

Klasifikasi Jenis Buah Durian dengan Metode *K-Nearest Neighbor*

Asminah¹, Dedi Setiadi^{1,*}, Tri Susanti¹

¹ Program Studi Teknik Informatika; Institut Teknologi Pagar Alam;
Jl. Masik Siagim No.75 Kelurahan Karang Dalo Kecamatan Dempo Tengah Kota Pagar Alam;
e-mail: asminahamar@gmail.com, dedisetiadi1212@gmail.com, trisantisubagyo8@gmail.com

* Korespondensi: e-mail: dedisetiadi1212@gmail.com

Diterima: 05 November 2023; Review: 20 November 2023; Disetujui: 08 Desember 2023

Cara sitasi: Asminah, Setiadi D, Susanti T. 2023. Klasifikasi Jenis Buah Durian dengan Metode *K-Nearest Neighbor*. Bina Insani ICT Journal. Vol 10 (2): 176-187.

Abstrak: Durian adalah buah yang sangat terkenal baik di Indonesia ataupun di dunia, yang merupakan asli buah tropis, dimana saat ini pada proses klasifikasinya masih secara tradisional atau belum terkomputerisasi, yaitu berdasarkan pengalaman petani, dilihat dari warna dan bentuknya, yang kadang-kadang masih terdapat kesalahan pada proses klasifikasinya. Penelitian ini bertujuan untuk memudahkan proses klasifikasi buah durian dengan menggunakan metode *k-nearest neighbor*, Sistem ini dibangun menggunakan perangkat lunak MATLAB, dan menggunakan model *waterfall* dalam pengembangan sistemnya yang terdiri dari tahap analisis, tahap desain, tahap pengkodean, dan tahap pengujian. Metode pengujiannya menggunakan *confusion matriks* yang dibagi menjadi dua bagian, yaitu data *training* dan data *testing*. Hasil penelitian ini menunjukkan bahwa sistem klasifikasi jenis buah durian dengan menggunakan metode *k-nearest neighbor* dengan pengolahan gambar menggunakan *image processing*, berhasil mengenali 224 dari 240 data *training*, dengan 16 data yang tidak berhasil dikenali. Setelah pengujian dengan menggunakan *validasi holdout* dengan 60 data, sistem mencapai tingkat keberhasilan sebesar 93,3%.

Kata kunci: *Image Processing*;KNN;Klasifikasi;MATLAB;*Waterfall*

Abstract: Durian is a very famous fruit both in Indonesia and in the world, which is a native tropical fruit, where currently the classification process is still traditional or not yet computerized, namely based on the experience of farmers, seen from the color and shape, which sometimes still contains errors. in the classification process. This research aims to facilitate the process of classifying durian fruit using the *k-nearest neighbor* method. This system was built using MATLAB software, and uses a *waterfall* model in system development which consists of the analysis stage, design stage, coding stage and testing stage. The testing method uses a *confusion matrix* which is divided into two parts, namely training data and testing data. The results of this research show that the durian fruit type classification system using the *k-nearest neighbor* method with *image processing* using *image processing*, successfully recognized 224 of the 240 training data, with 16 data that were not recognized. After testing using *holdout validation* with 60 data, the system achieved a success rate of 93.3%.

Keywords: *Image Processing*;KNN;Classification;MATLAB;*Waterfall*

1. Pendahuluan

Dalam era modern saat ini, perkembangan teknologi informasi dan komunikasi berkembang sangat cepat, di mana komputer saat ini dapat ditingkatkan dengan menggunakan sistem kecerdasan buatan, yang mampu mempermudah tugas manusia. Kecerdasan Buatan itu sendiri memiliki kemampuan untuk meniru perilaku manusia, sehingga setiap tindakan yang dilakukannya dapat dianggap cerdas atau pintar. [1] *Artificial Intelligence* adalah disiplin ilmu berbasis komputer yang berfokus pada cara mengembangkan kemampuan mesin (komputer)

untuk melakukan tugas sebagaimana yang dilakukan oleh manusia. Penerapan *Artificial Intelligence*, baik yang kita sadari maupun tidak, telah menjadi bagian dari kehidupan sehari-hari kita. Banyak aplikasi yang telah mengadopsi *Artificial Intelligence* sebagai keunggulan utama dalam penggunaannya. [2] *Artificial Intelligence* dikonsepsi dengan baik sehingga dapat melakukan pekerjaan yang biasa dilakukan oleh manusia. *Artificial Intelligence* terdiri dari empat komponen dasar yang saling terintegrasi. Komponen-komponen ini meliputi: 1) Sistem pakar, yang digunakan untuk mengelola sistem dan sumber daya kinerja, 2) Pemecah masalah heuristik, yang digunakan untuk mengevaluasi sejumlah solusi yang terbatas dan mungkin memerlukan beberapa tahap pencarian untuk mencapai solusi yang mendekati optimal, 3) Pemrosesan bahasa alami (*Natural Language Processing* atau *NLP*) yang merupakan cabang dari kecerdasan buatan yang berkaitan dengan interaksi antara komputer dan bahasa manusia. Tujuan utama dari *NLP* adalah memungkinkan komputer untuk memahami, memproses, dan menghasilkan bahasa manusia dengan cara yang bermakna, yang memungkinkan komunikasi antara manusia dan mesin bisa terjadi dalam bahasa manusia, dan 4) Sistem pakar visualisasi, yang merujuk pada suatu sistem kecerdasan buatan yang dirancang khusus untuk memproses dan menganalisis data visual, seperti gambar, grafik, atau video. Sistem ini menggunakan algoritma dan model kecerdasan buatan untuk melakukan tugas-tugas tertentu dalam konteks visual yang memiliki kemampuan untuk mengenali bentuk dan fitur visual. Fungsi utama dari sistem pakar visualisasi melibatkan kemampuan untuk mengenali pola, objek, atau informasi dalam data visual. Ini dapat mencakup identifikasi objek dalam gambar, deteksi pola dalam grafik, atau bahkan interpretasi konten dalam video. Sistem ini seringkali memanfaatkan teknologi seperti pengolahan citra, jaringan saraf tiruan, atau metode pembelajaran mesin lainnya. Dengan kemampuannya yang terus berkembang, sistem pakar visualisasi digunakan dalam berbagai bidang dalam menyelesaikan tugas dan membantu manusia. [3] Salah satu pemanfaatan AI yaitu di bidang pertanian yang dapat digunakan untuk mengklasifikasi jenis buah durian.

