

KLASIFIKASI JENIS BUNGA MENGGUNAKAN PENGOLAHAN CITRA DIGITAL DAN ALGORITMA DECISION TREE

Hening Tri Haryumi¹, Dimas Pramudya Irawan², Sandi Apriliansyah³, Tazkiyah
Hani Aulia⁴, Nico Brando Rajagukguk⁵, Muhammad Muharrom⁶

Informatika^{1,2,3,4,5,6}, Universitas Bina Sarana Informatika^{1,2,3,4,5,6}
15230037@bsi.ac.id¹, 15230662@bsi.ac.id², 15230711@bsi.ac.id³,
15230013@bsi.ac.id⁴, 15230767@bsi.ac.id⁵, muhammad.muu@bsi.ac.id⁶

*Coressponding Author: 15230037@bsi.ac.id

ABSTRAK

Penelitian ini bertujuan untuk mengotomatisasi identifikasi spesies bunga melalui pengembangan model klasifikasi yang menggunakan Teknik Pengolahan Citra Digital serta algoritma Decision Tree. Pendekatan ini diharapkan dapat mengatasi keterbatasan metode klasifikasi tradisional yang cenderung subjektif dan membutuhkan waktu lama. Studi ini menerapkan metode eksperimental kuantitatif dengan menggunakan dataset berisi 926 gambar yang dibagi ke dalam tiga kategori, di mana data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Gambar digital diproses awal dengan penyesuaian ukuran menjadi 50×50 piksel, kemudian diubah menjadi vektor fitur berdimensi 7500 melalui proses flattening. Model Decision Tree Classifier dilatih menggunakan kriteria Entropy dengan pengaturan hyperparameter tertentu ($\text{max_depth}=10$). Hasilnya menunjukkan bahwa model mencapai tingkat akurasi klasifikasi sebesar 72,6% pada data uji. Kinerja ini diperkuat oleh skor F1-score rata-rata antar kelas (macro avg) sebesar 0,716. Temuan ini membuktikan bahwa integrasi pengolahan citra sederhana dengan Decision Tree merupakan pendekatan yang efektif, efisien, dan mudah diinterpretasikan untuk mengidentifikasi jenis bunga. Model akhir telah berhasil disimpan sebagai `img_model.pkl` dan siap digunakan dalam aplikasi identifikasi spesies di bidang botani atau pertanian modern.

Kata Kunci: -Klasifikasi Bunga, Pengolahan Citra Digital, Decision Tree, Akurasi, Machine Learning.

ABSTRACT

This research aims to automate the identification of flower species by developing a classification model based on Digital Image Processing techniques and the Decision Tree algorithm. This approach is expected to address the drawbacks of conventional classification methods, which are often subjective and time-intensive. The study utilized a quantitative experimental method with a dataset of 926 images divided into three classes, split into 80% for training and 20% for testing. Digital images underwent preprocessing by resizing to 50×50 pixels and flattening into a 7500-dimensional feature vector. The Decision Tree Classifier model was trained using the Entropy criterion with specified hyperparameters ($\text{max_depth}=10$). The results demonstrate that the model achieved a classification accuracy of 72.6% on the test dataset. This performance is backed by a macro-averaged F1-score across classes of 0.716. These findings confirm that combining simple

image processing with Decision Tree is an effective, efficient, and interpretable solution for flower species classification. The final model has been successfully saved as img_model.pkl and is ready for deployment in applications related to botany or modern agriculture.

Keywords: – Flower Classification, Digital Image Processing, Decision Tree, Accuracy, Machine Learning

1. Pendahuluan

Menggolongkan jenis bunga itu hal mendasar dalam ilmu taksonomi tumbuhan, budidaya, serta pengenalan pola gambar. Umumnya, ahli botani mengenali spesies bunga ini dengan cara melihat langsung, dan cara ini gampang dipengaruhi oleh pandangan pribadi serta kemampuan orang yang terbatas. Berkat kemajuan teknologi digital, pengolahan gambar kini menyajikan pilihan yang lebih adil dan cepat untuk membuat proses pengenalan bunga jadi otomatis. Teknik pembelajaran mesin, terutama cara Pohon Keputusan [1], dapat memperbaiki ketepatan pengelompokan dengan memakai ciri-ciri yang diambil dari gambar seperti warna[2], corak, wujud, dan unsur pandangan lainnya.

