

COLOR HISTOGRAM DAN SUPPORT VECTORE MACHINE UNTUK MENGLASIFIKASIKAN BIJI KOPI BERDASARKAN TINGKAT PEMANGGANGAN

Siti Khotimatul Wildah¹, Abdul Latif², Sarifah Agustiani³, Ali Mustopa⁴, Suharyanto⁵

^{1,3}Teknologi Komputer D3, Universitas Bina Sarana Informatika

²Sistem Informasi Kampus Kota Pontianak, Universitas Bina Sarana Informatika

⁴Teknik Informatika Kampus Kota Pontianak, Universitas Bina Sarana Informatika

⁵Sistem Informasi, Universitas Bina Sarana Informatika

Jalan Kramat Raya No.98, Senen, Jakarta Pusat, Indonesia, 10450, Indonesia

siti.ska@bsi.ac.id

ABSTRAK

Pemilihan biji kopi berdasarkan tingkat pemanggangan menjadi faktor kunci dalam menentukan rasa kopi yang dihasilkan. Untuk mengklasifikasikan biji kopi berdasarkan tingkat pemanggangan, biasanya penilaian dilakukan berdasarkan warna dan bentuk fisik biji kopi. Namun, variabilitas lingkungan dan kondisi individu dapat memengaruhi ketepatan penilaian tersebut. Oleh karena itu, diperlukannya deteksi otomatis mengenai klasifikasi biji kopi tersebut. Metode *computer vision* dengan teknik *image classification* dapat memberikan solusi terkait permasalahan tersebut. Penelitian ini memperkenalkan metode *image processing* menggunakan *color histogram* sebagai teknik ekstraksi fitur untuk merepresentasikan ciri-ciri visual biji kopi dalam berbagai tingkatan pemanggangan. Selanjutnya, *Support Vector Machine* (SVM) digunakan sebagai algoritma klasifikasi yang mampu mengelompokkan biji kopi berdasarkan distribusi warna yang dihasilkan dari *histogram* tersebut. Pendekatan ini bertujuan untuk meningkatkan ketepatan klasifikasi biji kopi berdasarkan tingkat pemanggangan, dengan mengurangi pengaruh faktor subjektif individu serta meningkatkan konsistensi dalam proses pengklasifikasian. Hasil eksperimen menunjukkan bahwa penggabungan *color histogram* dengan SVM mampu memberikan metode yang andal dan akurat dalam mengklasifikasikan biji kopi berdasarkan tingkat pemanggangan dengan nilai akurasi sebesar 98,95%.

Kata kunci: Biji Kopi, SVM, Color Histogram, Image Classification

1. PENDAHULUAN

Kopi merupakan salah satu komoditas yang paling umum dikonsumsi di seluruh dunia, dan memiliki berbagai karakteristik yang membuatnya menjadi minuman yang unik [1]. Biji kopi digunakan dalam penyajian minuman kopi, yang merupakan minuman kedua paling banyak dikonsumsi di dunia [2].

Kualitas kopi yang unggul dapat dikenali melalui aroma yang khas dan cita rasa yang unik. Parameter-parameter untuk menilai kualitas kopi yang unggul diukur selama proses pemanggangan biji kopi [3].

Pemilihan tingkat pemanggangan biji kopi merupakan salah satu tahap kritis dalam industri kopi karena berpengaruh langsung pada cita rasa, aroma, dan karakteristik kopi yang dihasilkan.

Proses pemanggangan biji kopi adalah tahap dimana biji kopi hijau dipanaskan pada suhu tinggi tertentu, menyebabkan reaksi kimia seperti hidrolisis, polimerisasi, dan pirolisis. Perubahan warna yang terbentuk selama pemanggangan disebabkan oleh melanoidin, yang merupakan hasil dari reaksi Maillard dan karamelisasi. Perubahan warna ini memiliki pengaruh yang signifikan terhadap kualitas rasa yang dihasilkan oleh biji kopi [4].

Oleh karena itu, penting bagi produsen kopi untuk dapat mengklasifikasikan biji kopi berdasarkan tingkat pemanggangan dengan akurasi yang tinggi.

Penentuan harga dan potensi ekspor kopi di negara-negara produsen sangat bergantung pada penilaian kualitas kopi. Saat ini, evaluasi fisik dan kualitas rasa menjadi metode utama untuk menilai kualitas kopi [5].

Secara tradisional, penilaian tingkat pemanggangan biji kopi dilakukan secara visual oleh manusia berdasarkan warna dan bentuk fisik biji kopi. Namun, metode ini rentan terhadap subjektivitas dan variasi individu dalam menilai tingkat pemanggangan. Selain itu, faktor-faktor lingkungan seperti pencahayaan dan kelelahan juga dapat memengaruhi ketepatan penilaian.

