

SENASTIKA Universitas Malikussaleh

ANALISIS KINERJA KOMBINASI ALGORITMA K- NEAREST NEIGHBOR (KNN) DAN PRINCIPAL COMPONENT ANALYSIS (PCA) PADA KLASIFIKASI GAMBAR SPESIES BURUNG

Ari Wardana^{*1}, Munirul Ula², Sujacka Retno^{*3}

^{1,3}(Teknik Informatika, Universitas Malikussaleh, Lhokseumawe, Aceh, Indonesia)

²(Teknologi Informasi, Universitas Malikussaleh, Lhokseumawe, Aceh, Indonesia)

Email: 1ari.200170010@mhs.ac.id , 2munirulula@unimal.ac.id , 3sujacka@unimal.ac.id,

Abstrak

Bagi masyarakat umum, mengetahui lebih detail tentang berbagai jenis burung menjadi sulit karena banyaknya spesies yang ada dan kesamaan yang dimiliki oleh setiap spesies burung dalam hal ukuran tubuh, warna, dan bentuk fisiknya. Membedakan spesies burung bukanlah tugas yang mudah karena membutuhkan kemampuan, waktu, dan biaya yang besar untuk mempelajari setiap jenis burung yang ada. Maka dari itu, penelitian ini bertujuan untuk mengembangkan sistem pengolahan citra untuk pengklasifikasian jenis burung khususnya burung yang berada di daerah Aceh dengan kombinasi algoritma *K-Nearest Neighbor* (K-NN) dan *Principal Component Analysis* (PCA). Ekstraksi fitur yang digunakan adalah ekstraksi fitur berdasarkan warna dan bentuknya. Algoritma K-NN dapat mengelompokkan objek tertentu dengan mempertimbangkan jarak terpendek dari objek tersebut. Berdasarkan kriteria terbaik, maka digunakan metode PCA untuk memperkecil dan mempertahankan sebagian besar data yang relevan dari ciri aslinya. Berdasarkan hasil pengujian, diperoleh nilai akurasi sebesar 82,50%, presisi sebesar 83,06%, *recall* sebesar 82,50%. Penggunaan kombinasi algoritma K-NN dan PCA pada klasifikasi citra jenis burung pada penelitian yang telah dilakukan terbukti mampu meningkatkan akurasi yang lebih besar dibandingkan dengan hanya menggunakan algoritma K-NN saja.

Keywords: *spesies burung, k-nearest neighbors, klasifikasi gambar, principal component analysis.*

1. PENDAHULUAN

Negara Indonesia dikenal sebagai negara yang kaya akan keanekaragaman flora dan fauna, termasuk memiliki kekayaan yang sangat tinggi dalam hal jenis hewan dan tumbuhan. Indonesia sering disebut sebagai *Megabiodiversity* karena kekayaan ini. Berbagai jenis satwa di negara ini memiliki peran penting dalam kehidupan sosial, ekonomi, dan budaya masyarakat. Salah satu kekayaan fauna yang dimiliki Indonesia adalah burung. Terdapat 1.539 jenis burung di Indonesia [1].

Burung termasuk dalam kelompok hewan vertebrata yang bereproduksi melalui perkawinan, serta memiliki bulu, warna, dan suara yang beragam dan menarik. Keberagaman jenis burung di suatu wilayah sangat dipengaruhi oleh faktor cuaca yang mendukung, variasi jenis tumbuhan, dan keadaan habitat tempat burung tersebut hidup [2].

Aceh memiliki keanekaragaman hayati burung yang sangat beragam dan berpotensi untuk dikembangkan sebagai peluang usaha di bidang ekowisata. Sejak tahun 2007, Lembaga Cicem Nanggro (LCN) telah melakukan berbagai penelitian terkait keragaman spesies burung dan habitatnya di Aceh, serta bekerja sama secara intens dengan berbagai Lembaga Konservasi, NGO, dan kelompok pengamat burung baik nasional maupun internasional. Berdasarkan penelitian tersebut, diketahui bahwa di Hutan Kota Banda Aceh terdapat 20 spesies burung dari 10 famili. Beberapa spesies burung yang ditemukan antara lain Elang-alap besar, Cekakak suci, Kepinis rumah, Walet, Kuntul karang, Kuntul kecil, Kokokan laut, Kowak malam abu-abu, Perkatut Jawa, Merpati, Bondol tunggir putih, Kolibri kelapa, Madu polos, Madu sriganti, Cabai bunga api, Burung gereja, Kutulang, Merbah cerucuk, dan Kipas belang [3].

Bagi masyarakat umum, mengetahui lebih detail tentang berbagai jenis burung menjadi sulit karena banyaknya spesies yang ada dan kesamaan yang dimiliki oleh setiap spesies burung dalam hal ukuran tubuh, warna, dan bentuk fisiknya. Membedakan spesies burung bukanlah tugas yang mudah karena membutuhkan kemampuan, waktu, dan biaya yang besar untuk mempelajari setiap jenis burung yang ada.

Spesies burung dapat dikenali dari penampilannya karena setiap spesies umumnya memiliki perbedaan dalam bentuk dan warna. Pemrosesan citra digital melibatkan tahapan pembentukan, pengolahan, serta analisis gambar untuk mendapatkan informasi yang lebih bermanfaat dari citra tersebut. Salah satu fungsi dari pemrosesan citra adalah klasifikasi gambar [4].

Penelitian yang dilakukan oleh Nuraini dkk dengan judul kombinasi Algoritma K-NN dan PCA pada klasifikasi gambar spesies ikan mendapatkan hasil dengan tingkat akurasi sebesar 85% dibandingkan dengan menggunakan metode K-NN saja .

Penelitian yang dilakukan Adenugraha dkk dengan judul Klasifikasi Kematangan Buah Pisang Ambon Menggunakan Metode KNN dan PCA Berdasarkan Citra RGB dan HSV, dari hasil penelitian didapatkan pengukuran jarak terdekat, menggunakan nilai $k = 5$. Yang di peroleh akurasi sebesar 90,9 % .

Penelitian yang dilakukan Raysyah dkk dengan judul Klasifikasi Tingkat Kematangan Buah Kopi Berdasarkan Deteksi Warna Menggunakan Metode KNN Dan PCA, dengan mengukur jarak tetangga terdekat dan menggunakan nilai $K = 3$. Hasil penelitian menunjukkan akurasi sebesar 97,77%, di mana 44 data uji diklasifikasikan secara akurat, sementara 1 data diklasifikasikan secara tidak akurat.

Dari penelitian-penelitian diatas yang telah dilakukan oleh peneliti sebelumnya menunjukkan bahwa menggunakan metode K-NN dan PCA mampu memberikan akurasi yang cukup tinggi dibandingkan hanya menggunakan metode K-NN saja.

