya sebagai referensi atau gambaran untuk penelitian ini. Berikut bagan keseluruhan perancangan sistem pada gambar 3:



Gambar 3 Flowchart Program

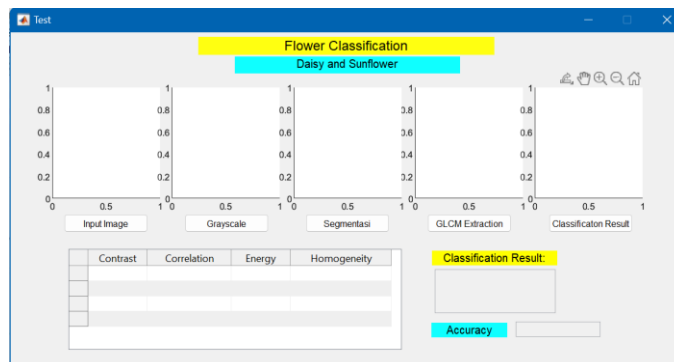


Gambar 4 Use Case Diagram

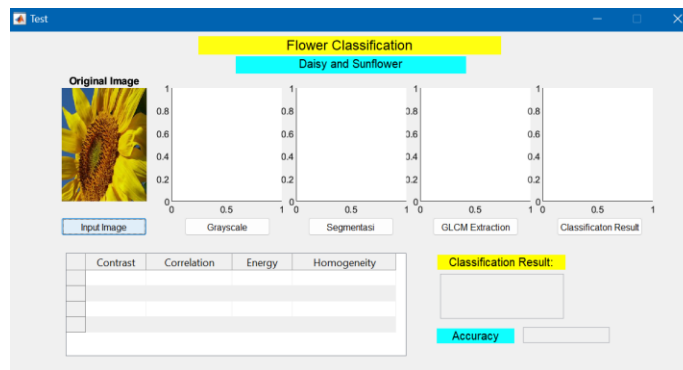
3.2 Penjelasan proses GUI:

3.2.1 Tampilan awal GUI

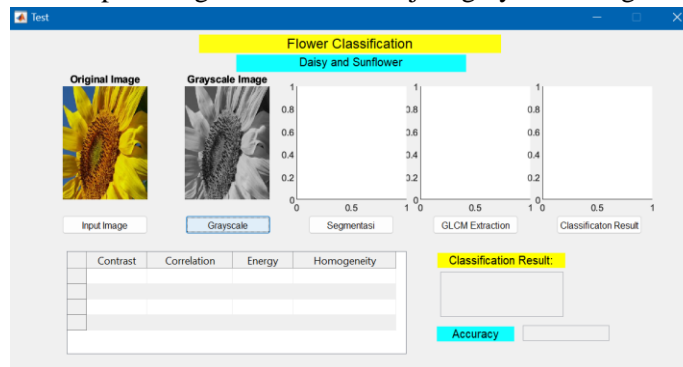
Berisi 4 bagian untuk gambar, empat tombol, yakni tombol *input image*, tombol *grayscale*, tombol *GLCM extraction*, dan tombol untuk menunjukkan *classification result*, tabel yang berisi hasil ekstraksi GLCM, dan tempat munculnya hasil klasifikasi dan akurasi.



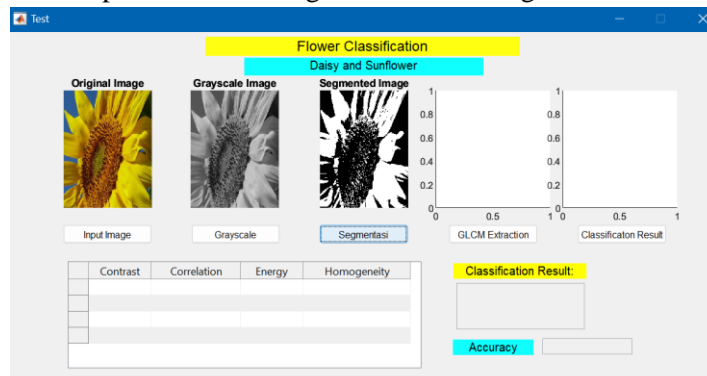
3.2.2 User dapat melakukan input citra yang diinginkan untuk diklasifikasikan dengan menekan tombol *Input Image*.



