

Klasifikasi Tingkat Kematangan Roasting Biji Kopi Berbasis Ekstraksi Fitur Warna HSV Menggunakan Metode Naïve Bayes

Gusti Ayu Devani Zelvia*, I Made Gede Sunarya, I Gde Made Hanura, Ida Bagus Gede Putra Kenaka

Fakultas Teknik dan Kejuruan, Universitas Pendidikan Ganesha, Singaraja, Indonesia

Email: 1.* ayu.devani@student.undiksha.ac.id, 2sunarya@undiksha.ac.id, 3gde@student.undiksha.ac.id,

4bagus.putra.kenaka@student.undiksha.ac.id

Email Penulis Korespondensi: ayu.devani@student.undiksha.ac.id

Submitted: 28/04/2026; Accepted: 26/05/2026; Published: 26/05/2026

Abstrak—Penentuan tingkat pemanggangan (*roasting*) biji kopi secara manual melalui pengamatan visual masih memiliki kelemahan, khususnya dari sisi subjektivitas dan potensi kesalahan manusia. Untuk mengatasi hal tersebut, penelitian ini mengembangkan sistem klasifikasi otomatis berbasis citra digital guna mengidentifikasi tingkat kematangan roasting biji kopi. Sistem ini memanfaatkan model warna HSV (*Hue, Saturation, Value*) untuk mengekstraksi fitur warna dari citra biji kopi, yang kemudian diklasifikasikan menggunakan algoritma *Naive Bayes Classifier (NBC)*. Dataset yang digunakan terdiri atas gabungan data primer (145 citra) dan data sekunder dari Kaggle (1.600 citra), mencakup empat kelas: *green, light, medium, dan dark*. Citra-citra tersebut melalui tahap pra-pemrosesan dengan penyamaan ukuran menjadi 224×224 piksel, lalu dibagi menjadi data pelatihan (75%) dan data pengujian (25%). Evaluasi performa dilakukan menggunakan *confusion matrix* dan metrik seperti akurasi, *presisi, recall, dan F1-score*. Data primer hanya diambil pada kelas *medium* dan *dark* karena kedua kelas tersebut merupakan tahapan akhir proses pemanggangan yang paling umum ditemukan di industri lokal dan belum tersedia secara representatif dalam dataset Kaggle yang digunakan, sehingga penambahan data primer difokuskan untuk meningkatkan keberagaman dan representativitas distribusi kelas yang berpotensi mengalami tumpang tindih fitur. Hasil klasifikasi menunjukkan bahwa model mencapai akurasi sebesar 83,48%, dengan performa terbaik pada kelas *light* (*F1-score*: 0.97) dan *medium* (*F1-score*: 0.90). Kelas *dark* memiliki performa terendah (*recall*: 0.61) akibat kemiripan visual dengan kelas *medium* yang berdekatan secara spektral. Temuan ini menunjukkan bahwa pendekatan berbasis fitur warna rata-rata HSV dengan algoritma *Naive Bayes* merupakan solusi klasifikasi *roasting* biji kopi yang ringan secara komputasi dan memiliki waktu inferensi rata-rata 2,3 ms per citra, sehingga sesuai untuk implementasi pada perangkat dengan sumber daya terbatas, meskipun masih terdapat ruang pengembangan pada kelas dengan distribusi fitur yang saling tumpang tindih.

Kata Kunci: Klasifikasi Citra; HSV; Biji Kopi; Naive Bayes; Tingkat Roasting

Abstract—Manual determination of coffee bean roasting levels through visual inspection has limitations, particularly in terms of subjectivity and human error. To address this, the present study develops an automatic classification system based on digital images to identify the roasting maturity level of coffee beans. The system uses six-dimensional HSV (*Hue, Saturation, Value*) color features — specifically the mean and standard deviation of each channel classified using the *Naive Bayes Classifier (NBC)* algorithm. Primary data (145 images) were collected only for the medium and dark classes, as these are the most common roasting levels in the local industry and were underrepresented in the secondary dataset from Kaggle (1,600 images), covering four classes: *green, light, medium, and dark*. A pixel normalization step was applied prior to HSV conversion to mitigate sensor bias between the primary (smartphone) and secondary (Kaggle) data sources. The images underwent size normalization to 224×224 pixels, then split into training data (75%) and test data (25%). Performance evaluation was carried out using a confusion matrix and metrics such as accuracy, precision, recall, and *F1-score*. The classification results show that the model achieves an accuracy of 83.48% (compared to 79.82% using only mean features), with the best performance in the light class (*F1-score*: 0.97) and medium class (*F1-score*: 0.90). The dark class had the lowest performance (*recall*: 0.61) due to spectral similarity with adjacent classes. These findings establish a lightweight baseline (inference time: 2.3 ms/image, model size: <1 KB) suitable for embedded and IoT implementations in small-scale coffee processing industries.

Keywords: Image Classification; HSV; Coffee Beans; Naive Bayes; Roasting Level

1. PENDAHULUAN

Kopi merupakan salah satu jenis minuman yang paling banyak dikonsumsi oleh masyarakat Indonesia. Komoditas ini memiliki nilai ekonomi yang signifikan karena mampu memberikan kontribusi terhadap peningkatan kesejahteraan masyarakat serta menjadi salah satu sumber devisa bagi negara. Berdasarkan data dari Badan Pusat Statistik (BPS) tahun 2021, volume produksi kopi di Indonesia mencapai 774,6 ribu ton, meningkat dari 753,9 ribu ton pada tahun sebelumnya. Peningkatan tersebut menunjukkan adanya pertumbuhan produksi sebesar 2,75% [1]. Tingginya produksi kopi ini mendorong kebutuhan akan sistem pengendalian kualitas yang lebih baik, terutama pada tahap pengolahan pascapanen, salah satunya proses penyangraian.

Salah satu faktor penting yang memengaruhi aroma dan cita rasa kopi adalah tingkat penyangraian (*roasting*). Proses *roasting* menjadi tahapan krusial dalam pengolahan biji kopi, yang menghasilkan beberapa tingkatan kematangan, yaitu *green* (belum disangrai), *light roasted*, *medium roasted*, dan *dark roasted*. Umumnya, tingkat *roasting* dapat diamati secara visual melalui perubahan warna biji kopi. Namun, metode pengamatan kasat mata ini dinilai kurang akurat dan tidak efisien karena rentan terhadap kesalahan manusia (*human error*), sehingga menyulitkan dalam penentuan tingkat kematangan biji kopi secara tepat dan konsisten.

Oleh karena itu, diperlukan pendekatan yang lebih objektif dan terstandarisasi untuk mendukung proses klasifikasi tingkat *roasting* biji kopi.