Tujuan riset ini adalah membuat alat penggolong bunga dengan memakai olah gambar dan cara Pohon Keputusan supaya bisa mengatasi kelemahan cara lama. Tahapan riset ini meliputi pengambilan ciri gambar, melatih model, dan menguji seberapa baik kinerja alat itu dari segi ketepatan dan kegunaannya di lapangan. Alat yang dibuat ini diharapkan dapat dipakai untuk banyak keperluan, termasuk mengenali jenis tanaman di alam liar dan sistem otomatis dalam pertanian modern yang canggih.

Susunan penulisan laporan penelitian ini diatur begini: Bagian 2 mengulas kajian tentang teknik olah gambar dan cara Pohon Keputusan. Bagian 3 menjelaskan cara penelitian yang mencakup pengumpulan data, pembersihan data gambar awal, dan pembuatan model. Bagian 4 menampilkan hasil uji coba serta pembahasan kinerja model. Bagian 5 menyampaikan kesimpulan dan saran untuk penelitian selanjutnya. Lewat riset ini, diharapkan bisa memberi kontribusi pada pengembangan teknologi pengenalan berdasarkan gambar di bidang ilmu data dan botani.

2. Kajian Pustaka

2.1 Pengolahan Citra Digital Dalam Klasifikasi Bunga

Pengolahan citra digital [3] telah menjadi pendekatan utama dalam klasifikasi jenis bunga karena kemampuannya mengekstrak fitur visual seperti warna, tekstur, dan bentuk dari gambar. Teknik ini memungkinkan identifikasi spesies bunga secara otomatis [4], mengatasi keterbatasan klasifikasi manual yang subjektif dan memakan waktu. Salah satu metode umum adalah ekstraksi fitur menggunakan algoritma seperti Histogram of Oriented Gradients (HOG) atau Scale-Invariant Feature Transform (SIFT), yang kemudian digunakan sebagai input untuk model klasifikasi.

Pengolahan citra digital telah terbukti mampu mencapai akurasi hingga 85% dalam klasifikasi bunga iris ketika menggunakan fitur warna dan bentuk sebagai dasar identifikasi. Pendekatan tersebut memanfaatkan dataset Iris dari UCI Machine Learning Repository dan menerapkan segmentasi citra untuk memisahkan objek bunga dari latar belakang. Tahapan pra-pemrosesan seperti normalisasi dan denoising berperan penting dalam mengurangi gangguan visual yang dapat mempengaruhi kualitas ekstraksi fitur.

Pada metode berbasis Convolutional Neural Networks (CNN)[5], akurasi klasifikasi bunga dapat mencapai 92%, meskipun tantangan seperti ukuran dataset yang kecil dapat menyebabkan model

mengalami overfitting. Temuan ini menunjukkan bahwa pengolahan citra digital merupakan pendekatan yang efektif untuk klasifikasi bunga, namun memerlukan dataset yang lebih beragam agar model dapat melakukan generalisasi dengan lebih baik.

2. 2 Algoritma Decision Tree dalam Klasifikasi

Algoritma pohon keputusan merupakan salah satu metode umum dalam klasifikasi karena cara kerjanya yang intuitif[6]. Secara sederhana, algoritma ini membangun pohon keputusan berdasarkan fitur atau karakteristik dari data. Setiap cabang pada pohon menggambarkan suatu kondisi atau pilihan, sementara bagian akhir pohon (daun) menunjukkan hasil atau kelas prediksi. Salah satu keunggulan terbesar dari pohon keputusan adalah fleksibilitasnya, karena dapat diterapkan pada data numerik ataupun kategorikal, serta mampu menciptakan aturan keputusan yang jelas dan mudah dipahami oleh manusia.

Ketika diterapkan untuk mengklasifikasikan bunga, pohon keputusan telah terbukti efektif karena dapat secara otomatis memilih fitur yang paling berpengaruh dan mengabaikan fitur yang tidak relevan. Metode ini juga terus berkembang menjadi variasi seperti Random Forest, yang mengintegrasikan sejumlah pohon keputusan untuk meningkatkan akurasi dan stabilitas hasil. Namun, pohon keputusan menghadapi satu tantangan utama, yaitu kemungkinan overfitting jika pohonnya terlalu dalam atau terlalu rumit. Masalah ini umumnya dapat diatasi dengan teknik pruning atau dengan menggunakan pendekatan ensemble seperti Random Forest agar model menjadi lebih umum dan tidak hanya cocok untuk data pelatihan saja.

