Terakreditasi Peringkat 4
Keputusan Direktur Jenderal Pendidikan Tinggi, Riset, dan Teknologi, No. 225/E/KPT/2022
Masa berlaku dari Volume 8 Nomor 1 Tahun 2022 sampai dengan Volume 12 Nomor 2 Tahun 2026

Published online at: **https://ejournal.raharja.ac.id/**

JOURNAL CERITA:
Creative Education of Research in Information Technology and Artificial informatics

Vol. 12 No. 2 (2026) 282-290 e-ISSN: 2655 - 2574

Klasifikasi Citra Biji Kopi Dengan Fitur Statistik RGB dan Algoritma SVM

Raisah Nurul Faridah[1], Anggraini Puspita Sari[*2], Hendra Maulana[3]
[1,2,3]Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya, Indonesia
Email: 21081010105@student.upnjatim.ac.id[1]; **anggraini.puspita.if@upnjatim.ac.id[*2]**;
hendra.maulana.if@upnjatim.ac.id[3]

Faridah, R. N., Sari, A. P., & Maulana, H. (2026). Klasifikasi Citra Biji Kopi Dengan Fitur Statistik RGB dan Algoritma SVM. *Journal Cerita: Creative Education of Research in Information Technology and Artificial Informatics*, 12(2), 282-290

DOI: https://doi.org/10.33050/cerita.v12i2. 3887

ABSTRAK

Perbedaan warna dan tekstur pada biji kopi memengaruhi cita rasa serta nilai jual produk, sehingga klasifikasi jenis biji kopi menjadi tahap penting dalam menjaga mutu dan konsistensi. Penelitian ini bertujuan mengembangkan sistem klasifikasi otomatis untuk tiga jenis biji kopi, yaitu Arabika, Robusta, dan Excelsa berbasis citra digital. Fitur warna diekstraksi dari kanal merah, hijau, dan biru dengan menghitung nilai statistik berupa rata-rata, simpangan baku, dan kemencengan. Data fitur digunakan sebagai masukan untuk model klasifikasi berbasis *Support Vector Machine*. Evaluasi dilakukan terhadap empat skenario pengujian dengan kombinasi rasio pembagian data dan jenis kernel. Hasil terbaik diperoleh pada rasio 70:30 menggunakan kernel linier, dengan akurasi sebesar 96,44% dan nilai F1-score sebesar 0,96. Sebaliknya, kernel fungsi radial menghasilkan akurasi lebih rendah, berkisar antara 72% hingga 73%. Temuan ini menunjukkan bahwa fitur warna statistik mampu digunakan secara efektif dalam klasifikasi biji kopi secara otomatis dan berpotensi diterapkan dalam proses penyortiran biji kopi di tingkat industri.

Kata kunci: klasifikasi biji kopi, warna RGB, citra digital, pembelajaran mesin, *Support Vector Machine*

Received: 20-02-2026 | Accepted: 20-06-2026 | Published Online: 08-08-2026

ABSTRACT

Variations in color and texture among coffee beans affect their flavor and market value, making bean classification a vital step in maintaining product quality and consistency. This study aims to develop an automatic classification system for three coffee bean types, there are Arabica, Robusta, and Excelsa based on digital images. Color features were extracted from the red, green, and blue channels by calculating statistical values: mean, standard deviation, and skewness. The extracted features were used as input for a Support Vector Machine classification model. Evaluation was carried out using four testing scenarios, combining different data split ratios and kernel types. The best result was achieved with a 70:30 train-test ratio using a linear kernel, yielding an accuracy of 96.44% and an F1-score of 0.96. In contrast, the radial basis function kernel produced lower accuracy, ranging from 72% to 73%. These results indicate that statistical color features can be effectively applied for automatic coffee bean classification and have potential for use in industrial sorting processes.

Keywords: *coffee bean classification, RGB color, digital image, machine learning, Support Vector Machine*

I. PENDAHULUAN

Kopi merupakan salah satu komoditas unggulan Indonesia yang tidak hanya dikonsumsi sebagai minuman, tetapi juga digunakan dalam berbagai industri seperti makanan, kosmetik, dan farmasi (Maharani & Nuryana, 2023). Dengan produksi yang mencapai 758,7 ribu ton pada tahun 2023 (Badan Pusat Statistik, 2024), kopi memainkan peran penting dalam perekonomian nasional. Pertumbuhan tanaman kopi dipengaruhi oleh berbagai faktor, termasuk jenis tanah, curah hujan, dan ketinggian tempat (Sembiring N, et al., 2023). Selain nilai ekonominya, kopi juga memiliki pengaruh budaya yang kuat, terutama dengan meningkatnya konsumsi di kalangan muda dan menjamurnya kedai kopi di kota-kota besar (Fitriani, 2023).

