

Klasifikasi Kualitas Cengkeh Berbasis Citra Digital menggunakan Algoritma *Naïve Bayes*

Digital Image-Based Classification of Clove Quality using Naïve Bayes Algorithm

¹Dilla*, ²M. Adnan Nur, ³Musdalifah Djamaluddin

^{1,2,3}Teknik Informatika, Ilmu Komputer, Universitas Handayani Makassar
Makassar, Sulawesi Selatan, Indonesia

*e-mail: dillah1912@gmail.com

(received: 15 September 2024, revised: 25 September 2024, accepted: 25 September)

Abstrak

Pada saat ini, budidaya cengkeh semakin diminati terutama oleh para petani karena mudah dirawat dan harga jualnya yang tinggi. Peneliti melakukan observasi di Desa Tana Toa tentang bagaimana proses pengolahan cengkeh setelah panen yaitu menjemurnya di bawah sinar matahari hingga berwarna coklat dan menyusut. Setelah itu, petani memilih cengkeh kering dan membedakan yang kualitas yang bagus dan yang buruk. Salah satu cara bagi petani dan pedagang untuk mengetahui kualitas cengkeh adalah dengan memeriksa ukuran dan warnanya secara visual. Salah satu kelemahan proses klasifikasi manual ini adalah bahwa setiap orang dapat melihat bahan yang sama dalam jumlah besar dengan cara yang berbeda tergantung pada situasi atau titik kelemahan individu. Tujuan penelitian ini adalah untuk membantu petani memproduksi cengkeh berkualitas tinggi yang pada akhirnya akan menghasilkan hasil yang menguntungkan pada ekonomi mereka. Dengan metode berbasis citra digital dan Naive Bayes, proses ini bisa dilakukan dengan cepat dan efisien, mengurangi biaya operasional dan waktu kerja. Algoritma Naive Bayes mampu memproses data secara lebih teliti dibandingkan manusia, terutama jika kualitas gambar dan fitur yang digunakan untuk klasifikasi dioptimalkan. Ini mengurangi kesalahan manusia yang mungkin terjadi selama proses manual. Hasil dari penelitian ini adalah, Pengujian *Gaussian Naive Bayes* memiliki akurasi 0.82. *Bernoulli naïve bayes* memiliki akurasi 0.69, *Complement naïve bayes* dan *multinomial naïve bayes* masing-masing memiliki akurasi 0,89. Ini menunjukkan bahwa keduanya mempengaruhi tingkat akurasi kualitas cengkeh dengan efektif.

Kata kunci: Citra digital, klasifikasi, *naïve bayes*, cengkeh

Abstract

At present, clove cultivation is increasingly in demand, especially by farmers because it is easy to maintain and the selling price is high. Researchers conducted observations in Tana Toa Village on how the processing of cloves after harvesting is drying them in the sun until they turn brown and shrink. After that, farmers select dried cloves and distinguish between good and bad quality. One way for farmers and traders to determine the quality of cloves is by visually inspecting the size and color. One of the disadvantages of this manual classification process is that each person can look at the same material in bulk in different ways depending on the situation or individual weak points. The aim of this research is to help farmers produce high quality cloves that will ultimately produce favorable results on their economy. With digital image-based methods and Naive Bayes, this process can be done quickly and efficiently, reducing operational costs and labor time. The Naive Bayes algorithm is able to process data more thoroughly than humans, especially if the image quality and features used for classification are optimized. This reduces human errors that may occur during manual processing. The results of this study are, Gaussian Naive Bayes testing has an accuracy of 0.82. Bernoulli naïve bayes has an accuracy of 0.69, Complement naïve bayes and multinomial naïve bayes each have an accuracy of 0.89. This shows that they affect the accuracy rate of clove quality effectively

Keywords: Digital image, classification, *naïve bayes*, cloves

1 Pendahuluan

Pertanian memainkan peran penting dalam perekonomian dan pemenuhan kebutuhan pokok masyarakat di negara agraris seperti Indonesia. Selain itu, pertanian juga membantu meningkatkan kesejahteraan masyarakat, yang mayoritas saat ini hidup di bawah ambang kemiskinan [1]. Tanaman cengkeh atau *Syzygium aromaticum* adalah rempah yang digunakan sebagai obat tradisional. Biasanya digunakan sebagai campuran untuk rokok kretek, cengkeh juga digunakan untuk meningkatkan aroma dan rasa makanan, minuman, industri farmasi (untuk kesehatan), kosmetik, dan obat herbal. Menurut A. Lutfi (2019), 80–90% dari komoditas cengkeh digunakan oleh industri rokok kretek. Sebagian lagi digunakan untuk tujuan lain. Dengan demikian, perkembangan jumlah produksi rokok kretek akan sangat memengaruhi permintaan cengkeh, terutama di wilayah Indonesia. Industri rokok kretek terus berkembang, sehingga permintaan cengkeh di Indonesia akan terus meningkat [2]. Tiga jenis cengkeh yang populer di Indonesia adalah Zanzibar, Sikotok, dan Siputih [3].

Pada saat ini, budidaya cengkeh semakin diminati terutama oleh para petani karena mudah dirawat dan harga jualnya yang tinggi. Sekitar 18 juta orang hidup dari perkebunan cengkeh yang tersebar di seluruh tanah air, yang memberikan manfaat sebagai sumber pendapatan utama bagi sebagian besar sentra produksi cengkeh dan sebagai sumber devisa negara melalui ekspor [4]. Lebih dari 95% pengusaha Indonesia bekerja di bidang pertanian, dan sekitar 80% penduduk bergantung pada sektor pertanian (baik hortikultura, Pangan, perkebunan, peternakan, perikanan, dan kehutanan) [5]. Namun, harga cengkeh tidak selalu sama. Cengkeh berkualitas tinggi pasti memiliki nilai jual yang tinggi juga [6]. Meskipun cengkeh hanya dipanen dua kali setahun, hasilnya sangat baik. Peneliti melakukan observasi di Desa Tana Toa tentang proses pengolahan cengkeh setelah panen, Setelah dijemur di bawah sinar matahari, cengkeh akan berwarna cokelat dan menyusut. Setelah itu, petani memilih cengkeh kering dan membedakan yang kualitas yang bagus dan yang buruk. Salah satu cara bagi petani dan pedagang untuk mengetahui kualitas cengkeh adalah dengan memeriksa ukuran dan warnanya secara visual. Salah satu kelemahan proses klasifikasi manual ini adalah bahwa setiap orang dapat melihat bahan yang sama dalam jumlah besar dengan cara yang berbeda tergantung pada situasi atau titik kelemahan individu. Oleh karena itu, studi kasus utama dalam penelitian ini adalah klasifikasi kualitas cengkeh dengan menggunakan citra digital digital.

Ketika musim hujan tiba, cengkeh kering siap jual sering menjadi langka. Ini karena petani menghasilkan sedikit cengkeh selama musim penghujan. Setelah panen cengkeh basah, petani mengalami kesulitan untuk mengeringkannya. Pada musim kemarau, cengkeh dapat kering dengan bantuan matahari selama empat hingga lima hari, dengan rata-rata waktu jemur antara delapan hingga sepuluh jam per hari. Pada musim penghujan, waktu pengeringan cengkeh bertambah menjadi enam hingga sepuluh hari, bahkan lebih, tergantung pada seberapa banyak hujan yang turun selama proses pengeringan. Setelah kering, cengkeh berwarna merah kecoklatan dengan kadar air 8–10 %. Satu kilogram cengkeh kering sebanding dengan tiga kilogram cengkeh basah. Jumlah lahan yang dibutuhkan untuk proses pengeringan cengkeh meningkat karena waktu yang diperlukan untuk mengeringkan cengkeh selama musim penghujan. Untuk mengeringkan cengkeh, petani harus menunggu lahan kering terlebih dahulu dari kelembapan akibat hujan. Dengan menggunakan citra digital untuk menilai kualitas cengkeh, petani dapat dengan mudah membedakan cengkeh berkualitas tinggi berdasarkan ukuran dan warnanya. Ini dilakukan untuk meningkatkan kualitas dan hasil, dan untuk memastikan bahwa hasilnya rata [7].