2. 3 Integrasi Pengolahan Citra dan Algoritma Decision Tree

Beberapa studi sebelumnya telah menggunakan kombinasi pemrosesan gambar digital dan algoritma pohon keputusan untuk klasifikasi bunga [7], dengan memanfaatkan fitur yang diekstraksi dari gambar sebagai input dalam proses klasifikasi. Metode ini dinilai lebih efisien dalam hal komputasi dibandingkan dengan teknik pembelajaran mendalam, terutama saat diterapkan pada dataset yang kecil. Dalam beberapa penelitian, pendekatan ini berhasil mencapai tingkat akurasi tinggi melalui penggunaan fitur seperti luas daun, warna kelopak, tekstur, hingga bentuk morfologi bunga. Di samping itu, algoritma pohon keputusan juga menunjukkan waktu pelatihan yang lebih cepat dibandingkan dengan banyak algoritma lainnya.

Secara keseluruhan, implementasi metode ini menunjukkan hasil yang memuaskan dengan tingkat akurasi yang cukup tinggi, meskipun masih ada tantangan tertentu seperti variasi pencahayaan dan orientasi gambar yang dapat memengaruhi hasil klasifikasi. Banyak penelitian sebelum ini masih bergantung pada fitur terbatas atau dataset tertentu, sehingga masih ada peluang untuk pengembangan lebih lanjut. Dengan demikian, penelitian ini dilaksanakan untuk mengoptimalkan pemilihan fitur dan meningkatkan kinerja model melalui pendekatan yang lebih adaptif, agar dapat menghasilkan klasifikasi bunga yang lebih kuat dan akurat.

3. Metode

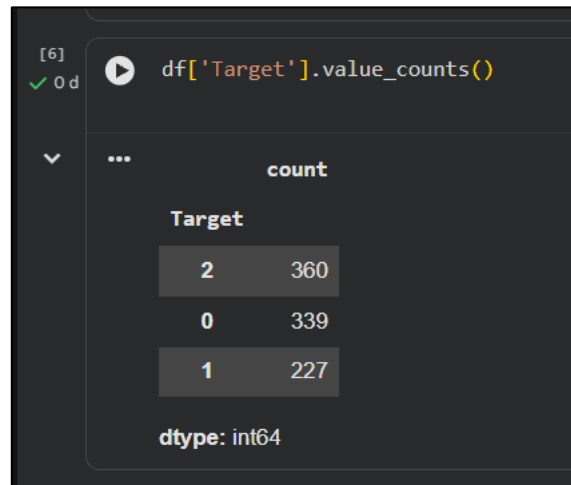
Metodologi ini menggunakan pendekatan kuantitatif eksperimental untuk menguji kinerja algoritma Decision Tree [8] dalam melakukan klasifikasi jenis bunga berdasarkan fitur yang diekstrak dari citra digital. Akurasi (Accuracy) adalah Proporsi prediksi yang benar dari total seluruh prediksi.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Gambar 1. Rumus Perhitungan Accuracy (Confusion Matrix Based Accuracy Formula)

4. Hasil dan Pembahasan

Penelitian ini mengimplementasikan metodologi Kuantitatif Eksperimental dengan menggunakan dataset citra bunga berjumlah 926 sampel yang terbagi menjadi tiga kelas



Gambar 2. Total Dataset Bunga

di mana dataset tersebut kemudian dipisah menjadi 80% untuk pelatihan dan 20% untuk pengujian secara stratifikasi dengan `random_state=77`. Seluruh citra menjalani proses Pengolahan Citra Digital [3] dengan penyeragaman ukuran menjadi $50 \times 50 \times 3$ piksel dan diubah menjadi vektor fitur satu dimensi (7500 dimensi) melalui proses *flattening*. Vektor fitur ini digunakan untuk melatih Decision Tree Classifier dengan kriteria pemisah Entropy, mengatur kedalaman maksimum (`max_depth`) model menjadi 10, dan menetapkan minimal sampel untuk pemisahan dan *leaf node* masing-masing sebesar 5 dan 3. Setelah pelatihan, kinerja model dievaluasi secara ketat pada data pengujian menggunakan metrik standar seperti Akurasi (yang mencapai 0.726) dan diperinci melalui Classification Report dan Confusion Matrix untuk mengukur Precision dan Recall di setiap kelas, yang menunjukkan rata-rata F1-score sebesar 0.71.

