

PENERAPAN LIGHTGBM MENGGUNAKAN EKSTRAKSI FITUR RUANG WARNA HSV UNTUK KLASIFIKASI REMPAH RIMPANG

Achmad Fahmi Al Hafidz, Eva Yulia Puspaningrum, Afina Lina Nurlaili

Informatika, Universitas Pembangunan Nasional Veteran Jawa Timur
Jalan Raya Rungkut Madya, Gunung Anyar, Surabaya, Jawa Timur, Indonesia
evapuspaningrum.if@upnjatim.ac.id

ABSTRAK

Rempah rimpang memiliki peran penting dalam dunia kuliner dan industri pangan, namun identifikasi jenis rempah secara manual seringkali menemui hambatan karena kemiripan visual antar jenis. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi otomatis yang akurat dan efisien menggunakan metode Light Gradient Boosting Machine (LGBM) dengan fitur utama berupa representasi warna HSV (Hue, Saturation, Value). Dataset yang digunakan terdiri dari 2.250 citra rempah yang telah melalui proses preprocessing seperti resizing ke ukuran 512x512 piksel, penghapusan latar belakang, normalisasi, dan augmentasi. Hasil evaluasi menunjukkan bahwa penggunaan fitur HSV dengan rasio pembagian data 70:30 memberikan performa yang cukup optimal, dengan akurasi mencapai 82,67%. Temuan ini menegaskan bahwa fitur HSV bisa diterapkan dalam membedakan jenis rempah rimpang dan dapat menjadi pendekatan yang andal untuk membangun sistem klasifikasi otomatis, terutama dalam mendukung proses identifikasi cepat di industri pangan dan sektor UMKM.

Kata kunci : *Klasifikasi, Rempah Rimpang, Decision Tree, Light Gradient Boosting Machine (LGBM), Ruang Warna HSV*

1. PENDAHULUAN

Rempah rimpang seperti jahe, kunyit, kencur, lengkuas, dan temulawak punya peran penting dalam industri kuliner dan kesehatan karena kandungan senyawa bioaktifnya. Contohnya, jahe merah yang mengandung *gingerol* dan *shogaol* terbukti efektif dalam mengurangi nyeri pada penderita *rheumatoid arthritis* [1]. Sekarang ini, makin banyak orang beralih ke pengobatan herbal karena lebih alami dan terjangkau [2]. Tapi sayangnya, proses identifikasi rempah secara visual masih sering keliru karena kemiripan warna, bentuk, dan tekstur antar jenis rimpang [3].

Penelitian ini berfokus pada klasifikasi otomatis jenis rimpang berdasarkan karakteristik visualnya. Identifikasi rimpang seperti jahe, kunyit, kencur, lengkuas, dan temulawak masih banyak dilakukan secara manual, yang dinilai kurang efisien dan rawan kesalahan akibat kemiripan visual antar jenis [4]. Sistem klasifikasi otomatis dapat meningkatkan akurasi identifikasi, mengurangi kesalahan distribusi, serta menjaga kualitas produk. Studi sebelumnya menunjukkan bahwa pendekatan *machine learning*, seperti yang dilakukan oleh Agustian dan Maimunah, efektif dalam mengklasifikasikan rimpang berdasarkan warna [5]. Selain mempercepat proses identifikasi, sistem ini juga bermanfaat untuk pelatihan karyawan baru tanpa harus mengandalkan pengalaman manual [4].

Beberapa penelitian sebelumnya mendukung pemanfaatan model klasifikasi berbasis citra dalam berbagai *domain*. Muhathir dkk. menunjukkan bahwa fitur bentuk HOG efektif dalam klasifikasi rempah, dengan model *XGB Classifier* mencatat akurasi tertinggi sebesar 80,6% [6]. Sementara itu, Jayasooriya dan Arachchi (2022) membuktikan bahwa

ruang warna HSV sangat efektif dalam diagnosis penyakit daun padi menggunakan algoritma LightGBM, karena HSV lebih mencerminkan persepsi warna manusia secara alami [7]. Berlandaskan temuan sebelumnya, penelitian ini mengembangkan sistem klasifikasi otomatis untuk rempah rimpang menggunakan algoritma LightGBM dengan fokus utama pada ekstraksi fitur HSV. Evaluasi dilakukan dengan berbagai rasio pembagian data, dan rasio 70:30 digunakan sebagai konfigurasi utama untuk menguji kemampuan model dalam mengenali karakteristik visual rempah berdasarkan representasi warna HSV. Pendekatan ini dipilih karena HSV dianggap mampu merepresentasikan informasi warna secara lebih konsisten dan relevan terhadap persepsi visual manusia.

2. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Beberapa penelitian sebelumnya telah menunjukkan efektivitas model klasifikasi berbasis citra, khususnya yang menggunakan *algoritma Light Gradient Boosting Machine* (LightGBM) dan metode ekstraksi fitur visual. Muhathir menerapkan metode *Histogram of Oriented Gradient* (HOG) untuk klasifikasi rempah dan mencatat akurasi tertinggi sebesar 80,6% dengan XGBoost, sementara LightGBM memperoleh 77,3% karena lebih optimal untuk dataset besar [6]. Jayasooriya dan Arachchi menunjukkan bahwa fitur warna HSV efektif dalam diagnosis penyakit daun padi menggunakan LightGBM, dengan akurasi mencapai 81%, menegaskan bahwa HSV merepresentasikan warna sesuai persepsi manusia [7]. Goswami juga berhasil menerapkan LightGBM dengan kombinasi fitur warna dari berbagai ruang warna (termasuk HSV) dan tekstur

untuk klasifikasi kanker mulut, mencapai akurasi 99,25% pada klasifikasi biner [8]. Berdasarkan temuan tersebut, penelitian ini mengembangkan sistem klasifikasi rempah rimpang berbasis LightGBM dengan fokus pada fitur HSV. Evaluasi dilakukan pada rasio data 70:30 untuk mengkaji efektivitas fitur HSV dalam membedakan jenis rempah berdasarkan karakteristik visual warna yang lebih sesuai dengan persepsi manusia.

2.2 Rimpang

Rimpang merupakan bagian tanaman yang memiliki nilai penting dalam kehidupan sehari-hari, terutama di bidang kuliner dan kesehatan, dengan contoh seperti jahe, kunyit, kencur, lengkuas, dan temulawak yang sering dimanfaatkan karena kandungan bioaktifnya. Namun, identifikasi visual terhadap rimpang menjadi tantangan dalam konteks pengolahan citra digital, mengingat kemiripan warna, bentuk, dan teksturnya. Penelitian-penelitian sebelumnya menunjukkan bahwa fitur visual seperti ruang warna (HSV) dapat meningkatkan akurasi klasifikasi objek alami, termasuk rimpang [2][9]. Penerapan algoritma seperti *Light Gradient Boosting Machine* (LightGBM) juga terbukti efisien dalam menangani data berdimensi tinggi dan memproses fitur secara optimal dalam berbagai konteks klasifikasi [10]. Lebih lanjut, berbagai jenis rimpang memiliki karakteristik visual unik yang dapat dimanfaatkan dalam sistem klasifikasi berbasis citra digital, seperti warna oranye cerah pada temulawak, permukaan kasar pada kunyit, hingga bentuk membulat pada kencur [11]. Dengan dukungan teknologi pengolahan citra digital, klasifikasi rimpang dapat dilakukan secara otomatis, cepat, dan akurat, sekaligus berpotensi terintegrasi dengan perangkat *mobile* atau sistem *IoT* untuk mendukung proses industri herbal dan pertanian modern secara lebih efisien dan berkelanjutan.

2.3. Color Space HSV

Hue, *Saturation*, dan *Value* (HSV) merupakan model warna yang efektif dalam pengolahan citra karena lebih menyerupai cara manusia memandang warna, sehingga cocok untuk tugas klasifikasi dan pengenalan objek. Diperkenalkan oleh A. R. Smith pada tahun 1978, HSV terdiri dari tiga komponen: *Hue* (H) yang menunjukkan warna dasar seperti merah atau hijau, *Saturation* (S) yang mengukur tingkat kejenuhan warna, dan *Value* (V) yang merepresentasikan kecerahan seperti yang ditunjukkan pada persamaan 1 sampai 3. Kombinasi ketiganya memberikan representasi warna yang lebih intuitif dan membantu membedakan objek dengan lebih akurat. Efektivitas HSV telah dibuktikan dalam berbagai studi, seperti identifikasi kematangan daun teh oleh Auliasari dkk. [12].

$$H = \cos^{-1} \frac{\frac{1}{2}(2R - G - B)}{\sqrt{(R - G)^2 - (R - G)(G - B)}} \quad (1)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (2)$$

$$V = \max(R, G, B) \quad (3)$$

H : Menghitung warna dasar gambar RGB
S : Mengukur sejauh mana warna jenuh.
V : Menentukan tingkat kecerahan warna.

Color moments merupakan metode ekstraksi fitur yang penting dalam klasifikasi citra, karena mampu merepresentasikan karakteristik warna secara statistik melalui tiga metrik utama: *mean* (warna dominan), standar *deviasi* (keragaman warna), dan *skewness* (asimetri distribusi warna). Ketiga metrik ini memberikan informasi yang cukup untuk membedakan objek berdasarkan perbedaan warna secara efektif, sehingga sering digunakan dalam berbagai aplikasi seperti identifikasi objek dan pengenalan pola. Penerapannya pada ruang warna HSV telah terbukti efektif, sebagaimana ditunjukkan dalam penelitian Jayasooriya dkk. yang mengembangkan model klasifikasi penyakit daun padi [7]. Secara matematis, ketiga atribut ini dihitung dari nilai-nilai warna pada piksel citra berukuran $M \times N$ untuk memberikan gambaran menyeluruh terhadap distribusi warna dalam sebuah gambar. rumus bisa dilihat pada persamaan 4 sampai 6

$$\mu_i = \frac{\sum_{j=1}^{M.N} X_{ij}}{M.N} \quad (4)$$

$$\sigma_i = \sqrt{\frac{1}{M.N} \sum_{j=1}^{M.N} (X_{ij} - \mu_i)^2} \quad (5)$$