Berdasarkan hal tersebut, diperlukan adanya pengembangan metode yang lebih objektif dan konsisten dalam mengklasifikasikan tingkat pemanggangan biji kopi. Penggunaan teknologi *computer vision* dapat digunakan untuk klasifikasi otomatis yang berfungsi untuk mengenali objek atau informasi secara mandiri tanpa intervensi manusia. Teknologi *computer vision*, yang melibatkan *image classification* dan *machine learning* [6], memungkinkan sistem untuk melakukan identifikasi biji kopi tersebut.

Prosesnya mirip dengan memberikan mata kepada komputer untuk memahami dan mengenali objek di sekitarnya, dengan memanfaatkan berbagai algoritma dan teknik yang telah dipelajari dari data yang telah ada. Hal ini memungkinkan sistem untuk

memproses informasi visual, mengklasifikasikan objek, dan membuat keputusan berdasarkan informasi yang diperoleh dari citra yang dianalisis secara otomatis.

Dalam proses pengolahan citra, terdapat serangkaian langkah yang digunakan untuk mengidentifikasi biji kopi dengan menggunakan pola atau karakteristik khusus dari biji tersebut. Proses ini dikenal sebagai ekstraksi fitur, yang bertujuan untuk mengurangi kompleksitas data citra agar mempermudah analisis dan pengolahan citra lebih lanjut [7].

Ekstraksi fitur merupakan langkah penting dalam *image processing* dan *computer vision*. Pada tahap ini, citra diubah menjadi bentuk angka atau huruf, yang kemudian digunakan untuk mengenali pola dalam proses klasifikasi [8]. Jumlah fitur yang diambil mempengaruhi seberapa baik citra diklasifikasikan karena fitur yang baik akan membuat mesin yang belajar lebih baik dalam memahami pola dalam data yang digunakan.

Color histogram merupakan salah satu metode fitur ekstraksi yang dapat mengekstraksi informasi mengenai warna dari suatu citra. Teknik ini digunakan untuk menganalisis sebaran warna pada objek yang terlihat dengan cara menghitung jumlah piksel dari setiap bagian rentang warna. Hal ini dapat membantu dalam memperoleh nilai yang mencerminkan distribusi warna dalam citra tersebut [9].

Sedangkan, untuk mengklasifikasikan biji kopi berdasarkan representasi visual yang dihasilkan dari *color histogram*, metode klasifikasi yang digunakan yaitu *support vector machine* (SVM). SVM merupakan salah satu algoritma *machine learning* yang efektif dalam klasifikasi data. Dalam metode SVM, dataset diwakili sebagai titik-titik dalam ruang multidimensi, di mana setiap titik data memiliki koordinat tertentu sesuai dengan nilainya. Algoritma ini kemudian mencari *hyperplane* (bidang pemisah) untuk memisahkan dengan baik antara dua kelas data tersebut [10].

Oleh karena itu, penelitian ini bertujuan untuk menerapkan *color histogram* dan SVM guna mengembangkan metode yang lebih objektif dan akurat dalam mengklasifikasikan biji kopi berdasarkan tingkat pemanggangan. Sehingga, diharapkan dapat memberikan solusi yang lebih handal dalam industri kopi untuk meningkatkan konsistensi dan akurasi dalam pemilihan biji kopi.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian yang dilakukan oleh Dede Herman Suryana dan Wahyu Kusuma Raharja mengenai penerapan kecerdasan buatan untuk mengklasifikasikan tingkat kematangan biji kopi selama proses pemanggangan. pada penelitian ini metode yang digunakan adalah *convolution neural network* (CNN) dimana hasil yang diperoleh yaitu nilai

akurasi sebesar 94,79% untuk data pelatihan dan sebesar 85,83% untuk data uji [3].

Penelitian yang dilakukan oleh Ade Febri Syah Putra, Jayanta, dan Mayanda Mega Santoni untuk menentukan tingkat kematangan biji kopi berdasarkan hasil pemanggangan dengan memanfaatkan metode identifikasi warna RGB dan klasifikasi jarak minimum. Data latih dan uji telah dibagi menggunakan validasi silang k-fold dengan k=5, dan hasil rata-rata akurasi yang dihasilkan sebesar 89.63% [11].

Selanjutnya penelitian dilakukan oleh T H Nasution dan U Andayani mengenai pengenalan tingkat pemanggangan biji kopi yang sudah dipanggang menggunakan *image processing* dan *neural network*. Hasil yang diperoleh pada penelitian ini menunjukkan akurasi sebesar 97,5% [12].

Penelitian yang dilakukan oleh Jun Noel C. Sarino, Melvin M. Bayas, Edwin R. Arboleda, Emeline C. Guevarra dan Rhowel M. Delloso mengenai klasifikasi derajat pemanggangan biji kopi menggunakan *image processing* dan *neural network*. Hasil penelitian menunjukkan bahwa metode yang diusulkan mampu mengidentifikasi tingkat pemanggangan biji kopi dengan akurasi sebesar 97,22% [13].

