

Analisis Efektivitas Fusi Fitur Multimodal dalam Klasifikasi Citra Daun Herbal

<http://dx.doi.org/10.28932/jutisi.v11i3.12262>

Riwayat Artikel

Received: 23 Juni 2025 | Final Revision: 01 Desember 2025 | Accepted: 01 Desember 2025

Creative Commons License 4.0 (CC BY – NC)



Riki Riyandi ^{#1}, Sumarsono^{*2}

[#] Program Studi Magister Informatika, Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta
Jl. Laksda Adisucipto, Kabupaten Sleman, 55281, Indonesia

¹123206051030@student.uin-suka.ac

²sumarsono@uin-suka.ac

^{*}Corresponding author: 23206051030@student.uin-suka.ac.id

Abstrak — Penelitian ini mengevaluasi strategi *feature fusion* yang menggabungkan representasi bentuk (Histogram of Oriented Gradients/HOG), tekstur (*Gabor Filter*), dan semantik (Convolutional Neural Network atau CNN berbasis MobileNetV2) untuk klasifikasi citra daun herbal. Hasil eksperimen menunjukkan bahwa fitur semantik dari CNN tunggal secara konsisten menghasilkan performa tertinggi (akurasi 91,0% dengan *Support Vector Machine* atau SVM), jauh mengungguli HOG, Gabor, maupun strategi fusi. Kontribusi utama penelitian ini adalah temuan bahwa penggabungan fitur, khususnya dengan HOG yang berdimensi tinggi, justru menurunkan performa secara signifikan akibat fenomena *curse of dimensionality*. Penelitian ini menunjukkan bahwa keunggulan CNN atas fusi fitur menekankan pentingnya kualitas representasi semantik dibanding sekadar kuantitas fitur gabungan. Temuan ini merekomendasikan penggunaan model CNN tunggal yang efisien dan akurat, membuka potensi aplikasi praktis di bidang agrikultur digital dan identifikasi tanaman herbal secara *real-time*.

Kata kunci— CNN; *curse of dimensionality*; daun herbal; ekstraksi fitur; fusi fitur.

Effectiveness Analysis of Multimodal Feature Fusion in Herbal Leaf Classification

Abstract — This study aims to evaluate the performance of leaf image classification models based on feature fusion strategies that integrate shape (Histogram of Oriented Gradients/HOG), texture (Gabor Filter), and semantic (Convolutional Neural Network or CNN-MobileNetV2) representations. Experimental results show that single semantic features from CNN consistently yielded the highest performance (91.0% accuracy with Support Vector Machine or SVM), significantly outperforming HOG, Gabor, and fusion strategies. The main contribution is the finding that feature fusion, particularly with high-dimensional HOG, drastically reduced performance due to the curse of dimensionality. This study demonstrates that the superiority of CNN over fusion highlights the importance of semantic representation quality over the mere quantity of combined features. These findings recommend the use of efficient and accurate single CNN models, opening practical potential for digital agriculture and real-time herbal plant identification.

Keywords— CNN; *curse of dimensionality*; feature extraction; feature fusion; herbal leaves.

I. PENDAHULUAN

Strategi feature fusion dalam klasifikasi daun secara umum dibagi ke dalam tiga level: level fitur (feature-level fusion), level keputusan (decision-level fusion), dan level representasi semantik [1], [2]. Meskipun fungsi fitur populer dan secara teoritis menjanjikan, studi empiris menunjukkan hasil yang kontradiktif. Pendekatan fusi berbasis concatenation atau integrasi langsung dari vektor fitur sering kali menimbulkan masalah pada dimensi tinggi, yang memicu fenomena curse of dimensionality dan berpotensi menurunkan akurasi model [3], [4], [5].

Dalam beberapa studi, penggabungan fitur tekstur dan Convolutional Neural Network (CNN) justru menurunkan performa, disebabkan oleh dominasi fitur konvensional yang memiliki dimensi jauh lebih besar dibandingkan fitur hasil ekstraksi CNN [6], [7]. Kesenjangan (gap) inilah yang menjadi fokus utama: perlunya evaluasi mendalam untuk memahami kapan dan mengapa fusi fitur gagal, terutama dalam konteks klasifikasi daun herbal yang menuntut efisiensi. Oleh karena itu, eksplorasi model fusi fitur yang tidak hanya mempertimbangkan jenis fitur, tetapi juga skala, dimensi, dan bobot kontribusinya menjadi kebutuhan penting dalam pengembangan sistem klasifikasi berbasis citra.

Penelitian ini bertujuan untuk mengevaluasi kinerja model klasifikasi daun herbal berbasis fusi fitur dari tiga kategori representasi: bentuk (HOG), tekstur (Gabor Filter), dan semantik (CNN). Penelitian ini juga menganalisis efek samping dari strategi fusi terhadap performa klasifikasi pada kondisi multi-kelas. Dataset terdiri dari sepuluh jenis daun herbal dengan variasi morfologis yang tinggi. Hal ini menjadikan eksperimen ini relevan bagi pengembangan sistem identifikasi tanaman herbal praktis, yang dapat digunakan dalam aplikasi lapangan di bidang botani, farmasi, maupun konservasi spesies [8], [9], [10].

Kontribusi utama dari penelitian ini mencakup: Evaluasi performa klasifikasi citra daun dengan pendekatan fitur tunggal dan fusi fitur (feature fusion), Analisis pengaruh dimensi fitur terhadap akurasi model dalam konteks curse of dimensionality, Rekomendasi strategi fusi fitur berbasis efisiensi, stabilitas, dan kemampuan generalisasi.

II. METODE PENELITIAN

A. Rancangan Penelitian

Penelitian ini dirancang sebagai eksperimen kuantitatif yang bertujuan untuk mengevaluasi dan membandingkan performa klasifikasi citra daun menggunakan tiga pendekatan ekstraksi fitur utama, yaitu bentuk (HOG), tekstur (Gabor Filter), dan semantik (CNN), serta strategi fusi dari ketiganya. Penelitian dilakukan melalui pendekatan feature-level fusion dengan metode konkatenasi vektor fitur. Pengujian dilakukan terhadap data uji terpisah untuk mengevaluasi generalisasi model, dengan matrik evaluasi mencakup akurasi, presisi, recall, F1-score, serta analisis ROC dan confusion matrix [9] [10], [11]. Secara garis besar, tahapan penelitian ini dilakukan secara sistematis mulai dari akuisisi data, pra-pemrosesan, ekstraksi fitur, hingga evaluasi model, sebagaimana diilustrasikan pada Gambar 1.