Algoritma *K-Nearest Neighbor* (K-NN) dapat digunakan untuk klasifikasi karena kesederhanaannya dalam penerapan. Algoritma ini bekerja dengan mengelompokkan data berdasarkan titik data yang paling dekat dengan data yang ingin diprediksi [5] [6]. K-NN merupakan algoritma berbasis instance yang sederhana, sehingga tidak memerlukan proses pelatihan yang rumit. Oleh karena itu, algoritma ini cocok untuk diterapkan pada data pelatihan yang terbatas dan mampu memberikan hasil yang baik karena berfokus pada informasi spesifik.

Namun, algoritma K-NN memiliki kelemahan dalam menangani nilai menyimpang, sehingga rentan terhadap variabel yang tidak mengandung informasi. Jadi, perlu untuk meningkatkan dan menambahkan metode berdasarkan fitur yang dibentuk oleh objek dan untuk mengurangi dan mempertahankan informasi yang tepat dari karakteristik aslinya untuk menghasilkan kinerja yang optimal. Salah satu algoritma yang dapat digunakan untuk menyelesaikan masalah reduksi data adalah *Principal Component Analysis* (PCA). Metode ini dapat digunakan untuk mengurangi dan mempertahankan sebagian besar informasi yang relevan dari fitur asli sesuai dengan kriteria optimal. PCA dapat menyederhanakan data melalui transformasi linier sehingga membentuk sistem yang memiliki koordinat baru yang menghasilkan varians maksimum [4].

Maka dari itu penelitian ini bertujuan untuk membuat model K-NN dan PCA dengan akurasi yang baik. kemudian model tersebut diimplementasikan kedalam sistem untuk mengklasifikasikan spesies burung. Ini dapat memudahkan bagi masyarakat awam yang ingin mengetahui satwa khususnya burung agar dapat lebih cepat dan mudah dalam membedakan jenis-jenis burung. Dalam penelitian ini terdapat 4 jenis burung yang akan di klasifikasikan yaitu burung cekakak suci, kuntul kecil, kutilang, dan cabai bunga api.

2. METODE PENELITIAN

2.1 Klasifikasi

Kata "klasifikasi" berasal dari istilah *classification* dalam bahasa Inggris, yang berasal dari akar kata *to classify*, yang berarti menggolongkan dan menempatkan benda-benda serupa dalam satu kategori. Dalam konteks perpustakaan, klasifikasi merujuk pada proses pengelompokan bahan pustaka, seperti buku, secara sistematis dalam kelompok atau kelas tertentu berdasarkan ciri-ciri khusus untuk memudahkan pencarian [7].

2.2 Citra Digital

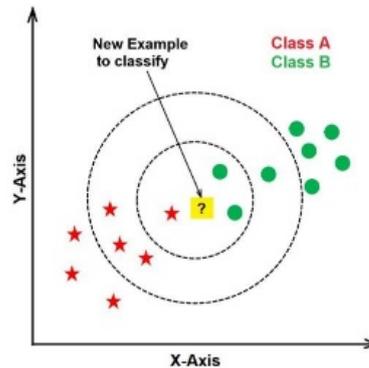
Citra digital adalah gambar dua dimensi yang dihasilkan dari gambar analog kontinu melalui proses sampling, di mana gambar analog dibagi menjadi N baris dan M kolom untuk membentuk gambar diskrit. Citra digital merupakan gambar yang dapat diproses oleh komputer, di mana data yang disimpan adalah angka-angka yang menunjukkan intensitas pada setiap piksel [8].

2.3 Pengolahan Citra Digital

Citra merupakan salah satu bentuk informasi yang diperlukan manusia selain teks, suara dan video. Informasi yang terkandung dalam sebuah citra dapat diinterpretasikan berbeda-beda oleh manusia satu dengan yang lain. Pengolahan Citra Digital (*Digital Image Processing*) merupakan disiplin ilmu yang mempelajari teknik dalam mengolah citra, citra yang dimaksud adalah merupakan gambar diam (foto) atau gambar yang bergerak (seperti video yang direkam). Sedangkan arti digital adalah pengolahan citra/gambar dilakukan menggunakan komputer secara digital [9].

2.4 Algoritma K-NN

Algoritma *K-Nearest Neighbor* (KNN) adalah sebuah metode yang digunakan untuk mengklasifikasikan objek berdasarkan data pembelajaran yang memiliki jarak paling dekat dengan objek tersebut. *K-Nearest Neighbor* didasarkan pada konsep 'learning by analogy' [4].



Gambar 1 Ilustrasi klasifikasi menggunakan K-NN

Nilai jarak pada metode *K-NN* dapat dihitung dengan menggunakan rumus *Euclidean Distance*. Cara ini sederhana dan dapat memberikan akurasi yang baik terhadap hasil klasifikasi. Adapun rumus Euclidean Distance seperti pada rumus di bawah ini [10].

$$d(xy) = \sqrt{\sum (x_i - y_i)^2} \quad ni=1 \quad (1)$$

2.5 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) merupakan suatu titik dimensi dari kumpulan kelompok data yang terdiri dari variable-variabel yang saling terkait. Ada juga yang mengatakan bahwa PCA adalah suatu metode untuk mengurangi kepadatan kumpulan data atau menyederhanakan kumpulan data. Dari kedua pengertian dapat disimpulkan bahwa PCA merupakan suatu teknik yang digunakan untuk menyederhanakan kumpulan data yang saling terkait. Perhitungan PCA menggunakan kumpulan data yang telah distandarisasi. PCA tersusun dari nilai *Eigen* dan kelompok variable data awal. Nilai *eigen* dibagi menjadi dua jenis yaitu *Eigenvalue* dan *Eigenvector*. Rumus untuk menghitung nilai *eigen* terdapat pada persamaan berikut [11].

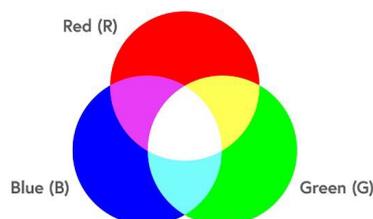
$$A \cdot v = \lambda \cdot v \quad (2)$$

2.6 Ekstraksi

Ekstraksi fitur adalah proses penting dalam analisis citra yang bertujuan untuk mengidentifikasi dan mengenali ciri khas dari citra yang sedang dipelajari. Ada berbagai metode yang dapat digunakan untuk melakukan ekstraksi fitur pada citra, di antaranya adalah ciri amplitudo, ciri histogram, matriks *co-occurrence*, gradien, deteksi tepi, spektrum *Fourier*, *wavelet*, fitur warna, filter *Gabor*, blok *overlapping* dan *non-overlapping*, serta fraktual [12].