Durian atau *durio zibethinus murr* merupakan buah yang sangat terkenal di Indonesia ataupun dunia. Buah ini juga sering disebut sebagai raja buah, termasuk dalam keluarga *bombacaceae* dan banyak tumbuh di wilayah-wilayah tropis seperti Indonesia dan negara tetangga seperti Malaysia dan Thailand. [4] Buah durian merupakan salah satu buah yang memiliki nilai ekonomi yang tinggi di berbagai negara di Asia Tenggara, termasuk juga Kota Pagar Alam sebagai salah satu daerah penghasil durian terbesar di provinsi Sumatera Selatan. Berdasarkan observasi yang dilakukan bahwa dalam menentukan jenis durian masih secara manual berdasarkan karakteristiknya, sehingga mengakibatkan kesulitan dan membutuhkan keahlian khusus dalam menentukan jenis durian. Oleh karena itu, dalam perkembangan teknologi informasi dan komunikasi bahwa untuk mengklasifikasi buah durian menjadi sangat penting dalam mempermudah proses ini. Salah satu metode yang telah terbukti berhasil dalam masalah klasifikasi adalah metode *k-nearest neighbor*.

Metode *k-nearest neighbor* merupakan bentuk pendekatan yang digunakan untuk mengelompokkan objek-objek menurut data pembelajaran yang memiliki jarak paling dekat dengan objek yang akan dikelompokkan. [5] Algoritma ini terkait dengan proses klasifikasi objek berdasarkan jarak di antara mereka. Dalam algoritma ini, data pembelajaran yang diperoleh diproyeksikan ke dalam ruang berdimensi tertentu. Ruang tersebut kemudian dibagi-bagi menjadi bagian-bagian yang akan digunakan sebagai pengklasifikasi data rujukan. Setiap titik dalam ruang ini ditandai dengan kelas c tertentu. Kelas c adalah hasil klasifikasi yang paling umum ditemukan di antara k tetangga terdekat dari titik tersebut. Metode *K-NN* ini bersifat "*supervised*", yang berarti bahwa ketika sebuah instansi baru (*query instance*) harus diklasifikasikan, itu dilakukan berdasarkan mayoritas kategori yang ditemukan di antara k tetangga terdekat dari instansi tersebut. [6] Metode ini digunakan untuk mengklasifikasi jenis buah durian dengan bantuan teknologi berbasis komputer yang melibatkan pemrosesan citra digital.

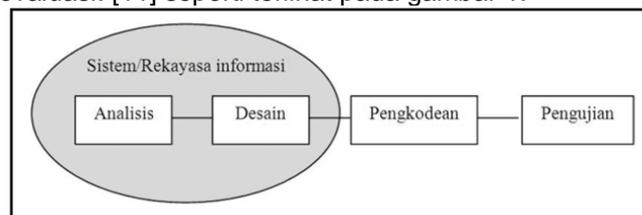
Citra digital atau *image processing* adalah sebuah bidang studi yang mengeksplorasi berbagai aspek terkait dengan peningkatan mutu gambar (seperti meningkatkan kontras, mengubah warna, memulihkan citra), mengubah gambar (meliputi rotasi, translasi, transformasi geometri, dan perubahan skala), dengan tujuan memudahkan interpretasi oleh manusia atau mesin (komputer). [7] Pemrosesan citra bertujuan membantu manusia dalam mengidentifikasi atau mengkategorikan objek secara efisien, dengan kecepatan, ketepatan, dan kemampuan

untuk memproses sejumlah besar data secara bersamaan. [8] masukannya adalah citra dan keluarannya juga citra tapi dengan kualitas lebih baik daripada citra masukannya.

Dari penelitian yang dilakukan oleh [9] dan hasil penelitian menunjukkan bahwa metode klasifikasi KNN yang diajukan mampu mengklasifikasikan jambu biji dengan tingkat akurasi sebesar 91.25% saat menggunakan nilai tetangga terbaik $K=3$. Fitur-fitur yang diekstraksi untuk tujuan klasifikasi jambu biji mencakup r, g, b, luas cacat, energi, homogenitas, dan kontras. Dan juga berdasarkan penelitian [10] dari percobaan klasifikasi tingkat kematangan jeruk nipis menggunakan fitur warna dengan metode *k-nearest neighbor*, dapat disimpulkan bahwa menggunakan ekstraksi fitur mean RGB dan menerapkan metode *k-nearest neighbor* efektif dalam mengklasifikasikan tingkat kematangan jeruk nipis berdasarkan ciri-ciri warna. Oleh sebab itu peneliti menggunakan metode K-NN pada data buah durian, karena penggunaan *k-nearest neighbors* pada dataset dapat memiliki beberapa alasan, terutama ketika karakteristik data dan tujuan analisis sesuai dengan kelebihan metode K-NN. Berikut adalah beberapa alasan penggunaan K-NN, yaitu tidak memerlukan pelatihan yang signifikan, tidak memerlukan proses pelatihan yang kompleks. Dalam konteks *dataset* yang sedikit, di mana tidak ada banyak data untuk dilatih, K-NN dapat menjadi pilihan yang cepat dan sederhana.