Salah satu daerah penghasil kopi unggulan di Jawa Timur adalah Wonosalam, Kabupaten Jombang. Wilayah ini dikenal sebagai sentra perkebunan kopi dengan lingkungan dataran tinggi yang mendukung pertumbuhan varietas Arabika, Robusta, dan Excelsa secara optimal (Agastya & Ariyani, 2023). Setiap jenis kopi memiliki karakteristik fisik seperti warna dan tekstur yang berbeda, yang berpengaruh terhadap cita rasa serta nilai jual produk (Amanina & Saraswati, 2022; Maharani & Nuryana, 2023). Oleh karena itu, klasifikasi jenis biji kopi menjadi aspek penting dalam menjaga mutu dan meningkatkan daya saing di pasar nasional maupun internasional.

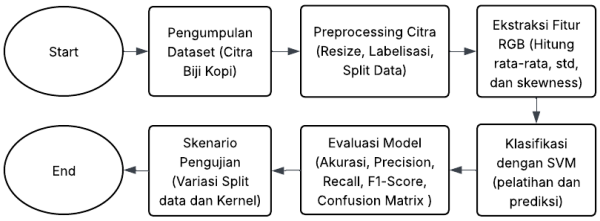
Saat ini, klasifikasi biji kopi umumnya masih dilakukan secara manual melalui pengamatan visual. Metode ini bergantung pada pengalaman tenaga ahli, sehingga bersifat subjektif dan dapat menimbulkan inkonsistensi (Yunita, 2022). Seiring berkembangnya teknologi, pendekatan berbasis citra digital dan machine learning mulai diterapkan untuk mengatasi keterbatasan tersebut (Mardisa et al., 2022). Pemrosesan gambar merupakan bagian penting dari sistem otomatisasi karena dapat mempercepat dan mempermudah proses identifikasi visual, termasuk dalam klasifikasi biji kopi (Karitas Darson, Yandhika Surya Akbar Gumilang, 2023).

Berbagai penelitian sebelumnya telah memanfaatkan pendekatan ini untuk klasifikasi objek visual. Penelitian yang dilakukan Sihombing menerapkan fitur warna HSV dan algoritma SVM untuk mengklasifikasikan buah kopi, dan menunjukkan bahwa SVM mampu menangani data nonlinier secara efektif dengan akurasi tinggi (Sihombing & Buulolo, 2021). Hal ini mendukung pemanfaatan SVM sebagai algoritma utama dalam klasifikasi citra.

Penelitian ini mengusulkan metode klasifikasi jenis biji kopi menggunakan fitur statistik dari kanal warna RGB, yaitu nilai rata-rata, standar deviasi, dan *skewness* dari masing-masing kanal R, G, dan B. Fitur-fitur ini diekstrak langsung dari citra biji kopi tanpa melalui proses konversi ruang warna atau reduksi dimensi, menjadikannya pendekatan yang lebih sederhana dan efisien. Algoritma *Support Vector Machine* (SVM) digunakan sebagai metode klasifikasi karena keunggulannya dalam menangani data multikelas dan kompleksitas nonlinier. Penelitian ini bertujuan menghasilkan model klasifikasi otomatis yang ringan namun akurat, dengan potensi implementasi langsung dalam sistem grading atau seleksi biji kopi di lingkungan industri.

II. METODE PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan utama, mulai dari pengumpulan dataset citra biji kopi, *preprocessing* gambar, ekstraksi fitur statistik dari kanal warna RGB, klasifikasi menggunakan algoritma *Support Vector Machine* (SVM), hingga pengujian dan evaluasi. Rangkaian tahapan ini dirancang seperti pada Gambar 1 untuk menghasilkan sistem klasifikasi otomatis yang sederhana namun akurat dalam membedakan jenis biji kopi berdasarkan informasi warna pada citra digital.



Gambar 1. Alur Penelitian
Sumber: diolah dari data primer

A. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini merupakan data primer berupa citra biji kopi yang diambil langsung dari perkebunan kopi di Wonosalam, Kabupaten Jombang, Jawa Timur. Wilayah ini dipilih karena dikenal sebagai sentra produksi kopi dengan kondisi geografis yang mendukung budidaya berbagai varietas berkualitas tinggi. Penelitian ini memfokuskan pada tiga jenis biji kopi, yaitu Arabika, Robusta, dan Excelsa, dengan masing-masing jenis berjumlah 250 citra, sehingga keseluruhan citra berjumlah 750 citra.