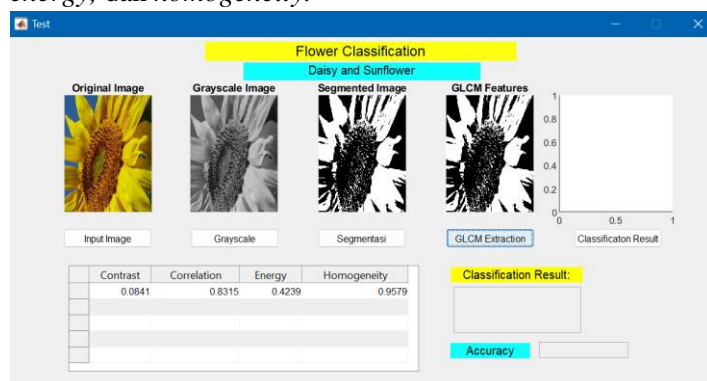
3.2.3 User dapat mengonversi citra menjadi grayscale dengan menekan tombol *Grayscale*



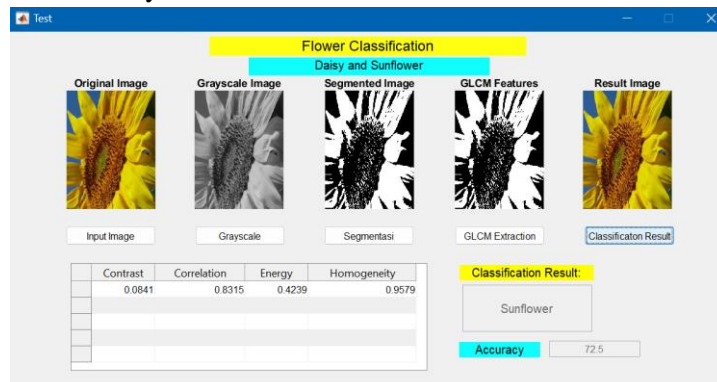
3.2.4 User dapat melakukan segmentasi citra dengan menekan tombol *Segmentasi*



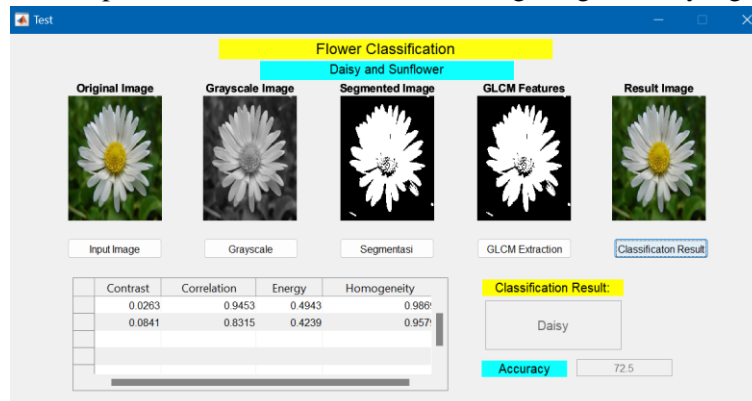
3.2.5 User dapat melakukan ekstraksi tekstur citra dengan menekan tombol *GLCM* dan akan muncul hasil ekstraksi dengan 4 fitur, yakni *contrast*, *correlation*, *energy*, dan *homogeneity*.