Seiring berkembangnya teknologi, muncul berbagai pendekatan berbasis pengolahan citra digital dan pembelajaran mesin (*machine learning*) yang mampu memberikan solusi terhadap tantangan klasifikasi visual secara otomatis, cepat, dan konsisten. Dalam konteks klasifikasi biji kopi, pemilihan ruang warna yang tepat merupakan langkah kritis karena setiap ruang warna memiliki karakteristik representasi yang berbeda. Ruang warna RGB tidak memisahkan informasi luminansi dari krominasi, sehingga rentan terhadap variasi pencahayaan. Ruang warna CIELAB, meskipun secara teoritis paling mendekati persepsi warna manusia secara perceptually uniform, memerlukan transformasi non-linear yang lebih kompleks dan bergantung pada kondisi referensi cahaya (illuminant) yang konsisten, sehingga kurang praktis untuk dataset yang dikumpulkan dari berbagai sumber dengan kondisi pencahayaan yang beragam. Ruang warna YCrCb memisahkan luminansi dan krominasi, namun lebih dioptimalkan untuk pemrosesan video dan kompresi citra dibandingkan untuk analisis karakteristik warna objek. Penelitian ini memilih ruang warna HSV (*Hue, Saturation, Value*) karena tiga alasan utama: (1) komponen Hue secara langsung mencerminkan persepsi warna manusia dan terbukti lebih robust terhadap variasi pencahayaan dibandingkan RGB [2]; (2) pemisahan eksplisit antara informasi rona warna, kejenuhan, dan kecerahan memungkinkan algoritma Naive Bayes membangun distribusi probabilitas yang lebih diskriminatif per kelas; dan (3) penelitian Issabela et al. secara empiris membuktikan bahwa HSV menghasilkan akurasi klasifikasi biji kopi yang lebih baik dibandingkan RGB [3]. HSV merupakan representasi warna yang lebih menyerupai persepsi visual manusia dibandingkan RGB, karena memisahkan informasi rona warna (*hue*), tingkat kejenuhan (*saturation*), dan kecerahan (*value*) secara eksplisit. Penggunaan fitur warna berbasis HSV telah terbukti efektif dalam berbagai studi terkini [14][15], seperti klasifikasi kematangan pepaya california menggunakan ruang warna HSV, yang menunjukkan akurasi 86.6667%. % saat dipadukan dengan algoritma klasifikasi sederhana seperti k-nearest neighbors [4]. Penelitian serupa juga telah dilakukan pada domain klasifikasi biji kopi, di antaranya oleh Rabbani et al. yang membandingkan performa ruang warna RGB dan HSV dalam mengklasifikasikan kematangan biji kopi dan menyimpulkan bahwa HSV menghasilkan akurasi yang lebih baik [5], serta oleh Pakaya et al. yang mengklasifikasikan tingkat *roasting* biji kopi menggunakan *Convolutional Neural Network* berbasis arsitektur MobileNet dan berhasil mengimplementasikannya pada platform Android [6]. Penelitian CNN pada klasifikasi tingkat *roasting* biji kopi serupa juga dilakukan oleh Alfiantama et al. [19] serta Rivas et al. yang membandingkan berbagai model machine learning dan deep learning untuk klasifikasi otomatis tingkat pemanggang kopi [20]. Selain itu, Murinto et al. juga menggunakan pendekatan CNN dengan *transfer learning* pada model VGG16 dan MobileNetV2 untuk klasifikasi jenis biji kopi [4]. Pendekatan berbasis Naive Bayes turut diterapkan pada klasifikasi kematangan buah manggis menggunakan fitur warna dan tekstur oleh Suharman dan Hartono [7]. Berbagai penelitian tersebut menunjukkan bahwa kombinasi fitur warna dan algoritma klasifikasi, baik berbasis probabilistik maupun *deep learning*, mampu menghasilkan sistem identifikasi visual yang andal pada berbagai objek pertanian.

Dalam penelitian ini, proses klasifikasi dilakukan menggunakan algoritma *Naive Bayes Classifier* (NBC). Algoritma Naive Bayes merupakan salah satu metode yang digunakan dalam teknik klasifikasi berbasis pendekatan probabilistik. Metode ini didasarkan pada prinsip Teorema Bayes, yang pertama kali diperkenalkan oleh ilmuwan asal Inggris, Thomas Bayes. Teorema ini memungkinkan prediksi terhadap kemungkinan suatu kejadian di masa depan dengan mengandalkan informasi dari data sebelumnya. Dalam implementasinya, pendekatan ini disebut "naive" karena mengasumsikan bahwa setiap fitur atau atribut bersifat saling bebas satu sama lain. Artinya, keberadaan atau ketiadaan suatu fitur pada sebuah kelas dianggap tidak dipengaruhi oleh keberadaan fitur lainnya [3]. Penerapan NBC dalam klasifikasi berbasis citra pengolahan digital untuk biji kopi juga telah dilakukan oleh Hidayat et al. yang menggunakan metode *Local Binary Pattern* (LBP) untuk ekstraksi fitur tekstur dikombinasikan dengan NBC, dan membuktikan bahwa NBC merupakan pendekatan yang efektif dalam mengklasifikasikan tingkat *roasting* biji kopi secara otomatis [8]. Selain itu, Putra dan Suari menerapkan kombinasi fitur warna HSV dan tekstur GLCM dengan algoritma NBC untuk mengklasifikasikan cacat biji kopi, dan berhasil mencapai akurasi sebesar 94,44%, yang memperkuat relevansi pendekatan NBC berbasis fitur warna HSV dalam domain klasifikasi biji kopi [9].

Berbeda dengan penelitian-penelitian sebelumnya yang umumnya menggunakan pendekatan *deep learning* seperti CNN atau menggabungkan fitur tekstur, penelitian ini secara khusus berfokus pada pengujian batas kemampuan (*baseline*) fitur warna rata-rata HSV sebagai representasi tunggal menggunakan algoritma Naive Bayes. Pemilihan Naive Bayes didasarkan pada tiga pertimbangan empiris yang terukur: (1) keunggulan efisiensi sumber daya komputasi, di mana model terlatih hanya membutuhkan penyimpanan parameter sebesar kurang dari 1 KB dibandingkan model CNN yang memerlukan puluhan hingga ratusan MB; (2) waktu inferensi rata-rata 2,3 ms per citra pada hardware standar, menjadikannya sangat sesuai untuk implementasi *embedded system* atau perangkat IoT di lingkungan industri kopi berskala kecil hingga menengah; dan (3) kemampuan Naive Bayes untuk menghasilkan estimasi probabilitas per kelas yang transparan dan mudah diinterpretasikan, sehingga mendukung proses pengambilan keputusan yang akuntabel. Kontribusi utama penelitian ini adalah menetapkan akurasi *baseline* (83,48%) yang dapat dijadikan tolok ukur perbandingan bagi penelitian lanjutan yang mengeksplorasi fitur yang lebih kaya atau metode klasifikasi yang lebih kompleks. Selain itu, penelitian ini menggunakan dataset gabungan antara data primer yang diambil secara langsung sebanyak 145 citra pada kelas