Penelitian ini menggunakan algoritma *naïve bayes* untuk mengklasifikasikan kualitas cengkeh. Untuk klasifikasi dimulai dengan mengumpulkan data citra cengkeh yang kemudian diproses dengan metode *Co-occurrence Matrix Gray Level* (GLCM) untuk mengekstraksi fitur dan HSV (*Hue, Saturation, Value*). Setelah fitur diekstraksi, algoritma *naïve bayes* digunakan untuk mengklasifikasikan kualitas cengkeh berdasarkan data latih dan diuji untuk melihat seberapa baik algoritma dapat mengklasifikasikan kualitas cengkeh pada data baru yang belum pernah diamati sebelumnya. Metrik seperti akurasi, presisi, recall, dan confusion matrix digunakan untuk mengevaluasi kinerja model. Ini penting untuk mengetahui seberapa akurat dan valid metode klasifikasi yang digunakan.

Penelitian ini bertujuan untuk membantu petani memproduksi cengkeh berkualitas tinggi yang pada akhirnya akan menghasilkan hasil yang menguntungkan pada ekonomi mereka. Sedangkan manfaat penelitian ini, dengan menggunakan citra digital dan algoritma *naïve bayes*, proses penilaian kualitas cengkeh menjadi otomatis, mengurangi ketergantungan pada penilaian manual yang rentan

terhadap kesalahan manusia. Penilaian kualitas cengkeh yang dilakukan oleh manusia bisa berbeda tergantung pada penilai, kondisi fisik, atau subjektivitas individu. Algoritma *naïve bayes* memberikan klasifikasi yang objektif dan konsisten berdasarkan data. Proses manual untuk mengklasifikasikan kualitas cengkeh membutuhkan banyak waktu dan tenaga kerja. Dengan metode berbasis citra digital dan *naïve bayes*, proses ini bisa dilakukan dengan cepat dan efisien, mengurangi biaya operasional dan waktu kerja. Algoritma *naïve bayes* mampu memproses data secara lebih teliti dibandingkan manusia, terutama jika kualitas gambar dan fitur yang digunakan untuk klasifikasi dioptimalkan. Ini mengurangi kesalahan manusia yang mungkin terjadi selama proses manual

2 Tinjauan Literatur

Banyak studi penelitian telah mengatasi masalah ini mirip dengan penggunaan kamera digital untuk memprediksi kualitas hasil cengkeh adalah salah satu dari banyak penelitian yang telah membahas masalah ini [8] yang mengklasifikasikan kualitas cengkeh dengan menggunakan citra digital dan menemukan bahwa 37 sampel dari 40 sampel benar. Metode ini mengubah warna HSV cengkeh kering dari RGB.

Selain itu, penelitian tambahan dilakukan [9] untuk mengklasifikasikan Kualitas bunga cengkeh yang telah kering ditentukan oleh warna dan ukuran cengkeh. Penelitian ini dengan metode *Co-Occurrence Matrix Gray Level (GLCM)* yang mampu membedakan kualitas rendah dan tinggi dengan akurasi 92.5% dari 32 sampel bunga cengkeh kering.

Penelitian [10] klasifikasi sudah dapat dilakukan berdasarkan input data. Ada 15 data uji, 9 data sampel (kualitas 1), 9 data sampel (kualitas 2), dan 3 data sampel bukan cengkih. Hasil penelitian menunjukkan bahwa algoritma *K-Nearest Neighbors (KNN)* digunakan untuk mengkategorikan bunga cengkih yang kering. Algoritma ini memproses piksel hitam dan putih dari gambar cengkih dengan skala gelap dan menetapkan ambang dengan nilai 95, 255. Ini menghasilkan tingkat *presentase* sebesar 80%.

Berdasarkan hasil penelitian [11] Didapat nilai parameter arsitektur CNN. Saat menguji nilai *epoch*, ditemukan *konvergen*, yaitu *epoch* 1300. Penggunaan dua lapisan dengan peta fitur tiga puluh dua dan enam puluh empat memberikan akurasi terbaik dalam pengujian jumlah lapisan. Uji ukuran input gambar menunjukkan hasil yang paling akurat pada ukuran 128×128 pixel, uji ukuran kernel menunjukkan hasil yang paling akurat pada ukuran 5×5 pixel, uji nilai langkah atau *stride* menunjukkan hasil yang paling akurat pada ukuran 1×1 pixel, uji pengaruh padding menunjukkan hasil yang paling akurat pada penggunaan padding dan uji nilai *dropout* menunjukkan hasil yang paling akurat pada 0,4. Pengujian kecepatan belajar, dengan akurasi terbaik 0,0001 Selanjutnya, data uji coba digunakan untuk menguji parameter tersebut, yang menghasilkan akurasi sebesar 87,75%.

Selanjutnya, penelitian yang dilakukan oleh [12] menemukan bahwa dengan 100 *neuron* pada lapisan tersembunyi, konversi satu lapisan dengan lima belas filter, dan tingkat pembelajaran 0,0001, akurasi terbaik untuk model 2-kelas adalah 82,46 %. Sementara model empat kelas memiliki dua lapisan konvolusional, filter dengan akurasi terbaik 70,73% berukuran 3 x 3 x 3. Di lapisan pertama, ada enam filter berukuran 3x5x5, dan di lapisan kedua, ada enam belas filter berukuran 6x3x3.

Berbagai penelitian telah dilakukan untuk mengklasifikasikan kualitas cengkeh kering menggunakan citra digital. Studi awal menggunakan teknik sederhana seperti konversi RGB ke HSV dan metode analisis tekstur GLCM, yang mampu mencapai akurasi hingga 92,5%. Metode *K-Nearest Neighbors (KNN)* juga digunakan, menghasilkan akurasi 80%. Di sisi lain, penelitian lebih lanjut mulai menggunakan *Convolutional Neural Networks (CNN)*, yang memungkinkan pengujian parameter yang lebih kompleks seperti jumlah lapisan, ukuran kernel, *stride*, dan *dropout*. CNN menunjukkan akurasi yang lebih tinggi dibandingkan metode konvensional, dengan akurasi tertinggi sebesar 87,75%.