Gambar 2. Sampel Dataset
Sumber: diolah dari data primer

Sebelum pengambilan gambar, seluruh sampel biji kopi disangrai hingga *dark roast* guna menyamakan tingkat kematangan dan menonjolkan perbedaan visual antar varietas seperti pada Gambar 2. Kamera smartphone POCO X6 Pro dijadikan alat pengambilan citra. Citra diambil pada jarak 10 cm dari objek dengan pembesaran 1.6x. Pengambilan dilakukan dalam pencahayaan seragam untuk menjaga konsistensi intensitas warna. Ukuran asli citra adalah 1920 × 2560 piksel dan akan diproses lebih lanjut pada tahap *preprocessing* sebelum diekstrak fitur statistiknya.

B. Preprocessing

Setiap citra biji kopi yang diperoleh dari lapangan diproses melalui tahap *preprocessing* untuk menyamakan ukuran dan formatnya. Seluruh gambar diubah ukurannya menjadi 256 × 256 piksel guna menyederhanakan proses ekstraksi fitur serta mempercepat proses komputasi. Label kelas diambil secara otomatis berdasarkan nama folder tempat citra disimpan.

Selanjutnya, dataset dibagi menjadi dua subset menggunakan metode stratified split, yaitu data untuk pelatihan dan untuk pengujian. Teknik stratifikasi digunakan untuk memastikan distribusi proporsional kelas (Arabika, Robusta, Excelsa) pada kedua subset, sehingga model dapat dilatih dan diuji secara seimbang.

Gambar 3. Contoh Hasil *Preprocessing*
Sumber: diolah dari data primer

Contoh hasil preprocessing citra dapat dilihat pada Gambar 3. Setiap gambar telah diubah ukurannya menjadi 256×256 piksel untuk menyeragamkan dimensi masukan ke proses ekstraksi fitur.

C. Ekstraksi Fitur Statistik RGB

Agar citra dapat diproses oleh sistem komputer, informasi warna harus diubah menjadi nilai numerik melalui model matematis atau sistem koordinat warna tertentu yang dikenal sebagai ruang warna (Syafi'i et al., 2021). Representasi warna menjadi aspek krusial dalam pengolahan citra, karena warna merupakan salah satu fitur utama dalam identifikasi objek berbasis visual.

Ruang warna adalah model matematis yang digunakan untuk menampilkan dan mengatur warna dalam citra digital. Setiap ruang warna memiliki pendekatan berbeda dalam merepresentasikan warna, tergantung pada tujuan penggunaannya. Dalam bidang pengolahan citra, ruang warna digunakan untuk mendukung tugas-

tugas seperti klasifikasi, segmentasi, dan deteksi fitur visual (Syafi'i et al., 2021).

Model warna RGB (*Red, Green, Blue*) merupakan sistem warna paling umum digunakan dalam perangkat digital seperti kamera, sensor citra, dan layar monitor. RGB bekerja dengan prinsip aditif, yaitu dengan menggabungkan tiga warna dasar dalam berbagai intensitas untuk menghasilkan jutaan variasi warna (Putra et al., 2024). Jika ketiga kanal memiliki intensitas maksimum (255, 255, 255), maka warna putih terbentuk, sedangkan kombinasi nol (0, 0, 0) menghasilkan warna hitam. Karena format RGB tersedia langsung dari perangkat pencitraan, model ini sering dijadikan dasar dalam ekstraksi fitur warna, termasuk dalam penelitian klasifikasi biji kopi.

Fitur statistik digunakan untuk merepresentasikan informasi warna dalam citra biji kopi secara numerik. Tiga jenis fitur yang diekstrak dari setiap kanal warna RGB (*Red, Green, Blue*) adalah: rata-rata, standar deviasi, dan *skewness*.

$$X_{mean} = \frac{1}{N}\sum_{i=1}^{N} X_i \quad (1)$$

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \quad (2)$$

$$Skewness = \frac{n}{(n-1)(n-2)}\sum_{i=1}^{N}\left(\frac{x_i - \bar{x}}{\sigma}\right)^3 \quad (3)$$

Di mana X_i adalah nilai piksel ke-i dan Nadalah jumlah piksel. Nilai rata-rata menunjukkan rata-rata intensitas warna suatu kanal dan dihitung menggunakan persamaan 1. Standar deviasi menggambarkan sebaran atau variasi intensitas piksel terhadap rata-rata yang dihitung menggunakan persamaan 2. Sementara itu, *skewness* mengukur kemiringan distribusi intensitas dan dihitung menggunakan persamaan 3.