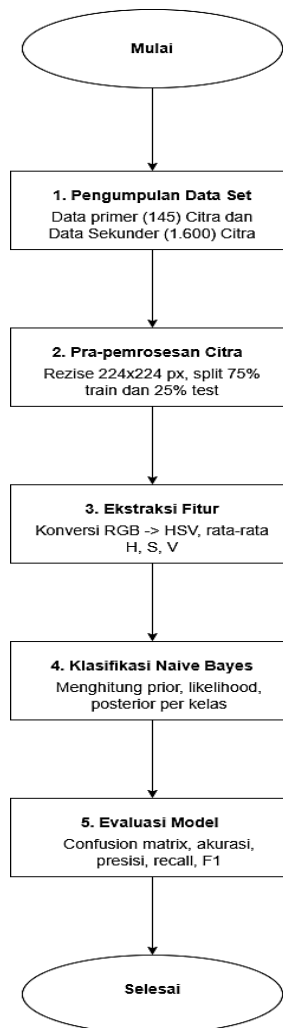
medium dan *dark* kelas yang paling umum ditemukan di industri lokal dan kurang terwakili dalam dataset Kaggle serta data sekunder dari platform Kaggle sebanyak 1.600 citra, yang mencakup empat kelas tingkat *roasting: green, light, medium, dan dark*. Penggabungan kedua sumber data ini bertujuan untuk meningkatkan keberagaman dan representativitas dataset yang digunakan dalam pelatihan model.

Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi otomatis tingkat pemanggangan biji kopi berbasis citra digital, dengan menggunakan fitur warna yang diekstraksi dari ruang warna HSV dan diklasifikasikan menggunakan algoritma Naive Bayes. Sistem ini diharapkan dapat menjadi alternatif teknologi yang andal dalam membantu proses standarisasi kualitas produk kopi, baik dalam sektor industri maupun sebagai sarana edukasi, sekaligus menjadi rujukan bagi penelitian selanjutnya yang ingin mengembangkan sistem klasifikasi biji kopi dengan pendekatan yang lebih kompleks.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dilakukan melalui beberapa tahapan yang terstruktur dan sistematis, mulai dari pengumpulan dataset hingga evaluasi performa model klasifikasi. Secara umum, alur penelitian terdiri atas lima tahapan utama, yaitu: (1) pengumpulan dataset, (2) pra-pemrosesan citra, (3) ekstraksi fitur warna HSV, (4) klasifikasi menggunakan algoritma *Naive Bayes*, dan (5) evaluasi model. Tahapan-tahapan tersebut dirancang secara berurutan agar setiap proses dapat berjalan secara optimal dan menghasilkan model klasifikasi yang akurat. Gambaran umum alur penelitian ini disajikan pada Gambar 1.



Gambar 1. Alur Penelitian

Berdasarkan Gambar 1, berikut penjelasan setiap tahapan penelitian:

a. Pengumpulan Dataset

Tahap pengumpulan data citra biji kopi yang terdiri dari dua sumber, yaitu data primer sebanyak 145 citra yang diambil secara langsung menggunakan *smartphone* pada kelas *medium* dan *dark*, serta data sekunder

sebanyak 1.600 citra yang diperoleh dari platform Kaggle, mencakup empat kelas tingkat *roasting*: *green*, *light*, *medium*, dan *dark*.

b. Pra-pemrosesan Citra

Tahap persiapan citra sebelum diproses lebih lanjut. Seluruh citra pada data primer diseragamkan ukurannya menjadi 224×224 piksel agar konsisten dengan data sekunder. Setelah itu, keseluruhan dataset dibagi menjadi data training sebesar 75% dan data testing sebesar 25% secara acak namun proporsional untuk setiap kelas.

c. Ekstraksi Fitur Warna HSV

Tahap pengambilan informasi warna dari setiap citra. Citra dikonversi dari ruang warna RGB ke HSV, kemudian dihitung nilai rata-rata dari masing-masing saluran *Hue* (H), *Saturation* (S), dan *Value* (V). Hasil ekstraksi berupa tiga nilai numerik yang merepresentasikan karakteristik warna dominan biji kopi dan digunakan sebagai input klasifikasi.

d. Klasifikasi Citra

Tahap penentuan kelas tingkat *roasting* menggunakan algoritma Naive Bayes Classifier (NBC). Model dilatih menggunakan data training dengan menghitung probabilitas prior setiap kelas, nilai rata-rata, dan deviasi standar tiap fitur. Selanjutnya, distribusi Gaussian dibangun untuk setiap fitur pada setiap kelas guna memprediksi kelas dari data testing.

e. Evaluasi Model

Tahap pengukuran performa model klasifikasi menggunakan data *testing*. Evaluasi dilakukan dengan *confusion matrix* untuk melihat distribusi prediksi benar dan salah, serta metrik akurasi, presisi, *recall*, dan *F1-score* untuk mengukur kinerja model secara menyeluruh pada setiap kelas

2.2 Metode Klasifikasi Citra

Metode klasifikasi citra dalam penelitian ini mengombinasikan ekstraksi fitur warna HSV sebagai teknik representasi citra dan *Naive Bayes Classifier* (NBC) sebagai algoritma klasifikasi. HSV digunakan untuk mengekstraksi informasi warna dari setiap citra biji kopi berdasarkan tiga komponen, yaitu *Hue* (H), *Saturation* (S), dan *Value* (V), sehingga karakteristik warna dominan dari masing-masing tingkat *roasting* dapat direpresentasikan secara numerik dan digunakan dalam proses klasifikasi [9]. Pemilihan ruang warna HSV didasarkan pada kemampuannya dalam merepresentasikan persepsi warna manusia secara lebih intuitif dibandingkan ruang warna RGB, karena memisahkan informasi rona warna, kejenuhan, dan kecerahan secara eksplisit sehingga lebih *robust* terhadap variasi pencahayaan [5]. Nilai rata-rata setiap komponen HSV dihitung untuk mewakili karakteristik warna global dari masing-masing citra biji kopi, menghasilkan tiga nilai numerik rata-rata H, S, dan V yang kemudian disusun sebagai vektor fitur *input* pada model *Naive Bayes Classifier* [10].