2.2. Computer vision

Computer vision adalah cabang ilmu komputer yang bertujuan untuk memberikan kemampuan pada mesin atau komputer untuk memahami, menganalisis, dan menafsirkan visual dari dunia nyata. Bidang ini berusaha meniru kemampuan visual manusia dengan menggunakan teknologi dan algoritma untuk mengolah gambar atau video, memungkinkan komputer untuk mengenali objek, pola, dan konteks dalam gambar atau video yang diprosesnya [14].

2.3. Color Histogram

Color histogram adalah salah satu teknik yang sering digunakan untuk mengekstraksi ciri-ciri warna dari sebuah citra karena mampu mewakili informasi tentang warna dari berbagai sudut pandang yang berbeda [15].

2.4. Support Vector Machine

SVM (*Support Vector Machine*) merupakan sebuah metode dalam *data mining* yang termasuk dalam kategori *supervised learning*. Fokus utama dari metode ini adalah menemukan *hyperplane* terbaik dari sekumpulan *hyperplane* yang ada. SVM berupaya mencari fungsi pemisah yang optimal untuk memisahkan dua set data dari dua kelas yang berbeda [16].

2.5. Evaluasi Model

Evaluasi merupakan langkah akhir dari proses Knowledge Discovery in Database (KDD). Pada tahap ini, bertujuan untuk mengungkapkan pengetahuan yang telah ditemukan secara lebih mudah dimengerti dengan menggunakan penyajian yang menarik.

Metode pengujian yang digunakan pada penelitian ini diantaranya mencari nilai *precision*, *recall*, *f-measure*, *accuracy* dan *confusion matrix* [17].

2.5.1. Precision

Precision merupakan rasio antara jumlah data positif yang diprediksi dengan benar terhadap total data yang diprediksi sebagai positif, yang mencerminkan perbandingan antara data yang benar dan salah yang diprediksi sebagai positif [18].

$$Precision = \frac{TP}{TP + PF} \quad (1)$$

2.5.2. Recall

Recall merupakan proporsi dari data positif yang telah diidentifikasi secara benar dibandingkan dengan total data yang sebenarnya positif dan yang telah diidentifikasi dengan benar bersama dengan data positif yang tidak teridentifikasi (salah negatif) [18].

$$All Recall = \frac{TP}{TP + FN} \quad (2)$$

2.5.3. Accuracy

Akurasi digunakan untuk mengevaluasi keakuratan dari setiap kernel [18]. Hal ini merujuk pada tingkat kebenaran atau ketepatan dari hasil yang diperoleh dibandingkan dengan nilai sebenarnya dalam suatu proses pengukuran atau prediksi.

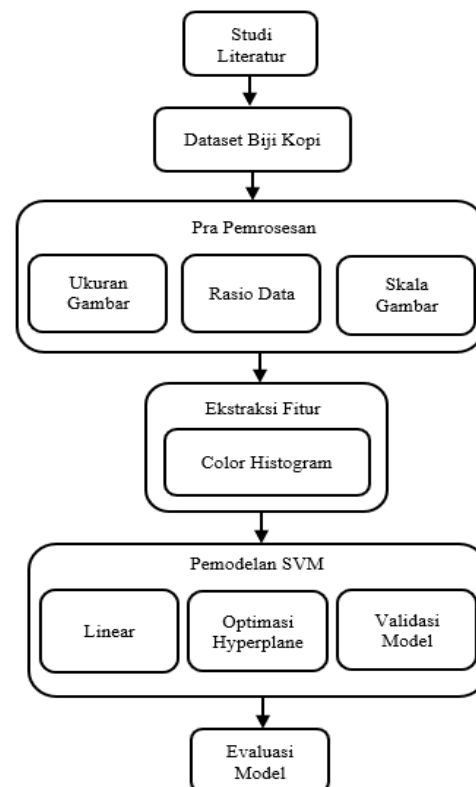
$$Accuracy = \frac{\sum TP}{Total Data} \times 100\% \quad (3)$$

2.5.4. Confusion Matrix

Confusion matrix, atau disebut juga *error matrix*, merupakan metode yang digunakan untuk mengukur keakuratan dalam klasifikasi atau *supervised learning*. Dalam metode ini, terdapat empat kemungkinan hasil antara prediksi dan realita yaitu *true positive* (TP) untuk data yang diprediksi dengan benar sebagai positif, *true negative* (TN) untuk data negatif yang diprediksi dengan benar, *false positive* (FP) merupakan kesalahan dalam mendeteksi data negatif sebagai positif, serta *false negative* (FN) menunjukkan kesalahan dalam mendeteksi data positif sebagai negatif [19].