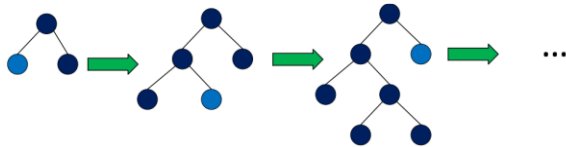
$$S_i = \sqrt[3]{\frac{1}{M.N} \sum_{j=1}^{M.N} (X_{ij} - \mu_i)^2} \quad (6)$$

μ_i : Rata-rata nilai gambar i
 X_{ij} : Nilai warna HSV baris ke-i dan kolom ke-j
 σ_i : Standar deviasi gambar i
 S_i : *Skewness* gambar i

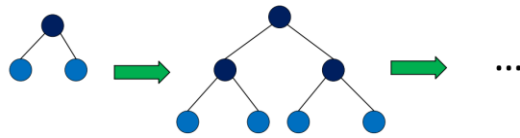
2.4. Light Gradient Boosting Machine

Light Gradient Boosting Machine (LightGBM) adalah algoritma klasifikasi berbasis pohon keputusan yang efisien dan dirancang untuk menangani data dalam skala besar dengan kecepatan tinggi dan konsumsi memori yang rendah. Algoritma ini menggunakan teknik *histogram-based decision tree learning* untuk mempercepat proses pelatihan dan mengurangi kompleksitas komputasi. Dalam penerapannya, LightGBM memiliki sejumlah parameter penting seperti *learning rate*, *max depth*, dan *num_leaves* yang dapat disesuaikan untuk mengoptimalkan kinerja model. Dalam konteks klasifikasi rempah rimpang, LGBM terbukti mampu memanfaatkan representasi fitur warna seperti HSV secara efektif untuk membedakan pola visual antar kelas. Kemampuannya dalam menangani berbagai

jenis data fitur dan menghindari *overfitting* membuatnya menjadi pilihan yang andal untuk klasifikasi citra dalam penelitian ini.



Gambar 1. Leaf-wise tree growth in LGBM



Gambar 2. Level-wise tree in other boosting machine[13]

Gambar ini membandingkan dua metode pertumbuhan pohon dalam algoritma boosting. Gambar 1 menunjukkan metode *leaf-wise* yang digunakan oleh LightGBM, di mana pembagian dilakukan pada daun dengan potensi error terbesar, menghasilkan pohon yang lebih dalam dan efisien. Sebaliknya, Gambar 2 memperlihatkan metode *level-wise* yang membagi pohon secara merata di setiap level, lebih stabil namun cenderung membutuhkan waktu komputasi lebih besar dibanding *leaf-wise*, terutama pada data besar dan kompleks. Persamaan model LGBM bisa dilihat pada persamaan 7.

$$F_m(x) = F_{m-1}(x) + \eta \cdot h_m(x) \quad (7)$$

$F_m(x)$: Prediksi model pada iterasi ke- m
 F_{m-1} : Prediksi model pada iterasi sebelumnya
 η : *learning rate*
 $h_m(x)$: Fungsi prediksi pada iterasi m

Log-loss digunakan sebagai fungsi kerugian untuk klasifikasi biner, dengan nilai lebih rendah menandakan prediksi lebih akurat. Rumusnya ada pada Persamaan 8.

$$L(y, \hat{y}) = - \sum_{i=1}^N \left[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \right] \quad (8)$$

y_i : Label kelas yang sebenarnya (0 atau 1)
 $\hat{y}_i$: Probabilitas prediksi kelas 1

Dalam klasifikasi biner, fungsi *sigmoid* digunakan untuk mengubah output model menjadi probabilitas antara 0 dan 1. Fungsi ini ditunjukkan pada Persamaan 9.

$$\hat{y} = \frac{1}{1 + e^{-F_m(x)}} \quad (9)$$

$\hat{y}$: Probabilitas prediksi kelas 1
 F_m : Output model

Probabilitas ini digunakan untuk menentukan kelas prediksi. Jika $\hat{y}$ lebih besar dari 0.5, model akan memprediksi kelas 1; jika $\hat{y}$ kurang dari 0.5, model memprediksi kelas 0.

2.5. Confusion Matrix

Confusion matrix digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan hasil prediksi terhadap data asli. Matriks ini menunjukkan jumlah *true positive*, *true negative*, *false positive*, dan *false negative*, yang membantu mengidentifikasi akurasi serta potensi kesalahan model. Tabel 1 menyajikan penjelasan mengenai *confusion matrix*.