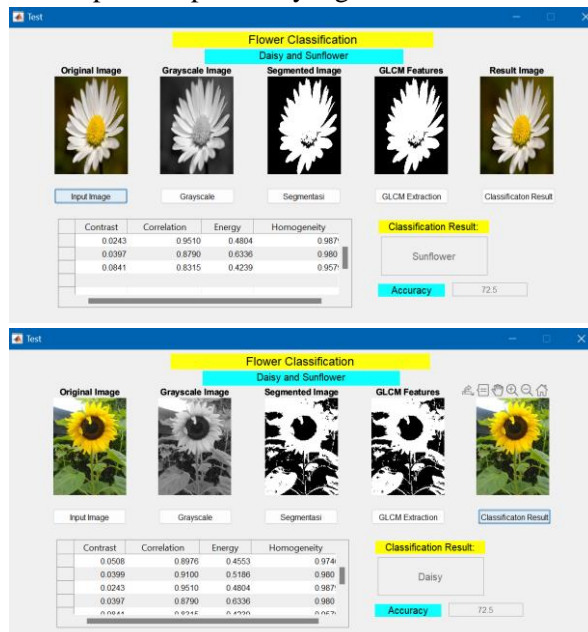
3.2.6 User dapat melakukan klasifikasi citra tersebut adalah bunga daisy atau matahari dengan menekan tombol *Classification Result*. Maka akan muncul hasil klasifikasinya beserta akurasi.

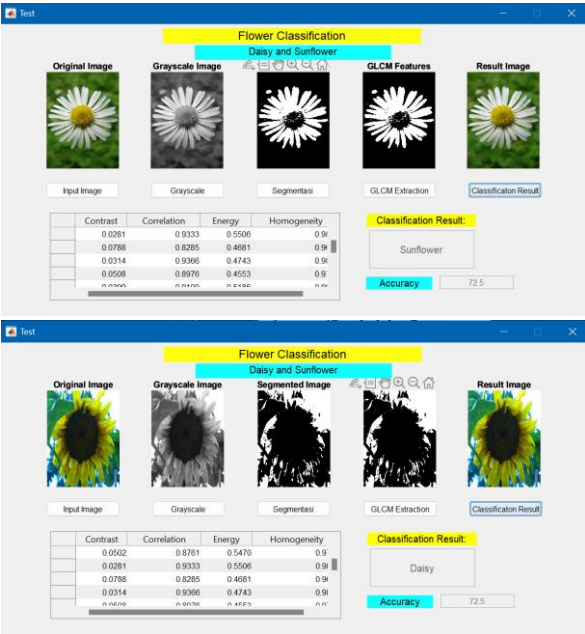


3.2.7 User dapat melakukan hal tersebut berulang dengan citra yang berbeda-beda.



3.3 Beberapa hasil prediksi yang keliru





3.4 Data Citra

Citra yang digunakan dalam penelitian ini adalah Bunga Daisy dan Bunga Matahari (*sunflower*) yang diperoleh dari <https://www.kaggle.com/datasets/marquis03/flower-classification>. Dari sumber tersebut, hanya digunakan citra *sunflowers* dan *daisy* pada folder *train*. Total citra yang digunakan hanya sebanyak 160 citra, dengan masing-masing data citra dibagi dua yaitu data citra latih dan data citra uji. Untuk citra latih digunakan sebanyak 120, dimana terbagi dua kelas yaitu 60 bunga daisy dan 60 bunga matahari dan untuk citra uji digunakan sebanyak 20 untuk setiap kelas.

3.5 Pre-processing

Pada tahap ini peneliti melakukan *pre-processing* terhadap citra. Proses yang dilakukan adalah mengubah ukuran citra menjadi 320x240 *pixels* dapat dilihat pada Tabel 1. Lalu mengubah menjadi citra *grayscale*. Setelah itu, dilakukan segmentasi citra dengan menggunakan *Otsu Color Threshold*.

| Sebelum <i>resize</i> | Sesudah <i>resize</i> |
|-----------------------|-----------------------|
| | |

Tabel 1 Resized Image

3.6 GLCM

Ekstraksi citra tekstur bunga pada penelitian ini akan dilakukan menggunakan metode GLCM (*Gray-Level Co-occurrence Matrix*). GLCM merepresentasikan hubungan antara 2-pixel yang bertetangga (*neighboring pixels*) yang memiliki intensitas keabuan (*grayscale intensity*) dengan sudut default, yakni 0° , 45° , 90° , dan 135° . Pada penelitian ini, peneliti menggunakan 4 fitur ekstraksi GLCM, yakni *contrast*, *correlation*, *energy*, dan *homogeneity*.