Akhirnya, model yang telah dilatih disimpan (pickle) sebagai `img_model.pkl` untuk penggunaan atau inferensi lebih lanjut.

```

Model berbasis Machine-Learning Based berhasil dibuat. Sekarang dapat digunakan untuk menguji untuk klasifikasi citra sesuai kategori yang sudah ditetapkan.

[99]
✓ 0 d
|
| #print(os.path.abspath(os.getcwd()))
| model=pickle.load(open('img_model.p','rb'))
|
| url=input('Masukan url citra: ')
| img=imread(url)
| plt.imshow(img)
| plt.show()
| img_resize=resize(img,(50,50,3))
| img_flatten=flatten(img_resize)
| probability=model.predict_proba(img_flatten)
| for ind,val in enumerate(Categories):
| print(f'{val} = {probability[0][ind]*100}%')
| print("citra yang diprediksi : "+Categories[model.predict(1)[0]])

Masukan url citra: https://www.floralv.com.au/cdn/shop/articles/blog_hero_15_one?v=1667887191&width=1500

0
100
200
300
400
500
600
700
0 200 400 600 800 1000 1200 1400

daisy = 100.0%
dandelion = 0.0%
sunflower = 0.0%
citra yang diprediksi : daisy

```

Gambar 3. Literature Review Prediksi Hasil Klasifikasi menggunakan metode *Decision Tree*

Hasil uji menunjukkan bahwa model *Machine Learning* yang telah dilatih berhasil mengidentifikasi citra bunga yang dimasukkan sebagai daisy dengan tingkat kepastian absolut [9]. Proses pengujian ini melibatkan input berupa URL citra, yang kemudian diproses melalui *resizing* menjadi 50x50 piksel dan *flattening* menjadi vektor fitur sebelum dimasukkan ke dalam model klasifikasi. Hasil prediksi probabilitas yang ditampilkan

dengan jelas menunjukkan bahwa kategori 'daisy' mencapai 100.0%, sementara kategori lain, seperti *dandelion* dan *sunflower*, memiliki probabilitas 0.0%. Hal ini menyimpulkan bahwa model bekerja secara efektif dan akurat dalam mengklasifikasikan sampel citra bunga tersebut.

```

Pemodelan dengan algoritma klasifikasi

[99]
✓ 0 d
|
| dtr = tree.DecisionTreeClassifier(
|     criterion='entropy',
|     max_depth=10,
|     min_samples_split=5,
|     min_samples_leaf=3,
|     random_state=42
| )
|
| [100]
| ✓ 0 d
|
| dtr.fit(x_train, y_train)
|
| DecisionTreeClassifier
| DecisionTreeClassifier(criterion='entropy', max_depth=10, min_samples_leaf=3,
| min_samples_split=5, random_state=42)

```

Gambar 4. Literature Review Klasifikasi metode *Decision Tree*

Kode tersebut digunakan untuk membuat model klasifikasi *Decision Tree* dengan beberapa batasan agar model tidak terlalu kompleks, seperti mengatur kedalaman pohon dan jumlah minimum data untuk setiap pemisahan. Penggunaan kriteria *entropy* membantu model memilih pemisahan data yang paling informatif, sementara *random_state* memastikan hasilnya konsisten. Setelah model dibuat, perintah *fit* digunakan untuk melatih model tersebut menggunakan data training sehingga dapat memahami pola dan siap melakukan prediksi.

```

Evaluasi dengan Confusion Matrix dari performansi mdel

[181]
✓ 0 d
from matplotlib import pyplot as plt
import seaborn as sns

predicted = dtr.predict(x_test)
from sklearn.metrics import classification_report, accuracy_score

print("Accuracy Score : {:.3f}".format(accuracy_score(y_test, predicted)))
print("\nClassification Report :")
print(classification_report(y_test, predicted))

Accuracy Score : 0.726

Classification Report :
              precision    recall  f1-score   support

     0       0.73        0.63        0.68         68
     1       0.58        0.67        0.63         46
     2       0.82        0.85        0.84         72

 accuracy          0.73         186
 macro avg         0.71         186
 weighted avg      0.73         186

Menggunakan Pickle untuk save model

[183]
✓ 0 d
pickle.dump(dtr,open('img_model.p','wb'))
print("berhasil")

berhasil