2. Metode Penelitian

Metodologi Penelitian adalah suatu pendekatan untuk memahami obyek yang menjadi fokus ilmu tertentu melalui kegiatan sistematis untuk menyelesaikan masalah dengan menerapkan metode ilmiah. Penelitian ini menggunakan metode kuantitatif karena studi kasus melibatkan data yang memerlukan perhitungan matematis, khususnya dalam statistika. Pendekatan yang digunakan dalam penelitian ini adalah model *waterfall*, adalah kerangka kerja *Software Development Life Cycle* atau SDLC, yang merupakan proses atau bentuk siklus hidup dari suatu pengembangan *software* atau aplikasi. SDLC terdiri dari beberapa langkah atau tahapan yang ditempuh oleh tim pengembang perangkat lunak untuk merancang, mengembangkan, menguji, dan menerapkan perangkat lunak dengan tujuan menghasilkan produk perangkat lunak berkualitas tinggi. Tahapan-tahapan umum dalam SDLC melibatkan perencanaan, studi kelayakan, desain, implementasi (pengkodean), pengujian, implementasi, pemeliharaan, dan evaluasi. [11] seperti terlihat pada gambar 1.



Sumber : Penelitian (2023)

Gambar 1. Metode SDLC

Analisis

Pengumpulan data dilakukan secara cermat dan mendalam untuk merinci persyaratan perangkat lunak sehingga perangkat lunak dapat memahami kebutuhan yang diinginkan oleh pengguna. [12] Pengumpulan data dilakukan dengan cara observasi secara langsung ke objek penelitian dan wawancara langsung ke petani dan dinas pertanian kota Pagar Alam untuk mengetahui beberapa jenis dan contoh durian yang ada di Pagar Alam, selain itu pengumpulan data juga dilaksanakan dengan cara dokumentasi untuk mengumpulkan foto-foto jenis durian, guna untuk mengetahui jenis durian apa saja yang akan di klasifikasikan dan mengetahui perangkat lunak apa yang dibutuhkan oleh petani dan dinas pertanian kota Pagar Alam.

Dari hasil pengumpulan data diperoleh beberapa jenis durian yang ada di kota Pagar Alam ataupun di Indonesia, tetapi dalam batasan penelitian ini, digunakan hanya 2 jenis durian saja yaitu durian belimbing dan durian montong, karena dua jenis durian ini yang banyak terdapat di kota Pagar Alam. Jenis durian belimbing ini memiliki bentuk buah bulat segi belimbing, warna buah hijau tua, duri buah tumpul, berat buah 2-3 kg, kulit buah tebal, tebal daging buah 1,4 cm, ciri khusus warna daging buah kuning keemasan, rasa daging buah manis sekali, tekstur halus, aroma tidak terlalu harum, produksi 200-400 buah per pohon serta tahan terhadap hama penggerek buah dan durian ini berbuah sepanjang musimnya. Adapun data durian jenis belimbing bisa dilihat pada gambar 2.



Sumber : Hasil Penelitian (2023)

Gambar 2. Durian jenis belimbing

Sementara itu, durian montong merupakan tumbuhan yang berasal dari Thailand. Durian montong umumnya memiliki umur yang panjang, dengan kayu yang tebal, tumbuh tegak, dan memiliki banyak cabang. Tanaman ini memiliki daun yang bertangkai pendek dengan warna kekuningan, dan permukaan daunnya bersinar, tetapi bagian bawahnya tampak agak suram. Bunga durian montong muncul di batang dan cabang yang sudah besar, dengan warna putih coklat yang mengkilap, dan biasanya berbunga pada bulan Januari. Daging buah durian montong berwarna kuning, tebal, kering, dan berlemak. Informasi lebih lanjut mengenai durian jenis montong dapat ditemukan pada gambar 3.



Sumber : Penelitian (2023)

Gambar 3. Durian jenis montong

Setelah data durian didapatkan selajutnya yakni *resizing* ukuran gambar durian. *Resizing* dilakukan untuk 225 data menyamakan semua ukuran gambar durian, Ukuran *image* durian yang digunakan ialah 150x150 pixel, 150 sebagai panjang dan lebar dari gambar yang di *resizing*. serta 25 data diubah percentage menjadi 500x500, 25 data diubah *pixel* nya menjadi 1000x1000, dan 25 data di ubah angelnya 180 derajat. sehingga nantinya data durian mempunyai ukuran tersendiri. Tahap *resizing* dilakukan dengan menggunakan aplikasi *photoshop* dengan cara menginput satu per satu hasil foto durian dari kamera *smartphone* kedalam aplikasi *photoshop* dan diubah ukuran *pixel* gambar buah durian sesuai dengan yang telah ditentukan.

Tahap selanjutnya yaitu *grayscale image* durian untuk melihat komposisi bentuk dari durian yang menyerupai keasliannya. *Grayscale* merupakan kumpulan warna *pixel* yang terdapat pada rentang gradasi warna putih dan warna hitam.[13] Tahap *grayscale* dalam pengolahan citra merujuk pada konversi citra dari representasi warna penuh (RGB atau yang lainnya) menjadi representasi skala keabuan (*grayscale*). Citra *grayscale* hanya menggunakan tingkat keabuan dari putih ke hitam tanpa mengandung informasi warna. Melakukan konversi pada citra asli menjadi *grayscale* yakni dengan perintah inputan “img_gray” dan menggunakan function file “rgb2gray” dalam program yang terisi formula matematika yang dapat memproses input data untuk menghasilkan output sebagai citra *grayscale*. Adapun coding untuk melakukan *grayscale* pada MATLAB sebagai berikut :

```
Img_gray = rgb2gray(Img);
```

Pada tahap ekstraksi ciri ini, dua neuron input digunakan, proses ekstraksi ciri di dalam jaringan saraf dapat melibatkan pembobotan dan transformasi dari nilai-nilai input ini untuk menciptakan representasi fitur yang lebih abstrak atau relevan. Ini sering dilakukan melalui penggunaan bobot dan fungsi aktivasi di dalam lapisan-lapisan tersembunyi (*hidden layers*)

jaringan saraf. yakni korelasi dan energi. Proses ekstraksi ciri pada kedua neuron tersebut menggunakan gambar yang telah dikonversi dari citra asli ke skala abu-abu. Hasil dari proses ini dapat dilihat pada gambar 4.