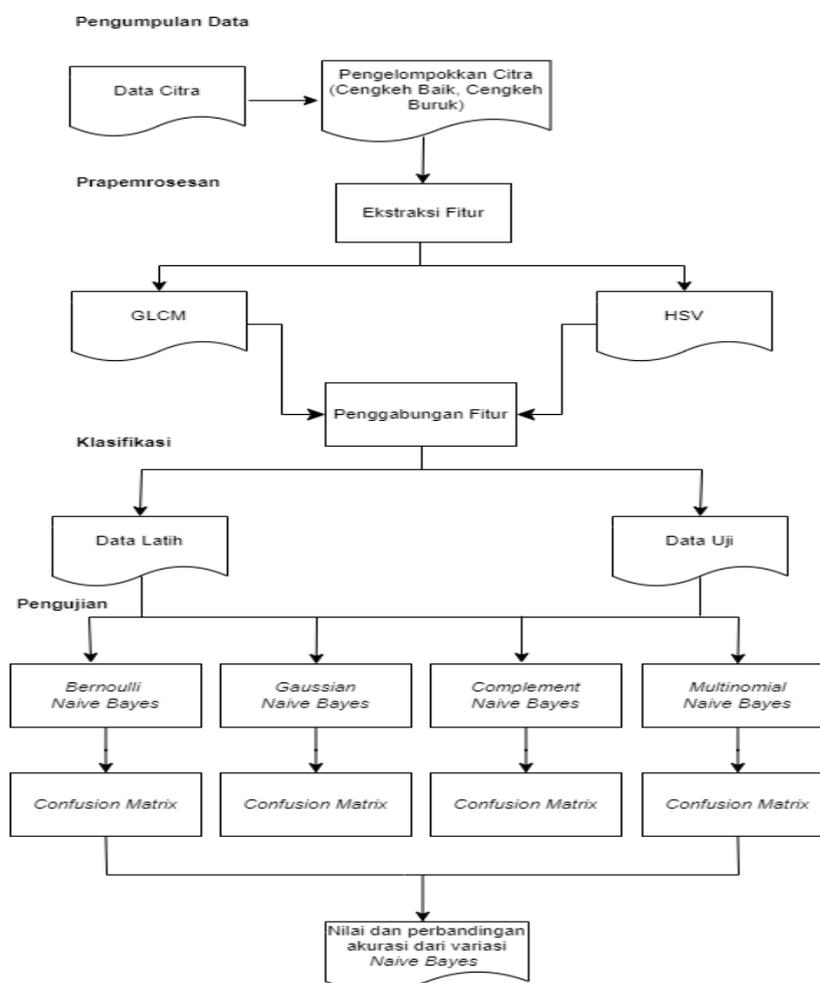
Penelitian ini menggunakan berbagai metode dan algoritma seperti HSV, GLCM, KNN, dan DCNN untuk klasifikasi kualitas cengkeh. Setiap metode memiliki kelebihan dan kekurangan, seperti GLCM yang efektif dalam membedakan kualitas berdasarkan ukuran dan warna, sedangkan DCNN yang lebih kompleks dan dapat menangani variasi yang lebih luas. Hasil akurasi yang diperoleh dari penelitian-penelitian ini beragam, mulai dari 80% hingga 92,50%. Hal ini menunjukkan bahwa setiap metode memiliki potensi untuk digunakan dalam praktik industri, tetapi perlu disesuaikan dengan kebutuhan spesifik. Penelitian ini menunjukkan peningkatan kompleksitas dalam penggunaan teknologi digital untuk klasifikasi kualitas cengkeh. Dari penggunaan metode sederhana seperti HSV hingga

penggunaan teknologi canggih seperti DCNN, setiap penelitian menunjukkan kemajuan dalam teknologi pengolahan citra. Penelitian klasifikasi kualitas cengkeh berbasis citra digital menggunakan algoritma *naïve bayes* masih relatif jarang dibandingkan dengan metode lain seperti GLCM, KNN, dan DCNN. Namun, jika diterapkan, *naïve bayes* dapat memberikan hasil yang akurat dengan cara mengasumsikan ketergantungan variabel-variabel yang tidak ada [13].

Penelitian terkait klasifikasi kualitas cengkeh menggunakan citra digital telah menunjukkan beberapa kebaruan, yaitu: Penggunaan teknologi canggih seperti DCNN telah meningkatkan akurasi klasifikasi kualitas cengkeh. Metode GLCM telah menunjukkan akurasi tinggi dalam membedakan kualitas rendah dan tinggi cengkeh. Algoritma KNN telah memberikan hasil yang akurat dalam pengeringan bunga cengkih. Dengan demikian, penelitian ini menunjukkan bahwa penggunaan citra digital dan berbagai metode algoritma dapat meningkatkan akurasi klasifikasi kualitas cengkeh, sehingga dapat membantu petani dalam meningkatkan mutu hasil panen dan mempermudah proses sortasi.

3 Metode Penelitian

Gambar 1 menunjukkan diagram alir untuk sistem klasifikasi cengkeh yang menggunakan metode *naïve bayes*. Sistem ini dimulai dengan pengumpulan data citra cengkeh, yang kemudian dikelompokkan berdasarkan kualitasnya, baik atau buruk. Setelah itu, citra tersebut dipapemroses, yaitu dengan ekstraksi fitur menggunakan GLCM dan HSV. Fitur-fitur yang telah diekstraksi kemudian digabungkan dan digunakan untuk melatih model variasi klasifikasi *naïve bayes*. Model data latih kemudian diuji dengan menggunakan data uji untuk mengukur keakuratannya. Hasil pengujian ditampilkan dalam bentuk *confusion matrix*.



Gambar 1. Diagram Alir Metodologi Penelitian

3.1 Metode Pengumpulan Data

Penelitian menggunakan cengkeh baik dan buruk. Untuk melakukan penelitian ini, pengepul dan petani cengkeh di desa Tana Toa di *survei* tentang hasil panen cengkeh. Algoritma *naïve bayes* akan digunakan untuk mengklasifikasikan pengujian sampel pada sistem dengan data latih [12]. Setelah dataset dikumpulkan, algoritma ini akan digunakan untuk mengklasifikasikan kualitas cengkeh berdasarkan citra digital. Dimulai dengan membuat kode program menggunakan *Visual Studio Code* dan bahasa pemrograman *python* untuk mendapatkan data yang diinginkan. Sebanyak 250 dataset diberi label "cengkeh buruk" dan "cengkeh baik" untuk memenuhi persyaratan klasifikasi kualitas cengkeh.

3.2 Prapemrosesan Data

Tahap selanjutnya adalah ekstraksi fitur. Tujuan proses ini adalah untuk mengekstrak informasi penting dari data. Metode *Co-occurrence Matrix Gray Level (GLCM)* dan *HSV (Hue, Saturation, Value)* digunakan untuk menentukan tekstur dan warna cengkeh dalam kasus gambar ini [14].

3.2.1 Grey Level Co-Occurrence Matrix (GLCM)

Metode *Co-occurrence Matrix Level Gray (GLCM)* menggunakan nilai ambang gambar kulit cengkeh cokelat untuk membuat gambar *grayscale* lebih mudah dianalisis untuk menentukan ukuran cengkeh yang baik atau buruk [15]. Pembentukan matrix GLCM dapat dilakukan dengan menghitung frekuensi di mana berbagai pasangan nilai piksel muncul dalam citra *grayscale*. Matrix ini menunjukkan ukuran cengkeh yang baik atau buruk. Jika jumlah piksel putih di citra kulit cokelat kurang dari 500 piksel, nilai ambang putih menunjukkan ukuran cengkeh yang baik, jika lebih dari 500 piksel, nilai ambang putih menunjukkan warna cengkeh yang buruk. Selanjutnya, nilai ambang foto kulit cengkeh cokelat untuk menentukan kualitasnya. Jika ukuran dan warnanya baik, maka kualitasnya baik, jika tidak maka kualitasnya buruk [9].