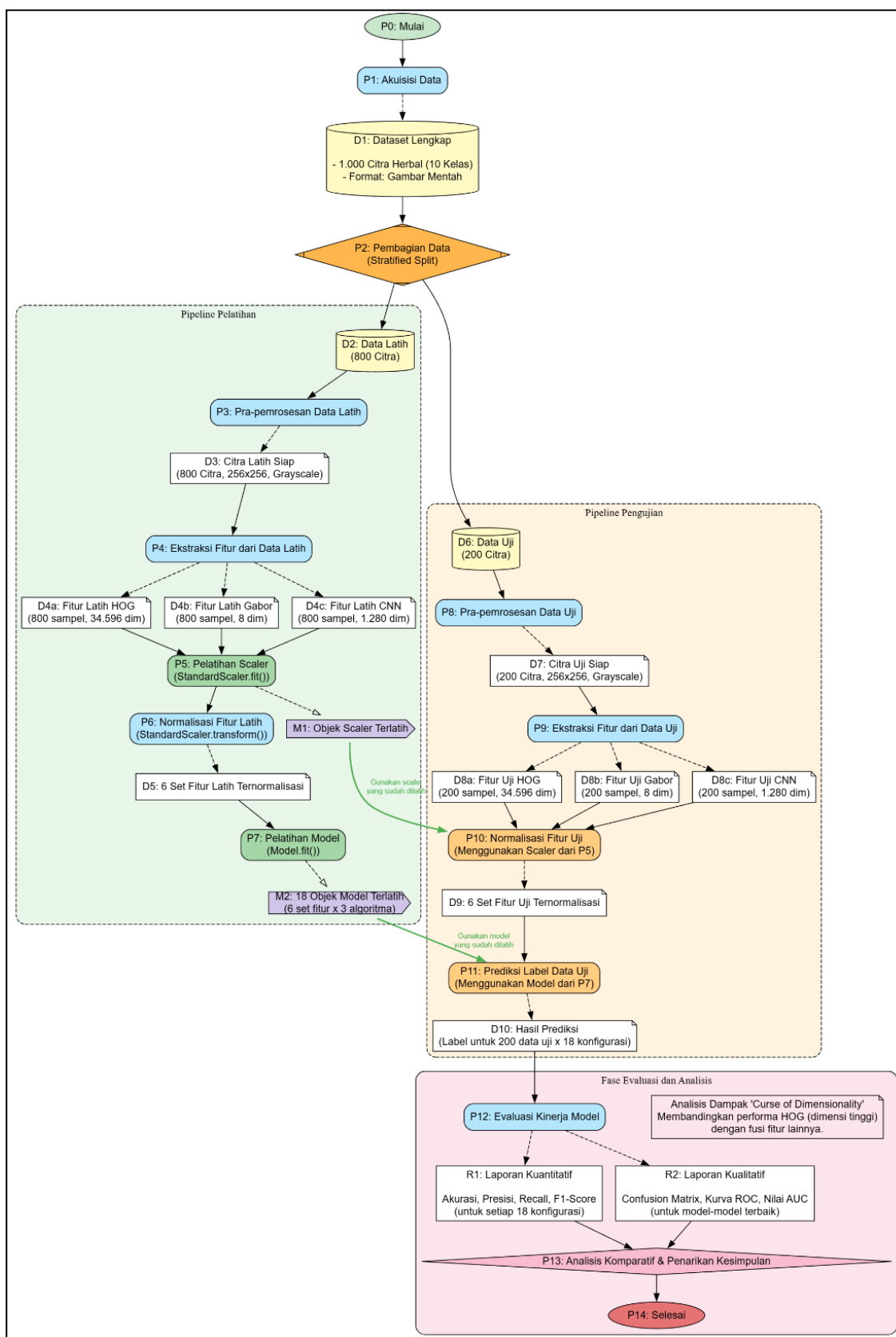
B. Data Penelitian

Data yang digunakan dalam penelitian ini berupa 1.000 citra daun herbal berwarna yang dikumpulkan dari sumber dataset publik. Dataset terdiri atas 10 kelas tanaman herbal tropis, seperti daun mint, jambu biji, kemangi, sirih, dan kunyit, dengan masing-masing kelas berjumlah 100 citra. Citra-citra tersebut telah dikurasi berdasarkan resolusi minimum 256×256 piksel dan label taksonomi yang valid. Seluruh data dibagi menjadi 800 data latih dan 200 data uji secara stratifikasi. Untuk meminimalkan variabilitas warna yang tidak relevan terhadap bentuk dan tekstur, seluruh citra dikonversi ke grayscale. Setelah itu, citra diubah ukurannya menjadi 256×256 piksel untuk menjaga konsistensi spasial. Khusus untuk model CNN, citra dikonversi kembali ke RGB guna memenuhi persyaratan input layer dari arsitektur MobileNetV2 [11]. Pengubahan ukuran ini (sering disebut squashing), meskipun tidak mempertahankan aspek rasio asli, merupakan praktik umum untuk memastikan dimensi input yang seragam bagi banyak arsitektur deep learning. Diasumsikan bahwa distorsi proporsional yang terjadi dapat ditoleransi dan dipelajari oleh model CNN selama proses pelatihan.

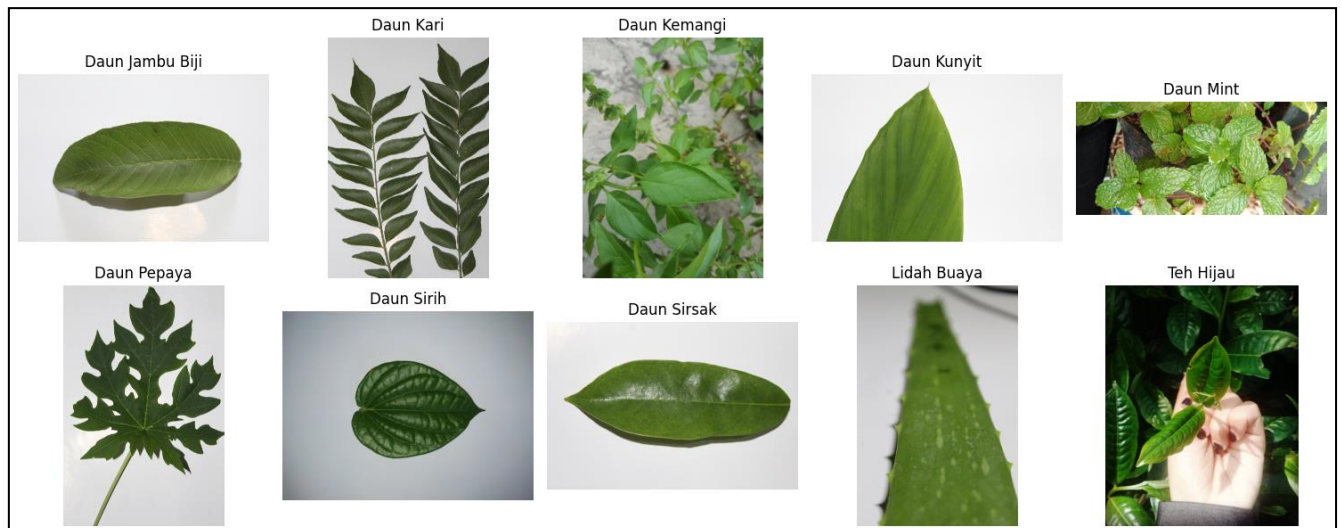
Contoh representatif dari masing-masing kelas tanaman disajikan pada Gambar 2 untuk memberikan gambaran visual mengenai variasi bentuk, tekstur, dan struktur morfologis yang menjadi objek klasifikasi dalam penelitian ini.