	1	2
1	0.8507	0.3785
2	0.8958	0.3660
3	0.9311	0.4436
4	0.9251	0.4414
5	0.9187	0.2938
6	0.9655	0.3034
7	0.9468	0.3450
8	0.9646	0.3186
9	0.9380	0.3925
10	0.9451	0.3498
11	0.9546	0.3781
12	0.9685	0.2560
13	0.9433	0.5245

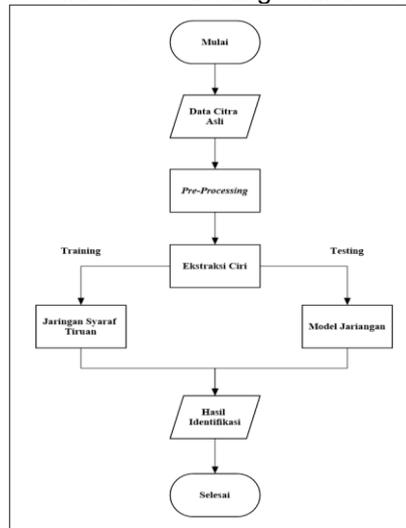
	1	2
1	0.8895	0.4810
2	0.9498	0.4634
3	0.9533	0.4809
4	0.9098	0.6193
5	0.9252	0.4547
6	0.9082	0.4141
7	0.9721	0.4476
8	0.9502	0.5860
9	0.8503	0.5408
10	0.9539	0.4550
11	0.9075	0.6390
12	0.9736	0.5416
13	0.9468	0.4087

Sumber : Hasil Penelitian (2023)

Gambar 4. (a) Ekstraksi data latih (b) Ekstraksi data uji

Desain

Perancangan perangkat lunak merupakan serangkaian langkah yang melibatkan perencanaan pembuatan program perangkat lunak, mencakup elemen-elemen seperti struktur data, arsitektur perangkat lunak, representasi antarmuka, dan prosedur pengkodean. [14] Pada tahap ini, kebutuhan perangkat lunak yang telah dianalisis dipindahkan ke representasi desain agar dapat diubah menjadi program pada tahap selanjutnya. Desain perangkat lunak yang dihasilkan dalam tahap ini melibatkan pembuatan *flowchart*, yang menjelaskan alur menggunakan simbol-simbol yang sederhana dan mudah dimengerti, [15] digunakan untuk menggambarkan proses kerja sistem pada klasifikasi jenis durian. *Flowchart* akan memudahkan dalam proses pembuatan implementasi sistem. Yang terlihat terlihat pada gambar 5. berikut.



Sumber : Hasil Penelitian (2023)

Gambar 5. *Flowchart* sistem klasifikasi buah durian

Pengkodean

Desain yang sudah dibuat pada tahap sebelumnya harus diimplementasikan menjadi suatu perangkat lunak atau aplikasi yang bisa berjalan sesuai dengan diinginkan. Pada tahap ini, peneliti melakukan pengembangan pada program berbasis komputer yang sesuai dengan struktur desain yang telah dirancang pada tahap sebelumnya, [16] dengan cara membuat program atau *coding* yang sesuai dengan perancangan yang telah dibuat dengan menggunakan *tools* atau aplikasi MATLAB karena MATLAB memiliki sejumlah keunggulan dan fasilitas yang membuatnya sangat cocok untuk pengolahan citra. Berikut adalah beberapa alasan mengapa MATLAB sering digunakan dalam *image processing* :

Keberagaman Fungsi dan Toolbox, MATLAB menyediakan *toolbox* khusus untuk pengolahan citra, seperti *Image Processing Toolbox*. *Toolbox* ini menyediakan berbagai fungsi dan algoritma siap pakai untuk berbagai tugas pengolahan citra, memudahkan pengguna dalam mengimplementasikan solusi tanpa harus membangun semuanya dari awal.

Kemudahan Implementasi, MATLAB memiliki sintaks yang mudah dipahami dan diterapkan, sehingga cocok untuk keperluan pengolahan citra yang sering kali melibatkan operasi matriks dan vektor. Ini membuatnya ramah pengguna dan efisien untuk mengimplementasikan algoritma pengolahan citra.

Visualisasi, MATLAB menyediakan alat visualisasi yang kuat. Setelah pengolahan citra dilakukan, pengguna dapat dengan mudah memvisualisasikan hasilnya menggunakan fungsi-fungsi visualisasi MATLAB. Hal ini penting untuk memahami dan memeriksa efek dari berbagai operasi pada citra.

Fleksibilitas, MATLAB merupakan platform yang sangat fleksibel, memungkinkan pengguna untuk mengintegrasikan berbagai metode pengolahan citra, melakukan eksperimen, dan mengadaptasi solusi sesuai kebutuhan proyek.

Pemrosesan Paralel dan GPU, MATLAB mendukung pemrosesan paralel dan penggunaan GPU (*Graphics Processing Unit*), yang memungkinkan percepatan signifikan dalam eksekusi algoritma pengolahan citra, terutama pada data yang besar.

Pengujian

Pengujian dilaksanakan dengan tujuan mempertahankan mutu perangkat lunak yang dibuat dan untuk mengoptimalkan efisiensi biaya produksi. [17] Pengujian berfokus pada aspek logika dan fungsional perangkat lunak, dan memeriksa bahwa setiap komponen telah diuji. Tujuan pengujian adalah untuk mengurangi kemungkinan kesalahan (*error*) dan memastikan bahwa hasil keluaran sesuai dengan harapan.

3. Hasil dan Pembahasan

Hasil dari penelitian ini yaitu suatu sistem berbasis komputer untuk mengklasifikasi buah durian menjadi dua jenis berdasarkan batasan yang telah ditentukan, dengan menggunakan aplikasi MATLAB sebagai alat bantu. Sistem klasifikasi adalah salah satu program komputer yang diterapkan dalam konteks penelitian ilmu pengetahuan dan teknologi.