2.7 Warna RGB

RGB merupakan ruang warna yang tersusun dari tiga channel warna yakni channel merah (*Red*), channel hijau (*Green*), dan channel biru (*Blue*). Setiap channel mempunyai intensitas warna yang berbeda, nilai intensitas warna minimum adalah nol (0) dan nilai intensitas warna maksimum adalah 255 (8 bit). Setiap piksel citra RGB memiliki variasi warna sejumlah 16.777.216 (256 x 256 x 256) Representasi warna RGA dapat dilihat pada Gambar 2 berikut [13].



Gambar 2 Warna RGB

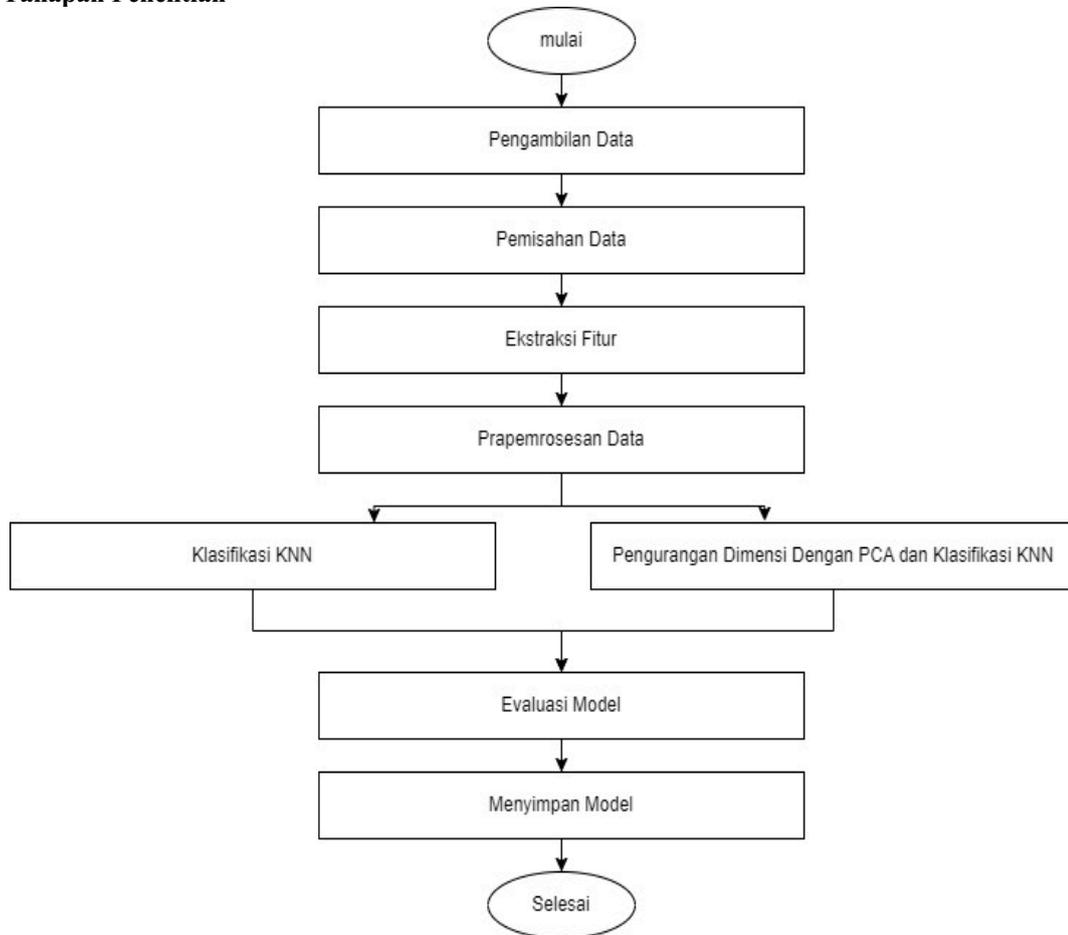
2.8 Warna HSV

HSV merupakan definisi dari *hue*, *saturation*, *value*. *Hue* merupakan ukuran dari jenis warna yang ada, sedangkan *saturation* merupakan keberwarnaannya dimana semakin berwarna semakin besar nilai saturasinya, namun semakin pudar berarti semakin rendah nilai saturasinya, dan *value* merupakan nilai kecerahan dari suatu warna [14].

2.9 Python

Pengertian Python (bahasa pemrograman) merupakan bahasa pemrograman tinggi yang bisa melakukan eksekusi sejumlah instruksi multi guna secara langsung (interpretatif) dengan metode Object Oriented Programming dan juga menggunakan semantik dinamis untuk memberikan tingkat keterbacaan syntax. Sebagai bahasa pemrograman tinggi, python dapat dipelajari dengan mudah karena telah dilengkapi dengan manajemen memori otomatis [15].

2.9 Tahapan Penelitian



Gambar 3 Tahapan Peneleitian

1. Pengambilan Data

Data gambar pada penelitian ini digunakan sebagai data set, data set yang digunakan di ambil dari Google. Dalam penelitian ini yang akan di lakukan adalah mengklasifikasikan gambar spesies burung. Spesies burung yang digunakan antara lain 4 jenis burung yaitu: Cekakak Suci, Kuntul Kecil, Kutilang, dan Cabai Bunga Api. Gambar yang digunakan berekstensi JPG, total data gambar yang digunakan 400 gambar yang setiap kelas berisi 100 gambar.

Tabel 1 Sampel data untuk kumpulan data

| Jenis Burung | Gambar Burung |
|--------------|--|
| Cekakak Suci |  |
| Kuntul Kecil |  |
| Kutilang |  |



2. Pemisahan Data

Dataset yang telah diproses kemudian dibagi menjadi dua bagian: 80% digunakan untuk pelatihan model, dan 20% digunakan untuk pengujian. Pemisahan ini dilakukan untuk memastikan bahwa model dapat diuji dengan data yang belum pernah dilihat sebelumnya, sehingga akurasinya dapat dievaluasi dengan baik.

3. Ekstraksi Fitur

Konversi citra ke format biner menggunakan metode thresholding Otsu dimulai dengan mengubah citra berwarna menjadi citra skala abu-abu. Proses ini dilakukan untuk menyederhanakan data dengan menghilangkan informasi warna dan hanya fokus pada intensitas piksel. Konversi ini penting untuk mempermudah ekstraksi fitur bentuk dan struktur objek dalam citra.

fitur warna diekstraksi dari gambar RGB asli, informasi warna dari gambar asli tetap penting untuk analisis yang lebih komprehensif. Fitur warna dihitung dengan menggunakan histogram warna, yang mencakup distribusi intensitas untuk setiap saluran warna (merah, hijau, dan biru). Histogram ini memberikan representasi kuantitatif dari warna yang ada dalam gambar, yang berguna untuk proses klasifikasi berbasis warna. Ekstraksi fitur warna RGB dilakukan setelah citra dikonversi ke format RGB dari format BGR yang digunakan oleh OpenCV.