Tabel 1. Confusion Matrix

Confusion Matrix		Nilai Aktual	
		Positif	Negatif
Nilai Prediksi	Positif	TP	FP
	Negatif	FN	TN

Confusion matrix terdiri dari empat komponen utama: *True Positive* (TP), ketika model benar mengklasifikasikan gambar sebagai rempah yang sesuai; *True Negative* (TN), saat model benar mengenali gambar bukan sebagai rempah yang diprediksi; *False Positive* (FP), ketika model salah mengklasifikasikan gambar rempah lain sebagai target; dan *False Negative* (FN), saat model gagal mengenali gambar rempah yang benar. Keempat komponen ini menjadi dasar dalam menghitung metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*, yang membantu mengukur dan meningkatkan kinerja model klasifikasi.

3. METODE PENELITIAN

3.1. Pengumpulan Data

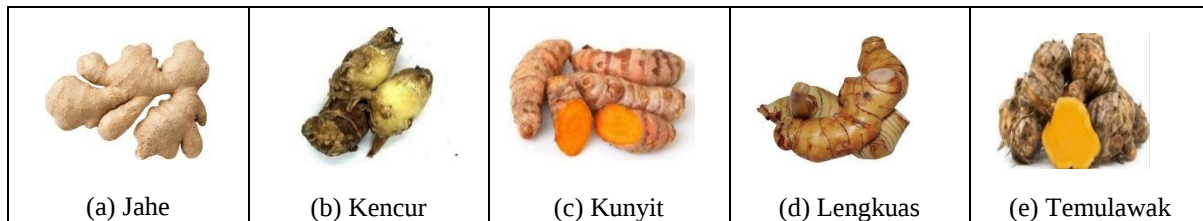


Gambar 3. Pengambilan Dataset Primer

Dataset yang digunakan terdiri dari 2.750 gambar lima jenis rempah rimpang: jahe, kencur, kunyit, lengkuas, dan temulawak. Data sekunder diperoleh

dari dua dataset *Kaggle* yang dilengkapi dengan hasil *web scraping*, masing-masing kelas berjumlah 250 gambar. Sementara itu, data primer berjumlah 200 gambar per kelas, diambil secara manual menggunakan kamera *smartphone* 12 MP dalam *studio box mini* dengan variasi sudut dan posisi objek. Seluruh gambar dikumpulkan dan disiapkan dalam satu folder yang ada pada gambar 5 untuk tahap penelitian selanjutnya.

Gambar 3 menunjukkan komposisi pencahayaan dari lampu *LED studio box mini* yang difokuskan tepat di atas objek untuk meminimalkan cahaya berlebih dan menghasilkan citra rempah yang tajam dalam hal warna, tekstur, dan bentuk. Hal ini mendukung akurasi pelatihan dan pengujian model klasifikasi dengan memberikan kualitas gambar yang konsisten. Label (a) hingga (e) menampilkan contoh sampel jahe, kencur, kunyit, lengkuas, dan temulawak. Seluruh data ini akan dimodifikasi sesuai kebutuhan, termasuk data primer rempah rimpang yang ditampilkan berikutnya.



Gambar 4. Sampel Dataset Rempah Rimpang

Label (a) memiliki warna coklat kekuningan dan permukaan berkerut; label (b) berwarna putih keabuan dengan tekstur halus; label (c) berwarna kuning cerah dengan bentuk pendek melengkung; label (d) berwarna coklat kemerahan dengan permukaan sedikit bersisik; dan label (e) berwarna coklat gelap dengan bentuk besar, membulat, dan bertekstur kasar. Seluruh data ini digunakan untuk membangun model klasifikasi, dengan *preprocessing* seperti *resizing* dan normalisasi sebelum tahap ekstraksi fitur dan pelatihan.

3.2. Preprocessing Data

Tahap awal penelitian dimulai dengan pengumpulan citra rempah rimpang dalam format RGB dari dataset *Kaggle*, *web scraping*, dan data primer, lalu disusun dalam satu folder. *Preprocessing* dilakukan untuk memastikan kualitas data yang optimal, dimulai dengan *background removal* agar fokus hanya pada objek rempah. Selanjutnya, dilakukan *resizing* ke ukuran 512×512 piksel untuk standarisasi dan efisiensi komputasi, lalu normalisasi dengan mengubah nilai piksel ke rentang 0–1 agar model lebih stabil dalam memproses data. Setelah itu, dataset dibagi (*split data*) menjadi data latih dan uji menggunakan rasio 80:20, 70:30, dan 60:40 untuk mengevaluasi kinerja model secara objektif. Terakhir, dilakukan augmentasi dengan teknik seperti rotasi, *flipping*, penskalaan, pemotongan, perubahan pencahayaan, dan *shearing* untuk memperkaya variasi data dan meningkatkan akurasi serta ketahanan model klasifikasi terhadap perubahan kondisi citra.