3. METODE PENELITIAN

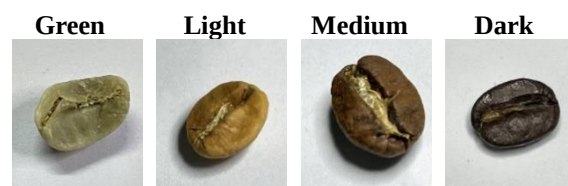
Pada proses metodologi penelitian untuk menentukan pemilihan biji kopi berdasarkan tingkat kematangan, terdapat serangkaian langkah yang dilakukan mulai dari studi literatur, penggunaan dataset, pra pemrosesan citra, ekstraksi fitur, klasifikasi, sampai dengan evaluasi penelitian. Adapun langkah penelitian dapat dilihat pada Gambar 1.



Gambar 1. Metode Penelitian

Berdasarkan metode penelitian pada gambar 1, langkah awal yang dilakukan pada penelitian ini adalah melakukan studi literatur yang mendukung proses penelitian, mengidentifikasi, dan menganalisis masalah yang mendukung penelitian mengenai biji kopi tersebut.

Tahap kedua yang dilakukan yaitu penggunaan dataset biji kopi yang akan diujikan. Dataset ini memiliki 4 jenis biji kopi yang menggambarkan tingkat kematangan dari masing-masing biji kopi yaitu green, light, medium dan dark.



Gambar 2. Dataset Biji Kopi

Langkah selanjutnya yaitu tahap pra-pemrosesan dilakukan untuk menyiapkan data sebelum dimasukkan ke dalam model yang akan digunakan. Langkah ini mencakup perubahan ukuran citra untuk memungkinkan proses pemodelan yang lebih cepat. Langkah berikutnya adalah menentukan pembagian antara data latih dan data uji yang akan digunakan dalam proses pemodelan, dengan menggunakan perbandingan 80% data latih dan 20% data uji. Selanjutnya, dilakukan penyesuaian skala pada citra.

Setelah data dipersiapkan maka tahap selanjutnya adalah melakukan ekstraksi fitur. Pada tahapan ini

metode yang akan digunakan adalah metode ekstraksi fitur *color histogram*. *Color Histogram* merupakan metode yang digunakan untuk menggambarkan variasi warna dalam suatu gambar dengan menghitung jumlah kemunculan setiap warna. Teknik ini memungkinkan pembentukan representasi warna dengan cepat, namun memerlukan proses kuantifikasi warna di mana komponen warna dibagi ke dalam interval tertentu [20].

Hasil dari ekstraksi fitur kemudian masuk ke tahap *modelling* atau klasifikasi. Klasifikasi yang akan dilakukan pada penelitian ini menggunakan metode *support vectore machine* dimana metode ini dimulai dari penentuan kernel atau parameter yang sesuai dengan data. Pada penelitian ini kernel yang digunakan yaitu linear. Kemudian dilakukan *hyperplane* terbaik yang memiliki jarak terbesar dari data-data kelas yang berbeda. Dimana *hyperplane* ini merupakan garis atau bidang yang memisahkan antara data-data dari setiap kelas atau kategori. Sedangkan margin mengacu pada jarak antara *hyperplane* dengan titik data terdekat dari masing-masing kelas [21].

Setelah pelatihan, model dievaluasi menggunakan data validasi (data yang belum dilihat oleh model selama pelatihan) untuk melihat kinerja dari model yang diterapkan pada dataset tersebut. Kemudian dilakukan evaluasi model untuk melihat nilai akurasi yang dihasilkan dari model yang diujikan. Untuk menggambarkan kinerja dari model klasifikasi, dibuat tabel *confusion matrix* dan *classification report*.

4. HASIL DAN PEMBAHASAN

Citra biji kopi yang digunakan dalam penelitian ini memiliki 4 kelas biji kopi dengan berbagai macam tingkat proses pemanggangan biji kopi yang dilakukan di "Bona Coffee," yang terletak di JJ Mall Jatujak. Terdapat empat tingkatan pemanggangan yang dilakukan pada biji kopi tersebut. Biji kopi hijau (kelas *green*) atau yang belum dipanggang berasal dari Laos Typica Bolaven (arabica) sebanyak 400 citra. Biji kopi Laos Typica Bolaven dipanggang secara ringan atau kelas *light* (arabica) sebanyak 400 citra. Kemudian, Doi Chaang (arabica) dipanggang secara sedang (kelas *medium*) sebanyak 400 citra, sementara Brazil Cerrado dipanggang secara gelap atau kelas *dark* (arabica) sebanyak 400 citra [22] total keseluruhan citra yang digunakan sebanyak 1600 citra.

Pengujian dilakukan dimulai dari tahapan pra pemrosesan, ekstraksi fitur, klasifikasi dan evaluasi model. Persentase data uji dan data latih yang digunakan pada penelitian ini menggunakan rasio 8:2 sehingga data yang digunakan sebanyak 320 citra untuk data uji dan 1280 citra untuk data latih. Pengujian dilakukan dengan menggunakan bantuan tools *jupyter notebook* dengan bahasa pemrograman *python*.