4. Prapemrosesan Data

Setelah fitur diekstraksi, data diflatten menjadi vektor satu dimensi dan distandarisasi menggunakan StandardScaler. Standarisasi dilakukan agar semua fitur berada pada skala yang seragam, yang membantu meningkatkan efektivitas algoritma KNN.

5. Pengurangan Dimensi dengan PCA

Untuk mengurangi kompleksitas data dan mempercepat komputasi, PCA diterapkan pada data fitur untuk mengurangi dimensi. Berbagai jumlah komponen diuji untuk menemukan jumlah komponen yang optimal. PCA membantu dalam menjaga fitur yang paling penting untuk klasifikasi.

Dalam penelitian ini, PCA diterapkan setelah proses standarisasi fitur untuk mengurangi dimensi data gambar yang telah diproses. Langkah pertama adalah menginisialisasi objek PCA dengan parameter `n_components`, yang menentukan jumlah komponen utama yang ingin dipertahankan. Komponen utama ini adalah kombinasi linier dari fitur asli yang menjelaskan sebagian besar variabilitas dalam data.

6. Klasifikasi dengan Algoritma KNN

Dalam proses ini, model KNN dibangun dengan memilih nilai `k`, yaitu jumlah tetangga terdekat yang akan dipertimbangkan dalam menentukan kelas dari setiap data pengujian. Model KNN dilatih menggunakan data pelatihan yang telah distandarisasi. Selama pelatihan, model ini menyimpan semua data pelatihan beserta labelnya. Ketika model dihadapkan pada data baru, model tersebut menghitung jarak antara data baru tersebut dengan semua data pelatihan yang telah disimpan, kemudian memilih `k` jumlah tetangga terdekat berdasarkan jarak tersebut. Kelas dari data baru ditentukan berdasarkan mayoritas kelas di antara `k` tetangga terdekat.

7. Evaluasi Model

Tahap evaluasi ini dilakukan untuk menilai efektivitas model yang dibuat, untuk memberikan evaluasi model yang dibangun menggunakan matriks kebingungan dengan mencari nilai presisi, ingatan, dan akurasi. Nilai-nilai ini dapat diperoleh dari persamaan berikut:

$$\text{presisi} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recal} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Prediksi}} \quad (5)$$

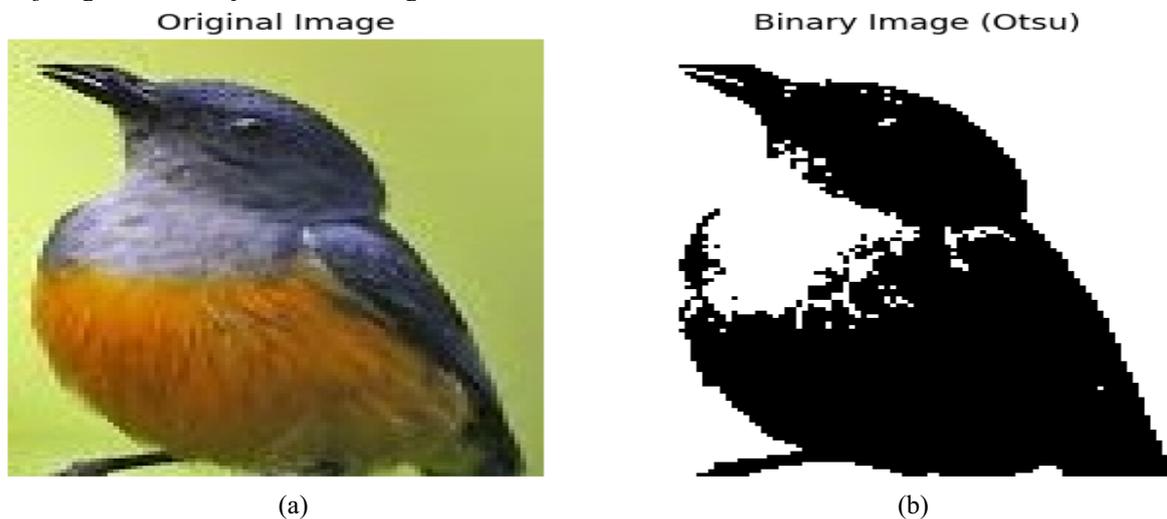
8. Penyimpanan Model

Tahap Selanjutnya model yang telah dilatih dan dievaluasi tersebut disimpan dalam format file yang dapat dengan mudah dimuat kembali dengan ekstensi `.joblib`. Ini memungkinkan model digunakan untuk melakukan klasifikasi pada data baru yang belum ada dalam dataset asli tanpa perlu melatih ulang model dari awal. Dengan menyimpan model dapat menghemat waktu dan sumber daya serta memanfaatkan hasil pelatihan yang telah dicapai

3. HASIL DAN PEMBAHASAN

Untuk mengklasifikasikan citra spesies burung menggunakan kombinasi algoritma K-Nearest Neighbor (K-NN) dan Principal Component Analysis (PCA) langkah pertama adalah menyiapkan dataset terlebih dahulu, di mana data tersebut nantinya akan digunakan sebagai pelatihan dan pengujian model yang sedang dibangun. Dataset diperlukan dalam pengenalan pola dan pembelajaran agar dapat menemukan informasi dalam gambar sehingga proses klasifikasi dapat dilakukan. Total data gambar yang digunakan adalah 100 gambar setiap jenis burung, kemudian data gambar dibagi menjadi data pelatihan dan data pengujian. Persentase distribusi yang digunakan adalah 80:20. Sehingga data pelatihan yang digunakan adalah 80 gambar dan data pengujian adalah 20 gambar. Jenis burung yang digunakan adalah burung cekakak suci, kuntul kecil, kutilang, dan cabai bunga api.

Pemodelan yang digunakan dalam penelitian ini adalah pemodelan dengan bahasa python untuk pelatihan dan pengujian. Dimulai dengan Konversi citra ke format biner menggunakan metode thresholding Otsu dimulai dengan mengubah citra berwarna menjadi citra skala abu-abu. Proses ini dilakukan untuk menyederhanakan data dengan menghilangkan informasi warna dan hanya fokus pada intensitas piksel. Selanjutnya, metode Otsu diterapkan untuk menentukan nilai threshold optimal yang membagi citra menjadi dua kelompok: piksel latar belakang dan objek. Dengan menggunakan fungsi `cv2.threshold()` yang menggabungkan opsi `cv2.THRESH_BINARY` dan `cv2.THRESH_OTSU`, citra biner dihasilkan, di mana piksel di atas threshold diatur menjadi putih dan piksel di bawahnya menjadi hitam. Konversi ini penting untuk mempermudah ekstraksi fitur bentuk dan struktur objek dalam citra, Gambar 3 menampilkan contoh gambar dari beberapa spesies burung yang muncul setelah konversi menjadi gambar biner yaitu untuk mengekstraksi fitur bentuk



Gambar 4 (a) gambar asli (b) gambar biner

Kemudian fitur warna diekstraksi dari gambar RGB asli, informasi warna dari gambar asli tetap penting untuk analisis yang lebih komprehensif. Fitur warna dihitung dengan menggunakan histogram warna, yang mencakup distribusi intensitas untuk setiap saluran warna (merah, hijau, dan biru). Histogram ini memberikan representasi kuantitatif dari warna yang ada dalam gambar, yang berguna untuk proses klasifikasi berbasis warna. Ekstraksi fitur warna RGB dilakukan setelah citra dikonversi ke format RGB dari format BGR yang digunakan oleh OpenCV.