C. Ekstraksi Fitur dan Strategi Fusi Fitur

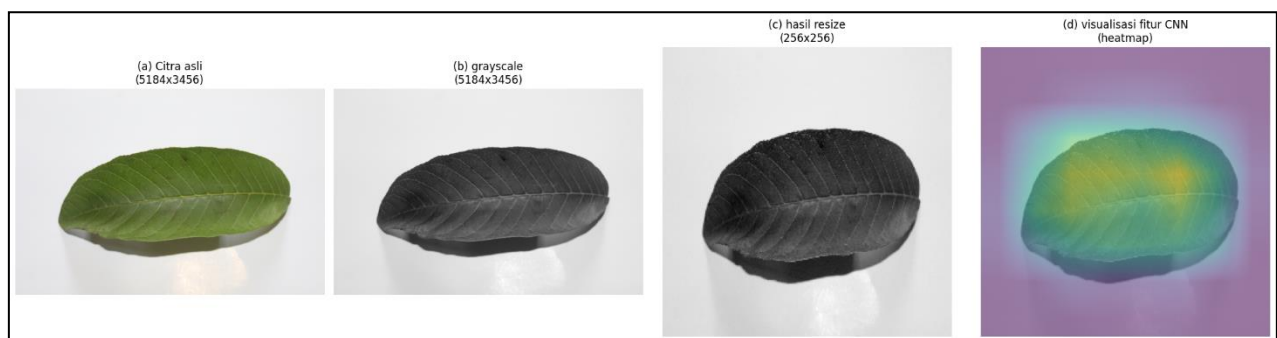
Gambar 3 menunjukkan transformasi citra dari tahap pra-pemrosesan hingga tahap representasi akhir. Dapat dilihat bahwa meskipun informasi warna dieliminasi, struktur bentuk dan tekstur tetap terjaga secara signifikan.



Gambar 1. Diagram Alur Proses Penelitian



Gambar 2. Contoh citra asli dari masing-masing kelas daun herbal



Gambar 3. Transformasi citra: (a) Citra asli, (b) grayscale, (c) hasil resize 256x256, (d) visualisasi fitur CNN (heatmap).

Tiga metode ekstraksi fitur digunakan (dijelaskan di Tabel 1). Perhitungan teknis untuk dimensi fitur yang dihasilkan adalah sebagai berikut:

1. HOG: Menggunakan *input* citra 256×256, ukuran sel 8×8, dan blok 2×2.
 - Jumlah sel: $(256 / 8) \times (256 / 8) = 32 \times 32$
 - Jumlah blok (dengan *overlap*): $(32 - 2 + 1) \times (32 - 2 + 1) = 31 \times 31 = 961$ blok.
 - Fitur per blok: $(2 \times 2 \text{ sel}) \times 9 \text{ orientasi} = 36$ fitur.
 - Total Dimensi: $961 \text{ blok} \times 36 \text{ fitur/blok} = 34.596$.
2. Gabor: Menggunakan 4 orientasi filter. Dari setiap respons, 2 statistik (*mean*, *std*) dihitung.
 - Total Dimensi: $4 \text{ orientasi} \times 2 \text{ statistik} = 8$.
3. CNN: Model MobileNetV2 (*include_top=False*) memiliki 1.280 *filter* (atau *channel*) pada lapisan konvolusional terakhir. *GlobalAveragePooling2D* merata-ratakan nilai di setiap *filter*, menghasilkan vektor akhir berdimensi 1.280.

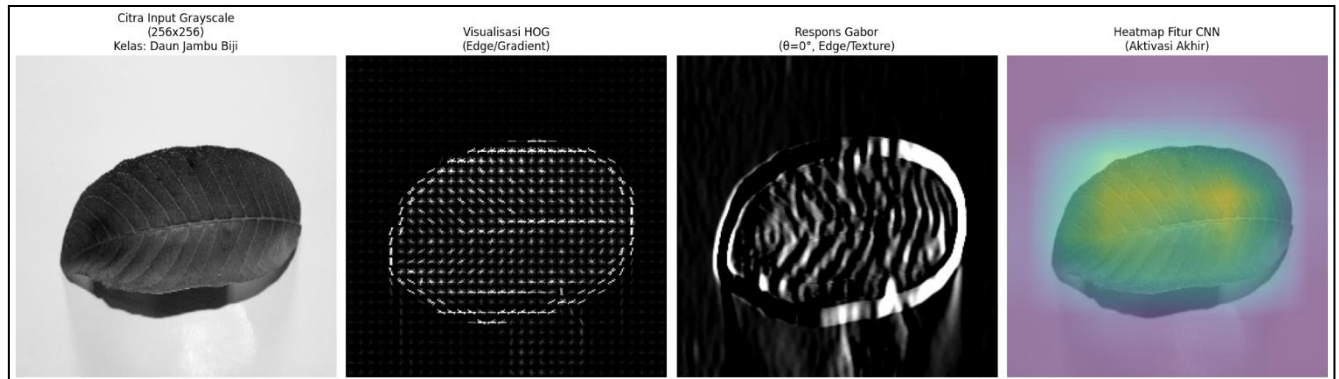
Semua fitur kemudian dinormalisasi menggunakan teknik *standard scaling*, di mana parameter statistik dihitung hanya dari data latih untuk menghindari kebocoran informasi (*data leakage*) [13] [19].

TABEL 1.
KONFIGURASI EKSTRAKSI FITUR DAN DIMENSI VEKTOR HASIL Masing-masing METODE

Kategori	Metode	Justifikasi & Parameter Utama	Dimensi Vektor
Fitur Tunggal			
HOG	Bentuk	Konfigurasi standar (Orientasi: 9 bin, Sel: 8×8, Blok: 2×2)	34.596
Gabor Filter	Tekstur	Parameter umum analisis tekstur (4 orientasi, Statistik: <i>mean</i> & <i>std</i>)	8

Kategori	Metode	Justifikasi & Parameter Utama	Dimensi Vektor
CNN (MobileNetV2)	Semantik	Model pra-latih <i>lightweight</i> (weights='imagenet', <i>GlobalAveragePooling2D</i>)	1.280
Fusi Fitur			
HOG + Gabor	Bentuk + Tekstur	Konkatenasi vektor	34.604
HOG + CNN	Bentuk + Semantik	Konkatenasi vektor	35.876
HOG + Gabor + CNN	Bentuk + Tekstur + Semantik	Konkatenasi vektor	35.884

Gambar 4 menunjukkan visualisasi hasil ekstraksi fitur dari citra yang sama. HOG menangkap struktur batas yang dominan, Gabor menyoroti pola tekstur arah tertentu, sementara CNN menghasilkan peta aktivasi yang lebih abstrak dan semantik.



Gambar 4. Perbandingan visual hasil ekstraksi fitur dari metode HOG, Gabor, dan CNN.

D. Algoritma Klasifikasi

Fitur-fitur yang dihasilkan, baik dari metode tunggal maupun kombinasi, diuji menggunakan tiga algoritma klasifikasi yang memiliki pendekatan berbeda: *Support Vector Machine* (SVM) dengan *kernel radial basis function* (RBF), *K-Nearest Neighbors* (KNN) dengan nilai $k = 5$, serta *Random Forest* (RF) dengan 100 pohon keputusan. Pemilihan algoritma ini didasarkan pada pertimbangan representatif dari model *margin-based*, *proximity-based*, dan *ensemble tree-based*. Seluruh model dilatih pada data latih dan diuji pada data uji yang telah dibagi secara stratifikasi.