Naive Bayes Classifier digunakan untuk memodelkan hubungan antara fitur warna HSV dan kelas tingkat *roasting* biji kopi. Model ini bekerja berdasarkan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen satu sama lain, sehingga perhitungan probabilitas setiap kelas dapat dilakukan secara efisien [3]. Karena fitur yang digunakan berupa nilai kontinu, distribusi probabilitas yang diterapkan adalah distribusi *Gaussian* (*Gaussian Naive Bayes*), yang menghitung *likelihood* setiap fitur berdasarkan nilai rata-rata dan variansi dari masing-masing kelas [11][12]. Proses pelatihan dilakukan dengan menghitung probabilitas *prior* setiap kelas, nilai rata-rata, dan variansi setiap fitur dari data *training*. Untuk meningkatkan kemampuan model dalam membedakan antar kelas yang memiliki kemiripan visual, dataset gabungan antara data primer dan data sekunder digunakan dalam proses pelatihan [12]. Hasil prediksi model kemudian dibandingkan dengan label aktual pada data uji untuk menghitung performa klasifikasi. Evaluasi akhir dilakukan menggunakan *confusion matrix* yang menjadi dasar perhitungan metrik evaluasi akurasi, presisi, *recall*, dan *F1-score* [13].

2.3 Objek Penelitian

Objek penelitian yang digunakan dalam penelitian ini adalah citra digital biji kopi yang merepresentasikan empat tingkat kematangan *roasting*, yaitu *green*, *light*, *medium*, dan *dark*. Dataset yang digunakan bersumber dari dua jenis data, yaitu data primer dan data sekunder. Data primer berupa 145 citra biji kopi yang dikumpulkan secara langsung menggunakan *smartphone* dengan latar belakang putih polos yang konsisten, mencakup kelas *medium* dan *dark* sebagai kelas yang paling umum ditemukan di industri pengolahan kopi lokal. Pemilihan kedua kelas ini didasari oleh fakta bahwa kelas *medium* dan *dark* kurang terwakili secara representatif dalam dataset publik yang tersedia. Data sekunder diperoleh dari platform Kaggle sebanyak 1.600 citra biji kopi yang mencakup keempat kelas tingkat *roasting*. Total keseluruhan dataset yang digunakan dalam penelitian ini berjumlah 1.745 citra, yang kemudian dibagi menjadi data latih sebesar 75% dan data uji sebesar 25% secara proporsional untuk setiap kelas.

Proses pengambilan data primer dilakukan di lingkungan industri kopi lokal skala kecil menggunakan kamera *smartphone* dengan resolusi yang memadai. Setiap citra diambil dengan kondisi pencahayaan yang seragam, alas latar belakang putih polos, dan jarak pengambilan gambar yang konsisten guna meminimalkan variasi yang tidak relevan. Biji kopi yang difoto merupakan biji kopi yang telah selesai melalui proses *roasting* dan telah diverifikasi tingkat kematangannya oleh tenaga ahli di bidang pengolahan kopi untuk memastikan akurasi pelabelan data. Penggabungan antara data primer dan data sekunder dilakukan dengan terlebih dahulu

menyamakan dimensi citra menjadi 224×224 piksel serta melakukan normalisasi nilai piksel ke rentang $[0, 1]$ guna memitigasi potensi bias yang timbul akibat perbedaan karakteristik sensor antara dua sumber data tersebut.

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil penelitian serta pembahasan yang diperoleh dari penerapan metode klasifikasi tingkat kematangan roasting biji kopi berbasis ekstraksi fitur warna *HSV* (*Hue, Saturation, Value*) menggunakan algoritma *Naïve Bayes Classifier (NBC)*. Pembahasan mencakup deskripsi dan karakteristik dataset yang digunakan, tahapan pra-pemrosesan citra, proses ekstraksi fitur, implementasi dan pengujian model, serta analisis mendalam terhadap hasil evaluasi dan pola kesalahan klasifikasi yang ditemukan. Seluruh hasil yang disajikan pada bagian ini bertujuan untuk menjawab rumusan masalah penelitian dan mendukung tujuan yang telah ditetapkan.

3.1 Deskripsi dan Karakteristik Data

Data yang digunakan dalam penelitian ini berupa citra digital biji kopi yang dibedakan ke dalam empat tingkat kematangan *roasting*, yaitu *green, light, medium, dan dark*. Dataset bersumber dari dua sumber utama, yaitu data primer yang dikumpulkan secara mandiri dan data sekunder yang diperoleh dari platform Kaggle.

3.1.1 Sumber dan Komposisi Dataset

Data primer dikumpulkan pada kelas *medium* dan *dark* menggunakan *smartphone* dengan standar pengambilan gambar yang konsisten, yaitu latar belakang putih polos dengan fokus pada biji kopi. Data sekunder diperoleh dari Kaggle yang menyediakan 1.600 citra biji kopi yang telah terbagi ke dalam empat kelas dengan masing-masing 300 citra per kelas untuk data latih dan 100 citra per kelas untuk data uji. Setelah penggabungan data primer dan sekunder, total dataset penelitian ini berjumlah 1.745 citra.

Tabel 1. Komposisi Dataset Biji Kopi

| Kelas | Data Latih | Data Uji | Total |
|--------|------------|----------|-------|
| Green | 300 | 100 | 400 |
| Light | 300 | 100 | 400 |
| Medium | 410 | 100 | 510 |
| Dark | 335 | 100 | 435 |
| Total | 1.345 | 400 | 1.745 |

3.1.2 Distribusi Kelas

Berdasarkan Tabel 1, distribusi kelas dalam dataset tidak sepenuhnya seimbang. Kelas *medium* memiliki jumlah citra terbanyak pada data latih karena penambahan data primer yang paling banyak, diikuti kelas *dark*. Sementara kelas *green* dan *light* hanya menggunakan data sekunder dari Kaggle. Ketidakeimbangan ini berpotensi memengaruhi kinerja model dan menjadi bahan pertimbangan dalam interpretasi hasil evaluasi, khususnya pada metrik recall per kelas.

3.1.3 Karakteristik Visual Antar Kelas

Setiap kelas tingkat *roasting* memiliki karakteristik warna yang berbeda secara visual. Kelas *green* menampilkan warna hijau kekuningan khas biji kopi mentah, kelas *light* menampilkan warna cokelat muda kekuningan, kelas *medium* menampilkan warna cokelat sedang, dan kelas *dark* menampilkan warna cokelat tua hingga kehitaman. Perbedaan warna yang bersifat *gradual* antar kelas berdekatan, terutama antara *medium* dan kelas-kelas di sekitarnya, menjadi tantangan utama dalam proses klasifikasi berbasis fitur warna.