Hasil

Berdasarkan hasil penelitian, sistem yang menggunakan metode *K-Nearest Neighbor* untuk mengklasifikasikan jenis durian dengan menggunakan *image processing* telah diimplementasikan dan diuji melalui metode *holdout validation*. Sistem ini memberikan petunjuk tentang cara mengklasifikasi jenis durian melalui sistem yang telah dikembangkan. Untuk mendapatkan hasil klasifikasi yang tepat sesuai dengan yang diharapkan, menggunakan metode *k-nearest neighbor* dengan mengumpulkan data-data awal, melakukan pemrosesan, dan menghasilkan klasifikasi sesuai dengan inputan yang dilakukan oleh *user*. Dengan sistem klasifikasi yang berbasis komputer ini, orang biasa atau non-ahli dapat mengatasi masalah yang kompleks, yang sebelumnya hanya dapat dipecahkan oleh para ahli. Dalam penelitian ini melibatkan 300 data dari buah durian, dengan 240 data digunakan sebagai data *training* dan 60 data digunakan sebagai data *testing*. Dan dari 240 data *training* tersebut, dua jenis durian yang telah ditentukan diwakili oleh masing-masing 120 data buah durian. Model *trainingdx* digunakan dalam penelitian ini, diuji dengan variasi jumlah *hidden layer* dan jumlah *neuron* pada setiap lapisan untuk mencapai akurasi optimal. Model ini juga digunakan dalam pengujian data tunggal untuk menunjukkan hasil dari proses klasifikasi yang dilakukan.

Pembahasan

Pada sektor ini, akan dibicarakan tentang tampilan *interface visual* dari sistem klasifikasi jenis buah durian berbasis komputer dengan memanfaatkan jaringan syaraf tiruan, yang dapat dilihat pada ilustrasi pada gambar 6.



Sumber : Hasil Penelitian (2023)

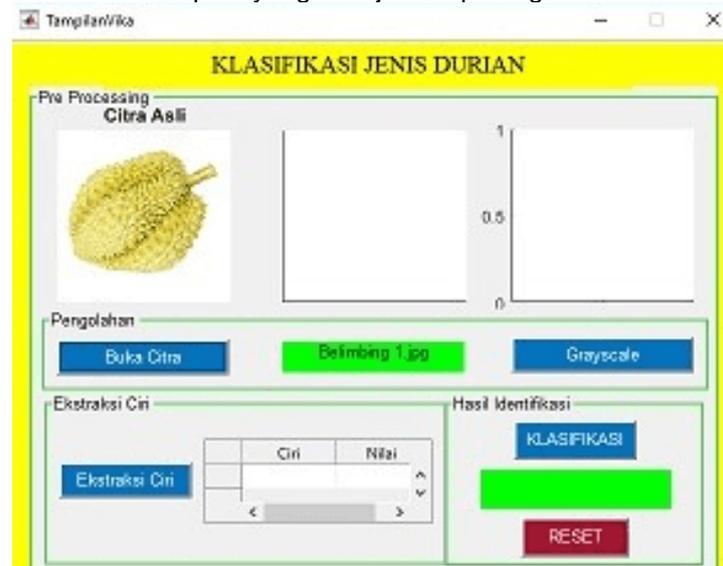
Gambar 6. Tampilan *user interface* keseluruhan

Bagian pra-pemrosesan

Pra-pemrosesan data adalah langkah penting dalam mempersiapkan data untuk analisis atau pelatihan model klasifikasi. Data berupa gambar dari buah durian yang terdiri dari berbagai jenis, yang difoto secara manual dengan menggunakan kamera resolusi tinggi, agar hasilnya bagus. Kemudian foto diedit dengan menggunakan aplikasi *photoshop* agar semua gambar memiliki ukuran yang sama. Dalam segmen pra-pemrosesan ini, terdapat tiga sumbu (2D) yang menampilkan citra asli dan citra yang telah diproses untuk hasil ekstraksi ciri.

Bagian pemrosesan

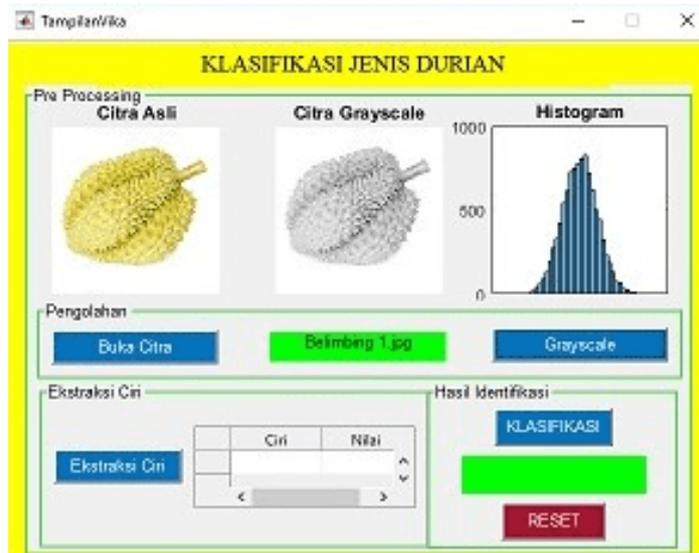
Pada bagian pemrosesan ini, pengguna diminta untuk memilih file citra asli gambar buah durian yang akan diproses klasifikasi. Pengguna akan diarahkan ke antarmuka untuk memilih file citra asli buah durian, yang diambil dari penyimpanan media baik di computer atau tempat penyimpanan lainnya. Setelah pemilihan file citra asli buah durian selesai, maka program akan menampilkan antarmuka *user* seperti yang ditunjukkan pada gambar 7.



Sumber : Hasil Penelitian (2023)

Gambar 7. Tampilan *user interface* citra asli

Setelah program menerima masukan dari file citra asli buah durian, selanjutnya *user* di bagian pemrosesan menekan tombol "*Grayscale*" untuk memperoleh nilai ekstraksi ciri, selanjutnya menampilkan antarmuka *user* oleh program, sebagaimana terlihat pada gambar 8.