Proses ekstraksi dilakukan dengan memisahkan setiap kanal warna dari citra, kemudian menghitung nilai statistik pada masing-masing kanal. Total terdapat sembilan fitur numerik untuk setiap citra, yaitu seperti pada Tabel 1. Fitur-fitur ini disusun dalam bentuk tabel (dataframe) dan digunakan sebagai input untuk proses klasifikasi.

Tabel 1. Daftar Fitur Statistik RGB

Nama Fitur	Deskripsi
R_mean	Rata-rata intensitas warna merah pada citra
R_std	Standar deviasi intensitas warna merah (mengukur variasi piksel merah)
R_skew	*Skewness* intensitas warna merah (mengukur kemiringan distribusi piksel)
G_mean	Rata-rata intensitas warna hijau pada citra
G_std	Standar deviasi intensitas warna hijau
G_skew	*Skewness* intensitas warna hijau
B_mean	Rata-rata intensitas warna biru pada citra
B_std	Standar deviasi intensitas warna biru
B_skew	*Skewness* intensitas warna biru

Sumber: diolah dari data primer

D. Klasifikasi dengan SVM

Support Vector Machine (SVM) merupakan algoritma *supervised learning* yang banyak digunakan untuk tugas klasifikasi dan dikenal memiliki performa tinggi dibandingkan algoritma lain (Ramadhani, M. A et al., 2025). Prinsip utama SVM adalah mencari *hyperplane* optimal yang memisahkan data antar kelas dengan margin maksimum (Ria Suci Nurhalizah, Rian Ardianto, 2024). Dalam kasus klasifikasi biner, SVM membagi dua kelompok data menggunakan garis atau bidang pemisah, dengan memperhitungkan jarak terjauh antara titik-titik terdekat dari masing-masing kelas (*support vectors)*. Untuk data yang tidak dapat dipisahkan secara linier, SVM menggunakan fungsi kernel guna memetakan data ke ruang berdimensi lebih tinggi agar dapat dipisahkan secara optimal.

Setelah proses ekstraksi fitur, klasifikasi jenis biji kopi dilakukan menggunakan algoritma *Support Vector Machine* (SVM). Model dilatih menggunakan data hasil ekstraksi fitur statistik RGB, yang terdiri dari sembilan fitur numerik untuk setiap citra. Setiap citra diberi label kelas

berdasarkan jenis biji kopi, yaitu Arabika, Robusta, atau Excelsa.

Pelatihan dilakukan menggunakan data latih, sementara proses prediksi dan evaluasi dilakukan pada data latih dan data uji untuk mengukur kinerja model. Dalam penelitian ini, digunakan beberapa konfigurasi kernel dan pembagian data yang berbeda sebagai bagian dari skenario pengujian. Pendekatan ini memungkinkan evaluasi model secara menyeluruh terhadap stabilitas, akurasi, dan kemampuannya dalam mengklasifikasikan data baru.

Secara matematis, tujuan dari SVM adalah memaksimalkan margin antara dua kelas data. Fungsi *hyperplane* dinyatakan dengan persamaan 4.

$$w^T x + b = 0 \tag{4}$$

dengan *w* adalah vector bobot, *x* adalah vector fitur, dan b adalah bias. Untuk mencapai margin maksimum, SVM menyelesaikan permasalahan optimasi dengan persamaan 5.

$$\min_{w,b} \frac{1}{2}\|w\|^2 \qquad dengan\ syarat$$

$$y_i(w^T x_i + b) \geq 1, \quad \forall i \tag{5}$$

Untuk data yang tidak dapat dipisahkan secara linier, SVM menggunakan fungsi kernel dengan persamaan 6 untuk kernel linier dan persamaan 7 untuk kernel RBF.

$$K(x_i, x_j) = x_i^T x_j \tag{6}$$

$$K(x_i, x_j) = exp\left(-\gamma \|x_i, x_j\|^2\right) \tag{7}$$

Seluruh proses mulai dari memasukkan citra, preprocessing, ekstraksi fitur, pelatihan model, hingga evaluasi dilakukan menggunakan bahasa pemrograman *Python*. Pengolahan citra dilakukan dengan pustaka *OpenCV*, sementara algoritma klasifikasi SVM dan evaluasi metrik diimplementasikan menggunakan pustaka *scikit-learn.*