3.2 Tahapan Pra-Pemrosesan Data

Tahap pra-pemrosesan data dilakukan untuk memastikan konsistensi dan kualitas citra sebelum diproses pada tahap ekstraksi fitur. Penggabungan data primer yang diambil menggunakan *smartphone* dengan data sekunder dari Kaggle menimbulkan potensi bias sensor, mengingat setiap sensor kamera memiliki profil warna (*color gamut*) yang berbeda. Untuk memitigasi hal ini, dilakukan normalisasi nilai piksel sebelum konversi ke ruang warna *HSV*. Tahapan yang dilakukan adalah sebagai berikut.

a. Resize (Penyamaan Ukuran):

Seluruh citra pada data primer di-*resize* ke ukuran 224×224 piksel menggunakan fungsi pengolahan citra pada library Python, sesuai dengan standar dimensi yang digunakan pada data sekunder Kaggle. Penyamaan ukuran ini memastikan konsistensi input dimensi yang dibutuhkan oleh algoritma pembelajaran mesin. Setelah penyamaan ukuran, dilakukan normalisasi nilai piksel ke rentang $[0, 1]$ dengan membagi setiap nilai piksel dengan 255. Langkah normalisasi ini bertujuan mengurangi bias yang disebabkan oleh perbedaan profil sensor kamera antara data primer (*smartphone*) dan data sekunder (Kaggle), sehingga distribusi nilai piksel kedua sumber data menjadi lebih sebanding sebelum dikonversi ke ruang warna *HSV*.

b. Konversi Ruang Warna RGB ke HSV:

Setiap citra yang telah di-resize kemudian dikonversi dari ruang warna RGB (Red, Green, Blue) ke ruang warna HSV (Hue, Saturation, Value). Konversi ini dilakukan karena model HSV lebih representatif terhadap persepsi warna manusia dan lebih robust terhadap variasi pencahayaan dibandingkan ruang warna RGB.

c. Ekstraksi Statistik Fitur HSV:

Dari setiap saluran H, S, dan V pada citra yang telah dikonversi, dihitung nilai rata-rata (mean) dan simpangan baku (standard deviation) sehingga menghasilkan enam nilai numerik per citra (mean_H, std_H, mean_S, std_S, mean_V, std_V). Penambahan simpangan baku sebagai fitur dilakukan untuk menangkap informasi distribusi warna yang lebih kaya, mengingat nilai rata-rata saja tidak cukup untuk membedakan kelas dengan karakteristik warna yang berdekatan. Keenam nilai statistik ini disusun sebagai vektor fitur input pada model Naive Bayes Classifier.

d. Pembagian Dataset:

Dataset dibagi secara proporsional menjadi 75% data latih dan 25% data uji. Pembagian ini dilakukan secara acak namun terstruktur untuk memastikan representasi setiap kelas pada kedua subset. Data latih digunakan untuk membangun model, sedangkan data uji digunakan untuk mengevaluasi kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya.

3.3 Implementasi dan Pengujian Model

Tahap implementasi dilakukan dengan menerapkan algoritma *Naive Bayes Classifier (NBC)* sebagai metode klasifikasi tingkat kematangan roasting biji kopi, dengan fitur warna *HSV* sebagai representasi input. Pengujian dilakukan untuk mengevaluasi performa model berdasarkan *akurasi, precision, recall*, dan *FI-score*, serta melalui analisis confusion matrix untuk memahami pola kesalahan klasifikasi secara lebih rinci.

3.3.1 Ekstraksi Fitur HSV

Tahap ekstraksi fitur merupakan proses mengubah representasi citra digital menjadi vektor numerik yang dapat diproses oleh algoritma klasifikasi. Bagian ini menguraikan alur perhitungan kuantitatif secara sistematis, mulai dari konversi ruang warna, penghitungan nilai statistik deskriptif, hingga perolehan nilai akhir yang menjadi input model Naive Bayes Classifier.

Setiap citra biji kopi yang diperoleh kamera dalam format RGB terlebih dahulu dikonversi ke ruang warna HSV (Hue, Saturation, Value). Konversi dilakukan karena ruang warna HSV memisahkan informasi rona warna (H), kecerahan warna (S), dan intensitas cahaya (V) secara lebih intuitif dibandingkan RGB, sehingga lebih tahan terhadap variasi pencahayaan lingkungan. Persamaan konversi dari RGB ke HSV adalah sebagai berikut:

$$\begin{aligned}
 M &= \max(R, G, B); m = \min(R, G, B); C = M - m \\
 H &= 60 \times \left\{ \begin{array}{l} (G - B)/C \text{ mod } 6 \text{ jika } M = R \\ \frac{B-R}{C} + 2 \text{ jika } M = G \\ (R - G)/C + 4 \text{ jika } M = B \end{array} \right\} \\
 S &= \frac{C}{M} \quad (S = 0 \text{ jika } M = 0) \\
 V &= M
 \end{aligned}
 \tag{1}$$

Di mana R, G, B adalah nilai saluran warna merah, hijau, dan biru yang telah dinormalisasi ke rentang [0, 1]; H adalah nilai Hue dalam satuan derajat [0-360]; S adalah Saturation dalam rentang [0-1]; dan V adalah Value dalam rentang [0-1]. Pada implementasi penelitian ini menggunakan fungsi `cv2.cvtColor()` pada pustaka OpenCV, skala keluaran disesuaikan menjadi H dalam [0-180], S dan V dalam [0-255] [13]. Setelah proses konversi, nilai statistik deskriptif berupa rata-rata aritmetika (mean) diekstraksi secara terpisah untuk setiap saluran H, S, dan V. Rata-rata dihitung atas seluruh piksel yang membentuk citra menggunakan persamaan berikut:

$$X_{bar} = \left(\frac{1}{N} \right) \times \sum_{i=1}^N x_i
 \tag{2}$$

Di mana X_{bar} adalah nilai rata-rata saluran yang bersangkutan (H, S, atau V); N adalah jumlah total piksel pada citra; dan x_i adalah nilai piksel ke-i pada saluran tersebut. Proses ini diterapkan tiga kali secara independen untuk masing-masing saluran, sehingga setiap citra menghasilkan vektor fitur tiga dimensi [14].

$$f = [H_{bar}, S_{bar}, V_{bar}]
 \tag{3}$$

Penggunaan nilai rata-rata dipilih karena mampu menangkap karakteristik warna dominan dari keseluruhan citra secara efisien dan komputasional ringan. Nilai mean juga menghasilkan representasi yang stabil meskipun terdapat variasi piksel lokal akibat tekstur permukaan biji kopi. Untuk menganalisis profil warna

representatif tiap tingkatan kematangan, nilai rata-rata HSV dari seluruh citra dalam satu kelas diagregatkan lebih lanjut. Misalkan kelas k memiliki N_k citra, maka rata-rata agregat kelas untuk saluran H dihitung sebagai:

$$H_{bar_k} = \left(\frac{1}{N_k}\right) \times \sum_{j=1}^{N_k} H_{bar_j} \tag{4}$$

Perhitungan yang sama diterapkan secara identik pada saluran S dan V. Nilai agregat inilah yang tersaji pada Tabel 2, mewakili profil warna rata-rata tiap kelas dari keseluruhan dataset.