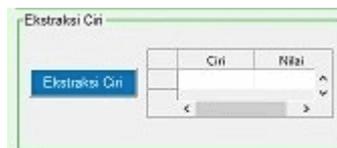


Sumber : Hasil Penelitian (2023)

Gambar 8. Tampilan *user interface grayscale*

Bagian Ekstraksi Ciri

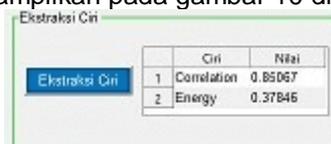
Di dalam bagian ekstraksi ciri ini, dimana *user* diminta untuk menekan tombol "Ekstraksi Ciri". Setelah itu akan ditampilkan ciri-ciri dan nilai-nilai dari citra asli yang telah diproses pada tahap sebelumnya. Pada tahap ini ada dua jenis ciri dari ekstraksi ciri, yaitu korelasi dan energi. Dan berikut adalah tampilan antarmuka *user* untuk bagian ekstraksi ciri yang dapat dilihat pada gambar 9.



Sumber : Penelitian (2023)

Gambar 9. Tampilan *user interface* ekstrasi ciri

Setelah tombol "Ekstraksi Ciri" ditekan, selanjutnya program akan menampilkan tampilan antarmuka *user* sebagaimana ditampilkan pada gambar 10 dibawah ini.



Sumber : Hasil Penelitian (2023)

Gambar 10. Tampilan *user interface* ekstrasi ciri lanjutan

Bagian Hasil

Pada bagian hasil klasifikasi yang ditampilkan, jika masih kosong atau tidak menunjukkan hasil yang diinginkan untuk gambar buah durian yang dimasukkan ke dalam program MATLAB, pengguna selanjutnya menekan tombol "Klasifikasi". Hal ini bertujuan untuk mengidentifikasi jenis buah durian yang sedang diklasifikasi. Di samping tombol "Klasifikasi", terdapat juga opsi "Reset" yang dapat digunakan untuk menghapus dan memulai kembali proses pada gambar durian berikutnya. Berikut adalah tampilan setelah pengguna menekan tombol "Klasifikasi", yang dapat dilihat pada ilustrasi gambar 12 di bawah ini.



Sumber : Hasil Penelitian (2023)

Gambar 11. Hasil indentifikasi

Pengujian

Pada tahap pengujian dengan menggunakan *holdout validation*, [18] yang merupakan suatu metode pengujian model yang umum digunakan dalam pembelajaran mesin (*machine learning*) untuk mengevaluasi kinerja suatu model pada *dataset* yang tidak digunakan selama proses pelatihan. Metode ini melibatkan *dataset* yang ada dibagi menjadi dua bagian utama: satu bagian digunakan untuk melatih model (*training set*), sementara bagian yang lainnya disimpan dan tidak digunakan selama pelatihan, kemudian digunakan untuk menguji model setelah pelatihan (*testing set*), dengan cara membagi data-data yang ada menjadi dua bagian yang masing-masing, yaitu data *training* sebesar 80 % dan data *testing* sebesar 20 %. Metode ini membantu mengidentifikasi apakah model yang telah dilatih dapat menggeneralisasi dengan baik pada data baru yang tidak terlibat selama pelatihan. *Validasi holdout* sering digunakan untuk menghindari *overfitting*, yaitu kondisi di mana model terlalu sesuai dengan data pelatihan dan kinerjanya menurun saat diterapkan pada data baru. Sehingga uji validasi *holdout* yang digunakan untuk mengukur akurasi pada sistem klasifikasi akan terjadi dua kali.

Evaluasi yang dilakukan pada data *training* dan juga pada data *testing*, dilakukan agar memperoleh nilai ketepatan dari proses klasifikasi pada perangkat lunak yang telah dikembangkan dengan menggunakan dataset yang telah tersedia. [19] Pengujian pada data *training* dilakukan dengan mengimplementasikan arsitektur dari jaringan syaraf tiruan dengan data buah durian yang digunakan yaitu sebanyak 240 gambar dari 2 jenis buah durian belimbing dan buah durian montong. Dan pada pengujian data *testing* yang akan di ujikan sebanyak 60 gambar jenis buah durian.

Evaluasi model performa menggunakan *confusion matrix*

Confusion matrix adalah matrik evaluasi untuk masalah klasifikasi pada *machine learning* di mana hasil *outputnya* terdiri dari dua buah kelas atau bisa lebih.. [20] *Confusion matrix* adalah suatu metode yang digunakan dalam evaluasi kinerja model klasifikasi pada *machine learning*. Ini memberikan gambaran *holistik* tentang seberapa baik model dapat memprediksi kelas target tertentu dan memberikan informasi tentang jenis kesalahan yang dilakukan oleh model. *Confusion matrix* terdiri dari empat elemen utama :

True Positive (TP), Jumlah pengamatan positif yang benar-benar diklasifikasikan dengan benar oleh model. Ini berarti model secara tepat mengidentifikasi instance yang sebenarnya positif.

True Negative (TN), Jumlah pengamatan negatif yang benar-benar diklasifikasikan dengan benar oleh model. Ini berarti model secara tepat mengidentifikasi instance yang sebenarnya negatif.

False Positive (FP), Jumlah pengamatan negatif yang keliru diklasifikasikan sebagai positif oleh model. Ini juga dikenal sebagai Type I error atau kesalahan tipe I.

False Negative (FN), Jumlah pengamatan positif yang keliru diklasifikasikan sebagai negatif oleh model. Ini juga dikenal sebagai Type II error atau kesalahan tipe II.