3.7 K-Nearest Neighbors

3.5.1 Pelatihan K-NN

Pada tahap ini, fitur atau *input* yang digunakan ada empat yang berasal dari ekstraksi fitur tekstur GLCM yang terdiri dari *Contrast*, *Correlation*, *Energy*, dan *Homogeneity*. Metode yang digunakan untuk pelatihan ini adalah K-NN dengan bahasa pemrograman matlab. Pada proses ini, digunakan *function fitcknn* dari matlab.

3.5.2 Pengujian K-NN

Tahapan setelah pelatihan adalah pengujian yang dilakukan pada seluruh data uji yang telah disediakan. Data uji ini berjumlah 60 yang dibagi 2 kelas, yakni Daisy dan Sunflower. UI sistem yang digunakan seperti pada bagian 3.1 *Alur Perancangan* untuk menguji jenis bunga berdasarkan tekstur menggunakan ekstraksi fitur GLCM dengan klasifikasi K-NN dengan perhitungan jarak yang digunakan adalah *euclidean*.

Setelah diperoleh hasil tabel tersebut, peneliti menggunakan metode *euclidean* untuk perhitungan jarak. Kemudian, dilakukan pengujian lagi untuk memperoleh nilai k yang dapat dilihat pada Tabel 2.

| Nilai k | | | | |
|-----------|---------|---------|---------|---------|
| Akurasi | $k = 1$ | $k = 3$ | $k = 5$ | $k = 7$ |
| | 72.5 % | 67.5% | 65 % | 67.5 % |

Tabel 2 Pengujian Nilai k

Setelah pengujian nilai k , dengan gambar yang sama digunakan $k = 1$ dan diperoleh *confusion matrix* yang ditunjukkan oleh Tabel 3.

| | | Predicted | |
|--------|------------|-----------|------------|
| | | Daisy | Sunflowers |
| Actual | Daisy | 14 | 6 |
| | Sunflowers | 5 | 15 |

Tabel 3 *Confusion Matrix*

Selanjutnya peneliti memperoleh pula Tabel 4 yang berisi *accuracy*, *precision*, dan *recall*.

| | |
|-----------|---------|
| | $k = 1$ |
| Accuracy | 72.5 |
| Precision | 73.68 |
| Recall | 70 |

Tabel 4 Nilai *Accuracy*, *Precision*, *Recall*

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan tentang klasifikasi jenis bunga ke dalam dua kelas, yakni Daisy dan Sunflower serta dilakukan fitur ekstraksi tekstur GLCM, diperoleh hasil yang cukup baik dengan *accuracy* sebesar 72,5% saat menggunakan nilai $k = 1$. Pada penelitian ini, ekstraksi fitur GLCM cukup membantu dalam mengekstrak informasi tekstur yang sulit dilihat secara kasual dan dalam memahami pola dan tekstur yang sulit dipahami manusia.

5. SARAN

Saran-saran untuk penelitian lebih lanjut adalah sebaiknya mencoba mempertimbangkan variasi sudut ataupun fitur GLCM yang digunakan untuk penelitian supaya mendapat hasil yang lebih optimal, serta mencoba menambahkan proses dengan ekstraksi warna.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih kepada kedua peneliti yang sudah membuat jurnal ini dengan sepenuh hati dan kepada Pak Novan Wijaya yang senantiasa membantu sehingga jurnal ini menjadi lebih baik. Peneliti juga mengucapkan terima kasih kepada teman-teman atas dukungan dan bantuannya sehingga project ini bisa selesai dengan baik.