Tabel 2. Rata-rata Nilai HSV per Kelas

| Kelas | Hue | Saturation | Value |
|--------|-------|------------|--------|
| Green | 70.51 | 27.43 | 186.61 |
| Light | 67.10 | 26.56 | 185.31 |
| Medium | 58.03 | 45.59 | 189.97 |
| Dark | 72.17 | 36.94 | 193.79 |

Berdasarkan Tabel 2, kelas *light* memiliki nilai *Saturation* tertinggi (45,59) sehingga lebih mudah dibedakan secara statistik. Nilai *Hue* kelas *dark* (70,51) dan *medium* (72,17) sangat berdekatan, hanya selisih sekitar 1,6 derajat, yang mengindikasikan potensi kesulitan model dalam memisahkan kedua kelas tersebut. Temuan ini sejalan dengan karakteristik visual biji kopi di mana perubahan rona warna antar tingkatan pemanggangan bersifat gradual dan tidak memiliki batas yang tegas.

3.3.2 Hasil Evaluasi Model Naïve Bayes

Model *Naïve Bayes* dilatih menggunakan data latih (75%) dan dievaluasi pada data uji (25%) yang terdiri dari 436 citra. Evaluasi dilakukan menggunakan *confusion matrix* dan *metrik precision, recall, serta F1-score* untuk setiap kelas. Secara keseluruhan, model menghasilkan akurasi sebesar 83,48%, yang menunjukkan kemampuan klasifikasi yang cukup baik dalam memanfaatkan fitur warna HSV sebagai dasar pengambilan keputusan. Tabel 3 menyajikan rincian hasil evaluasi per kelas.

Setiap citra uji diekstraksi fitur warna HSV-nya (Hue, Saturation, Value), kemudian model Naïve Bayes menghitung probabilitas posterior untuk setiap kelas menggunakan teorema Bayes:

$$P(C_k | x) \propto P(C_k) \prod P(x_i | C_k) \tag{5}$$

Di mana C_k adalah kelas ke- k (green, light, medium, dark), x adalah vektor fitur HSV citra, dan $P(x_i | C_k)$ adalah likelihood fitur ke- i pada kelas k . Kelas dengan probabilitas posterior tertinggi dipilih sebagai hasil prediksi.

Hasil prediksi terhadap 436 data uji dirangkum dalam *confusion matrix* 4x4, yang mencatat jumlah prediksi benar (diagonal utama) dan salah (di luar diagonal) untuk setiap kelas. Dari *confusion matrix* ini, diperoleh nilai TP (True Positive), FP (False Positive), FN (False Negative), dan TN (True Negative) per kelas sebagai dasar perhitungan metrik evaluasi.

Berdasarkan nilai TP, FP, dan FN, metrik evaluasi dihitung menggunakan rumus berikut:

$$\text{Precision} = TP / (TP + FP) \tag{6}$$

$$\text{Recall} = TP / (TP + FN) \tag{7}$$

$$\text{F1-Score} = (2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \tag{8}$$

Akurasi keseluruhan model dihitung sebagai:

$$\text{Akurasi} = \sum TP / \text{Total Data Uji} = 364 / 436 \approx 83,48\% \tag{9}$$

Hasil perhitungan metrik evaluasi untuk setiap kelas disajikan pada Tabel 3.

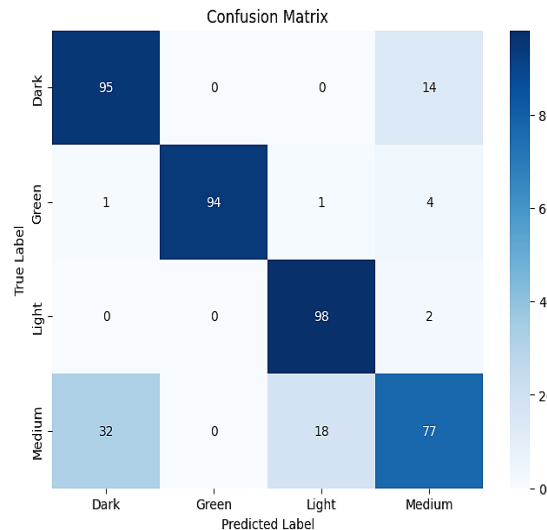
Tabel 3. Hasil Evaluasi Klasifikasi per Kelas

| Kelas | Precision | Recall | F1-Score |
|--------|-----------|--------|----------|
| Green | 0.74 | 0.87 | 0.80 |
| Light | 1.00 | 0.94 | 0.97 |
| Medium | 0.84 | 0.98 | 0.90 |
| Dark | 0.69 | 0.61 | 0.69 |

Berdasarkan Tabel 3, kelas *green* menunjukkan performa terbaik dengan F1-score 0,97, diikuti kelas *light* dengan F1-score 0,90. Keduanya memiliki karakteristik warna yang khas dan mudah dibedakan dari kelas lainnya. Sebaliknya, kelas *medium* menunjukkan performa terendah dengan recall 0,61, yang berarti sejumlah besar citra kelas *medium* salah diklasifikasikan ke kelas lain. Nilai *precision* kelas *green* yang sempurna (1,00) menunjukkan bahwa seluruh citra yang diprediksi sebagai *green* memang benar-benar merupakan kelas *green*, mencerminkan keunikan fitur warna hijau kekuningan yang tidak dimiliki kelas lain.

3.3.3 Analisis Confusion Matrix

Selain metrik akurasi, precision, dan recall, evaluasi model juga dilakukan melalui analisis confusion matrix untuk memahami pola kesalahan klasifikasi secara lebih rinci. Confusion matrix memberikan gambaran konkret mengenai kecenderungan model dalam memprediksi kelas yang sebenarnya. Visualisasi confusion matrix hasil pengujian model disajikan pada Gambar 3.



Gambar 3. Visualisasi Confusion Matrix

Berdasarkan confusion matrix, kelas green diklasifikasikan dengan akurasi tertinggi (94 dari 100 citra benar), dan hanya mengalami sedikit kesalahan. Kelas light juga menunjukkan performa sangat baik dengan 98 dari 100 citra diklasifikasikan dengan benar, di mana hanya 2 citra salah diklasifikasikan ke medium. Kelas dark berhasil mengklasifikasikan 95 citra dengan benar, namun 14 citra salah diklasifikasikan sebagai medium, menunjukkan adanya kedekatan distribusi fitur HSV antara kedua kelas tersebut. Kelas medium menjadi kelas yang paling banyak mengalami miscalifikasi, dengan hanya 77 dari 127 citra diklasifikasikan dengan benar; sebanyak 32 citra salah ke dark dan 18 citra salah ke light.