E. Evaluasi Model

Evaluasi performa model dilakukan menggunakan beberapa metrik klasifikasi, yaitu akurasi, *precision, recall*, dan *F1-score*. Masing-masing dihitung berdasarkan nilai yang diperoleh dari confusion matrix, yang mencerminkan distribusi prediksi model terhadap kelas Arabika, Robusta, dan Excelsa. Adapun persamaan yang digunakan adalah seperti pada persamaan 8 untuk akurasi, persamaan 9 untuk precision, dilanjutkan recall pada persamaan 10, dan F1-score pada persamaan 11.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{11}$$

Akurasi digunakan untuk mengukur keseluruhan proporsi prediksi yang tepat dari total data, sedangkan *precision* menunjukkan tingkat ketepatan prediksi terhadap masing-masing kelas. *Recall* mengukur sejauh mana model berhasil mengenali kelas yang sebenarnya, dan *F1-score* digunakan untuk menilai keseimbangan antara *precision* dan *recall*. Evaluasi dilakukan terhadap data latih dan data uji, dengan hasil yang juga divisualisasikan dalam bentuk *confusion matrix* untuk mengidentifikasi kelas yang rawan tertukar dan konsistensi model dalam proses klasifikasi.

F. Skenario Pengujian

Untuk mengevaluasi performa model klasifikasi secara menyeluruh, penelitian ini menggunakan empat skenario pengujian yang mengombinasikan variasi pada rasio pembagian data latih dan uji serta jenis kernel SVM. Dua rasio data digunakan, yaitu 80:20 dan 70:30, untuk mengamati pengaruh jumlah data latih terhadap kemampuan generalisasi model. Selain itu, dua jenis kernel SVM dibandingkan, yaitu linier kernel yang mengasumsikan pemisahan data secara linier, dan RBF (*Radial Basis Function*) kernel yang digunakan untuk memetakan data nonlinier ke ruang berdimensi lebih tinggi. Kombinasi ini menghasilkan empat skenario yang dirangkum dalam Tabel 2, yang bertujuan untuk mengidentifikasi konfigurasi terbaik dalam hal akurasi dan stabilitas klasifikasi.

Tabel 2. Skenario Pengujian

Skenario	Split Data	Jenis Kernel	Deskripsi
S1	80 : 20	*Linier*	Pengujian *baseline* dengan konfigurasi umum

S2	80 : 20	RBF	Pengujian kernel nonlinier pada rasio standar
S3	70: 30	*Linier*	Evaluasi generalisasi model dengan data latih lebih sedikit
S4	70 : 30	RBF	Pengujian kernel nonlinier dalam kondisi data terbatas

Sumber: diolah dari data primer

III. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil evaluasi model klasifikasi biji kopi yang dibangun menggunakan fitur statistik RGB dan algoritma *Support Vector Machine* (SVM). Evaluasi dilakukan berdasarkan empat skenario pengujian dimana terdapat kombinasi antara variasi rasio data latih dan jenis kernel. Setiap skenario dianalisis menggunakan metrik akurasi, *precision, recall*, dan *F1-score* untuk menilai performa model dalam mengenali tiga kelas biji kopi. Hasil dari tiap skenario dibandingkan untuk mengidentifikasi konfigurasi terbaik, diikuti dengan analisis dan pembahasan terhadap performa dan keterbatasan model.

A. Hasil Evaluasi

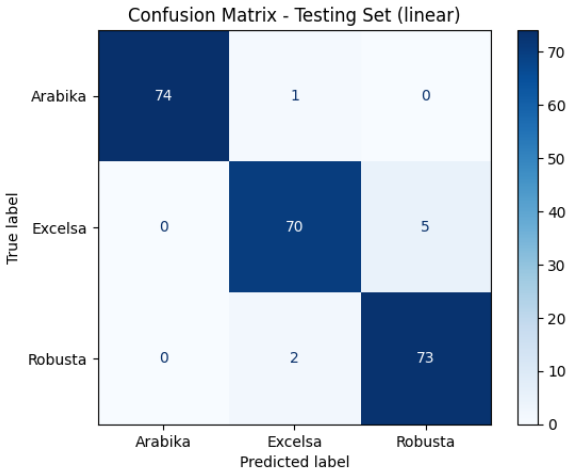
Hasil evaluasi model klasifikasi biji kopi berdasarkan empat skenario pengujian ditampilkan pada Tabel 4. Masing-masing skenario mengombinasikan variasi rasio pembagian data (80:20 dan 70:30) serta jenis kernel yang digunakan pada algoritma SVM (linier dan RBF). Evaluasi dilakukan terhadap data pengujian menggunakan metrik akurasi, *precision, recall*, dan *F1-score* rata-rata.