DAFTAR PUSTAKA

- [1] R. Firmansyah, "Implementasi Deep Learning Menggunakan Convolutional Neural Network Untuk Klasifikasi Bunga," vol. 10, p. 6, 2021.
- [2] P. Putra, A. M. H. Pardede, and S. Syahputra, "Analisis Metode K-Nearest Neighbour (Knn) Dalam Klasifikasi Data Iris Bunga," *J. Tek. Inform. Kaputama*, vol. 6, no. 1, pp. 297–305, 2022.
- [3] R. Rafie, "Klasifikasi Bunga Menggunakan Naïve Bayes Berdasarkan Fitur Warna Dan Texture," *J. Sains Komput. Dan Teknol. ...*, vol. 4, no. 1, pp. 90–94, 2021, [Online]. Available: <http://journal.umpalangkaraya.ac.id/index.php/jsakti/article/view/3173>
- [4] M. M. Chandra and Yoanita, "Klasifikasi Jenis Bunga Menggunakan Metode Svm Berdasarkan Citra Dengan Fitur Hsv," *J. Indones. Sos. Teknol.*, vol. 4, no. 2, pp. 255–264, 2023.
- [5] W. Gazali, H. Soeparno, and J. Ohliati, "Dalam Pengolahan Citra Digital," *J. Mat Stat*, vol. 12, pp. 103–113, 2012.
- [6] S. I. Syafi'i, R. T. Wahyuningrum, and A. Muntasa, "Segmentasi Obyek Pada Citra Digital Menggunakan Metode Otsu Thresholding," *J. Inform.*, vol. 13, no. 1, pp. 1–8, 2016, doi: 10.9744/informatika.13.1.1-8.
- [7] A. Tri Utami, "Implementasi Metode Otsu Thresholding untuk Segmentasi Citra Daun," *Fak. Komun. dan Inform. Univ. Muhammadiyah Surakarta*, 2017.
- [8] R. A. Rizal *et al.*, "Analisis Gray Level Co-Occurrence Matrix (GlcM) Dalam Mengenali Citra Ekspresi Wajah," *J. Mantik Augustus Manajemen, Teknol. Informatiak dan Komun.*, vol. 3, no. 2, pp. 31–38, 2019, [Online]. Available: <https://iocscience.org/ejournal/index.php/mantik/index>
- [9] N. Zulpe and V. Pawar, "GLCM textural features for Brain Tumor Classification," *Int. J. Comput. Sci.*, vol. 9, no. 3, pp. 354–359, 2012, [Online]. Available: <http://www.doaj.org/doaj?func=abstract&id=1158398>
- [10] K. Adi, C. E. Widodo, A. P. Widodo, R. Gernowo, A. Pamungkas, and R. A. Syifa, "Detection Lung Cancer Using Gray Level Co-Occurrence Matrix (GLCM) and Back Propagation Neural Network Classification," *J. Eng. Sci. Technol. Rev.*, vol. 11, no. 6, pp. 7–13, 2018, doi: 10.25103/jestr.112.02.
- [11] F. A. D. Aji Prasetya Wibawa, Muhammad Guntur Aji Purnama, Muhammad Fathony Akbar, "Metode-metode Klasifikasi," Malang, 2018.
- [12] S. Zhang, X. Li, M. Zong, X. Zhu, and R. Wang, "Efficient kNN classification with different numbers of nearest neighbors," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 29, no. 5, 2018, doi: 10.1109/TNNLS.2017.2673241.
- [13] E. Budianita, J. Jasril, and L. Handayani, "Implementasi Pengolahan Citra dan Klasifikasi K-Nearest Neighbour Untuk Membangun Aplikasi Pembeda Daging Sapi dan Babi Berbasis Web," *J. Sains dan Teknol. Ind.*, vol. 12, no. Vol 12, No 2 (2015): Juni 2015, pp. 242–247, 2015, [Online]. Available: <http://ejournal.uin-suska.ac.id/index.php/sitekin/article/view/1005>