Setelah tahap persiapan data, masuk kedalam tahap ekstraksi fitur menggunakan *color histogram*. Proses ekstraksi fitur ini melibatkan pengukuran

distribusi warna dalam citra. Beberapa ciri ekstraksi yang diperoleh dari *color histogram* meliputi:

- 1) Distribusi warna untuk menunjukkan seberapa sering warna tertentu muncul dalam citra. Hal ini memberikan informasi tentang keberadaan dan persebaran warna dalam gambar.
- 2) Frekuensi kemunculan warna untuk mendeskripsikan seberapa banyak dan seberapa sering setiap warna muncul dalam citra. Hal ini bertujuan untuk dapat memberikan gambaran tentang dominasi warna tertentu dalam citra.
- 3) Bentuk kurva histogram yang bertujuan untuk menunjukkan pola distribusi warna.
- 4) Informasi tentang kontras warna yang bertujuan untuk dapat memberikan petunjuk tentang perbedaan warna yang signifikan dalam citra, yang dapat berpengaruh pada deteksi objek atau fitur tertentu.

Setelah proses ekstraksi fitur dilakukan masuk ke tahap pemodelan atau klasifikasi dengan menggunakan algoritma *support vectore machine*. Pemilihan metode ekstraksi fitur *color histogram* dan *support vectore machine* dilakukan berdasarkan hasil komparasi dari pengujian yang dilakukan. Berikut merupakan pengujian yang dilakukan dengan menerapkan metode ekstraksi fitur *color histogram* dengan pengujian beberapa metode klasifikasi *machine learning* yang dapat dilihat pada tabel 1.

Tabel 1. Perbandingan Ekstraksi Fitur *Color Histogram*

No	Metode Klasifikasi	Nilai Akurasi
1	Naïve Bayes (NB)	89,89%
2	Classification and Regression Trees (CART)	95,52%
3	Linear Discriminant Analysis (LDA)	96,56%
4	K-Nearest Neighbors (KNN)	97,29%
5	Logistic Regression (LR)	97,70%
6	Random Forest (RF)	98,64%
7	Support Vectore Machine (SVM)	98,95%

Tabel 1. Menunjukkan penerapan metode ekstraksi fitur *color histogram* menghasilkan nilai akurasi yang sangat baik terutama dengan kombinasi metode klasifikasi *support vectore machine* (SVM) yaitu sebesar 98,95%. Hasil tersebut menunjukkan tingkat keberhasilan dalam mengklasifikasikan citra biji kopi berdasarkan fitur yang diekstraksi dari *color histogram*. Selain itu dilakukan penerapan metode ekstraksi fitur lainnya dengan menggunakan metode ekstraksi fitur *haralick*, yang dapat dilihat pada tabel 2.

Tabel 2. Perbandingan Ekstraksi Fitur *Haralick*

No	Metode Klasifikasi	Nilai Akurasi
1	Naïve Bayes (NB)	75,31%
2	Classification and Regression Trees (CART)	75,52%
3	Logistic Regression (LR)	76,97%
4	K-Nearest Neighbors (KNN)	78,64%

No	Metode Klasifikasi	Nilai Akurasi
5	Linear Discriminant Analysis (LDA)	78,95%
6	Support Vector Machine (SVM)	80,93%
7	Random Forest (RF)	82,08%

Pada tabel 2 penerapan metode ekstraksi fitur *haralick* pada dataset biji kopi menghasilkan nilai akurasi yang cukup baik namun tidak lebih baik dari penerapan ekstraksi fitur *color histogram*. Nilai akurasi yang dihasilkan berada pada rentang 70% hingga 80%. Selain itu dilakukan perbandingan lainnya dengan menggunakan metode ekstraksi fitur *hu-moment* yang dapat dilihat pada tabel 3.

Tabel 3. Perbandingan Ekstraksi Fitur *Hu-Moment*

No	Metode Klasifikasi	Nilai Akurasi
1	Naïve Bayes (NB)	32,91%
2	Classification and Regression Trees (CART)	34,47%
3	Support Vector Machine (SVM)	38,43%
4	Logistic Regression (LR)	38,75%
5	Linear Discriminant Analysis (LDA)	40,41%
6	K-Nearest Neighbors (KNN)	40,93%
7	Random Forest (RF)	43,95%

Pada tabel 3 penerapan metode ekstraksi fitur *hu-moment* pada dataset biji kopi menghasilkan nilai akurasi yang kurang baik dibandingkan dengan penerapan ekstraksi fitur *color histogram* dan *haralick* dimana nilai akurasi yang dihasilkan pada rentang 30 hingga 40%. Selain itu, perbandingan tambahan dilakukan dengan menerapkan metode ekstraksi fitur *zernike*, yang dapat dilihat pada tabel 4.