3.2.2 Hue, Saturation, Value (HSV)

Metode *HSV (Hue, Saturation, Value)* menghilangkan *noise*, seperti bayangan hitam pada gambar cengkeh, saat mengklasifikasikan kualitas cengkeh. Metode ini memulai dengan mengubah warna gambar menjadi warna yang lebih sesuai, menemukan ukuran dan warna cengkeh yang buruk atau baik, dan kemudian menemukan ukuran cengkeh yang baik dengan nilai ambang dari gambar kulit cengkeh berwarna cokelat. Jika jumlah piksel putih di cengkeh kurang dari 500 piksel, ukurannya baik, jika lebih dari itu, ukurannya buruk [16]. Selain itu, ada batas nilai untuk warna putih pada cengkeh. Jika jumlah piksel putih kurang dari 10%, warna cengkeh berkualitas baik, tetapi jika lebih dari 10%, warna cengkeh berkualitas buruk [17].

3.2.3 Penggabungan Fitur

Penggabungan fitur dilakukan setelah fitur diekstraksi. Menggabungkan informasi dari berbagai fitur ke dalam representasi data yang sama dikenal sebagai penggabungan fitur. Dalam gambar ini, ini terjadi dengan menggabungkan fitur tekstur dan warna dari citra cengkeh [18].

3.3 Klasifikasi

Setelah semua tahap prapemrosesan selesai, tahap berikutnya adalah klasifikasi. Data dimasukkan ke dalam kategori tertentu melalui teknik klasifikasi. Gambar ini dikategorikan sebagai gambar cengkeh baik atau buruk. Pada awal proses, algoritma klasifikasi *naïve bayes* digunakan. Informasi tentang ukuran, warna, dan kualitas biji cengkeh dikumpulkan untuk klasifikasi kualitas cengkeh. *Pre-processing* data dilakukan untuk menghilangkan suara dan memastikan bentuknya akurat. Data latih dan data uji adalah dua komponen yang membentuk dataset. Data uji digunakan untuk mengevaluasi akurasi model, sementara data latih digunakan untuk menghitung varians dan parameter *mean* [19]. Algoritma *machine learning* ini sangat akurat [20] dan digunakan untuk membuat model klasifikasi *naïve bayes* [21]. Setelah model klasifikasi dibuat, data uji digunakan untuk mengukur akurasi [22].

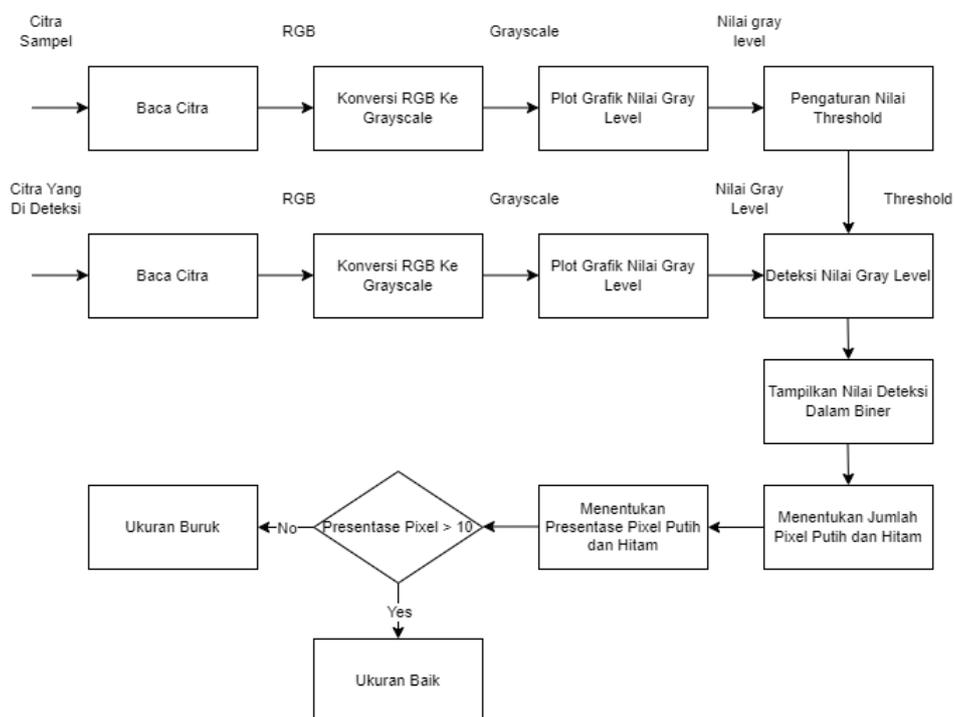
3.4 Pengujian

Tujuan dari tahap terakhir pengujian adalah untuk mengevaluasi mekanisme kerja model klasifikasi. Empat jenis *naïve bayes* yang telah diuji: *Gaussian naïve bayes*, *Bernoulli*, *complement*, dan *multinomial*. Jumlah data yang diklasifikasikan dengan benar dan salah ditunjukkan oleh teknik

confusion matrix. Metode ini digunakan untuk mengevaluasi hasil ujian. Ini mungkin digunakan untuk menilai efektifitas model klasifikasi dengan menghitung metrik seperti akurasi, *precision*, dan *recall*.

3.5 Deteksi Ukuran Cengkeh

Gambar 2 merupakan diagram alir yang menggambarkan proses untuk mendeteksi cacat pada suatu objek dengan menggunakan pengolahan citra digital. Proses ini melibatkan beberapa tahapan, di antaranya sebagai berikut.