3.3. Ekstraksi Fitur

Ekstraksi fitur pada penelitian ini difokuskan pada ruang warna HSV untuk memperoleh ciri visual dari citra rempah rimpang. HSV dipilih karena mampu merepresentasikan persepsi warna manusia dengan lebih baik dibandingkan model warna lainnya, serta

lebih stabil terhadap perubahan pencahayaan. Proses ini melibatkan konversi citra dari RGB ke HSV, kemudian dihitung sembilan fitur berdasarkan color moments, yaitu mean, standar deviasi, dan skewness dari masing-masing komponen hue, saturation, dan value. Fitur-fitur ini digunakan sebagai representasi visual utama dalam pelatihan dan pengujian model klasifikasi, karena terbukti memberikan hasil paling optimal dalam membedakan jenis rempah pada rasio data 70:30 yang digunakan dalam penelitian ini.

3.4. Pembangunan Model LGBM

Setelah ekstraksi fitur HSV, tahap berikutnya adalah membangun model klasifikasi menggunakan *Light Gradient Boosting Machine* (LGBM). Proses ini melibatkan pelatihan dengan data *training* dan pengujian dengan data *testing*, dengan tujuan mengklasifikasikan jenis rempah rimpang secara akurat. Evaluasi model dilakukan menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Beberapa skenario pengujian, seperti variasi *learning rate* dan jumlah iterasi, diterapkan untuk melihat pengaruh hyperparameter terhadap kinerja model.

Tabel 2. Hyperparameter LGBM

Hyperparameter	Nilai
<i>Learning Rate</i>	0.01, 0.05
<i>Num Leaves</i>	10, 20, 50
<i>Max Depth</i>	6, 10, 13
<i>n_estimators</i>	300, 500

Hyperparameter model LightGBM pada tabel 2 ini disetel menggunakan metode *Grid Search* untuk menemukan kombinasi terbaik dalam menghasilkan model klasifikasi yang optimal. *Learning Rate* (0.01 dan 0.05) mengontrol kecepatan pembelajaran model, sementara *Num Leaves* (10, 20, 50) menentukan jumlah maksimum daun dalam setiap pohon yang memengaruhi kompleksitas model. *Max Depth* (6, 10,

13) digunakan untuk membatasi kedalaman pohon agar menghindari *overfitting*, dan *n_estimators* (300, 500) menunjukkan jumlah pohon yang dilatih. Dengan melakukan pencarian kombinasi dari setiap nilai *hyperparameter* melalui *Grid Search*, model dapat dievaluasi secara menyeluruh untuk mendapatkan performa terbaik berdasarkan metrik seperti akurasi, *precision*, *recall*, dan *F1-score*.

3.5. Skenario Pengujian

Penelitian ini melibatkan 2 skenario pengujian yang difokuskan pada evaluasi performa ekstraksi fitur HSV dengan rasio data 70:30 dalam klasifikasi rempah rimpang menggunakan algoritma LightGBM. Kedua skenario ini menggunakan data dari sumber primer saja, dengan kombinasi parameter model yang berbeda untuk memahami pengaruhnya terhadap akurasi klasifikasi. *Hyperparameter* yang diuji meliputi *learning rate* (0.01 dan 0.05) serta jumlah iterasi (300 dan 500), dengan tujuan menemukan konfigurasi optimal yang mampu menyeimbangkan antara akurasi dan efisiensi pelatihan. Meskipun bersumber dari satu jenis data, variasi konfigurasi pelatihan ini memberikan gambaran menyeluruh mengenai performa fitur HSV dan sensitivitas model terhadap perubahan parameter. Hasil pengujian memperlihatkan bahwa fitur HSV mampu memberikan akurasi tinggi secara konsisten, menjadikannya salah satu fitur yang paling efektif untuk klasifikasi citra rempah rimpang.

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Dataset dalam penelitian ini terdiri dari dua sumber utama, yaitu dataset sekunder dan primer. Dataset sekunder dikumpulkan dari platform *Kaggle* seperti “*Indonesian Spices Dataset*” dan “*Empon Dataset*”, serta dilengkapi menggunakan teknik *web scraping* melalui *Selenium WebDriver* pada *Bing Images* hingga mencapai 250 gambar per kelas. Proses *scraping* dilakukan secara otomatis dan disimpan berdasarkan kategori rempah. Sementara itu, dataset primer diperoleh dengan pengambilan gambar langsung menggunakan kamera *smartphone* dan *studio box mini* seperti pada gambar 4, sebanyak 200 gambar per kelas. Meskipun beberapa rempah dalam kondisi kotor untuk meniru situasi nyata seperti di pasar, langkah ini bertujuan meningkatkan keberagaman citra dan ketahanan model. Seluruh data, baik primer maupun sekunder, disimpan secara sistematis dan digunakan dalam tahap preprocessing dan ekstraksi fitur.