E. Evaluasi Performa

Evaluasi performa dilakukan dengan mengukur nilai akurasi, presisi, *recall*, dan *F1-score*. Di samping metrik kuantitatif, analisis visual dilakukan melalui *confusion matrix* dan kurva ROC dengan pendekatan *One-vs-Rest* untuk setiap kelas. Eksperimen juga mencakup *tuning* parameter terbatas untuk model SVM dengan fitur HOG menggunakan *GridSearchCV*, guna mengetahui sensitivitas terhadap parameter C dan γ . Tujuan evaluasi ini adalah untuk mendapatkan gambaran menyeluruh tentang akurasi dan stabilitas model terhadap data uji, serta mengevaluasi potensi *overfitting* akibat redundansi fitur.

F. Platform Implementasi

Seluruh eksperimen diimplementasikan menggunakan bahasa pemrograman Python pada platform Google Colaboratory dengan akselerasi GPU T4. Pustaka yang digunakan mencakup OpenCV untuk pemrosesan citra, Scikit-learn untuk klasifikasi dan evaluasi metrik, serta TensorFlow dan Keras untuk ekstraksi fitur CNN. Spesifikasi pustaka lingkungan eksperimen dijelaskan secara ringkas pada Tabel 2.

TABEL 2.
TABEL SPESIFIKASI PERANGKAT KERAS DAN PUSTAKA PERANGKAT LUNAK YANG DIGUNAKAN SELAMA EKSPERIMEN.

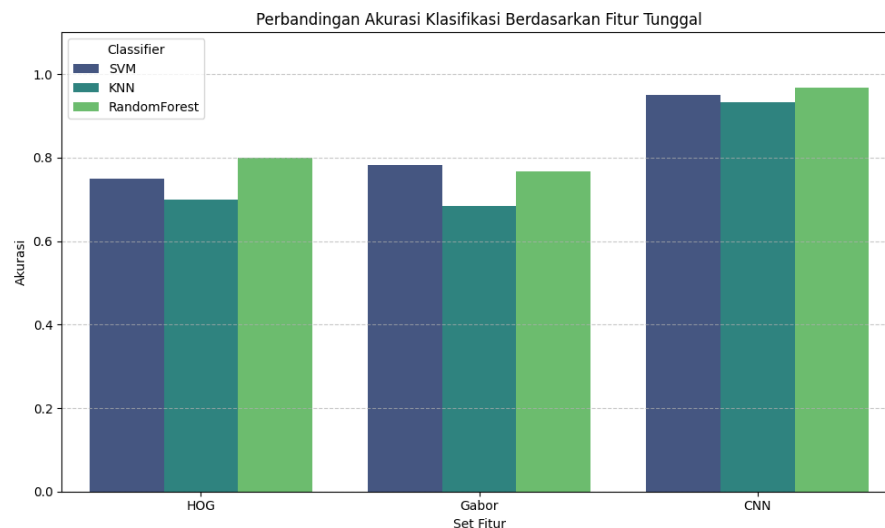
Kategori	Spesifikasi
CPU	2 vCPU Intel® Xeon® (2.30 GHz)
GPU	NVIDIA® Tesla K80 (12 GB GDDR5)
RAM	± 12.72 GB
Penyimpanan	Ephemeral SSD (± 107.7 GB tersedia)
Sistem Operasi	Ubuntu 18.04 LTS
Python	3.8.10

Kategori	Spesifikasi
TensorFlow	2.10.0
scikit-learn	1.0.2
Keras	2.10.0 (terintegrasi dalam TensorFlow)
OpenCV	4.6.0
NumPy	1.21.6
Matplotlib	3.2.2
Pandas	1.3.5

III. HASIL DAN PEMBAHASAN

A. Kinerja Model pada Fitur Tunggal

Pengujian awal dilakukan pada tiga jenis fitur tunggal. Hasil menunjukkan perbedaan performa yang sangat jelas. Model berbasis CNN secara konsisten menghasilkan performa terbaik, mencapai akurasi 91,0% (CNN+SVM), 89,5% (CNN+RF), dan 88,0% (CNN+KNN). Sebaliknya, fitur konvensional menunjukkan performa yang jauh lebih rendah, dengan akurasi maksimal hanya 53,0% (HOG+SVM) dan 46,5% (Gabor+SVM). Penyebab kesenjangan performa ini adalah keunggulan fundamental dari representasi semantik CNN. Tidak seperti HOG (bentuk) atau Gabor (tekstur) yang *hand-crafted*, CNN mampu mempelajari pola visual secara hierarkis dan adaptif langsung dari data mentah [13]. Fitur semantik CNN lebih *robust* terhadap variasi dan mampu menangkap esensi kelas daun, sementara fitur HOG dan Gabor kesulitan menggeneralisasi pola pada 10 kelas yang berbeda. Perbandingan visual kinerja akurasi ketiga metode ekstraksi fitur tersebut disajikan secara rinci pada Gambar 5.



Gambar 5. Perbandingan akurasi klasifikasi berdasarkan fitur Tunggal

Penyebab kesenjangan performa ini adalah keunggulan fundamental dari representasi semantik CNN. Hal ini sesuai dengan temuan sebelumnya bahwa CNN menunjukkan performa tinggi pada berbagai domain klasifikasi citra, mulai dari klasifikasi spesies ikan [12] hingga varietas tanaman pangan dan hias [13], [14]. Studi lain juga mengonfirmasi bahwa arsitektur deep learning (seperti VGG, Inception, atau MobileNet) secara konsisten mengungguli metode manual pada kasus daun mangga, anggur, dan tomat [15], [16], [17]. Secara visual, model CNN mampu mengekstraksi pola semantik kompleks dari citra daun yang tidak ditangkap oleh fitur tekstur maupun bentuk. Hal ini sesuai dengan temuan sebelumnya bahwa CNN menunjukkan performa tinggi pada domain klasifikasi citra tanaman

B. Evaluasi Fusi Fitur

Temuan ini menantang asumsi umum bahwa fusi fitur selalu meningkatkan performa. Eksperimen kemudian dilanjutkan dengan pengujian strategi fusi fitur: HOG+Gabor, HOG+CNN, dan HOG+Gabor+CNN. Hasil menunjukkan bahwa fusi HOG+CNN maupun HOG+Gabor+CNN justru menurunkan performa secara signifikan dibandingkan CNN saja. Misalnya, pada kombinasi HOG+CNN, akurasi SVM turun drastis dari 91.0% menjadi 57.5%, dan pada kombinasi HOG+Gabor+CNN hanya mencapai 57.5% pula. Penurunan ini terjadi juga pada KNN dan RF, yang menyiratkan bahwa penambahan fitur HOG yang berdimensi sangat tinggi (34.596) menyebabkan redundansi dan noise yang mengganggu pembelajaran model.