Tabel 4. Perbandingan Ekstraksi Fitur *Zernike*

No	Metode Klasifikasi	Nilai Akurasi
1	Classification and Regression Trees (CART)	29,58%
2	K-Nearest Neighbors (KNN)	31,97%
3	Naïve Bayes (NB)	35,00%
4	Random Forest (RF)	35,00%
5	Support Vector Machine (SVM)	35,20%
6	Logistic Regression (LR)	35,31%
7	Linear Discriminant Analysis (LDA)	35,31%

Pada tabel 4 penerapan metode ekstraksi fitur *zernike* pada dataset biji kopi menghasilkan nilai akurasi yang kurang baik dibandingkan dengan penerapan ekstraksi fitur lainnya. Perbandingan penerapan metode ekstraksi fitur pada penelitian ini setelah di evaluasi menghasilkan bahwa penerapan metode *color histogram* memberikan nilai akurasi yang paling baik dengan di kombinasikan metode klasifikasi *support vectore machine* serta pembagian rasio data sebesar 8:2 dengan hasil sebesar 98,95%.

Pengujian lain dilakukan dengan pembagian rasio data latih dan data uji 7:3 dan 9:1 dimana hasil dari pengujian tersebut menghasilkan nilai akurasi yang baik akan tetapi tidak sebaik hasil yang diberikan oleh pembagian rasio data 8:2. Pengujian dilakukan dengan skema penerapan metode ekstraksi fitur *color histogram* dan *support vectore machine* dimana hasil pengujian tersebut dapat dilihat pada tabel 5.

Tabel 5. Perbandingan Rasio Data

No	Rasio Data	Nilai Akurasi
1	70% Data Latih dan 30% Data Uji	98,88%
2	80% Data Latih dan 20% Data Uji	98,95%
3	90% Data Latih dan 10% Data Uji	98,88%

Sehingga dapat ditentukan setelah beberapa pengujian yang dilakukan mulai dari penerapan beberapa metode ekstraksi fitur, metode klasifikasi *machine learning* dan rasio data latih dan data uji, maka dapat ditentukan kombinasi metode ekstraksi fitur *color histogram* dan *support vectore machine* serta rasio data 8:2 menghasilkan nilai akurasi atau performa yang lebih baik dibandingkan dengan lainnya. Hasil dari kombinasi tersebut kemudian dievaluasi menggunakan *confusion matrix* yang merupakan sebuah tabel yang menunjukkan kinerja pengklasifikasi dengan membandingkan jumlah data label yang diklasifikasikan dengan label aslinya. Tabel *confusion matrix* hasil dari metode yang digunakan pada penelitian ini dapat dilihat pada Tabel 6.

Tabel 6. *Confusion Matrix*

Predicted Values	Green	42	0	0	0
	Light	0	41	0	0
	Medium	0	0	55	0
	Dark	1	0	1	52
		Green	Light	Medium	Dark
Actual Values					

Tabel 6 menunjukkan hasil jumlah prediksi yang benar pada setiap kelas. Pada tabel tersebut menunjukkan bahwa terdapat 42 citra yang diprediksi dengan benar untuk kelas *green*, 41 citra untuk kelas *light*, 55 citra terprediksi sebagai kelas *medium* dan 52 citra terprediksi sebagai kelas *dark*. Akan tetapi terdapat 1 citra terprediksi *dark* akan tetapi sebenarnya adalah *green* dan 1 citra terprediksi *dark* akan tetapi sebenarnya adalah *medium*. Secara keseluruhan data yang terprediksi benar lebih banyak dibandingkan data yang terprediksi salah.

Berdasarkan nilai pada *confusion matrix* yang telah diperoleh, dapat dihitung nilai akurasi, *recall*, dan presisi dapat dihitung. Di bawah ini merupakan proses perhitungan manual untuk mendapatkan nilai *accuracy*, *recall* dan *precision*.

1) Menghitung Nilai Accuracy

$$\begin{aligned} \text{Accuracy} &= \frac{\sum TP}{\text{Total Data}} \times 100\% \\ &= \frac{42+41+55+52}{192} \times 100\% \\ &= 98,95\% \end{aligned} \quad (4)$$

2) Menghitung Nilai Recall

$$\begin{aligned} \text{All Recall} &= \sum \frac{TP}{TP + FN} \\ &= \frac{TP(0)}{TP(0) + FN(0)} + \frac{TP(1)}{TP(1) + FN(1)} + \frac{TP(2)}{TP(2) + FN(2)} + \frac{TP(3)}{TP(3) + FN(3)} \\ &= \frac{42}{42+0} + \frac{41}{41+0} + \frac{55}{55+0} + \frac{52}{52+2} \\ &= 1 + 1 + 1 + 0,96 \\ &= 3,96 \\ \text{Recall} &= \frac{R(0) + R(1) + R(2) + R(3)}{\text{Jumlah Kelas}} \times 100\% \\ &= \frac{1 + 1 + 1 + 0,96}{4} \times 100\% \\ &= 99,07\% \end{aligned} \quad (5)$$