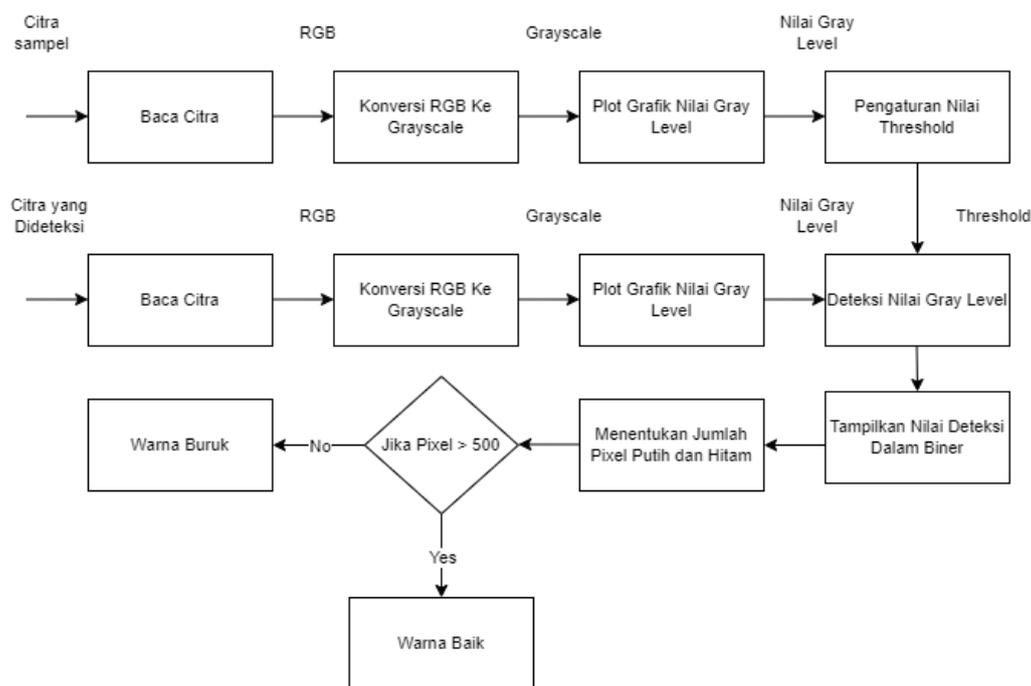


Gambar 2. Alur Deteksi Ukuran Cengkeh

Proses pengolahan gambar digital pada cengkeh ditunjukkan pada Gambar 2. Alur kerja pengolahan citra untuk mendeteksi objek menggunakan metode *thresholding*. Dimulai dari membaca citra yang ingin diproses. Mengubah citra dari format RGB ke Grayscale. Memplot grafik nilai gray level dari citra. Menentukan nilai ambang batas (*threshold*) untuk membedakan objek dari latar belakang. Mendeteksi nilai gray level dari citra dan membandingkannya dengan nilai *threshold*. Menampilkan hasil deteksi dalam bentuk biner (hitam putih) untuk menunjukkan objek yang terdeteksi. Setelah proses deteksi objek, flowchart akan melakukan percabangan untuk menentukan apakah ukuran objek baik atau buruk. Jika ukuran objek buruk ($\text{presentase pixel} < 10\%$), maka flowchart akan kembali ke langkah membaca citra. Jika ukuran objek baik ($\text{presentase pixel} > 10\%$), maka flowchart akan menghitung presentasi pixel putih dan hitam pada objek yang terdeteksi.

3.6 Deteksi Cacat/warna Putih Pada Cengkeh

Gambar 3 merupakan diagram alir yang menggambarkan langkah-langkah dalam proses deteksi warna pada citra digital



Gambar 3. Alur Proses Deteksi Ukuran Cengkeh

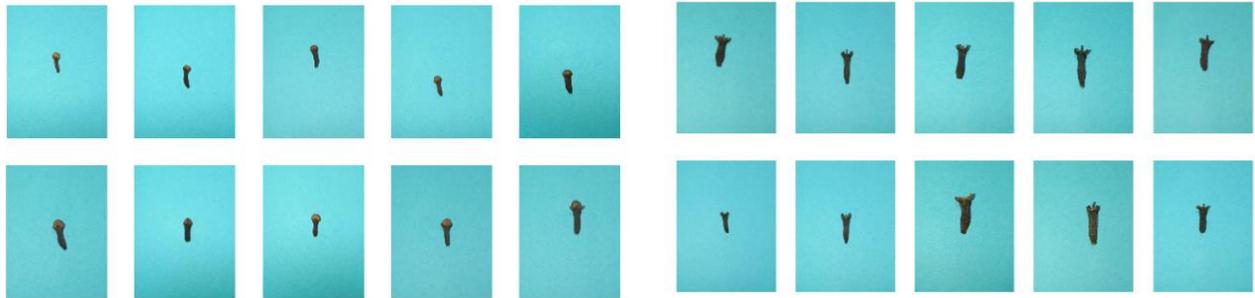
Gambar 3 menggambarkan diagram alir untuk mendeteksi warna pada gambar dengan metode *thresholding*. Alur prosesnya adalah sebagai berikut: Citra yang akan dideteksi warnanya dibaca terlebih dahulu. Citra yang dibaca diubah ke skala abu-abu. Nilai *grayscale* pada citra diplot dalam bentuk grafik. Deteksi warna pada citra dilakukan dengan menentukan batas nilai *gray level* yang menjadi pembeda antara warna putih dan hitam. Batas nilai *threshold* ditentukan untuk memisahkan warna putih dan hitam. Jumlah pixel yang terdeteksi sebagai warna putih dan hitam dihitung. Hasil deteksi warna ditampilkan dalam bentuk biner, yaitu 1 untuk putih dan 0 untuk hitam. Jika jumlah pixel yang terdeteksi sebagai warna putih lebih dari 500, maka citra tersebut dianggap memiliki warna yang baik. Jika jumlah pixel yang terdeteksi sebagai warna putih kurang dari 500, maka citra tersebut dianggap memiliki warna yang buruk. Diagram alir ini menunjukkan bagaimana metode *thresholding* dapat digunakan untuk mendeteksi warna pada gambar dan menentukan apakah gambar tersebut memiliki warna yang baik atau buruk.

4 Hasil dan Pembahasan

Dataset yang telah melewati tahap prapemrosesan, ekstraksi fitur kemudian akan diuji pada setiap variasi *naïve bayes* yang akan digunakan untuk melakukan klasifikasi kualitas cengkeh berbasis citra digital. Berikut adalah hasil dan pembahasan dari implementasi variasi *naïve bayes* pada klasifikasi kualitas cengkeh.

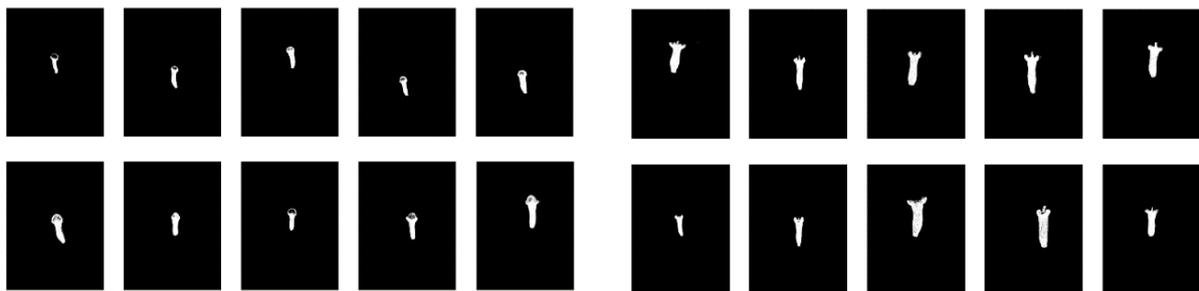
4.1 Hasil Deteksi Ukuran Cengkeh

Berikut ini adalah hasil deteksi ukuran cengkeh yang dilakukan untuk menilai kualitas dan klasifikasi cengkeh berdasarkan parameter fisik utamanya.



Gambar 4. Sampel A1-A10 dan B1-B10

Gambar 4 merupakan sampel cengkeh A1-A10 yang berada di sebelah kiri, menurut penilaian petani memiliki ukuran dan warna yang baik karena kepalanya utuh dan kulitnya bebas bercak putih atau warna cacat, namun sampel cengkeh B1-B10 di sebelah kanan memiliki kepala dan bunga terbuka, sehingga ukuran dan warnanya buruk.



Gambar 5. Hasil visualisasi gambar biner untuk mendeteksi ukuran cengkeh

Gambar 5 merupakan gambar kulit cengkeh berwarna coklat diubah menjadi *grayscale* dari gambar RGB, lalu nilai *grayscale* diplot pada grafik untuk melihat kisaran nilai. Selanjutnya, gambar sampel cengkeh yang akan diuji diubah dari RGB ke *grayscale*, dan kemudian kedua gambar dibandingkan. Setelah itu, nilai perhitungan error untuk kedua gambar dalam matriks dicocokkan, dan gambar biner menunjukkan hasilnya. Gambar biner dapat digunakan untuk mengetahui berapa banyak piksel putih dan hitam dan persentasenya.

4.2 Hasil Deteksi Warna Putih Pada Cengkeh

Gambar 6 menunjukkan hasil deteksi warna putih pada cengkeh. Tujuannya adalah untuk menentukan apakah perubahan warna menunjukkan kualitas atau kemungkinan kerusakan.



Gambar 6. Hasil citra biner yang mendeteksi warna cengkeh.