Setelah gambar diekstraksi bentuk dan warnanya kemudian diikuti dengan prapemrosesan data yaitu data diflatten menjadi vektor satu dimensi dan distandarisasi menggunakan StandardScaler. Standarisasi dilakukan agar semua fitur berada pada skala yang seragam, yang membantu meningkatkan efektivitas algoritma KNN. Ini penting untuk menjaga integritas evaluasi model, memastikan bahwa data pengujian diperlakukan secara seragam dengan data pelatihan, dan menghindari hasil yang bias atau tidak akurat.

Untuk mengurangi kompleksitas data dan mempercepat komputasi, PCA diterapkan pada data fitur untuk mengurangi dimensi. Berbagai jumlah komponen diuji untuk menemukan jumlah komponen yang optimal. PCA membantu dalam menjaga fitur yang paling penting untuk klasifikasi.

Dalam penelitian ini, PCA diterapkan setelah proses standarisasi fitur untuk mengurangi dimensi data gambar yang telah diproses. Langkah pertama adalah menginisialisasi objek PCA dengan parameter `n_components`, yang menentukan jumlah komponen utama yang ingin dipertahankan. Komponen utama ini adalah kombinasi linier dari fitur asli yang menjelaskan sebagian besar variabilitas dalam data.

Selanjutnya klasifikasi data dengan algoritma K-NN dalam proses ini, model KNN dibangun dengan memilih nilai `k`, yaitu jumlah tetangga terdekat yang akan dipertimbangkan dalam menentukan kelas dari setiap data pengujian. Model KNN dilatih menggunakan data pelatihan yang telah distandarisasi. Selama pelatihan, model ini menyimpan semua data pelatihan beserta labelnya. Ketika model dihadapkan pada data baru, model

tersebut menghitung jarak antara data baru tersebut dengan semua data pelatihan yang telah disimpan, kemudian memilih k jumlah tetangga terdekat berdasarkan jarak tersebut. Kelas dari data baru ditentukan berdasarkan mayoritas kelas di antara k tetangga terdekat.

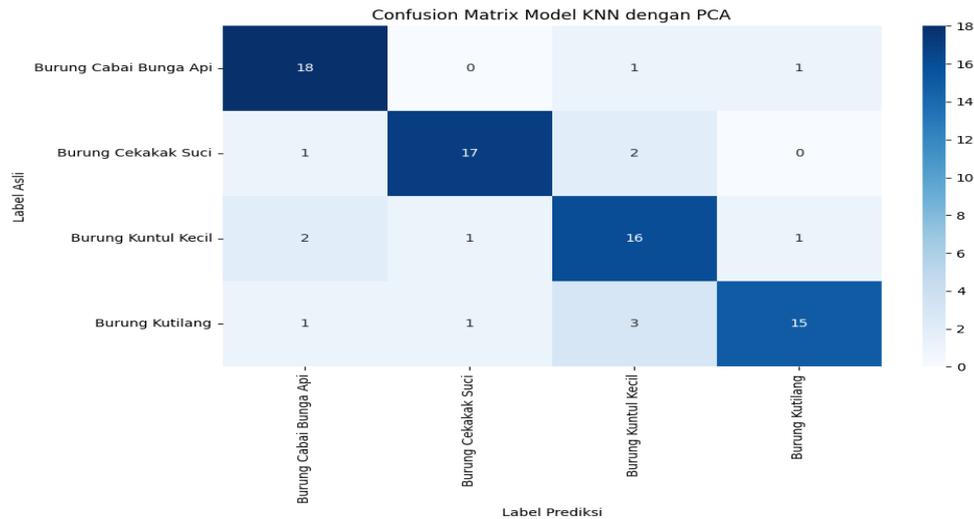
Berikut adalah hasil dari pengujian model dari kombinasi KNN dan PCA yang akan di tampilkan pada tabel 2 berikut ini.

Tabel 2 hasil pengujian dengan PCA

| No | n_components | n_neighbors | Akurasi |
|----|--------------|-------------|---------|
| 1 | 20 | 7 | 78.75% |
| 2 | 21 | 7 | 73.75% |
| 3 | 22 | 7 | 78.75% |
| 4 | 23 | 7 | 76.25% |
| 5 | 24 | 7 | 76.25% |
| 6 | 25 | 7 | 77.50% |
| 7 | 26 | 7 | 82.50% |
| 8 | 27 | 7 | 77.50% |
| 9 | 28 | 7 | 80.00% |
| 10 | 29 | 7 | 78.75% |
| 11 | 30 | 7 | 77.50% |
| 12 | 31 | 7 | 77.50% |
| 13 | 32 | 7 | 80.00% |
| 14 | 33 | 7 | 77.50% |
| 15 | 34 | 7 | 77.50% |
| 16 | 35 | 7 | 73.75% |
| 17 | 36 | 7 | 77.50% |
| 18 | 37 | 7 | 75.00% |
| 19 | 38 | 7 | 77.50% |
| 20 | 39 | 7 | 78.75% |
| 21 | 49 | 7 | 75.00% |

Hasil pengujian menunjukkan bahwa akurasi model bervariasi tergantung pada jumlah komponen utama yang digunakan untuk PCA%. Secara keseluruhan, kombinasi terbaik ditemukan pada n_components = 26 dan n_neighbors = 7, dengan akurasi mencapai 82.50%. Ini mengindikasikan bahwa menggunakan 26 komponen utama untuk PCA dan 7 tetangga terdekat memberikan keseimbangan optimal antara kompleksitas model dan performa akurasi.

Untuk mendapatkan penjelasan lebih detail mengenai hasil klasifikasi model KNN menggunakan PCA, dapat ditampilkan pada *convusion matrix* sebagai berikut.



Gambar 5 *convusion matrix* model KNN dan PCA

Untuk melihat keterangan lebih lanjut mengenai detail hasil klasifikasi masing-masing kelas, dapat diperhatikan pada tabel berikut ini.