3) Menghitung Nilai Precision

$$\begin{aligned} \text{All Precision} &= \sum \frac{TP}{TP + FP} \\ &= \frac{TP(0)}{TP(0) + FP(0)} + \frac{TP(1)}{TP(1) + FP(1)} + \frac{TP(2)}{TP(2) + FP(2)} + \frac{TP(3)}{TP(3) + FP(3)} \\ &= \frac{42}{42+1} + \frac{41}{41+0} + \frac{55}{55+1} + \frac{52}{52+0} \\ &= 0,97 + 1 + 0,98 + 1 \\ &= 3,95 \\ \text{Precision} &= \frac{P(0) + P(1) + P(2) + P(3)}{\text{Jumlah Kelas}} \times 100\% \\ &= \frac{0,97 + 1 + 0,98 + 1}{4} \times 100\% \\ &= \frac{0,97 + 1 + 0,98 + 1}{4} \times 100\% \\ &= 98,75\% \end{aligned} \quad (6)$$

Berdasarkan perhitungan tersebut terlihat nilai *accuracy* sebesar 98,95%, *recall* sebesar 99,07% dan *precision* sebesar 98,75% hasil tersebut menunjukkan tingkat keberhasilan dari penerapan metode yang digunakan. Sehingga pemilihan metode yang diujikan pada penelitian ini sesuai dengan kebutuhan data karena memberikan performa yang baik.

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang dilakukan dengan menerapkan metode ekstraksi fitur *color histogram* dan *support vector machine* (SVM) untuk mengklasifikasikan tingkat pemanggangan biji kopi, diperoleh kesimpulan bahwa kedua teknik tersebut memberikan hasil yang sangat baik dengan tingkat akurasi mencapai 98,95%. Hasil ini menunjukkan efektivitas teknologi ini dalam mengklasifikasikan biji kopi berdasarkan ciri visualnya terkait tingkat pemanggangan. *Color histogram* membantu dalam mengidentifikasi ciri visual biji kopi, sedangkan SVM mampu mengelompokkan biji kopi dengan baik. Penerapan teknologi ini dapat meningkatkan konsistensi dan obyektivitas dalam pemilihan biji kopi industri, mengurangi pengaruh subjektivitas manusia. Namun untuk pengembangan lebih lanjut, disarankan untuk memperluas dan memperkaya *dataset* dengan citra yang lebih beragam serta mencoba menerapkan metode *deep learning*.

DAFTAR PUSTAKA

- [1] Y. T. Chang, M. C. Hsueh, S. P. Hung, J. M. Lu, J. H. Peng, and S. F. Chen, "Prediction of specialty coffee flavors based on near-infrared spectra using machine- and deep-learning methods," *J. Sci. Food Agric.*, vol. 101, no. 11, pp. 4705–4714, 2021, doi: 10.1002/jsfa.11116.
- [2] K. Przybył *et al.*, "Application of Machine Learning to Assess the Quality of Food Products—Case Study: Coffee Bean," *Appl. Sci.*, vol. 13, no. 19, 2023, doi: 10.3390/app131910786.
- [3] D. Herman Suryana and W. Kusuma Raharja, "Applying Artificial Intelligence to Classify the Maturity Level of Coffee Beans During Roasting," *Int. J. Eng. Sci. Inf. Technol.*, vol. 3, no. 2, pp. 97–105, 2023, doi: 10.52088/ijesty.v3i2.461.
- [4] D. S. Leme, S. A. da Silva, B. H. G. Barbosa, F. M. Borém, and R. G. F. A. Pereira, "Recognition of coffee roasting degree using a computer vision system," *Comput. Electron. Agric.*, vol. 156, no. November 2018, pp. 312–317, 2019, doi: 10.1016/j.compag.2018.11.029.
- [5] V. Sari, F. Firdausi, and Y. Azhar, "Perbandingan Prediksi Kualitas Kopi Arabika dengan Menggunakan Algoritma SGD, Random Forest dan Naive Bayes," *Edumatic J. Pendidik. Inform.*, vol. 4, no. 2, pp. 1–9, 2020, doi: 10.29408/edumatic.v4i2.2202.
- [6] S. K. Wildah, A. Latif, A. Mustopa, and M. S.