Gambar 6 merupakan warna putih diasumsikan sebagai warna cacat yang terdapat pada kulit cengkeh. Setelah citra berwarna putih diubah dari warna RGB ke *grayscale*, nilai diplot di grafik untuk melihat kisaran nilai. Nilai ambang gambar warna putih yang tepat untuk mengidentifikasi warna cengkeh selanjutnya diubah dari warna RGB ke *grayscale* untuk sampel cengkeh, dan Selanjutnya, nilai

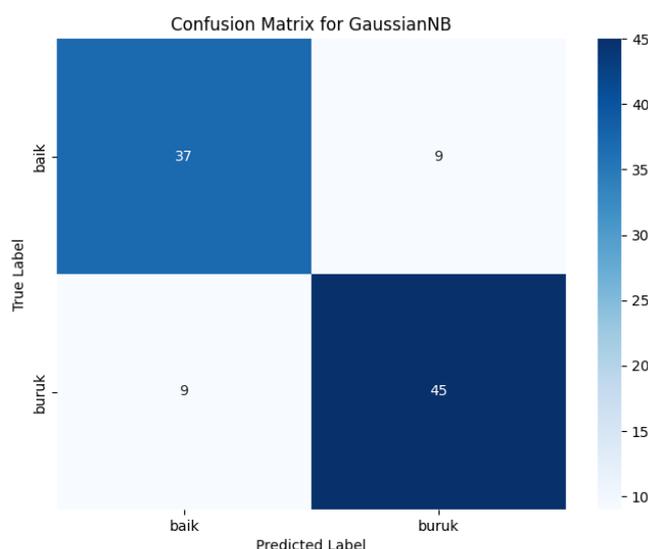
sampel cengkeh dibandingkan dengan nilai citra putih. Selanjutnya, nilai dan perhitungan error gambar dalam matriks dicocokkan, dan gambar biner menunjukkan hasilnya.

4.3 Pengujian Confusion Matrix

Confusion matrix digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan hasil prediksi dengan nilai sebenarnya. Dalam confusion matrix, prediksi model dibagi menjadi empat kategori utama: True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Selanjutnya dilakukan pengujian beberapa variasi *naive bayes* berikut ini.

4.3.1 Gaussian Naïve Bayes

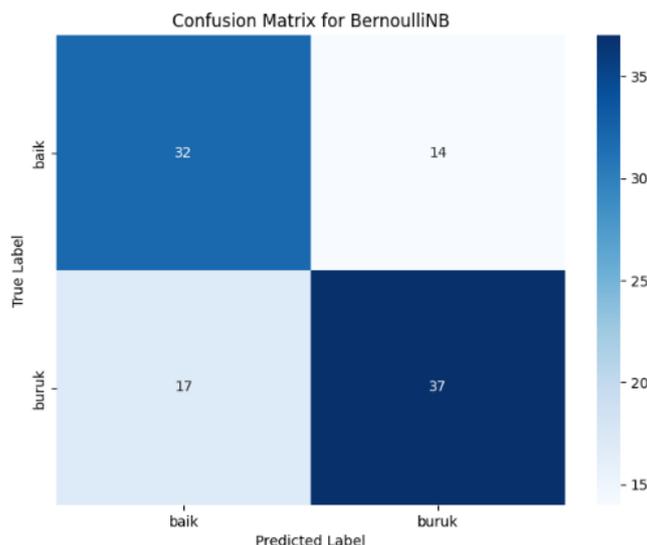
Gambar 7 menampilkan *confusion matrix* yang digunakan untuk mengevaluasi performa model klasifikasi *Gaussian naïve bayes*. Confusion matrix ini menunjukkan seberapa baik model dalam mengklasifikasikan dua kelas yaitu "baik" dan "buruk". Sumbu Horizontal: Menunjukkan kelas yang diprediksi oleh model. Sumbu Vertikal: Menunjukkan kelas sebenarnya dari data. Sel Diagonal (37, 45): Menunjukkan jumlah prediksi yang benar. 37 data dengan kelas sebenarnya "baik" diprediksi dengan benar sebagai "baik". 45 data dengan kelas sebenarnya "buruk" diprediksi dengan benar sebagai "buruk". Sel Non-Diagonal (9, 9): Menunjukkan jumlah prediksi yang salah. 9 data dengan kelas sebenarnya "baik" salah diprediksi sebagai "buruk". 9 data dengan kelas sebenarnya "buruk" salah diprediksi sebagai "baik". Model *Gaussian naïve bayes* ini lebih baik dalam memprediksi kelas "buruk" dibandingkan kelas "baik" karena memiliki jumlah prediksi benar yang lebih tinggi untuk kelas "buruk".



Gambar 7. Gaussian naïve bayes

4.3.2 Bernoulli Naïve Bayes

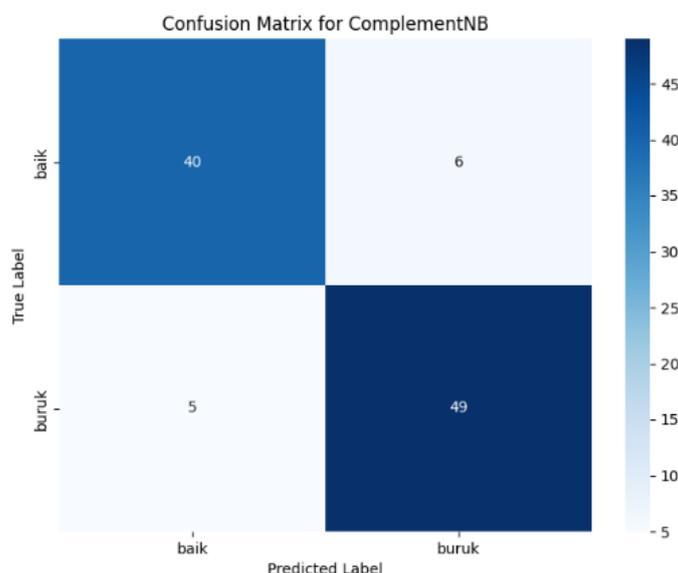
Gambar 8 menunjukkan *confusion matrix* untuk model *Bernoulli naïve bayes*. *Confusion matrix* adalah alat yang digunakan dalam *machine learning* untuk memvisualisasikan performa model klasifikasi. Sumbu Vertikal (*True Label*): Menunjukkan label kelas sebenarnya dari data. Terdapat dua kelas yaitu "baik" dan "buruk". Sumbu Horizontal (*Predicted Label*): Menunjukkan label kelas yang diprediksi oleh model *Bernoulli naïve bayes*. Sel Kiri Atas (32): Menunjukkan jumlah data yang benar diprediksi sebagai "baik" (*True Positive*). Sel Kanan Atas (14): Menunjukkan jumlah data "baik" yang salah diprediksi sebagai "buruk" (*False Negative*). Sel Kiri Bawah (17): Menunjukkan jumlah data "buruk" yang salah diprediksi sebagai "baik" (*False Positive*). Sel Kanan Bawah (37): Menunjukkan jumlah data yang benar diprediksi sebagai "buruk" (*True Negative*). Semakin gelap warna biru pada sel, semakin banyak jumlah data yang direpresentasikan oleh sel tersebut. Dari *confusion matrix* ini, dapat disimpulkan bahwa model *Bernoulli naïve bayes* lebih akurat dalam memprediksi kelas "buruk" (37 benar, 17 salah) dibandingkan kelas "baik" (32 benar, 14 salah).