Tabel 3. Hasil Evaluasi

Skenario	Split Data	Kernel	Akurasi	Precision	Recall	F1-Score
S1	80:20	Linier	0.9600	0.96	0.96	0.96
S2	80:20	RBF	0.7200	0.73	0.72	0.72
S3	70:30	Linier	0.9644	0.96	0.96	0.96
S4	70:30	RBF	0.7333	0.74	0.73	0.74

Sumber: diolah dari data primer

Secara umum, kernel linier menunjukkan performa yang jauh lebih stabil dan unggul dibandingkan kernel RBF di kedua rasio pembagian data. Skenario S3 (*split* 70:30, kernel linier) menghasilkan akurasi tertinggi sebesar 96,44%, dengan nilai *F1-score* rata-rata yang konsisten di angka 0.96. Sebaliknya, kernel RBF menghasilkan performa yang lebih rendah, khususnya pada kelas Arabika dan Excelsa, dengan nilai *precision* dan *recall* di bawah 0.70 pada skenario S2 dan S4. Hal ini menunjukkan bahwa untuk dataset biji kopi berbasis fitur statistik RGB, kernel linier lebih mampu memisahkan kelas secara efektif dibandingkan RBF.



Gambar 4. Evaluasi Metrik Skenario S3
Sumber: diolah dari data primer

Untuk memberikan gambaran lebih rinci terhadap hasil klasifikasi, Gambar 4 menampilkan confusion matrix dari skenario terbaik (S3) dengan kernel linear. Sebagian besar data diklasifikasikan secara benar, bisa dilihat

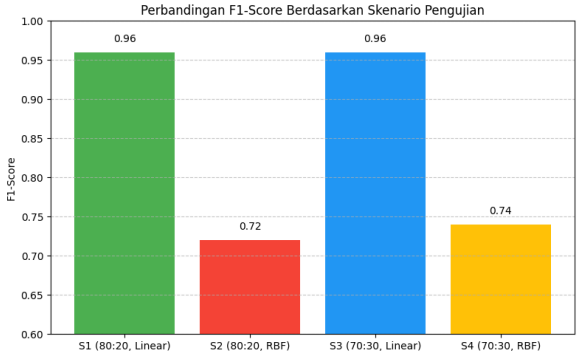
dari angka yang mendominasi pada diagonal utama.

B. Analisis Perbandingan

Berdasarkan hasil evaluasi pada Tabel 3, terlihat bahwa pemilihan kernel SVM memiliki dampak signifikan terhadap performa model. Kernel linier secara konsisten memberikan hasil yang lebih tinggi dibandingkan kernel RBF pada kedua rasio pembagian data. Skenario S1 dan S3 (kernel linier) menunjukkan nilai *F1-score* rata-rata sebesar 0.96, jauh di atas skenario S2 dan S4 (kernel RBF) yang hanya mencapai 0.72 dan 0.74. Hal ini mengindikasikan bahwa data fitur statistik RGB cenderung memiliki pola distribusi yang lebih sesuai untuk dipisahkan secara linier.

Dari sisi rasio data latih dan uji, perbedaan antara rasio 80:20 dan 70:30 tidak menunjukkan dampak signifikan terhadap performa model dengan kernel linier. Baik S1 maupun S3 mencatat nilai akurasi dan F1-score yang sangat mirip, yang menandakan bahwa model tetap mampu menjaga kestabilan meskipun data latih sedikit berkurang. Namun, pada kernel RBF, performa pada rasio 70:30 (S4) sedikit lebih baik daripada 80:20 (S2), meskipun perbedaannya tidak terlalu besar.

Untuk memperjelas perbandingan antar skenario, Gambar 3 menunjukkan visualisasi nilai F1-score rata-rata dari masing-masing skenario. Grafik ini menegaskan bahwa linier kernel secara konsisten unggul dan stabil dalam klasifikasi tiga jenis biji kopi berdasarkan fitur warna RGB.