- Maulana, "Klasifikasi Penyakit Daun Kopi Menggunakan Kombinasi Haralick , Color Histogram dan Random Forest Classification of Coffee Leaf Disease Using a Combination of Haralick , Color Histogram and Random Forest," vol. 11, no. 1, 2023, doi: 10.26418/justin.v11i1.60985.
- [7] E. H. Rachmawanto and H. P. Hadi, "Optimasi Ekstraksi Fitur Pada Knn Dalam Klasifikasi Penyakit Daun Jagung," *Dinamik*, vol. 26, no. 2, pp. 58–67, 2021, doi: 10.35315/dinamik.v26i2.8673.
- [8] M. S. Wibawa and K. D. P. Novianti, "Reduksi Fitur untuk Optimalisasi Klasifikasi Tumor Payudara Berdasarkan Data Citra FNA," *Konf. Nas. Sist. Inform.*, pp. 73–78, 2017.
- [9] S. Agustiani, Y. T. Arifin, A. Junaidi, S. K. Wildah, and A. Mustopa, "Klasifikasi Penyakit Daun Padi menggunakan Random Forest dan Color Histogram," vol. 10, no. 1, 2022.
- [10] J. Kusuma, A. Jinan, M. Zulkarnain Lubis, Rosnelly, and R. Rosnelly, "Komparasi Algoritma Support Vector Machine Dan Naive Bayes Pada Klasifikasi Ras Kucing," *J. Ilmu Komput. dan Teknol. Infrmasi*, vol. 14, no. 1, pp. 8–12, 2022.
- [11] M. Putra, Ade Febri Syah., Jayanta., Santoni, M., "Penentuan Level Kematangan Kopi Berdasarkan Hasil Roasting Menggunakan Metode Deteksi RGB dan Klasifikasi Minimum Distance," *Molecules*, vol. 2, no. 1, pp. 1–12, 2020, [Online]. Available: <http://clik.dva.gov.au/rehabilitation-library/1-introduction-rehabilitation%0Ahttp://www.scrip.org/journal/doi.aspx?DOI=10.4236/as.2017.81005%0Ahttp://www.scrip.org/journal/PaperDownload.aspx?DOI=10.4236/as.2012.34066%0Ahttp://dx.doi.org/10.1016/j.pbi.201>
- [12] T. H. Nasution and U. Andayani, "Recognition of Roasted Coffee Bean Levels using Image Processing and Neural Network," *J. Phys. Conf. Ser.*, vol. 755, no. 1, 2016, doi: 10.1088/1742-6596/755/1/011001.
- [13] J. N. C. Sarino, M. M. Bayas, E. R. Arboleda, E. C. Guevarra, and R. M. Dellosa, "Classification of coffee bean degree of roast using image processing and neural network," *Int. J. Sci. Technol. Res.*, vol. 8, no. 10, pp. 3231–3233, 2019.
- [14] T. A. Dompeipen, S. R. U. . Sompie, and M. E. . Najoan, "Computer Vision Implementation for Detection and Counting the Number of Humans," *J. Tek. Inform.*, vol. 16, no. 1, pp. 65–67, 2021.
- [15] M. G. T. Akbar and H. Leidiyana, "Klasifikasi pada Citra Bunga dengan Ekstraksi Fitur Color Histogram," *Format J. Ilm. Tek. Inform.*, vol. 12, no. 1, p. 54, 2023, doi: 10.22441/format.2023.v12.i1.007.
- [16] K. R. Kavitha, K. V. Aiswarya Rajan, and A. Pillai, "An Improved Feature Selection and Classification of Gene Expression Profile using SVM," *2019 2nd Int. Conf. Intell. Comput. Instrum. Control Technol. ICICICT 2019*, pp. 1033–1037, 2019, doi: 10.1109/ICICICT46008.2019.8993358.
- [17] Oryza Habibie Rahman, Gunawan Abdillah, and Agus Komarudin, "Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 1, pp. 17–23, 2021, doi: 10.29207/resti.v5i1.2700.
- [18] N. Arifin, U. Enri, and N. Sulistiyowati, "Penerapan Algoritma Support Vector Machine (SVM) dengan TF-IDF N-Gram untuk Text Classification," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 6, no. 2, p. 129, 2021, doi: 10.30998/string.v6i2.10133.
- [19] M. D. Purbolaksono, M. Irvan Tantowi, A. Imam Hidayat, and A. Adiwijaya, "Perbandingan Support Vector Machine dan Modified Balanced Random Forest dalam Deteksi Pasien Penyakit Diabetes," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 393–399, 2021, doi: 10.29207/resti.v5i2.3008.
- [20] C. P. Beauty and Y. A. Sari, "Temu Kembali Citra Makanan Menggunakan Color Histogram Dan Local Binary Pattern," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 6, pp. 5514–5520, 2019.
- [21] H. N. Irmada and Ria Astriratma, "Klasifikasi Jenis Pantun Dengan Metode Support Vector Machines (SVM)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 5, pp. 915–922, 2020, doi: 10.29207/resti.v4i5.2313.
- [22] S. Ontoum, T. Khemanantakul, P. Sroison, T. Triyason, and B. Watanapa, "Coffee Roast Intelligence," 2022, [Online]. Available: <http://arxiv.org/abs/2206.01841>