Tabel 3 Rincian skor masing-masing kelas model KNN dengan PCA

| Jenis Kelas | True Positive | False Negative | False Positive | True Negative |
|------------------------|---------------|----------------|----------------|---------------|
| Burung Cabai Bunga Api | 18 | 2 | 4 | 56 |
| Burung Cekakak Suci | 17 | 3 | 2 | 58 |
| Burung Kuntul Kecil | 16 | 4 | 6 | 54 |
| Burung Kutilang | 15 | 5 | 2 | 58 |

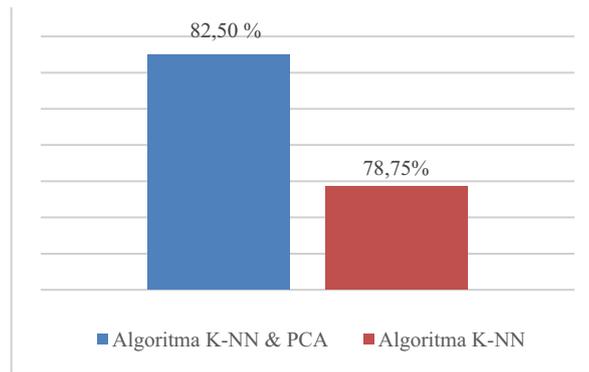
Kemudian langkah selanjutnya menampilkan hasil evaluasi model yaitu menentukan tingkat akurasi, recall, dan presisi. Lebih jelasnya dapat di lihat pada tabael 4 berikut ini.

Tabel 4 Hasil evaluasi model K-NN dan PCA

| Metrik Evaluasi | Skor |
|-----------------|--------|
| Akurasi | 82,50% |
| Presisi | 83,06% |
| Recall | 82,50% |

Secara keseluruhan, hasil pengujian menunjukkan bahwa penerapan PCA pada model kNN dapat meningkatkan akurasi dibandingkan dengan k-NN tanpa PCA. Meskipun kNN tanpa PCA menghasilkan akurasi yang bervariasi antara 71.25% dan 78.75%, penggunaan PCA dengan 26 komponen utama dan 7 tetangga terdekat menghasilkan akurasi tertinggi sebesar 82.50%. Ini menunjukkan bahwa pengurangan dimensi melalui PCA dapat membantu memperbaiki performa model dengan mengurangi *noise* dan menyaring fitur yang tidak relevan, terutama ketika dikombinasikan dengan parameter yang optimal.

Hasil akurasi untuk metode K-NN dan PCA kemudian dibandingkan hanya dengan menggunakan algoritma K-NN dan bukan pendekatan PCA yang digunakan untuk mengurangi data. Gambar 5 menunjukkan grafik perbandingan keakuratan kategorisasi gambar spesies ikan menggunakan metode K-NN saja vs kombinasi teknik K-NN dan PCA.

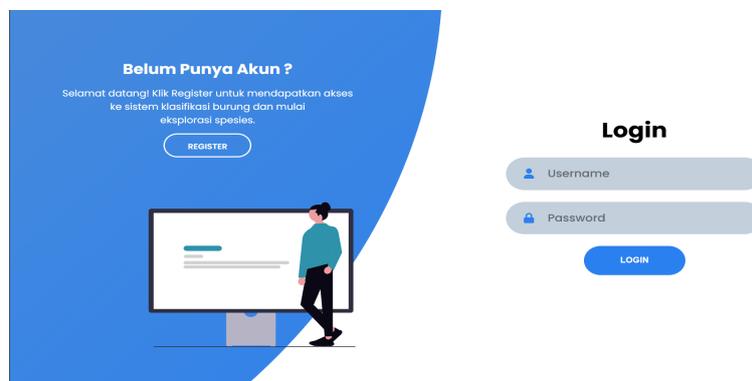


Gambar 5 Grafik perbandingan algoritma K-NN dan PCA dengan algoritma K-NN saja

Kemudian dari pemodelan di atas untuk menentukan spesies burung dilakukan implementasi. Proses implementasi sistem ini dilakukan menggunakan bahasa pemrograman Python yang diintegrasikan ke dalam sebuah aplikasi web menggunakan *framework* Django. Adapun pengimplementasian sistem tersebut, adalah sebagai berikut.

1. Halaman *Login* dan Register

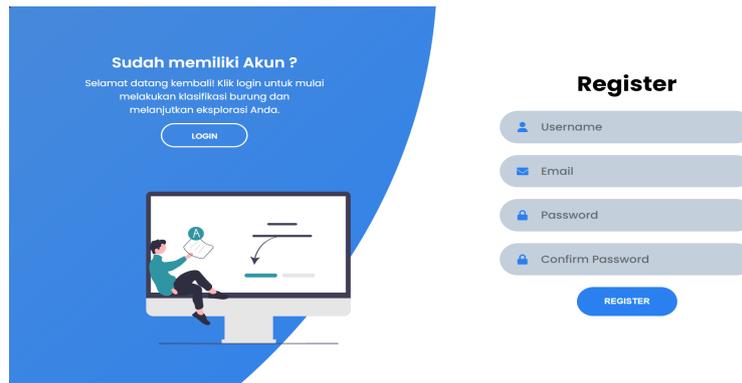
Pengguna dapat mengakses halaman login untuk masuk ke dalam sistem. Pada langkah awal, pengguna diminta untuk memasukkan informasi akun, yaitu nama pengguna dan kata sandi, guna memastikan keamanan dan otentikasi pengguna. Setelah informasi dimasukkan, sistem akan memeriksa kesesuaian data tersebut dengan yang tersimpan di basis data. Jika data yang dimasukkan valid, pengguna akan diarahkan ke halaman utama aplikasi. Namun, jika terdapat kesalahan dalam data yang dimasukkan, sistem akan memberikan notifikasi dan meminta pengguna untuk mengoreksi informasi tersebut.



Gambar 6 Halaman *login*

Dalam implementasi ini, basis data yang digunakan oleh Django adalah SQLite, yang merupakan basis data default Django. SQLite menyimpan data dalam satu file di sistem file, sehingga memudahkan pemindahan antara lingkungan pengembangan dan produksi.

Jika pengguna belum memiliki akun, pengguna dapat menekan tombol register yang ada di halaman *login*, yang akan mengarahkan pengguna ke halaman pendaftaran.



Gambar 7 Halaman register

2. Halaman *Home*

Halaman *Home* berisi penjelasan singkat mengenai aplikasi yang dibuat dalam penelitian ini, yaitu klasifikasi jenis burung menggunakan algoritma KNN.



Gambar 8 Halaman *Home*

Di bagian atas halaman terdapat menu navigasi yang berisi beberapa menu yaitu klasifikasi, riwayat, kontak dan *sign out*. Serta halaman navigasi utama yaitu mulai klasifikasi, yang juga dapat digunakan untuk mengakses halaman klasifikasi.

Jika digulirkan ke bawah, halaman *home* memuat konten yang berisi penjelasan singkat mengenai algoritma yang digunakan untuk penelitian.



Gambar 9 Halaman *Home* (Lanjutan)

Serta, halaman *home* memuat konten yang berisi penjelasan singkat mengenai 4 jenis burung yang akan diklasifikasi dalam penelitian ini.