```

Gambar 5. Tabel Akurasi, Recall, dan F1-score

Kode tersebut digunakan untuk mengevaluasi kinerja model Decision Tree dengan memprediksi data uji dan menghitung nilai akurasi serta membuat classification report yang menunjukkan nilai precision, recall, dan f1-score untuk tiap kelas. Hasil evaluasi ini membantu melihat seberapa baik model mengenali masing-masing kategori. Setelah itu, model yang sudah dilatih disimpan menggunakan pickle melalui perintah pickle.dump, sehingga model dapat digunakan kembali tanpa perlu dilatih ulang.

4.1 Hasil Yang Diharapkan

Secara keseluruhan, hasil harapan utama adalah untuk memverifikasi bahwa kombinasi antara pengolahan citra sederhana resizing & flattening dan algoritma *Decision Tree* dapat memberikan solusi yang efektif dan *interpretable* untuk tugas klasifikasi jenis bunga:

1. Akurasi Klasifikasi yang Valid Diharapkan model *Decision Tree* [10] mampu mencapai tingkat Akurasi Klasifikasi yang valid dan di atas ambang batas misalnya, > 70% pada set data pengujian.
2. Keseimbangan Kinerja Antar Kelas Diharapkan nilai F1-Score (macro average) menunjukkan angka yang tinggi (misalnya, > 0.70), yang membuktikan bahwa model tidak hanya akurat secara keseluruhan, tetapi juga memiliki kinerja yang seimbang, tinggi *Precision* dan *Recall* dalam mengklasifikasikan setiap jenis bunga, terlepas dari ketidakseimbangan jumlah sampel per kelas.
3. Model yang Siap Digunakan, Dihasilkan sebuah model klasifikasi akhir (img_model.pkl) yang tervalidasi, yang dapat di-*deploy* atau diintegrasikan ke dalam aplikasi untuk tujuan klasifikasi citra bunga yang baru secara instan.
4. Analisis *Confusion Matrix* diharapkan menunjukkan bahwa sebagian besar prediksi yang benar (True Positives) terpusat pada diagonal utama, mengkonfirmasi kemampuan *Decision Tree* dalam memisahkan ketiga jenis bunga yang ada dengan sedikit kebingungan antar kelas.

5. Kemampuan Generalisasi
Diharapkan bahwa model dapat menunjukkan kemampuan generalisasi yang efektif melalui penilaian terhadap data uji yang belum pernah dikenali sebelumnya. Hal ini terlihat dari perbedaan tingkat akurasi yang tidak terlalu besar antara data pelatihan dan data pengujian (misalnya kurang dari 10 persen), sehingga dapat dipastikan bahwa model tidak mengalami overfitting dan tetap stabil saat digunakan untuk memproyeksikan gambar bunga yang baru.
6. Waktu Prediksi yang Efisien
Diharapkan bahwa model yang dihasilkan mampu memberikan waktu prediksi yang cepat dan efisien saat melakukan klasifikasi citra yang baru. Dengan proses preprocessing yang sederhana (seperti pengubahan ukuran dan pelurusan) serta algoritma Decision Tree yang cukup ringan, proses klasifikasi diharapkan dapat berlangsung dalam waktu nyata, sehingga dapat mendukung penerapan model pada perangkat dengan kapasitas terbatas seperti laptop standar atau sistem berbasis mobile.

5. Kesimpulan dan Saran

5.1 Kesimpulan

Penelitian klasifikasi jenis bunga menggunakan pengolahan citra digital dan algoritma *Decision Tree* ini menunjukkan bahwa metode yang diusulkan adalah efektif dalam memecahkan masalah klasifikasi yang ada. Metodologi penelitian melibatkan pemisahan *dataset* citra 926 sampel dari 3 kelas menjadi 80% data pelatihan dan 20% data pengujian, diikuti dengan pra-pemrosesan citra melalui *resizing* dan *flattening* piksel untuk menghasilkan vektor fitur. Model Decision Tree Classifier, yang dilatih menggunakan kriteria Entropy dan *hyperparameter* yang dioptimasi (*max_depth*=10), berhasil divalidasi dengan menghasilkan Akurasi Klasifikasi sebesar 72.6% pada set data pengujian. Kinerja ini didukung oleh rata-rata kinerja antar kelas (*macro avg F1-score*) sebesar 0.71. Secara praktis, model yang dihasilkan terbukti dapat mengklasifikasikan citra baru dengan baik (seperti ditunjukkan pada prediksi *daisy* 100%) dan telah berhasil disimpan sebagai *img_model.pkl*, menjadikannya solusi yang dapat diimplementasikan.