4.2. Pemuatan Data

Proses pemuatan data ini bertujuan untuk membaca gambar dari dataset, mengonversinya ke format RGB, dan menyimpannya kembali ke direktori *1_load_rgb* sesuai label kelasnya. Dengan menggunakan *OpenCV*, gambar dari masing-masing subfolder kelas seperti *jahe*, *kunyit*, atau *kencur*,

terlebih dahulu dikonversi dari format BGR (*default OpenCV*) ke RGB agar sesuai dengan format warna standar untuk pemrosesan citra. Sistem ini juga mencatat jumlah gambar per kelas ke dalam struktur *dictionary*, yang berfungsi untuk memantau distribusi data agar tetap seimbang. Selain itu, proses ini membantu memverifikasi kualitas dan keberadaan gambar sebelum melanjutkan ke tahap preprocessing. Distribusi data divisualisasikan dalam bentuk diagram pie pada Gambar 5, menunjukkan total 2250 gambar dengan 450 gambar per kelas.

Total data gambar: 2250

Proporsi data per kelas:

- jahe: 450 gambar (20.00%)
- kencur: 450 gambar (20.00%)
- kunyit: 450 gambar (20.00%)
- lengkuas: 450 gambar (20.00%)
- temulawak: 450 gambar (20.00%)

Gambar 5. Load Dataset Rempah Rimpang

4.3. Pembangunan Model LGBM

Tahap pembangunan model LGBM (*Light Gradient Boosting Machine*) dilakukan dengan menggunakan data hasil ekstraksi fitur HSV sebagai input. Model ini dipilih karena efisien dalam menangani data berdimensi tinggi dan memiliki kemampuan mengurangi *overfitting* melalui metode *boosting*. Proses pelatihan dilakukan dengan data latih untuk mengenali pola visual dari masing-masing jenis rempah, kemudian dievaluasi menggunakan metrik akurasi, *precision*, *recall*, dan *f1-score* guna memastikan kinerja model optimal sebelum diuji pada data yang belum dikenal.

Tabel 3. *Hyperparameter Tuning* LGBM Terpilih

Fitur	Parameter			
	LR	NL	MD	NI
HSV	0.05	20	13	500
	0.05	20	10	500
	0.01	50	6	300
	0.05	50	13	300
	0.01	20	10	500

Penelitian ini mengevaluasi performa ekstraksi fitur HSV dengan rasio data 70:30 melalui empat skenario pengujian yang mengombinasikan variasi sumber data dan parameter model LightGBM. *Hyperparameter* yang diuji meliputi *learning rate* (0.01 dan 0.05) serta jumlah iterasi (300 dan 500), untuk menyeimbangkan antara akurasi dan efisiensi pelatihan. Berdasarkan Tabel 3, konfigurasi terbaik diperoleh dengan *learning rate* 0.05, *num leaves* 20, *max depth* 13, dan 500 iterasi, yang menunjukkan bahwa *tuning hyperparameter* berperan penting dalam mengoptimalkan kinerja klasifikasi rempah rimpang. Selain itu, penggunaan skema data primer dan gabungan primer-sekunder turut diuji untuk menilai

pengaruh variasi data terhadap kemampuan generalisasi model.

Tabel 4 Hasil Akurasi *Hyperparameter Tuning*

Fitur	Akurasi (%)	Runtime (Second)
HSV	82.912	312,66
	82.102	321,20
	80.963	389,82
	80.308	387,98
	80.219	356.42

Tabel 4 menampilkan performa model LightGBM berdasarkan kombinasi *hyperparameter* yang digunakan pada fitur HSV dalam klasifikasi rempah rimpang. Hasil terbaik diperoleh dengan akurasi mencapai 82,91% dan waktu proses yang efisien yaitu 312,66 detik, menjadikannya sebagai konfigurasi paling optimal dalam penelitian ini. Variasi pengujian lainnya menunjukkan bahwa meskipun terdapat sedikit perbedaan kombinasi *hyperparameter*, akurasi tetap berada di atas 80% dengan *runtime* yang masih relatif stabil. Hal ini menunjukkan bahwa fitur HSV tidak hanya mampu menghasilkan akurasi tinggi, tetapi juga efisiensi komputasi yang baik. Selanjutnya, pada Tabel 5, fitur HSV juga terbukti memberikan hasil terbaik di semua rasio data (80:20, 70:30, dan 60:40), dengan akurasi tertinggi sebesar 82,67% pada rasio 70:30. Evaluasi dilakukan pada dataset primer, memberikan gambaran objektif terhadap performa model dalam mengenali pola visual alami pada citra rempah. Dengan stabilitas performa di berbagai skenario, fitur HSV terbukti sebagai representasi paling informatif dan efisien untuk tugas klasifikasi rempah rimpang.