Gambar 5. Grafik Perbandingan F1-Score
Sumber: diolah dari data primer

Gambar 3 merupakan grafik batang yang menunjukkan perbandingan nilai *F1-score* dari keempat skenario pengujian. Visualisasi ini memperjelas bahwa kernel linier (S1 dan S3) secara konsisten lebih unggul daripada kernel RBF (S2 dan S4), dengan performa stabil meskipun data latih berkurang.

Secara keseluruhan, analisis ini menunjukkan bahwa model paling optimal diperoleh pada skenario S3, yaitu dengan rasio 70:30 dan kernel linier, karena memberikan akurasi tinggi, keseimbangan antar kelas, dan ketahanan performa terhadap variasi data latih.

C. Pembahasan Hasil

Hasil evaluasi dan analisis menunjukkan bahwa kombinasi fitur statistik RGB dan algoritma SVM mampu menghasilkan performa klasifikasi yang tinggi, khususnya saat menggunakan kernel linier. Tanpa perlu transformasi warna atau ekstraksi fitur kompleks, pendekatan ini mampu mencapai nilai *F1-score* hingga 0.96 pada dua skenario berbeda (S1 dan S3). Ini menegaskan bahwa informasi warna dasar dalam format RGB cukup representatif untuk membedakan jenis biji kopi Arabika, Robusta, dan Excelsa.

Kinerja rendah pada kernel RBF (S2 dan S4) mengindikasikan bahwa data fitur yang digunakan tidak memiliki pola nonlinier yang kuat, sehingga penggunaan kernel nonlinier justru menurunkan akurasi. Selain itu, meskipun jumlah data latih dikurangi pada rasio 70:30, performa model dengan kernel linier tetap tinggi, menandakan ketahanan model terhadap variasi ukuran data latih.

Secara praktis, pendekatan ini dapat diimplementasikan dalam sistem klasifikasi ringan di sektor industri kopi, misalnya pada proses grading otomatis. Namun, terdapat keterbatasan, seperti sensitivitas terhadap pencahayaan dan kesamaan visual antara kelas Excelsa dan Robusta yang dapat menyebabkan kesalahan klasifikasi. Pengembangan lebih lanjut dapat diarahkan pada penambahan fitur tekstur, penggunaan augmentasi data, atau integrasi model deep learning untuk meningkatkan ketahanan terhadap variasi kondisi nyata.

IV. KESIMPULAN

Penelitian ini mengusulkan metode klasifikasi jenis biji kopi berbasis citra digital menggunakan fitur statistik RGB dan algoritma *Support Vector Machine* (SVM). Fitur yang digunakan terdiri dari nilai rata-rata, standar deviasi, dan *skewness* dari kanal warna merah (R), hijau (G), dan biru (B), yang diolah sebagai input numerik dalam proses klasifikasi.

Hasil pengujian dari empat skenario

menunjukkan bahwa model dengan kernel linier memberikan performa terbaik. Pada skenario S3 (rasio data latih:uji 70:30), model menghasilkan akurasi 96,44% dan nilai F1-score rata-rata 0,96, sementara skenario S1 (80:20, linier) mencatat akurasi 96,00% dengan nilai F1-score 0,96. Sebaliknya, kernel RBF menunjukkan performa lebih rendah. Skenario S2 (80:20, RBF) menghasilkan akurasi 72,00% dan nilai F1-score rata-rata 0,72, sementara S4 (70:30, RBF) hanya sedikit lebih baik dengan akurasi 73,33% dan nilai F1-score 0,74.

Penelitian ini membuktikan bahwa fitur warna RGB dapat digunakan secara efektif dan efisien untuk membedakan jenis biji kopi Arabika, Robusta, dan Excelsa menggunakan model klasifikasi sederhana. Model juga terbukti stabil terhadap pengurangan data latih, tanpa penurunan performa yang signifikan. Dengan akurasi tinggi dan kompleksitas rendah, metode ini memiliki potensi implementasi pada sistem klasifikasi biji kopi otomatis di industri.