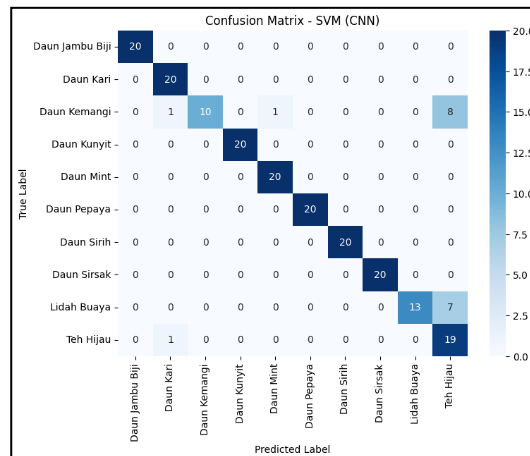
TABEL 3.
AKURASI MODEL PADA DATA UJI

Feature Set	Classifier	Accuracy	Precision (Weighted)	Recall (Weighted)	F1-Score (Weighted)
CNN	KNN	0.880	0.887416	0.880	0.873029
CNN	RandomForest	0.895	0.920207	0.895	0.890016
CNN	SVM	0.910	0.942030	0.910	0.908624
Gabor	KNN	0.480	0.449213	0.480	0.441301
Gabor	RandomForest	0.520	0.500289	0.520	0.488720
Gabor	SVM	0.465	0.435298	0.465	0.425737
HOG	KNN	0.435	0.514298	0.435	0.418063
HOG	RandomForest	0.520	0.598097	0.520	0.489066
HOG	SVM	0.530	0.604023	0.530	0.480768
HOG+CNN	KNN	0.480	0.521514	0.480	0.452905
HOG+CNN	RandomForest	0.820	0.850236	0.820	0.803140
HOG+CNN	SVM	0.575	0.623391	0.575	0.528703
HOG+Gabor	KNN	0.435	0.514298	0.435	0.418063
HOG+Gabor	RandomForest	0.510	0.534978	0.510	0.468993
HOG+Gabor	SVM	0.525	0.569709	0.525	0.471646
HOG+Gabor+CNN	KNN	0.480	0.521514	0.480	0.452905
HOG+Gabor+CNN	RandomForest	0.790	0.844821	0.790	0.774574
HOG+Gabor+CNN	SVM	0.575	0.623391	0.575	0.528703

Penurunan ini mengindikasikan bahwa penambahan fitur yang terlalu besar justru memperkenalkan redundansi dan noise yang tidak memperkaya informasi klasifikasi, melainkan memperburuk kinerja model [3], [14].

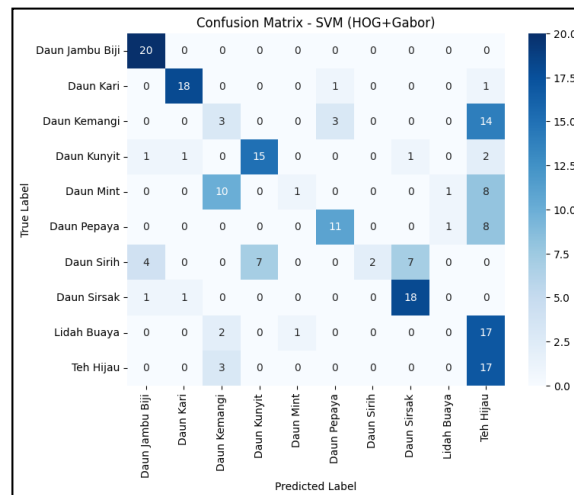
C. Visualisasi Hasil Klasifikasi

Visualisasi pada Gambar 6 menampilkan matriks konfusi untuk model terbaik (SVM + CNN), yang menunjukkan distribusi prediksi mendekati ideal. Mayoritas prediksi berada di diagonal utama, yang menandakan klasifikasi yang tepat untuk sebagian besar kelas daun. Beberapa kesalahan kecil terletak pada kelas dengan morfologi serupa seperti ‘Daun Kemangi’ dan ‘Lidah Buaya’ yang sebagian salah diklasifikasikan sebagai ‘Teh Hijau’.



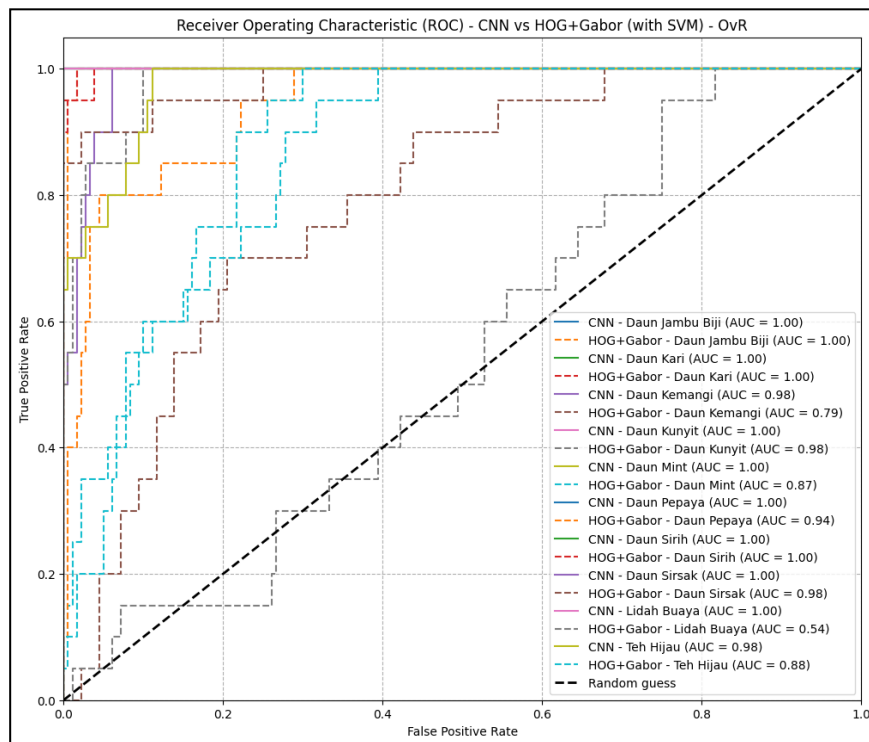
Gambar 6. Confusion Matrix SVM + CNN

Sebaliknya, seperti terlihat pada Gambar 7, matriks konfusi dari model SVM dengan fitur HOG+Gabor memperlihatkan distribusi kesalahan yang jauh lebih besar, dengan prediksi yang tersebar di luar diagonal utama. Sebagai contoh, hanya 3 dari 20 sampel ‘Daun Kemangi’ berhasil diklasifikasikan dengan benar. Kelas ‘Lidah Buaya’ tidak berhasil dikenali sama sekali, dan seluruhnya salah diklasifikasikan sebagai ‘Teh Hijau’.



Gambar 7. Confusion Matrix SVM + HOG+Gabor

Untuk analisis ROC, model SVM + CNN menghasilkan kurva ROC dengan AUC mendekati 1 pada hampir semua kelas, mengindikasikan separabilitas antar kelas yang sangat baik. Sebaliknya, ROC macro-average dari model SVM + HOG+Gabor hanya sedikit di atas garis diagonal, yang menandakan performa mendekati acak pada beberapa kelas. Perbedaan signifikan area di bawah kurva (AUC) antara kedua konfigurasi model tersebut dapat dilihat pada Gambar 8.