Tabel 5. Hasil Skenario Ekstraksi Fitur dan Rasio

Fitur	Rasio	Evaluasi (%)			
		Acc	Prec	Rec	F1
HSV	8020	81.25	81.80	80.90	81.10
	8020	80.80	81.35	80.60	80.80
	8020	82.00	82.55	81.45	81.75
	8020	80.95	81.50	80.70	80.90
	8020	80.35	80.90	80.30	80.50
	7030	81.10	81.65	80.80	81.00
	7030	82.15	82.70	81.60	81.90
	7030	82.67	83.19	82.04	82.31
	7030	81.55	82.10	81.10	81.35
	7030	82.50	83.00	81.90	82.20
	6040	80.50	81.05	80.40	80.60
	6040	81.70	82.25	81.20	81.45
	6040	81.40	81.95	81.00	81.20
	6040	82.30	82.85	81.75	82.00
	6040	80.65	81.20	80.50	80.70

4.4.1. Variasi Parameter LGBM

Tabel 6 menunjukkan pengaruh variasi *learning rate* terhadap performa model klasifikasi rempah rimpang dengan 500 iterasi. Pada pengaturan *learning rate* 0.05, model mencatatkan akurasi tertinggi sebesar 82.67%, dengan *precision* 83.19%, *recall* 82.04%, dan *F1-score* 82.31%, serta waktu komputasi yang efisien

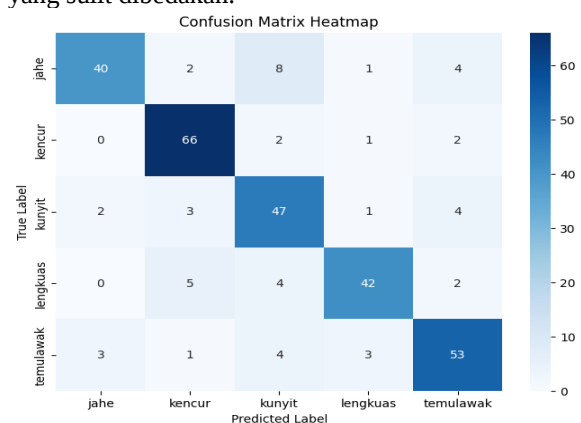
sebesar 3.21 detik. Sementara itu, pada *learning rate* 0.01, semua metrik mengalami penurunan akurasi turun menjadi 79.33% dan *F1-score* menjadi 79.07%, dengan waktu komputasi sedikit lebih lama yaitu 3.55 detik. Pengujian ini menggunakan parameter *num_leaves* = 20 dan *max_depth* = 13, serta dataset primer sebagai data uji. Hasil ini memperlihatkan bahwa *learning rate* yang lebih tinggi dalam batas wajar dapat memberikan performa yang lebih optimal untuk model LGBM.

Tabel 6. Hasil Variasi Parameter LGBM

LR	N_I	Evaluasi (%)			
		Acc	Prec	Rec	F1
500	0.05	82.67	83.19	82.04	82.31
500	0.01	79.33%	80.21	78.70	79.07

Gambar 6 menampilkan *Confusion Matrix Heatmap* yang menggambarkan distribusi hasil prediksi model terhadap masing-masing kelas rempah rimpang. Terlihat bahwa model memiliki akurasi tinggi pada kelas kencur, dengan 66 prediksi benar, menunjukkan efektivitas fitur HSV dalam menangkap warna khas kencur secara konsisten. Sementara itu, kelas jahe mencatat 40 prediksi benar, namun masih sering tertukar dengan kunyit dan temulawak karena kemiripan warna dan tekstur antar kelas. Kelas kunyit dan temulawak sendiri menunjukkan tingkat kesalahan yang lebih tinggi, terutama sering tertukar dengan lengkuas, yang memiliki pola visual serupa.

Tingginya ketertukaran antar kelas berwarna mirip menandakan bahwa HSV saja belum cukup untuk menghasilkan pemisahan kelas yang optimal dalam semua kasus. Meskipun fitur HSV sangat efektif untuk kelas dengan kontras warna jelas, performanya cenderung menurun saat menghadapi kelas yang memiliki rentang warna sempit dan tekstur serupa. Oleh karena itu, integrasi fitur tambahan seperti tekstur (GLCM) atau bentuk, serta strategi augmentasi data yang lebih selektif dan terarah, dapat menjadi solusi potensial untuk memperbaiki akurasi. Visualisasi heatmap ini memberikan wawasan penting yang dapat dimanfaatkan dalam penyusunan ulang strategi pelatihan model agar lebih adaptif terhadap kelas-kelas yang sulit dibedakan.

Gambar 6. *Confusion Matrix* Skema HSV (70:30)

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa model klasifikasi rempah rimpang yang dikembangkan berhasil mencapai performa tinggi dan efisien, pada konfigurasi penelitian ini menggunakan dataset primer dengan rasio data 70:30, learning rate 0.05, iterasi 500, jumlah leaves 20, max depth 13, serta ekstraksi fitur berbasis ruang warna HSV. Pada skenario ini, model mencapai akurasi 82.67%, precision 83.19%, recall 82.04%, dan F1-score 82.31% dengan waktu komputasi yang singkat, yaitu 3.21 detik. Keberhasilan ini menunjukkan bahwa HSV merupakan fitur yang relevan untuk klasifikasi rempah rimpang, dan model LightGBM mampu memproses data berdimensi tinggi secara optimal. Penelitian ini memberikan kontribusi nyata dalam pengembangan sistem klasifikasi otomatis yang dapat mendukung industri rempah dalam meningkatkan akurasi analisis kualitas produk. Untuk penelitian selanjutnya, disarankan untuk meningkatkan kualitas dataset dengan menambah jumlah gambar dan menggunakan citra primer yang konsisten, mengeksplorasi model lain seperti XGBoost, Random Forest, CNN, atau SVM, serta menambahkan fitur bentuk dan tekstur seperti Compactness, Aspect Ratio, Gabor Filter, GLCRM, atau LBP. Selain itu, implementasi sistem dalam platform mobile atau IoT, serta visualisasi feature importance untuk meningkatkan interpretabilitas model, juga direkomendasikan guna memperluas manfaat dan penerapan sistem secara praktis.