DAFTAR PUSTAKA

Agastya, A. A. R., & Ariyani, A. H. M. (2023). Strategi pengembangan kawasan agropolitan komoditas kopi di Kecamatan Wonosalam Kabupaten Jombang. *Agriscience*, *3*(3), 732–751. https://doi.org/10.21107/agriscience.v3i3.16757

Amanina, N. N., & Saraswati, G. W. (2022). Classification of Arabica Coffee Green Beans Using Digital Image Processing Using the K-Nearest Neighbor Method. *Journal of Applied Intelligent System*, *7*(2), 111–119. https://doi.org/10.33633/jais.v7i2.6449

Badan Pusat Statistik. (2024). *Statistik Produksi Kopi Indonesia 2024*, *Volume 8*, https://doi.org/5504006

Fitriani, D. (2023). Eksistensi budaya minum kopi dari era kolonial hingga era modern. *Daya Nasional*, *1*(3), 114–119. https://doi.org/10.26418/jdn.v1i3.70369

Karitas Darson, Yandhika Surya Akbar Gumilang, A. R. (2023). Citra Identifikasi Warna dengan Metode Nilai Hue Secara Realtime Berbasis Web Camera dan Raspberry Pi. *SNESTIK Seminar Nasional Teknik Elektro, Sistem Informasi, Dan Teknik Informatika*, 6. https://doi.org/10.31284

Maharani, H. S., & Nuryana, I. K. D. (2023). Role Of Gray Level Co-Occurrence Matrix for Convolution Neural Network Transfer Learning in Coffee Bean Classification. *(Journal of Informatics and Computer Science)*, *05*(1), 1–6. https://doi.org/10.26740/jinacs.v5n01.p1-6

Mardisa, R., Siregar, K., & Nasution, I. S. (2022). Klasifikasi Kualitas Fisik Kopi Beras Arabika menggunakan Pengolahan citra dengan Metode K-Nearest Neighbor (K-NN). *Jurnal Ilmiah Mahasiswa Pertanian*, *7*(2), 514–522. https://doi.org/10.17969/jimfp.v7i2.19896

Ramadhani, M. A., & Khumaidi, A. (2025). Implementasi Algoritma Support Vector Machine (SVM) Untuk Diagnosis Kesehatan Manusia Berbasis Web Application. *JURNAL NERS*, *9(1)*, 896–902. https://doi.org/10.31004/jn.v9i1.31481

Putra, R. P., Jumadi, J., & Lianda, D. (2024). Pengolahan Citra Digital Untuk Mengidentifikasi Tingkat Kematangan Buah Kelapa Sawit Berdasarkan Warna Rgb Dan Hsv Dengan Menggunakan Metode Self Organizing Map (SOM). *Jurnal Media Infotama*, *20*(1), 341149. https://doi.org/10.37676/jmi.v20i1.5354

Nurhalizah, R. S., Ardianto, R., & Purwono, P. (2024). Analisis Supervised dan Unsupervised Learning pada Machine Learning. *Jurnal Ilmu Komputer Dan Informatika (JIKI)*, *4* no. *1*, 61–72. https://doi.org/10.54082/jiki.168

Sembiring, Anita C & Tampubolon, J & Purnasari, N. (2023). Peningkatan Pengetahuan Petani Kopi Karo dalamPengolahan Pasca Panen Buah Kopi di DesaBuluhnaman Sumatera Utara. *Jurnal Mitra Prima*, *5*(2). http://jurnal.unprimdn.ac.id/index.php/mitra_prima/article/view/4270

Sihombing, H. A., & Buulolo, I. C. (2021). Pengenalan Buah Kopi Berdasarkan Parameter Warna Menggunakan Algoritma Backpropagation Dan Algoritma Support Vector Machine (Svm). In *Seminastika* (Vol. 3, Issue 1, pp. 26–32). https://doi.org/10.47002/seminastika.v3i1.234

Syafi'i, A. M., Ahadi, M. F., Rasyid, M. I., Adhinata, F. D., & Junaidi, A. (2021). Mendeteksi Kematangan Pada Buah Mangga Garifta Merah Dengan

Transformasi Ruang Warna HSI. *Journal of Applied Informatics and Computing*, *5*(2), 117–121. https://doi.org/10.30871/jaic.v5i2.3217

Yunita, R. D. (2022). Implementasi Metode Linear Discriminan Analysis Untuk Klasifikasi Biji Kopi. *Jurnal Teknologi Informatika Dan Komputer*, *8*(1), 27–39. https://doi.org/10.37012/jtik.v8i1.664

Zahra, N. Z., Farhanatussaidah, S., Afifah, N. N., Muthoharoh, L., Satria, A., & Manullang, M. C. T. (2026). Benchmarking Logistic Regression, SVM, Naive Bayes, and IndoBERT Fine-Tuning for Sentiment Analysis on Indonesian Product Reviews. *ArXiv Preprint ArXiv:2605.03439*.