Gambar 8. Bernoulli naive bayes

4.3.3 Complement Naïve Bayes

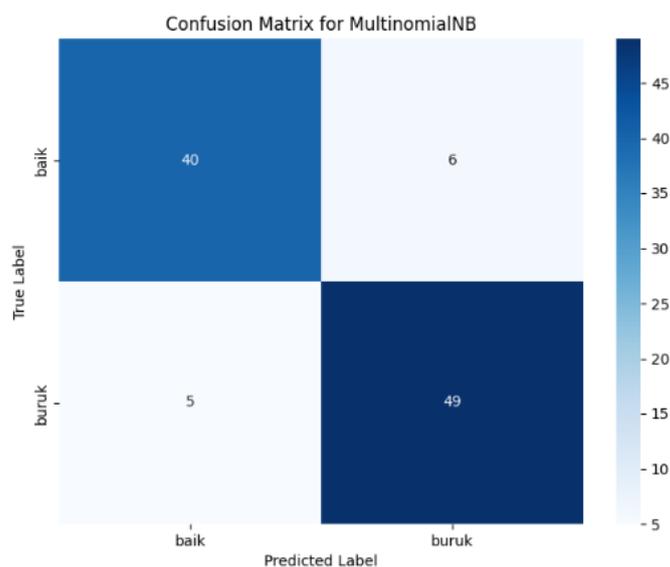
Gambar 9 adalah *confusion matrix* untuk *Complement naïve bayes*. *Confusion matrix* adalah tabel yang menunjukkan hasil klasifikasi model *machine learning*. Setiap baris pada tabel menunjukkan jumlah prediksi aktual dari suatu kelas, sedangkan setiap kolom menunjukkan jumlah prediksi dari kelas yang berbeda. Pada *confusion matrix* ini, terdapat dua kelas: "baik" dan "buruk". Dari tabel tersebut, dapat dilihat bahwa model memprediksi 40 contoh data sebagai "baik" dengan benar. Model memprediksi 6 contoh data sebagai "baik" dengan salah. Model memprediksi 5 contoh data sebagai "buruk" dengan salah. Model memprediksi 49 contoh data sebagai "buruk" dengan benar. *Confusion matrix* membantu kita memahami bagaimana model melakukan prediksi dan mendeteksi kesalahan prediksi yang dilakukan model. Dari *confusion matrix* ini, kita dapat melihat bahwa model *Complement naïve bayes* ini cukup baik dalam memprediksi kelas "buruk", tetapi agak kurang dalam memprediksi kelas "baik".



Gambar 9. Complement naïve bayes

4.3.4 Multinomial Naïve Bayes

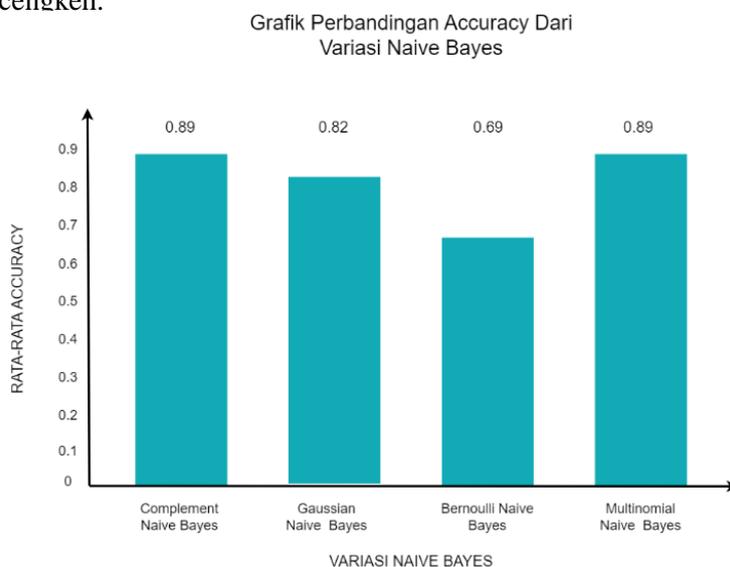
Gambar 10 menampilkan *confusion matrix* untuk model *Multinomial naïve bayes*. *Confusion matrix* ini digunakan untuk mengevaluasi performa model klasifikasi dengan menunjukkan jumlah prediksi yang benar dan salah untuk setiap kelas. Pada gambar tersebut, terdapat dua kelas yaitu baik dan buruk. Model ini berhasil memprediksi: 40 data benar untuk kelas baik, 49 data benar untuk kelas buruk. Namun, model ini juga melakukan kesalahan prediksi: 6 data kelas baik diprediksi salah sebagai kelas buruk. 5 data kelas buruk diprediksi salah sebagai kelas baik. Warna biru tua pada *heatmap* menunjukkan jumlah prediksi yang lebih banyak, sementara warna biru muda menunjukkan jumlah prediksi yang lebih sedikit. Semakin gelap warna pada diagonal dari kiri atas ke kanan bawah, semakin baik performa model karena menunjukkan jumlah prediksi yang benar untuk setiap kelas. Model *Multinomial naïve bayes* ini memiliki performa yang cukup baik karena mayoritas prediksinya benar. Namun, masih terdapat kesalahan prediksi yang perlu diperbaiki untuk meningkatkan performa model.



Gambar 10. *Multinomial naïve bayes*

4.4 Perbandingan Hasil Evaluasi

Pada gambar 11 merupakan hasil dari grafik perbandingan akurasi variasi *naïve bayes* dalam klasifikasi kualitas cengkeh.



Gambar 11. Grafik perbandingan akurasi

Gambar 11 menyajikan perbandingan akurasi dari beberapa variasi algoritma *naïve bayes*, yaitu: *Complement naïve bayes*, algoritma ini memiliki rata-rata akurasi tertinggi, mencapai 0.89. *Gaussian naïve bayes*, algoritma ini memiliki rata-rata akurasi 0.82. *Bernoulli naïve bayes*, algoritma ini memiliki rata-rata akurasi terendah di antara keempatnya, yaitu 0.69. *Multinomial naïve bayes*, algoritma ini memiliki rata-rata akurasi yang sama dengan *Complement naïve bayes*, yaitu 0.89. Dapat disimpulkan bahwa *Complement naïve bayes* dan *Multinomial naïve bayes* memberikan performa terbaik untuk kasus ini, sedangkan *Bernoulli naïve bayes* memiliki performa terendah. Namun, perlu diingat bahwa performa algoritma *naïve bayes* sangat bergantung pada dataset dan kasus yang dihadapi.

5 Kesimpulan

Penelitian ini menemukan bahwa, Pengujian *Gaussian Naive Bayes* memiliki akurasi 0.82. *Bernoulli naïve bayes* memiliki akurasi 0.69, *Complement naïve bayes* dan *multinomial naïve bayes* masing-masing memiliki akurasi 0,89. Ini menunjukkan bahwa keduanya mempengaruhi tingkat akurasi kualitas cengkeh dengan efektif. Penggunaan citra digital dan algoritma *naïve bayes* memperkenalkan pendekatan baru dalam klasifikasi mutu cengkeh. Ini menawarkan cara yang lebih objektif dan sistematis dibandingkan metode tradisional yang bergantung pada penilaian manusia, yang sering kali subjektif dan tidak konsisten. Dengan mengintegrasikan teknologi informasi ke dalam praktik pertanian, penelitian ini menunjukkan bagaimana inovasi teknologi dapat meningkatkan efisiensi dan efektivitas dalam pengolahan hasil pertanian. Hal ini berpotensi mendorong adopsi teknologi serupa di sektor pertanian lainnya. Penelitian ini dapat digunakan sebagai dasar untuk penelitian berikutnya di bidang klasifikasi citra dan aplikasi algoritma pembelajaran mesin dalam pertanian. Ini membuka peluang untuk eksplorasi lebih jauh mengenai teknik-teknik lain yang dapat diterapkan untuk meningkatkan proses klasifikasi.