Gambar 8. Kurva ROC OvR CNN vs HOG+Gabor, dengan SVM

D. Analisis Kegagalan Fusi: Curse of Dimensionality

Temuan empiris menunjukkan bahwa upaya penggabungan fitur HOG ke dalam konfigurasi fusi berdampak negatif, bahkan ketika dikombinasikan dengan CNN. Analisis ini mengarah pada identifikasi fenomena Curse of Dimensionality. Fitur HOG memiliki 34.596 dimensi, jauh melampaui dimensi CNN yang hanya 1.280. Kombinasi tersebut menyebabkan dominasi fitur kurang informatif yang membanjiri model dengan noise dan korelasi palsu, sehingga informasi penting dari CNN justru "tenggelam" dalam redundansi fitur HOG.

Secara matematis, CoD menyatakan bahwa penambahan dimensi akan meningkatkan kompleksitas ruang vektor secara eksponensial dan menurunkan kepadatan data di ruang fitur. Akibatnya, model klasifikasi menjadi lebih rentan terhadap *overfitting* dan kesalahan estimasi parameter.

Secara matematis-geometris, fenomena ini dapat dijelaskan sebagai berikut: Diasumsikan bahwa fraksi data f (misal $f = 0.01$ atau 1%) yang berada dalam *hypercube* di sekitar pusat data, dalam ruang fitur berdimensi n . Jika s adalah panjang sisi *hypercube* tersebut (sebagai persentase dari total rentang data), maka volume *hypercube* Adalah V yang dinyatakan dalam persamaan (1):

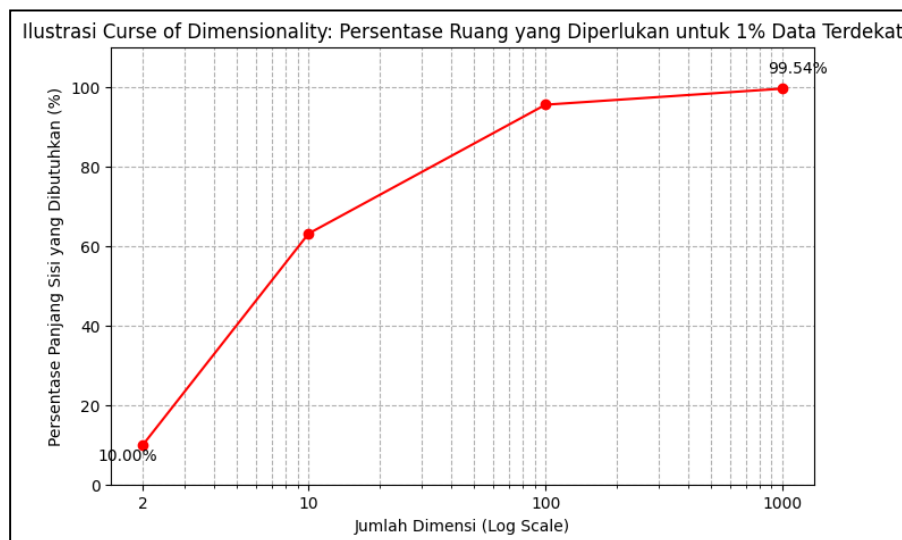
$$V = s^n \quad (1)$$

Untuk mendapatkan volume $V = f$, maka panjang sisi s yang dibutuhkan dihitung menggunakan Persamaan (2):

$$s = f^{(1/n)} \quad (2)$$

Pada 2 Dimensi ($n = 2$): $s = (0.01)^{(1/2)} = 0.1$. Artinya, hanya diperlukan cakupan sebesar 10,00% panjang sisi di setiap dimensi untuk menangkap 1% data. Namun pada 1000 Dimensi ($n = 1000$): $s = (0.01)^{(1/1000)} \approx 0.9954$. Artinya, diperlukan cakupan sebesar 99,54% panjang sisi di setiap dimensi hanya untuk menangkap 1% data terdekat yang sama.

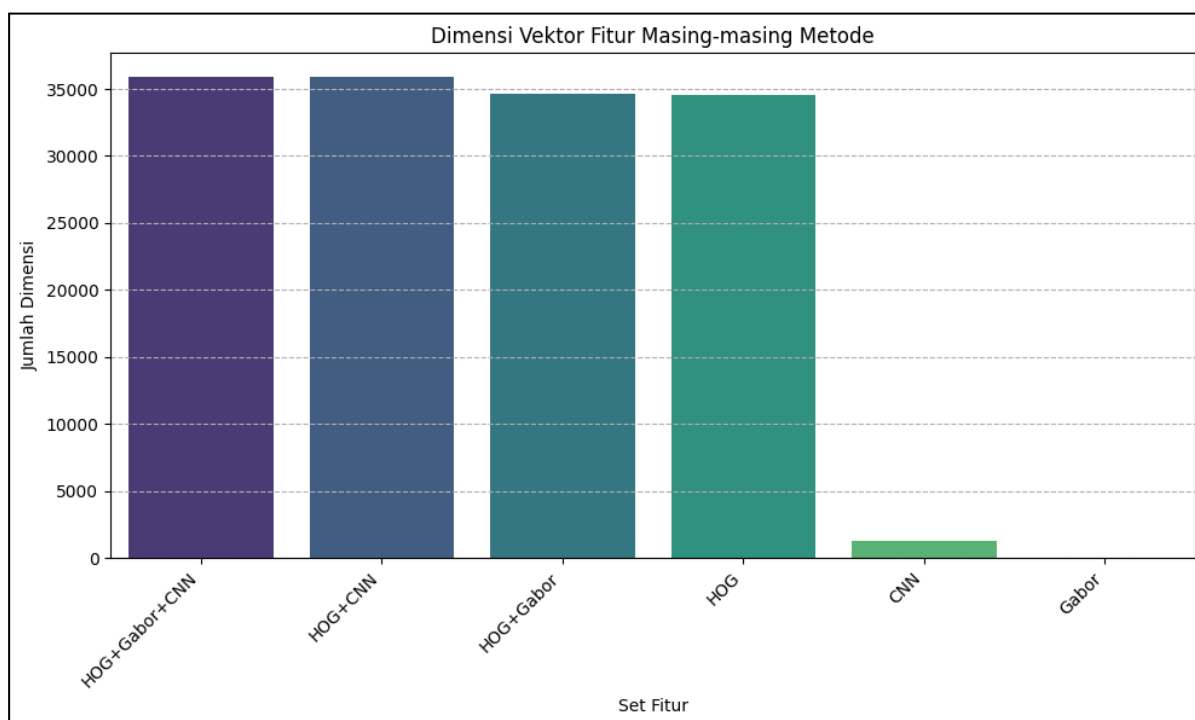
Ilustrasi pada Gambar 9 menunjukkan efek ini. Di ruang dimensi tinggi (seperti HOG dengan 34.596 dimensi), semua titik data menjadi 'jarang' dan 'terpisah' jauh satu sama lain (*sparse*). Fenomena ini membuat banyak algoritma tradisional (seperti KNN yang berbasis jarak atau *density*) menjadi kurang efektif karena konsep 'kedekatan' kehilangan makna.