Confusion matrix membentuk matriks dua dimensi yang menggambarkan distribusi prediksi model terhadap kelas yang sebenarnya. Dengan menggunakan empat elemen di atas, kita dapat menghitung beberapa metrik evaluasi kinerja model, seperti:

- Akurasi (*Accuracy*) : $(TP + TN) / (TP + TN + FP + FN)$
- Presisi (*Precision*) : $TP / (TP + FP)$
- *Sensitivitas* atau *Recall* : $TP / (TP + FN)$
- *Spesifisitas* : $TN / (TN + FP)$
- F1-Score : $2 * (Precision * Recall) / (Precision + Recall)$

Confusion matrix sangat penting untuk memahami kinerja model secara lebih mendalam daripada menggunakan metrik akurasi saja. Ini membantu mengidentifikasi di mana model cenderung membuat kesalahan dan dapat membantu penyesuaian atau perbaikan model untuk meningkatkan kinerjanya..

Data training

Tabel 1. Tabel evaluasi model *confusion matrix*

Confusion Matrix		Kelas hasil prediksi	
		Positive	Negative
Kelas sebenarnya	True	TP = 224	FP = 16
	False	FN = 0	TN = 0

Sumber : Hasil Penelitian (2023)

Rumus *confusion matrix* digunakan untuk menghitung akurasi, presisi, recall, dan skor F-1. Akurasi diukur sebagai perbandingan antara jumlah data yang terklasifikasi dengan benar

dan total data keseluruhan. Akurasi mencerminkan sejauh mana model mampu mengklasifikasikan data secara benar.

$$\begin{aligned}
 Accuracy &= \frac{TP+TN}{TP+TN+FP+FN} \times 100 \% \\
 &= \frac{224+0}{224+0+0+16} \times 100 \% \\
 &= \frac{224}{240} \times 100 \% \\
 Accuracy &= 93,3 \%
 \end{aligned}$$

Precision mengindikasikan sejauh mana keakuratan antara data yang diminta dan hasil prediksi yang diberikan oleh model.

$$\begin{aligned}
 Precision &= \frac{TP}{TP+FP} \times 100 \% \\
 &= \frac{224}{224+16} \times 100 \% \\
 Precision &= 93,3 \%
 \end{aligned}$$

Recall atau *sensitivity* mencerminkan keberhasilan model dalam mendeteksi kembali informasi.

$$\begin{aligned}
 Recall &= \frac{TP}{TP+FN} \times 100 \% \\
 &= \frac{224}{224+0} \times 100 \% \\
 Recall &= 100 \%
 \end{aligned}$$

Skor F-1 mencerminkan rasio rata-rata presisi dan recall yang diberi bobot. Meskipun akurasi cocok digunakan sebagai indikator performa algoritma pada dataset dengan jumlah *false negative* dan *false positive* yang mendekati (simetris), namun jika angka-angka tersebut tidak seimbang, disarankan untuk menggunakan skor F-1 sebagai acuan.

$$\begin{aligned}
 F-1 \text{ score} &= \frac{2 \times Recall \times Precision}{Recall + Precision} \\
 &= \frac{2 \times 1 \times 0,93}{1 + 0,93} \\
 &= \frac{1,86}{1,93} \\
 &= 0,96 * 100 \% \\
 F-1 \text{ score} &= 96 \%
 \end{aligned}$$

Dari evaluasi matriks konfusi pada model di atas, dapat disimpulkan bahwa akurasi mencapai 93,3%, presisi sebesar 93,3%, recall mencapai 100%, dan skor F-1 mencapai 96%, dengan pengujian menggunakan 240 data pelatihan untuk jenis buah durian..

Data testing

Tabel 2. Tabel evaluasi model *confusion matrix*

Confusion Matrix		Kelas hasil prediksi	
		Positive	Negative
Kelas sebenarnya	True	TP = 56	FP = 4
	False	FN = 0	TN = 0

Sumber: Hasil Penelitian (Tahun)

$$\begin{aligned}
 Accuracy &= \frac{TP+TN}{TP+TN+FP+FN} \times 100 \% \\
 &= \frac{56+0}{56+0+0+4} \times 100 \% \\
 &= \frac{56}{60} \times 100 \% \\
 Accuracy &= 93,3 \% \\
 Precision &= \frac{TP}{TP+FP} \times 100 \% \\
 &= \frac{56}{56+4} \times 100 \% \\
 Precision &= 93,3, \% \\
 Recall &= \frac{TP}{TP+FN} \times 100 \% \\
 &= \frac{56}{56+0} \times 100 \%
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall} &= 100 \% \\
 &= \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \\
 \text{F-1 score} &= \frac{2 \times 1 \times 0,93}{1 + 0,93} \\
 &= \frac{1,86}{1,93} \\
 &= 0,96 * 100 \% \\
 \text{F-1 score} &= 96 \%
 \end{aligned}$$

Dari hasil evaluasi matriks konfusi pada model di atas, dapat disimpulkan bahwa akurasi mencapai 93,3%, presisi sebesar 93,3%, recall mencapai 100%, dan skor F-1 mencapai 96%, dengan pengujian menggunakan 60 data *testing*.

4. Kesimpulan

Sistem klasifikasi ini membantu pengguna dalam melakukan pengklasifikasian jenis buah durian dengan menggunakan algoritma *k-nearest neighbor*. Selain itu, sistem klasifikasi ini dapat digunakan sebagai alat untuk menerapkan keahlian atau pengetahuan seorang ahli dalam menentukan klasifikasi jenis buah durian dengan algoritma *k-nearest neighbor*. Hasil dari program klasifikasi buah durian dengan algoritma *k-nearest neighbor*, yang dibandingkan dengan perhitungan manual, menunjukkan akurasi data *training* sebesar 93,3% dari total 240 data *training*. Setelah melalui proses klasifikasi, akurasi data *testing* mencapai 93,3% dari total 60 data *testing*. Dari model jaringan syaraf tiruan yang dibuat telah dapat mengidentifikasi beberapa jenis buah durian dengan sangat baik. Ini dapat dilihat dari hasil uji data yang tidak terlibat dalam proses pelatihan, di mana 56 data teridentifikasi dengan benar dan 4 data teridentifikasi dengan kesalahan.