Gambar 10 Halaman *Home* (Lanjutan)

3. Halaman Klasifikasi

Halaman klasifikasi berisi menu untuk mengunggah gambar burung, pengguna memilih file gambar yang akan diunggah dari folder lokal, kemudian mengklik tombol *upload*. Selanjutnya akan diarahkan ke halaman hasil klasifikasi.

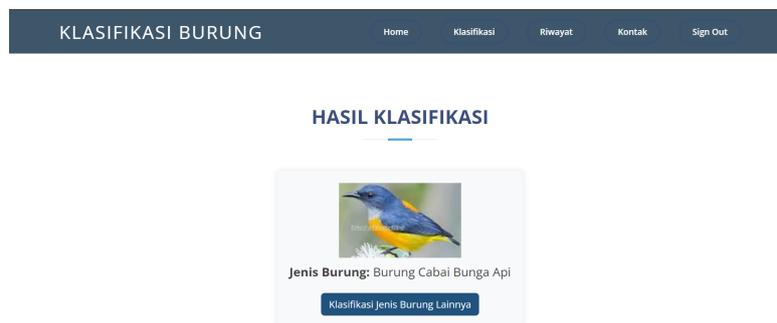


Gambar 11 Halaman Klasifikasi

Saat pengguna mengunggah gambar burung, gambar tersebut dikirim ke server Django untuk *preprocessing*. Setelah ekstraksi fitur, gambar dimasukkan ke model yang telah dilatih dan diekspor dengan joblib. Model kemudian mengklasifikasikan spesies burung, dan hasilnya ditampilkan di halaman klasifikasi.

4. Halaman Hasil Klasifikasi

Halaman hasil klasifikasi menampilkan hasil klasifikasi, berikut adalah tampilan dari halaman hasil klasifikasi.



Gambar 12 Halaman Hasil Klasifikasi

5. Halaman Riwayat Klasifikasi

Halaman riwayat klasifikasi berisi riwayat klasifikasi yang dilakukan sebelumnya. Terdapat tombol hapus pada masing-masing data riwayat yang digunakan untuk menghapus data.



Gambar 4. 1 Halaman Riwayat Klasifikasi

Selanjutnya dilakukan pengujian klasifikasi pada gambar masing-masing burung. Diperoleh hasil sebagai berikut.

Tabel 4. 1 Pengujian klasifikasi

| No | Gambar | Kelas Aktual | Kelas Klasifikasi |
|----|---|------------------------|------------------------|
| 1 |  | Burung Cekakak Suci | Burung Cekakak Suci |
| 2 |  | Burung Cekakak Suci | Burung Cekakak Suci |
| 3 |  | Burung Kuntul Kecil | Burung Kuntul Kecil |
| 4 |  | Burung Kuntul Kecil | Burung Kuntul Kecil |
| 5 |  | Burung Kutilang | Burung Kutilang |
| 6 |  | Burung Kutilang | Burung Kutilang |
| 7 |  | Burung Cabai Bunga Api | Burung Cabai Bunga Api |
| 8 |  | Burung Cabai Bunga Api | Burung Cabai Bunga Api |

Berdasarkan hasil pengujian klasifikasi menggunakan sistem django tersebut, dapat disimpulkan bahwa sistem dapat menerapkan dan mendapatkan hasil klasifikasi yang baik

4. DISKUSI

Penelitian ini membutuhkan referensi dari penelitian-penelitian sebelumnya sebagai rujukan dan pendukung dalam penyelesaian penelitian ini. Berikut beberapa referensi penelitian terdahulu yang berkaitan dengan penelitian ini diantaranya sebagai berikut [14] [16] [17] .

Penelitian yang di lakukan oleh Nuraini dkk dengan judul kombinasi Algoritma K-NN dan PCA pada klasifikasi gambar spesies ikan. Sebuah model untuk mengklasifikasikan foto spesies ikan telah dikembangkan menggunakan kombinasi algoritma *K-NN* dan *PCA* menggunakan Matlab. Proses ekstraksi fitur dalam model ini berfokus pada warna dan bentuk objek. Algoritma *K-NN* berperan dalam mengelompokkan item dengan mempertimbangkan jarak terdekat antar objek. Sementara itu, metode *PCA* digunakan untuk mengurangi dimensi data dengan tetap mempertahankan sebagian besar informasi penting dari fitur asli sesuai kriteria optimal. Integrasi *PCA* dengan *K-NN* bertujuan untuk meningkatkan efisiensi dan akurasi algoritma *K-NN*. Berdasarkan hasil pengujian, akurasi yang dicapai adalah 85%, yang menunjukkan hasil yang sangat baik. Penelitian juga menyimpulkan bahwa penggunaan gabungan algoritma *K-NN* dan *PCA* dalam klasifikasi gambar spesies ikan dapat meningkatkan akurasi sebesar 75% dibandingkan hanya menggunakan metode *K-NN* saja.

Sama halnya dengan penelitian sebelumnya, penelitian ini juga mengembangkan sistem klasifikasi menggunakan algoritma K-NN dan PCA dalam pengurangan dimensi agar akurasi yang didapatkan akan meningkat perbedaannya terdapat pada objek nya saja. Penelitian ini menggunakan objek burung untuk di teliti, penelitian ini mendapatkan hasil akurasi mencapai 82,50% Ini mengindikasikan bahwa menggunakan 26 komponen utama untuk PCA dan 7 tetangga terdekat memberikan keseimbangan optimal antara kompleksitas model dan performa akurasi.

5. KESIMPULAN

Penelitian ini telah mengeksplorasi dan mengembangkan sistem klasifikasi gambar burung menggunakan metode kombinasi *K-Nearest Neighbors* (KNN) dan *Principal Component Analysis* (PCA). Berikut adalah kesimpulan yang dapat ditarik dari penelitian ini. Sistem klasifikasi burung menggunakan *K-Nearest Neighbors* (KNN) dan *Principal Component Analysis* (PCA) dimulai dengan fitur diekstraksi dengan mengonversi gambar dari format BGR ke RGB, menghitung histogram warna, dan melakukan *flattening* untuk mendapatkan vektor satu dimensi. Fitur ini kemudian distandarisasi dengan *Standard Scaler*. Model KNN diimplementasikan dalam aplikasi berbasis web menggunakan *framework* Django. Penelitian ini menggunakan dataset yang terdiri dari 400 gambar burung, terbagi dalam empat kelas, dengan masing-masing kelas berisi 100 gambar. Data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. model KNN dievaluasi setelah menerapkan teknik PCA untuk mereduksi dimensi data. Secara keseluruhan, kombinasi terbaik ditemukan pada $n_components = 26$ dan $n_neighbors = 7$, dengan akurasi mencapai 82,50%. Ini mengindikasikan bahwa menggunakan 26 komponen utama untuk PCA dan 7 tetangga terdekat memberikan keseimbangan optimal antara kompleksitas model dan performa akurasi. Serta skor presisi dan *recall* masing-masing 83,06% dan 82,50%.