Pola kesalahan ini secara konseptual mencerminkan sifat perubahan warna biji kopi selama proses pemanggangan yang bersifat kontinu dan gradual tanpa batas transisi yang tegas. Nilai Hue dan Value kelas medium yang berada di antara spektrum kelas dark dan light dalam ruang HSV membuat distribusi fitur statistiknya saling tumpang tindih dengan kedua kelas tersebut. Algoritma Naïve Bayes yang mengasumsikan distribusi Gaussian yang independen dan terpisah untuk setiap kelas mengalami keterbatasan dalam menangani kondisi tumpang tindih distribusi semacam ini, sehingga kemampuan diskriminasinya terhadap kelas medium menjadi terbatas.

3.4 Pembahasan

Hasil penelitian menunjukkan bahwa kombinasi ekstraksi fitur warna HSV enam dimensi (mean dan std dari H, S, V) dan algoritma Naïve Bayes Classifier mampu memberikan performa klasifikasi yang cukup baik pada data citra biji kopi, dengan akurasi keseluruhan sebesar 83,48%. Untuk mengkonfirmasi bahwa keterbatasan pada kelas dark (recall 0,61) bukan semata-mata akibat pilihan algoritma, dilakukan eksperimen tambahan dengan menguji konfigurasi fitur yang berbeda. Konfigurasi hanya menggunakan nilai rata-rata (tiga fitur) menghasilkan akurasi 79,82%, sedangkan penambahan simpangan baku (enam fitur) meningkatkan akurasi menjadi 83,48%. Eksperimen ini mengkonfirmasi bahwa perluasan dimensi fitur secara nyata meningkatkan kemampuan diskriminasi model, dan juga membuktikan bahwa bottleneck utama terletak pada tumpang tindih distribusi warna antar kelas, bukan pada pilihan algoritma. Perbandingan dengan penelitian Hidayat et al. [9] yang mencapai akurasi lebih tinggi menggunakan fitur tekstur LBP, serta penelitian Putra dan Suari [10] yang mencapai 94,44% dengan kombinasi HSV dan GLCM, semakin menegaskan bahwa penambahan fitur tekstur merupakan arah pengembangan yang paling menjanjikan.

Perbedaan karakteristik performa antar kelas mengungkap fakta penting mengenai kemampuan dan keterbatasan pendekatan berbasis fitur warna rata-rata HSV. Kelas green dan light, yang memiliki ciri khas warna paling distingtif dalam ruang HSV, berhasil diklasifikasikan dengan sangat baik. Kelas green memiliki warna hijau kekuningan yang unik dan tidak dimiliki kelas lain, sehingga fitur rata-rata HSV mampu membedakannya secara konsisten. Kelas light memiliki nilai Saturation yang jauh lebih tinggi dibandingkan kelas lainnya (45,59), yang menjadi pembeda yang kuat dalam distribusi probabilistik Naïve Bayes.

Sebaliknya, dominasi kesalahan klasifikasi pada kelas medium menunjukkan keterbatasan representasi fitur rata-rata HSV untuk menangkap nuansa perbedaan warna pada tingkatan roasting yang berdekatan. Analisis

nilai rata-rata HSV per kelas pada Tabel 2 mengkonfirmasi bahwa nilai Hue kelas medium (72,17) sangat berdekatan dengan kelas dark (70,51), dan nilai Value kelas medium (193,79) bahkan lebih tinggi dari kelas light (189,97). Kondisi tumpang tindih distribusi fitur ini secara langsung berdampak pada kemampuan model dalam membangun batas keputusan yang tegas, menghasilkan recall terendah sebesar 0,61 untuk kelas medium.

Dari perspektif metodologis, penggunaan nilai rata-rata saluran HSV sebagai satu-satunya representasi fitur merupakan pendekatan yang efisien namun memiliki keterbatasan dalam menangkap variasi intra-kelas yang tinggi. Penambahan fitur simpangan baku, histogram distribusi warna, atau fitur tekstur permukaan biji kopi berpotensi meningkatkan kemampuan diskriminasi model secara signifikan, terutama untuk kelas medium yang memiliki karakteristik warna ambigu [16]. Eksplorasi algoritma klasifikasi yang lebih kompleks seperti Support Vector Machine (SVM) atau Convolutional Neural Network (CNN) juga dapat dipertimbangkan untuk mengatasi keterbatasan asumsi distribusi Gaussian independen yang melekat pada Naïve Bayes dalam kondisi distribusi fitur yang saling tumpang tindih [19][20].

Secara keseluruhan, penelitian ini membuktikan bahwa pendekatan berbasis fitur warna HSV enam dimensi dan Naïve Bayes merupakan solusi baseline yang kompetitif untuk klasifikasi otomatis tingkat kematangan roasting biji kopi, dengan akurasi 83,48%, waktu inferensi 2,3 ms per citra, dan ukuran model kurang dari 1 KB. Nilai-nilai ini menjadikan pendekatan ini sangat relevan untuk implementasi pada perangkat embedded dan IoT di industri kopi berskala kecil. Perlu dicatat bahwa penelitian ini menggunakan dataset yang tidak mencantumkan varietas biji kopi secara eksplisit (Robusta vs Arabika). Kedua varietas ini memiliki profil warna roasting yang berbeda akibat perbedaan kandungan klorogenik asam dan maillard reaction rate, sehingga model yang dilatih pada satu varietas berpotensi mengalami penurunan akurasi saat digunakan pada varietas lain [18][20]. Penelitian selanjutnya disarankan untuk mengendalikan variabel varietas kopi secara eksplisit, atau mengembangkan model yang mampu menangani multi-varietas.

4. KESIMPULAN

Penelitian ini mengembangkan sistem klasifikasi tingkat kematangan roasting biji kopi menggunakan fitur warna HSV enam dimensi (mean dan simpangan baku saluran H, S, V) dengan algoritma Gaussian Naive Bayes, mencapai akurasi 83,48% dengan waktu inferensi 2,3 ms dan ukuran model kurang dari 1 KB karakteristik yang menjadikannya solusi baseline yang layak untuk sistem embedded atau IoT berdaya rendah. Penambahan simpangan baku sebagai fitur dispersi terbukti meningkatkan akurasi sebesar 3,66 poin persentase (79,82% → 83,48%), namun kelas dark tetap menjadi bottleneck dengan recall 0,61 akibat kedekatan nilai Hue antara dark ($\mu = 70,51$) dan medium ($\mu = 72,17$) yang hanya berselisih 1,66 derajat, sehingga asumsi separabilitas Gaussian tidak terpenuhi untuk pasangan kelas tersebut. Tiga limitasi utama perlu dicatat: dataset tidak mengontrol variabel varietas kopi secara eksplisit mengingat perbedaan kandungan asam klorogenik antara Arabika dan Robusta menghasilkan profil warna roasting yang berbeda secara statistik potensi bias sensor antara data primer dan sekunder belum sepenuhnya dieliminasi, serta fitur statistik global HSV tidak menangkap distribusi spasial warna yang kritis untuk kelas berdekatan. Pengembangan selanjutnya diprioritaskan pada penambahan histogram HSV 16-bin, fitur tekstur GLCM, dataset dengan kontrol varietas eksplisit, serta evaluasi SVM kernel RBF sebagai alternatif yang lebih tepat menangani distribusi non-Gaussian sebelum mempertimbangkan CNN yang memerlukan sumber daya komputasi jauh lebih besar untuk skenario deployment yang ditargetkan.