Gambar 9. Visualisasi ilustratif efek curse of dimensionality pada ruang fitur HOG vs CNN

Fitur HOG dan Gabor (dengan parameter yang wajar) menghasilkan vektor fitur dengan dimensi yang relatif rendah hingga menengah (misal ratusan hingga ribuan dimensi). Algoritma seperti SVM dan KNN dapat bekerja dengan baik pada dimensi ini, meskipun KNN mungkin mulai kesulitan jika dimensinya terlalu tinggi atau datanya terlalu sedikit. Fitur yang diekstrak dari layer akhir CNN (misal dari layer GlobalAveragePooling2D seperti di MobileNetV2) biasanya memiliki dimensi yang jauh lebih tinggi (misal 1280 dimensi untuk MobileNetV2, atau ribuan/puluhan ribu untuk arsitektur lain). Meskipun CNN beroperasi di ruang fitur dimensi tinggi, mereka menggunakan arsitektur berjenjang (hierarchikal) dan proses ekstraksi fitur yang terstruktur (konvolusi, pooling) yang secara implisit belajar representasi data yang lebih diskriminatif dan mungkin lebih 'tahan' terhadap beberapa aspek curse of dimensionality dibandingkan algoritma sederhana pada ruang fitur mentah berdimensi tinggi. Heatmap CNN (visualisasi aktivasi spasial) menunjukkan area citra mana yang paling berkontribusi pada fitur akhir yang digunakan untuk klasifikasi. Ini memberikan wawasan spasial yang berbeda dari representasi global atau lokal berbasis grid seperti HOG. Singkatnya, HOG dan Gabor beroperasi di ruang fitur dimensi yang lebih dapat 'dikelola' secara tradisional, sementara CNN mengeksplorasi ruang fitur dimensi tinggi tetapi dengan cara yang berbeda, memanfaatkan struktur data visual.

Dimensi vektor fitur masing-masing metode ditampilkan pada Gambar 10. Dominasi HOG terlihat sangat signifikan dalam struktur fusi, yang menjadi indikasi potensi terjadinya gangguan dominasi fitur pada model klasifikasi.



Gambar 10. Dimensi fitur

E. Diskusi

Hasil penelitian ini menegaskan bahwa fitur semantik dari CNN memiliki keunggulan signifikan dibandingkan metode representasi bentuk maupun tekstur. Meskipun fitur berbasis tekstur seperti *Gabor Filter* dan GLCM (*Gray Level Co-occurrence Matrix*) populer digunakan dalam deteksi kualitas daun tembakau atau pola daun [18], [19], [20] fusi naif dengan fitur berdimensi tinggi terbukti tidak efektif dalam studi ini. . Fitur CNN yang diperoleh melalui *transfer learning* mampu menangkap pola visual kompleks secara hierarkis, dari detail spasial lokal hingga struktur global. Hal ini menjadikan CNN lebih adaptif terhadap variasi bentuk, rotasi, dan pencahayaan, yang sering menjadi tantangan dalam klasifikasi citra daun herbal [13][15].

Secara umum fusi fitur sering diklaim sebagai pendekatan yang mampu meningkatkan performa, hasil penelitian ini menunjukkan bahwa fusi yang dilakukan secara naif (*naïve concatenation*), khususnya dengan menggabungkan fitur berdimensi tinggi seperti HOG, justru menyebabkan penurunan kinerja. Dalam eksperimen ini, akurasi model SVM turun dari 91,0% (dengan fitur CNN tunggal) menjadi hanya 57,5% saat digabungkan dengan HOG. Ini adalah penurunan kinerja drastis lebih dari 30-40%. Hal ini menunjukkan adanya dominasi dimensi yang tidak proporsional, di mana HOG dengan dimensi lebih dari 34.000 “menenggelamkan” kontribusi informatif dari CNN yang hanya berdimensi 1.280. Fenomena ini merupakan manifestasi dari *curse of dimensionality* [2], [3].

Dari sudut pandang metodologis, hasil ini memberikan pesan penting bahwa strategi fusi fitur harus dilakukan secara selektif dan kontekstual. Penggabungan fitur harus mempertimbangkan kesetaraan skala, korelasi antar fitur, serta potensi redundansi informasi. Pendekatan seperti *feature selection*, *feature weighting*, atau penerapan *attention mechanism* menjadi sangat relevan untuk menghindari dominasi fitur yang tidak proporsional [16], [17].

Secara praktis, model berbasis CNN seperti MobileNetV2 sangat menjanjikan untuk implementasi di sistem *agritech* dan perangkat terbatas (*embedded systems* atau *smartphone*). Dengan dimensi fitur yang relatif kecil, waktu pelatihan cepat, dan akurasi tinggi, model ini sangat cocok untuk aplikasi lapangan, seperti membantu petani melakukan diagnosis penyakit daun atau identifikasi tanaman obat secara *real-time*. Pendekatan ini juga mendukung prinsip *Occam's Razor* dalam pengembangan *machine learning*: bahwa model yang lebih sederhana, bila memberikan performa sebanding atau lebih baik, seharusnya menjadi prioritas.

Dari sisi pengembangan penelitian lanjutan, hasil ini membuka ruang eksplorasi untuk strategi fusi multimodal yang mengintegrasikan fitur visual dari CNN dengan data non-visual. Pendekatan ini tidak hanya meningkatkan akurasi, tetapi juga memperluas konteks pemanfaatan sistem klasifikasi. Dengan demikian, penelitian ini menyarankan pergeseran fokus dari kuantitas fitur ke kualitas dan relevansi fitur dalam konteks tugas klasifikasi sangat cocok untuk aplikasi lapangan.

Temuan ini sejalan dengan riset terbaru yang merekomendasikan model *Deep Learning* efisien (seperti EfficientNet) untuk klasifikasi daun herbal dibandingkan pendekatan hibrida yang kompleks [21].

IV. SIMPULAN

Penelitian ini telah mengevaluasi performa model klasifikasi citra daun herbal dengan membandingkan tiga jenis representasi fitur utama: bentuk (HOG), tekstur (*Gabor Filter*), dan semantik (CNN berbasis MobileNetV2), serta mengkaji dampak strategi fusi fitur terhadap akurasi klasifikasi *multi-kelas*. Hasil eksperimen menunjukkan bahwa fitur CNN secara konsisten menghasilkan performa klasifikasi tertinggi pada ketiga algoritma yang diuji, dengan akurasi maksimum 91,0% pada model SVM. Sebaliknya, fitur HOG dan Gabor menunjukkan akurasi yang jauh lebih rendah.

Penggabungan fitur melalui strategi fusi naif (*naïve concatenation*) terbukti menurunkan performa secara signifikan. Hal ini disebabkan oleh dominasi dimensi tinggi dari HOG. Secara spesifik, *noise* dan redundansi dari 34.596 fitur HOG "menenggelamkan" kontribusi informatif dari 1.280 fitur semantik CNN, menyebabkan ketidakseimbangan masif dalam struktur fitur gabungan dan memicu fenomena *curse of dimensionality*.

Berdasarkan temuan ini, kontribusi konseptual utama studi ini adalah penegasan bahwa kualitas representasi fitur (semantik CNN) jauh lebih penting daripada kuantitas fitur (fusi naif). Model CNN ringan seperti MobileNetV2 terbukti menjanjikan untuk aplikasi lapangan *real-time*, karena lebih efisien dan akurat dibandingkan pendekatan fusi yang kompleks.