DAFTAR PUSTAKA

- [1] Nora Hayani, fenti Hasnani, Zulkarnaini, and Azwarni, "Efektivitas kompres hangat jahe merah dan kompres air hangat terhadap penurunan intensitas nyeri pada penderita rheumatoid arthritis," *J. SAGO gizi dan Kesehat.*, vol. 5, no. 2, pp. 284–392, 2024.
- [2] N. P. Batubara, D. Widiyanto, and N. Chamidah, "Klasifikasi rempah rimpang berdasarkan ciri warna rgb dan tekstur glcm menggunakan algoritma naive bayes," *Inform. J. Ilmu Komput.*, vol. 16, no. 3, p. 156, 2020, doi: 10.52958/iftk.v16i3.2196.
- [3] R. Setiawan and Hayatou Oumarou, "Classification of Rice Grain Varieties Using Ensemble Learning and Image Analysis Techniques," *Indones. J. Data Sci.*, vol. 5, no. 1, pp. 54–63, Mar. 2024, doi: 10.56705/ijodas.v5i1.129.
- [4] S. Sulayha and P. M. Kustiawan, "Study of Ethnopharmaceutical Plants with Anti-Inflammatory Properties In Loa Lepu Village, Kutai Kartanegara," *J. Serambi Eng.*, vol. 7, no. 3, pp. 3278–3284, 2022, doi: 10.32672/jse.v7i3.4215.
- [5] B. Agustian, "Machine Learning Approach for Rhizomes Classification Based on Color," 2021.
- [6] I. Technology, "Performance Comparison of Boosting Algorithms in Spices Classification Using Histogram of Oriented Gradient Feature Extraction," *J. Comput. Sci. Inf. Technol. Telecommun. Eng.*, vol. 4, no. 1, pp. 342–349, 2023, doi: 10.30596/jcositte.v4i1.13710.
- [7] G. R. I. L. Jayasooriya and S. M. Arachchi, "Diagnosis of Bacterial Leaf Blight, Brown Spots, and Leaf Smut Rice Plant Diseases using Light GBM," *Int. J. Comput. Appl.*, vol. 183, no. 48, pp. 53–58, 2022, doi: 10.5120/ijca2022921895.
- [8] B. Goswami, M. K. Bhuyan, S. Alfarhood, and M. Safran, "Classification of Oral Cancer Into Pre-Cancerous Stages From White Light Images Using LightGBM Algorithm," *IEEE Access*, vol. 12, pp. 31626–31639, 2024, doi: 10.1109/ACCESS.2024.3370157.
- [9] U. M. Area, "KLASIFIKASI REMPAH MENGGUNAKAN EKSTRAKSI FITUR HISTOGRAM OF ORIENTED GRADIENT (HOG) DENGAN METODE K-NEAREST NEIGHBOR (K-NN) SKRIPSI OLEH : MELISAH FAKULTAS TEKNIK KLASIFIKASI REMPAH MENGGUNAKAN EKSTRAKSI FITUR HISTOGRAM OF ORIENTED GRADIENT (HOG) DE," 2023.
- [10] B. Jehangir and S. R. Nayak, "GLCM Based LGBM classification for COVID-19 Analysis using Chest X-ray Images," *Proc. Conflu. 2022 - 12th Int. Conf. Cloud Comput. Data Sci. Eng.*, pp. 477–481, 2022, doi: 10.1109/Confluence52989.2022.9734124.
- [11] M. B. Sagita, A. Turchan, B. Utomo, A. Al Fauzi, and D. Fauziah, "Expression malondialdehyde (MDA) of brain after injury with the extract of kencur (*Kaempferia Galanga L.*)," *Int. J. Heal. Med. Sci.*, vol. 5, no. 1, pp. 114–121, 2022, doi: 10.21744/ijhms.v5n1.1848.
- [12] A. T. M. Simbolon, F. Jakfar, and A. Nugroho, "Analisis Perdagangan Intra Industri Komoditi Teh Indonesia," *J. Ilm. Mhs. Pertan.*, vol. 7, no. 1, pp. 141–149, 2022, doi: 10.17969/jimfp.v7i1.18970.
- [13] P. Szczuko, "Texture-based Finger Vein Presentation Attack Detection with Optimal Gray Level Co-Occurrence Matrix Features and Light-".