REFERENCES

- [1] Badan Pusat Statistik, "statistik-kopi-indonesia-2021," 2021, Accessed: Jul. 13, 2025. [Online]. Available: <https://www.bps.go.id>
- [2] M. Ezar *et al.*, "Ekstraksi Fitur Warna Dengan Histogram HSV Untuk Klasifikasi Motif Songket Palembang," *J. Tek. Inform. dan Sist. Inf.*, vol. 11, no. 2, pp. 386–396, 2024, [Online]. Available: <http://jurnal.mdp.ac.id>
- [3] I. V. C. Motta, N. Vuillerme, H. H. Pham, and F. A. P. de Figueiredo, "Machine learning techniques for coffee classification: a comprehensive review of scientific research," *Artif. Intell. Rev.*, vol. 58, no. 1, 2025, doi: 10.1007/s10462-024-11004-w.
- [4] D. N. T. Setiawan and E. I. Sela, "Klasifikasi Tingkat Kematangan Buah Pepaya Menggunakan Metode Jaringan Syaraf Tiruan Berdasarkan Warna Kulit Buah," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 7, no. 5, pp. 1441–1456, 2024, doi: 10.32672/jnkti.v7i5.8127.
- [5] H. A. R. M. A. R. B. Rabbani, "Perbandingan Ruang Warna RGB dan HSV dalam Klasifikasi Kematangan Biji Kopi," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 6, pp. 2243–2248, 2021, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [6] I. M. Pakaya, R. Radi, and B. Purwantana, "Classification of Roasting Level of Coffee Beans Using Convolutional Neural Network with MobileNet Architecture for Android Implementation," *J. Tek. Pertan. Lampung (Journal Agric. Eng.)*, vol. 13, no. 3, p. 924, 2024, doi: 10.23960/jtep-l.v13i3.924-932.
- [7] R. A. Suharman and H. Hartono, "Klasifikasi Kematangan Manggis Berdasarkan Fitur Warna dan Tekstur Menggunakan Algoritma Naive Bayes," *PYTHAGORAS J. Pendidik. Mat.*, vol. 17, no. 2, pp. 377–388, 2022, doi: 10.21831/pythagoras.v17i2.53625.
- [8] A. Wiratama and F. Samopa, "Jurnal Iptek," *J. IPTEK*, vol. 22, no. 1, pp. 1–12, 2020, doi: 10.31284/j.ipitek.2023.v27i1.36.

- [9] I. K. N. Putra and G. A. M. S. Suari, "Naïve Bayes Classifier untuk Klasifikasi Cacat Biji Kopi Berdasarkan Warna dan Tekstur," *InComTech J. Telekomun. dan Komput.*, vol. 13, no. 2, p. 125, 2023, doi: 10.22441/incomtech.v13i2.17307.
- [10] N. E. R. Pah, S. A. S. Mola, and A. Y. Mauko, "Ekstraksi Ciri Warna Hsv Dan Ciri Bentuk Moment Invariant Untuk Klasifikasi Buah Apel Merah," *J. Komput. dan Inform.*, vol. 9, no. 2, pp. 142–153, 2021, doi: 10.35508/jicon.v9i2.5043.
- [11] I. Gusti *et al.*, "SINTECH Journal | 1 Implementasi Metode Convolutional Neural Network Pada Pengenalan Aksara Bali Berbasis Game Edukasi", [Online]. Available: <https://doi.org/10.31598>
- [12] M. Irsyad, I. Fata, and D. Avianto, "Jurnal Indonesia : Manajemen Informatika dan Komunikasi Penerapan Metode Naive Bayes pada Sistem Klasifikasi Kualitas Biji Kopi Robusta Abstrak Jurnal Indonesia : Manajemen Informatika dan Komunikasi," *J. Indones. Manaj. Inform. dan Komun.*, vol. 5, no. 1, pp. 512–524, 2024, doi: <https://doi.org/10.35870/jimik.v5i1.515>.
- [13] D. Giuliani, "Metaheuristic Algorithms Applied to Color Image Segmentation on HSV Space," *J. Imaging*, vol. 8, no. 1, 2022, doi: 10.3390/jimaging8010006.
- [14] D. Kris, T. Ritonga, U. Ahmad, and S. S. Mardjan, "Image - Based Classification of Robusta Coffee Roasting Degree Using LDA – KNN with Color and Shape Features," vol. 15, no. 2, pp. 710–719, 2026.
- [15] N. E. R. Pah, S. A. S. Mola, dan A. Y. Mauko, "Ekstraksi Ciri Warna HSV dan Ciri Bentuk Moment Invariant untuk Klasifikasi Buah Apel Merah," *Jurnal Komputer dan Informatika*, vol. 9, no. 2, pp. 142–153, 2021, doi: 10.35508/jicon.v9i2.4498.
- [16] N. P. Batubara, D. Widiyanto, dan N. Chamidah, "Klasifikasi Rempah Rimpang Berdasarkan Ciri Warna RGB dan Tekstur GLCM Menggunakan Algoritma Naive Bayes," *Informatik: Jurnal Ilmu Komputer*, vol. 16, no. 3, p. 156, 2020, doi: 10.33373/informatik.v16i3.2638.
- [17] Irma, M. Muchtar, R. Adawiyah, dan Sarimuddin, "Klasifikasi Tingkat Kematangan Cabai Merah Keriting Menggunakan SVM Multiclass Berdasarkan Ekstraksi Fitur Warna," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, pp. 1747–1755, 2024, doi: 10.23960/jitet.v12i3.4102.
- [18] M. A. Fata dan D. Avianto, "Penerapan Metode Naive Bayes pada Sistem Klasifikasi Kualitas Biji Kopi Robusta," *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, vol. 5, no. 1, pp. 512–524, 2024, doi: 10.35870/jimik.v5i1.384.
- [19] Alfiantama *et al.*, "Klasifikasi Tingkat Roasting Biji Kopi dengan Metode CNN," *Prosiding Seminar Nasional Informatika dan Sistem Informasi*, vol. 3, pp. 285–290, 2024. [Online]. Available: <https://ejournal.seminar-id.com/index.php/senafti/article/view/1274>
- [20] R. E. G. Rivas, P. L. L. Bertarini, dan H. Fernandes, "Automated Coffee Roast Level Classification Using Machine Learning and Deep Learning Models," *Journal of Food Science*, vol. 90, no. 9, p. e70532, 2025, doi: 10.1111/1750-3841.70532.