Untuk penelitian lanjutan, disarankan eksplorasi strategi fusi fitur yang lebih adaptif dan kontekstual. Secara spesifik, dapat diuji penerapan teknik reduksi dimensi (seperti *Principal Component Analysis (PCA)* atau *Autoencoder*) untuk fitur HOG sebelum fusi, atau menerapkan mekanisme *feature weighting* dan *attention-based fusion* untuk menyeimbangkan kontribusi setiap jenis fitur.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Program Studi Magister Informatika, Universitas Islam Negeri Sunan Kalijaga Yogyakarta atas dukungan akademik dan fasilitas penelitian yang diberikan selama proses penyusunan karya ilmiah ini. Penghargaan juga diberikan kepada para dosen pembimbing dan penguji yang telah memberikan masukan konseptual, metodologis, serta kritik konstruktif yang memperkaya arah penelitian ini.

Ucapan terima kasih disampaikan pula kepada tim pengelola dataset citra daun herbal yang telah menyediakan sumber data terbuka sehingga memungkinkan penelitian ini dilakukan secara replikatif dan dapat diverifikasi. Tak lupa, apresiasi kepada rekan-rekan sejawat yang turut berdiskusi dalam pengembangan eksperimen dan penyusunan artikel ini.

DAFTAR PUSTAKA

- [1] Y. Peng, S. Zhao, and J. Liu, "Fused-Deep-Features Based Grape Leaf Disease Diagnosis," *Agronomy*, vol. 11, no. 11, p. 2234, 2021.
- [2] R. Yu, X. Yao, Y. Huo, and C. Chang, "Enhancing Skin Lesion Classification using Segmentation-based Global-Local Feature Fusion," in *Medical Imaging 2025: Image Processing*, O. Colliot and J. Mitra, Eds., SPIE, 2025.
- [3] J. M. Ramirez and H. Arguello, "Multiresolution Compressive Feature Fusion for Spectral Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 9900–9911, 2019.
- [4] J. Ramirez, H. Vargas, J. I. Martinez, and H. Arguello, "Subspace-Based Feature Fusion from Hyperspectral and Multispectral Images for Land Cover Classification," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, IEEE, Jul. 2021, pp. 3003–3006.
- [5] Z. Cai, J. Yang, H. Fang, T. Ji, Y. Hu, and X. Wang, "Research on Waste Plastics Classification Method Based on Multi-Scale Feature Fusion," *Sensors*, vol. 22, no. 20, p. 7974, 2022.
- [6] X. Wu, Z. Luo, and H. Xu, "Recognition of Pear Leaf Disease Under Complex Background Based on DBPNet and Modified mobilenetV2," *IET Image Process*, vol. 17, no. 10, pp. 3055–3067, 2023.
- [7] A. Riska, Purnawansyah, H. Darwis, and W. Astuti, "Studi Perbandingan Kombinasi GMI, HSV, KNN, Dan CNN Pada Klasifikasi Daun Herbal," *Indonesian Journal of Computer Science*, vol. 12, no. 3, 2023.
- [8] R. G. Guntara, "Ekstraksi Fitur Warna Citra Daun Untuk Klasifikasi Skala Klorofil Dan Rekomendasi Pemupukan," *Jurnal Minfo Polgan*, vol. 11, no. 1, pp. 15–22, 2022.
- [9] S. I. Fitriyati, M. Ma'rifat, N. Rahman, D. Nadila, A. I. Randani, and A. Ariani, "Kadar Zat Gizi, Daya Antioksidan, Dan Organoleptik Biskuit Berbasis Daun Kelor Dan Tulang Ikan Tuna," *Ghidza Jurnal Gizi Dan Kesehatan*, vol. 7, no. 2, pp. 273–288, 2023.
- [10] V. Pandiyaraju *et al.*, "Improved Tomato Leaf Disease Classification Through Adaptive Ensemble Models with Exponential Moving Average Fusion and Enhanced Weighted Gradient Optimization," *Front Plant Sci*, vol. 15, 2024.
- [11] M. E. Prasetyo, M. R. Faza, R. Pratama, S. N. H. Alhabsy, H. Purwanti, and A. P. A. Masa, "Klasifikasi Ragam Kendaraan Menggunakan Metode Convolutional Neural Network (Cnn)," *Adopsi Teknologi dan Sistem Informasi (ATASI)*, vol. 2, no. 2, pp. 142–148, 2023.
- [12] W. Arrank Tonapa, P. D.K. Manembu, and F. D. Kambey, "Klasifikasi Ikan Cakalang dan Tongkol Menggunakan Convolutional Neural Network," *Jurnal Teknik Informatika*, vol. 19, no. 01, pp. 31–36, 2024.
- [13] M. Rijal, A. M. Yani, and A. Rahman, "Deteksi Citra Daun Untuk Klasifikasi Penyakit Padi Menggunakan Pendekatan Deep Learning Dengan Model CNN," *Jurnal Teknologi Terpadu*, vol. 10, no. 1, pp. 56–62, 2024.
- [14] B. Suswati, "Implementasi Convolutional Neural Network Pada Klasifikasi Citra Daun Tanaman Hias," *Decode Jurnal Pendidikan Teknologi Informasi*, vol. 4, no. 2, pp. 554–565, 2024.
- [15] T. Ayu, V. Dwi, and A. E. Minarno, "Pendiagnosa Daun Mangga Dengan Model Convolutional Neural Network," *Cess (Journal of Computer Engineering System and Science)*, vol. 6, no. 2, p. 230, 2021.
- [16] N. Huda, A. Mahiruna, W. Sulistijanti, and R. C. N. Santi, "Analisis Performa Inceptionv3 Convolutional Network Pada Klasifikasi Varietas Daun Grapevine," *Jurnal Sains Komputer Dan Teknologi Informasi*, vol. 5, no. 2, pp. 47–53, 2023.

- [17] P. Palupiningsih, A. R. Sujiwanto, and R. R. B. P. Prawirodirjo, "Analisis Perbandingan Performa Model Klasifikasi Kesehatan Daun Tomat Menggunakan Arsitektur VGG, MobileNet, Dan Inception V3," *Jurnal Ilmu Komputer Dan Agri-Informatika*, vol. 10, no. 1, pp. 98–110, 2023.
- [18] M. Habibullah, H. Fahmi, and E. Herawati, "Penerapan Metode Segmentasi Gabor Filter Dan Algoritma Support Vector Machine Untuk Pendeteksian Penyakit Daun Tomat," *Jurnal Riset Mahasiswa Matematika*, vol. 2, no. 6, pp. 221–232, 2023.
- [19] S. Sriani and Y. Rizky, "Klasifikasi Kualitas Daun Tembakau Menggunakan GLCM (Gray Level Co-Occurrence Matrix) Dan SVM (Support Vector Machine)," *Jurnal Informatika Dan Teknik Elektro Terapan*, vol. 12, no. 3, 2024.
- [20] A. A. Paturrahman and I. G. P. S. Wijaya, "Analisis Pengenalan Pola Daun Berdasarkan Fitur Canny Edge Detection Dan Fitur GLCM Menggunakan Metode Klasifikasi K-Nearest Neighbor (kNN)," *Journal of Computer Science and Informatics Engineering (J-Cosine)*, vol. 5, no. 1, pp. 68–76, 2021.
- [21] R. P. S. Putra, C. S. K. Aditya, and G. W. Wicaksono, "Herbal Leaf Classification Using Deep Learning Model Efficientnetv2b0," *Jitk (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, vol. 9, no. 2, pp. 301–307, 2024.