Berdasarkan hasil penelitian, beberapa saran dapat diajukan untuk penelitian selanjutnya Untuk meningkatkan kualitas dan akurasi model, disarankan untuk memperluas ukuran dataset dengan menambahkan lebih banyak gambar burung, termasuk variasi yang lebih banyak dalam kondisi pencahayaan dan sudut pengambilan gambar. Data tambahan yang lebih beragam dapat membantu model dalam mengenali variasi yang lebih luas dan meningkatkan kemampuan generalisasi. Melakukan penyesuaian dan optimasi parameter model KNN dan PCA yang lebih mendalam, serta mengeksplorasi berbagai nilai k dan metrik lainnya, dapat membantu dalam menemukan konfigurasi yang paling efektif. Penelitian lebih lanjut dalam pemilihan parameter dapat meningkatkan hasil klasifikasi.

6. DAFTAR PUSTAKA

- [1] K. Kasmiruddin, T. Hidayat, And S. Yunita, "Keanekaragaman Jenis Burung Ekosistem Danau," *Bioedusainsjournal Pendidik. Biol. Dan Sains*, Vol. 5, No. 1, Pp. 197–206, Jun. 2022, Doi: 10.31539/Bioedusains.V5i1.3730.
- [2] Raihan Maulana, Raisya Dwi Zahra Putri, Sindy Fitriani Margareth Sihalo, And Sri Mulyana, "Implementasi Algoritma Convolutional Neural Network Dalam Mengklasifikasi Jenis Burung," *J. Creat. Student Res.*, Vol. 1, No. 6, Pp. 221–231, Nov. 2023, Doi: 10.55606/Jcsrpolitama.V1i6.2966.
- [3] Siti Zuhra Dan Samsul Kamal, "Keanekaragaman Jenis Burung Di Hutan Kota Banda Aceh," *Pros. Semin. Nas. Biot.*, Vol. 10, P. 2022, 2022, [Online]. Available: <https://jurnal.Ar-Raniry.Ac.Id/Index.Php/Pbiotik/Index>
- [4] R. Nuraini, A. Wibowo, B. Warsito, W. A. Syafei, And I. Jaya, "Combination Of K-Nn And Pca Algorithms On Image Classification Of Fish Species," *J. Resti (Rekayasa Sist. Dan Teknol. Informasi)*, Vol. 7, No. 5, Pp. 1026–1032, Aug. 2023, Doi: 10.29207/Resti.V7i5.5178.

- [5] N. H. Sujacka Retno, Rozzi Kesuma Dinata, "Evaluasi Model Data Chatbot Dalam Natural Language Processing Menggunakan K-Nearest Neighbor," *J. Comput. Sci. Inf. Technol.*, Vol. 5, No. 1, Pp. 75–83, 2024, [Online]. Available: [Http://Ejurnal.Umri.Ac.Id/Index.Php/Coscitech/Index](http://ejournal.umri.ac.id/index.php/coscitech/index)
- [6] R. K. Dinata, F. Fajriana, Z. Zulfa, And N. Hasdyna, "Klasifikasi Sekolah Menengah Pertama/Sederajat Wilayah Bireuen Menggunakan Algoritma K-Nearest Neighbors Berbasis Web," *Cess (Journal Comput. Eng. Syst. Sci.)*, Vol. 5, No. 1, P. 33, 2020, Doi: 10.24114/Cess.V5i1.14962.
- [7] Krismiartono.B, "Klasifikasi Kualitas Telur Bebek Menggunakan Metode K-Nearest Neighbor Berdasarkan Warna Kulit Telur," *Skripsi*, P. 9, 2021.
- [8] N. Zaid Munantri, H. Sofyan, And M. Yanu, "Aplikasi Pengolahan Citra Digital Untuk Identifikasi Umur Pohon," 2019.
- [9] Silvia Ratna, "Pengolahan Citra Digital Dan Histogram Dengan Phyton Dan Text Editor Phycharm," In *Technologia*, 2020.
- [10] S. P. Adenugraha, V. Arinal, And D. I. Mulyana, "Klasifikasi Kematangan Buah Pisang Ambon Menggunakan Metode Knn Dan Pca Berdasarkan Citra Rgb Dan Hsv," *J. Media Inform. Budidarma*, Vol. 6, No. 1, P. 9, Jan. 2022, Doi: 10.30865/Mib.V6i1.3287.
- [11] R. Yudistira, A. I. Meha, S. Yulianto, And J. Prasetyo, "Perubahan Konversi Lahan Menggunakan Ndvi, Evi, Savi Dan Pca Pada Citra Landsat 8 (Studi Kasus : Kota Salatiga)," 2019.
- [12] A. S. Aminatus Syarifah, Aditya Akbar Riadi, "Klasifikasi Tingkat Kematangan Jambu Bol Berbasis Pengolahan Citra Digital Menggunakan Metode K-Nearest Neighbor," *Jimp J. Inform. Merdeka Pasuruan*, Vol. 7, 2022.
- [13] H. Mubarak, S. Murni, And M. M. Santoni, *Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Tingkat Kematangan Buah Tomat Berdasarkan Fitur Warna*. 2021.
- [14] S. Raysyah, V. Arinal, And D. I. Mulyana, "Klasifikasi Tingkat Kematangan Buah Kopi Berdasarkan Deteksi Warna Menggunakan Metode Knn Dan Pca," *Sist. Inf. |*, Vol. 8, No. 2, Pp. 88–95, 2021.
- [15] M. K. Rahmadhika And A. M. Thantawi, "Rancang Bangun Aplikasi Face Recognition Pada Pendekatan Crm Menggunakan Opencv Dan Algoritma Haarcascade," *Ikra-Ith Inform. J. Komput. Dan Inform.*, Vol. 5, No. 1, Pp. 109–118, 2021.
- [16] S. 'Uyun Guslianto Safrida Ika, "Klasifikasi Kematangan Buah Sawit Berdasarkan Fitur Warna, Bentuk Dan Tekstur Menggunakan Algoritma K-Nn," *J. Edukasi Dan Penelit. Inform.*, Vol. 9, 2023.
- [17] S. & Dionta, "Pemodelan Pengolahan Citra Klasifikasi Jenis Mangga," *J. Sains Dan Teknol.*, Vol. 5, No. 1, Pp. 347–352, 2023, [Online]. Available: [Http://Journal.Universitaspahlawan.Ac.Id/Index.Php/Jpdk/Article/View/6721%0ahttp://Journal.Universitaspahlawan.Ac.Id/Index.Php/Jpdk/Article/Download/6721/5071](http://journal.universitaspahlawan.ac.id/index.php/jpdk/article/view/6721)