Referensi

- [1] M. Siahaan, C. H. Jasa, K. Anderson, M. V. Rosiana, S. Lim, and W. Yudianto, "Penerapan Artificial Intelligence (AI) Terhadap Seorang Penyandang Disabilitas Tunanetra," *Journal of Information System and Technology (JOINT)*, vol. 1, no. 2, pp. 186–193, 2020.
- [2] R. Rahmawati, F. Fahmi, C. Yusnar, Y. Yunus, and S. Amra, "Penerapan Deteksi Wajah Sebagai Implementasi Artificial Intelligence (AI) pada Guru Sekolah Muhammadiyah Kota Lhokseumawe," in *Prosiding Seminar Nasional Politeknik Negeri Lhokseumawe*, 2022, vol. 6, no. 1, pp. 64–69.
- [3] F. E. Jelahun, H. Y. Utang, Y. E. Jelahun, and L. Jehamat, "Menalar Skeptis Adopsi Artificial Intelligence (AI) di Indonesia: Sebuah Tinjauan Filsafat Ilmu Komunikasi," 2021.
- [4] D. Yudhayanti and M. Restiani, "Uji Mutu Tepung Biji Durian Sebagai Bahan Pangan Alternatif Berdasarkan Kadar Air Dan Kadar Abu Serta Cemaran Mikroba," *MEDFARM: Jurnal Farmasi dan Kesehatan*, vol. 8, no. 2, pp. 43–48, 2019.
- [5] A. A. WPR, F. Rozi, and F. Sukmana, "Prediksi penjualan produk unilever menggunakan metode k-nearest neighbor," *JIFI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, vol. 6, no. 1, pp. 155–160, 2021.
- [6] S. Ramadana, M. Diono, M. Susantok, and S. Ahdan, "Indoor location tracking pegawai berbasis Android menggunakan algoritma k-nearest neighbor: Indoor location tracking of employees based on Android using the k-nearest neighbor algorithm," *JITEL (Jurnal Ilmiah Telekomunikasi, Elektronika, Dan Listrik Tenaga)*, vol. 1, no. 1, pp. 51–58, 2021.
- [7] A. H. Hasibuan, T. Zebua, and R. K. Hondro, "Penerapan Metode Sobel Edge Detection dan Image Processing Untuk Mengetahui Diameter Apel Fuji Menggunakan Aplikasi Matlab," *JURIKOM (Jurnal Riset Komputer)*, vol. 7, no. 3, pp. 450–454, 2020.
- [8] F. F. Maulana and N. Rochmawati, "Klasifikasi Citra Buah Menggunakan Convolutional Neural Network," *Journal of Informatics and Computer Science (JINACS)*, vol. 1, no. 02, pp. 104–108, 2019.
- [9] T. Y. Prahudaya and A. Harjoko, "Metode Klasifikasi Mutu Jambu Biji Menggunakan Knn Berdasarkan Fitur Warna Dan Tekstur," *Jurnal Teknosains*, vol. 6, no. 2, pp. 113–123, 2017.
- [10] C. Paramita, E. H. Rachmawanto, and C. A. Sari, "Klasifikasi Jeruk Nipis Terhadap Tingkat Kematangan Buah Berdasarkan Fitur Warna Menggunakan K-Nearest Neighbor," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 4, no. 1, pp. 1–6, 2019.
- [11] A. A. Wahid, "Analisis metode waterfall untuk pengembangan sistem informasi," *J. Ilmu-*

- ilmu Inform. dan Manaj. STMIK*, no. November, pp. 1–5, 2020.
- [12] D. Maulana and I. Suryani, “Perancangan Sistem Informasi Pendaftaran Peserta Didik Baru Berbasis Web pada SMK Kosgoro Kota Bogor,” *Indonesian Journal on Software Engineering (IJSE)*, vol. 5, no. 1, pp. 9–18, 2019.
- [13] G. J. N. Putri, “METODE BACKGROUND SUBTRACTION UNTUK MONITORING OBYEK BERGERAK MELALUI KAMERA WEBCAM,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 3, no. 1, pp. 110–116, 2019.
- [14] H. Hardiyanto, A. Abdussomad, E. Haryadi, R. Sopandi, and A. Asep, “Penerapan Model Waterfall Dan Uml Dalam Rancang Bangun Program Pembelian Barangberorientasi Objek Pada PT. FUJITA INDONESIA,” *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, vol. 13, no. 4, pp. 4–11, 2019.
- [15] D. Rahayu and S. F. Rezky, “Perancangan Aplikasi Lowongan Kerja Berbasis Web dengan Menggunakan Metode Waterfall,” *SIKOM: Jurnal Sistem Informasi Komputer*, vol. 1, no. 1, pp. 25–36, 2023.
- [16] W. E. Jayanti and A. Hendini, “Pengembangan perangkat lunak pengujian kendaraan bermotor (Tanjidor) dengan model waterfall pada dinas perhubungan,” *Jurnal Khatulistiwa Informatika*, vol. 9, no. 1, 2021.
- [17] N. W. Rahadi and C. Vikasari, “Pengujian Software Aplikasi Perawatan Barang Miliki Negara Menggunakan Metode Black Box Testing Equivalence Partitions,” *Jurnal Infotekmesin*, vol. 11, no. 01, pp. 57–61, 2020.
- [18] J. Jamaludin, C. Rozikin, and A. S. Y. Irawan, “Klasifikasi Jenis Buah Mangga dengan Metode Backpropagation,” *Techné: Jurnal Ilmiah Elektroteknika*, vol. 20, no. 1, pp. 1–12, 2021.
- [19] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, “Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako,” *JURIKOM (Jurnal Riset Komputer)*, vol. 8, no. 6, pp. 219–225, 2021.
- [20] H. Hozairi, A. Anwari, and S. Alim, “Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes,” *Network Engineering Research Operation*, vol. 6, no. 2, pp. 133–144, 2021.