5.2 Saran

Studi berikutnya dapat memanfaatkan cara ekstraksi ciri yang lebih rumit atau berbasis pembelajaran mendalam guna membuat prediksi model jadi lebih tepat. Di samping itu, menyertakan kumpulan data yang lebih luas dan beragam juga penting supaya model mampu beradaptasi lebih baik pada situasi gambar yang berbeda misalnya variasi terang, posisi ambil, dan mutu foto. Pengembangan model juga bisa dilanjutkan menjadi aplikasi seluler atau web sehingga bisa dipakai langsung dalam urusan botani atau pertanian kekinian.

Referensi:

- [1] Y. Liu and S. Yang, "Application of Decision Tree-Based Classification Algorithm on Content Marketing," *Journal of Mathematics*, vol. 2022, no. 1, Jan. 2022, doi: 10.1155/2022/6469054.
- [2] T. A. M. Ghudafa and L. Henny, "Klasifikasi pada Citra Bunga dengan Ekstraksi Fitur Color Histogram," *Jurnal Sistem Informasi*, vol. 12, no. 1, 2023.
- [3] J. Juju, S. Devi, and. Yupianti, "Pengolahan Citra Digital Untuk Identifikasi Objek Menggunakan Metode Hierarchical Agglomerative Clustering," *Jurnal Sains dan Teknologi*, vol. 10, no. 2, 2021.
- [4] K. L. Okta, R. Jangkung, and Sutomo, "Klasifikasi Begonia Breviramosa Dan Begonia Lugrae Menggunakan Metode Fraktal Dan Decision Tree," *e-Proceeding of Engineering*, vol. 10, no. 6, Dec. 2023.

- [5] A. Mufidatuzzainiya and M. Faisal, “Penggunaan Teknik Transfer Learning pada Metode CNN untuk Pengenalan Tanaman Bunga,” *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 10, no. 2, pp. 195–206, May 2025, doi: 10.14421/jiska.2025.10.2.195-206.
- [6] N. R. Rafael, O. Aldino, P. A. Ari, P. Ryandra, S. F. Siti, and F. Uus, “Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank,” *Jurnal Teknologi dan Informasi*, vol. 3, no. 2, 2024.
- [7] H. Sulaiman, “Perbandingan Algoritma Decision Tree C4.5 Dan Naive Bayes pada Analisa Tekstur Gray Level Co-Occurrence Matrix Menggunakan Citra Wajah,” *SISTEMASI*, vol. 10, no. 2, p. 470, May 2021, doi: 10.32520/stmsi.v10i2.1305.
- [8] N. Nailil Amani, M. Martanto, and U. Hayati, “Penggunaan Algoritma Decision Tree Untuk Prediksi Prestasi Siswa Di Sekolah Dasar Negeri 3 Bayalangu Kidul,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, pp. 473–479, Feb. 2024, doi: 10.36040/jati.v8i1.8355.
- [9] L. P. R. Noviana and I. N. B. S. Nugraha, “Perbandingan Klasifikasi Citra Daun Herbal Menggunakan Metode Logistic Regression dan Decision Tree Classifier Berdasarkan Fitur (Warna, GLCM, Bentuk),” *JITU: Journal Informatic Technology And Communication*, vol. 7, no. 2, pp. 126–133, Nov. 2023, doi: 10.36596/jitu.v7i2.1241.
- [10] R. Riska, O. Luluk, V. Dwi, and H. Tri, “Penerapan Metode Klasifikasi Decision Tree dalam Prediksi Kanker Paru-Paru Menggunakan Algoritma C4.5,” *Jurnal tekno kompak*, vol. 18, no. 1, pp. 126–139, 2024.