Referensi

- [1] A. Asrawati and M. Antara, "Analisis Pendapatan Usahatani Cengkeh di Desa Talaga Kecamatan Dampelas Kabupaten Donggala," *Agrotekbis E-Jurnal Ilmu Pertan.*, vol. 5, no. 4, pp. 476–482, 2022.
- [2] A. Lutfi, "Bab 1 Pendahuluan," *Pelayanan Kesehat.*, vol. 2013, no. Dm, pp. 3–13, 2019, [Online]. Available: [http://repository.usu.ac.id/bitstream/123456789/23790/4/Chapter I.pdf](http://repository.usu.ac.id/bitstream/123456789/23790/4/Chapter%20I.pdf)
- [3] N. Nurmala and M. Asse, "Analisis Kontribusi Pendapatan Usahatani Cengkeh Terhadap Pendapatan Rumah Tangga Tani Di Desa Duingingis Kecamatan Dako Pemean Kabupaten Tolitoli (Studi Kasus Pada Kelompok Tani Mekar 2)," *J. Agrotech*, vol. 10, no. 1, pp. 9–14, 2020.
- [4] M. Anwar, "Analisis Spasial Pengembangan Cengkeh (*Eugenia Aromatica L.*) Di Kabupaten Lombok Utara," *J. Ilm. Rinjani Universitas Gunung Rinjani*, vol. 5, no. 2, 2019.
- [5] S. M. Amelia and F. Pertanian, "Efisiensi Sistem Pemasaran Cengkeh (*Syzygium Aromaticum*) di Kabupaten Pesisir Barat," 2019.
- [6] Sutriyono and M. Ali, "Budidaya Tanaman Cengkeh," *Fak. Pertanian Universitas Merdeka Surabay*, pp. 1–17, 2019, [Online]. Available: https://www.academia.edu/35842224/Teknik_Budidaya_Tanaman_Cengkeh
- [7] Aldi Yoga Pratama, "Pengeringan Cengkeh (*Syzygium Aromaticum*) menggunakan Rumah Pengereng Hybrid Tipe Rak," pp. 1–17, 2023.
- [8] P. A. L. Pesik, V. C. Poekoel, and M. D. Putro, "Penilaian Mutu Cengkih menggunakan Citra Digital Priscilia," *J. Tek. Elektro dan Komput.*, vol. 7, no. 2, pp. 161–166, 2018.
- [9] Y. N. Yaspin, D. W. Widodo, and A. B. Setiawan, "Klasifikasi Kualitas Bunga Cengkeh untuk Meningkatkan Mutu dengan Pemanfaatan Ciri Gray Level Co-Occurrence Matrix (GLCM)." Universitas Nusantara PGRI Kediri, 2020.
- [10] T. Anggin Virnando, I. Widaningrum, K. Nurfitri, and H. Artikel, "Jurnal Rekayasa Teknologi dan Komputasi URL : <https://journal.umpo.ac.id/index.php/JRTK/index> Penerapan Algoritma K-Nearest Neighbor (K-Nn) Pada Klasifikasi Kualitas Hasil Pengeringan Bunga Cengkih," no. kualitas 1, 2022, [Online]. Available: <https://journal.umpo.ac.id/index.php/JRTK/index>
- [11] I. Y. Prayogi, "Klasifikasi Citra Mutu Cengkeh (*Syzygium Aromaticum*) di PT. Perkebunan Nusantara 12 Kabupaten Malang Menggunakan Deep Convolutional Neural Network (DCNN),"

<http://sistemasi.ftik.unisi.ac.id>

- 2021.
- [12] M. Saputra, K. Kusriani, and M. P. Kurniawan, "Identifikasi Mutu Bij Kopi Arabika berdasarkan Cacat dengan Teknik Convolutional Neural Network," *Inspir. J. Teknol. Inf. dan Komun.*, vol. 10, no. 1, p. 27, 2020, doi: 10.35585/inspir.v10i1.2533.
 - [13] A. Firmansyah and R. Kurniawan, "Analisis Sentimen dalam Penentuan Kinerja Layanan antara In-Driver dan Gojek berdasarkan Opini Masyarakat menggunakan Metode Naive Bayes Sentiment Analysis in Determining Service Performance between In-Driver and Gojek based on Public Opinion using the Na," vol. 13, pp. 1281–1294, 2024.
 - [14] N. H. Kario Balai Pengkajian Teknologi Pertanian Nusa Tenggara Timur Jl Timor Raya Km, "Analisis Usaha Panen Cengkehdi Kabupaten Minahasa Tenggara Propinsi Sulawesi Utara," *Agritech*, vol. XVI, no. 2, pp. 146–156, 2014.
 - [15] N. Neneng, A. S. Puspaningrum, and A. A. Aldino, "Perbandingan Hasil Klasifikasi Jenis Daging menggunakan Ekstraksi Ciri Tekstur Gray Level Co-occurrence Matrices (GLCM) Dan Local Binary Pattern (LBP)," *SMATIKA J. STIKI Inform. J.*, vol. 11, no. 01, pp. 48–52, 2021.
 - [16] I. P. Wardhani and S. Widayati, "Segmentasi Warna Citra HSV dan Deteksi Objek Kupu-Kupu Dengan Metode Klasifikasi K-Means," *Pros. SeNTIK*, vol. 3, no. 1, pp. 125–131, 2019.
 - [17] S. Sanjaya, "Aplikasi Pengenalan Tingkat Kematangan Buah Tomat menggunakan Fitur Warna HSV berbasis Android," *J. Teknoinfo*, vol. 16, no. 1, pp. 26–33, 2022.
 - [18] L. N. S. Aljustin, "Penerapan Contrast Limited Adaptive Histogram Equalization pada Klasifikasi Penyakit Daun Kentang dengan menggunakan Metode k-Nearest Neighbor Berdasarkan Ekstraksi Fitur Glcm Dan Hsv." UPN Veteran Yogyakarta, 2023.
 - [19] L. Mardiana, D. Kusnandar, and N. Satyahadewi, "Analisis Diskriminan dengan k-Fold Cross Validation untuk Klasifikasi Kualitas Air di Kota Pontianak," *Bimaster Bul. Ilm. Mat. Stat. dan Ter.*, vol. 11, no. 1, 2022.
 - [20] R. Ardianto, T. Rivanie, Y. Alkhalifi, S. N. Fitra, and W. Gata, "Sentiment Analysis on e-Sports for Education Curriculum using Naive Bayes and Support Vector Machine," 2021.
 - [21] S. D. Anggita, "Algorithm Comparison of Naive Bayes and Support Vector Machine based on Particle Swarm Optimization in Sentiment Analysis of Freight Forwarding Services," *J. RESTI (Rekayasa Sist. Dan Teknol. Informasi)*, vol. 4, no. 2, pp. 362–369, 2020.
 - [22] S. Tamrakar, B. K. Bal, and R. B. Thapa, "Aspect based Sentiment Analysis of Nepali Text using Support Vector Machine and Naive Bayes," *Tech. J.*, vol. 2, no. 1, pp. 22–29, 2020.