Vol. 11, No. 2, Oktober 2026, Page. 670-688
https://journal.uin-alauddin.ac.id/index.php/instek
p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

Optimasi Hyperparameter Menggunakan Grid Search untuk Klasifikasi Penyakit Daun Cabai Berbasis Fitur HOG – SVM

Nurfadillah[*1], Nurhikma Arifin[2], Nourman S Irjanto[3]

[1,2,3]Teknik Informatika, Fakultas Teknik, Universitas Sulawesi Barat, Indonesia
Email: [1]nurfadila03092004@gmail.com, [2] nurhikma_arifin@unsulbar.ac.id,
[3]nourmansatyairjanto@unsulbar.ac.id

Abstrak

Tanaman cabai merupakan komoditas hortikultura penting di Indonesia, tetapi produktivitasnya dapat terganggu oleh penyakit daun. Identifikasi penyakit secara manual melalui pengamatan visual membutuhkan waktu dan bergantung pada pengalaman pemeriksa. Penelitian ini bertujuan mengoptimalkan *hyperparameter* model *Histogram of Oriented Gradients–Support Vector Machine* (HOG-SVM) menggunakan GridSearchCV dengan 3-*fold cross-validation* untuk mengklasifikasikan citra daun cabai ke dalam tiga kelas, yaitu Daun Sehat, *Leaf Curl*, dan *Yellowish*. Penelitian menggunakan 1.200 citra yang dibagi menjadi data pelatihan, validasi, dan pengujian dengan proporsi yang sama untuk setiap pendekatan. Selain HOG-SVM, penelitian ini mengevaluasi kombinasi fitur HOG dan warna *Hue, Saturation, Value* (HSV) pada SVM (HOG+HSV-SVM), serta *Vanilla Convolutional Neural Network* (CNN) menggunakan citra *grayscale* dan RGB. Kinerja model dievaluasi berdasarkan *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil optimasi HOG-SVM menghasilkan rata-rata *cross-validation accuracy* sebesar 97,38%, dengan *accuracy* dan *F1-score* pada data pengujian masing-masing sebesar 96,11% dan 96,12%. HOG+HSV-SVM memperoleh *accuracy* pengujian sebesar 96,67% dan *F1-score* sebesar 96,68%. Sementara itu, *Vanilla CNN* memperoleh *F1-score* sebesar 95,54% pada citra *grayscale* dan 98,33% pada citra RGB. Hasil ini menunjukkan bahwa HOG-SVM memiliki kinerja yang relatif sebanding dengan CNN *grayscale*, sedangkan CNN dengan masukan RGB memperoleh skor tertinggi dalam eksperimen ini. Temuan tersebut mengindikasikan bahwa informasi warna berpotensi membantu klasifikasi penyakit daun cabai.

Kata Kunci: penyakit daun cabai, klasifikasi citra, Histogram of Oriented Gradients, Support Vector Machine, Grid Search.

Abstract

Chili is an important horticultural commodity in Indonesia, but its productivity can be affected by leaf diseases. Manual disease identification through visual observation is time-consuming and depends on the examiner's experience. This study aims to optimize the hyperparameters of the Histogram of Oriented Gradients–Support Vector Machine (HOG-SVM) model using GridSearchCV with 3-fold cross-validation to classify chili leaf images into three classes: Healthy, Leaf Curl, and Yellowish. The study uses 1,200 images divided into training, validation, and testing sets using the same proportions for each approach. In addition to HOG-SVM, this study evaluates a combination of HOG and Hue, Saturation, and Value (HSV) color features with SVM (HOG+HSV-SVM), as well as Vanilla Convolutional Neural Network (CNN) models using grayscale and RGB images. Model performance is evaluated based on accuracy, precision, recall, and F1-score. The optimized HOG-SVM model achieves a mean cross-validation accuracy of 97.38%, with testing accuracy and F1-score of 96.11% and 96.12%, respectively. HOG+HSV-SVM achieves a testing accuracy of 96.67% and an F1-score of 96.68%. Meanwhile, Vanilla CNN obtains F1-scores of 95.54% for grayscale images and 98.33% for RGB images. These results indicate that HOG-SVM performs relatively comparably to grayscale CNN, while RGB CNN achieves the highest score in this experiment. The findings suggest that color information may contribute to chili leaf disease classification.

Keywords: *chili leaf disease, image classification, Histogram of Oriented Gradients, Support Vector Machine, Grid Search.*

This work is an open access article and licensed under a Creative Commons Attribution-NonCommercial ShareAlike 4.0 International (CC BY-NC-SA 4.0)

1. PENDAHULUAN

Tanaman cabai merupakan salah satu komoditas hortikultura yang banyak dibudidayakan. Kondisi tanaman dapat dipengaruhi oleh gangguan yang menimbulkan perubahan pada karakteristik visual daun,

seperti bentuk, tekstur, struktur, dan warna. Karakteristik tersebut dapat digunakan sebagai informasi untuk membedakan daun sehat dan daun yang menunjukkan gejala gangguan. Penelitian ini menggunakan tiga kategori citra yang tersedia dalam dataset, yaitu Daun Sehat, Leaf Curl, dan Yellowish [1].

Identifikasi penyakit daun cabai secara manual memiliki keterbatasan karena hasil pemeriksaan dapat dipengaruhi oleh pengetahuan dan pengalaman pemeriksa. Gejala yang muncul pada tahap awal juga dapat sulit dibedakan melalui pengamatan visual, sedangkan pemeriksaan dalam jumlah besar membutuhkan waktu. Perkembangan pengolahan citra digital dan kecerdasan buatan memberikan alternatif untuk menganalisis karakteristik visual daun secara sistematis sehingga dapat mendukung proses klasifikasi kondisi tanaman [2], [3].

Sejumlah penelitian telah menerapkan pendekatan *deep learning* untuk mendeteksi dan mengklasifikasikan penyakit daun cabai. Rosalina dan Wijaya [1] menerapkan metode *deep learning* untuk mendeteksi penyakit daun cabai, Fu et al. [4] mengembangkan model *Convolutional Neural Network* (CNN) untuk mengenali penyakit daun cabai, dan Pratap dan Suresh [5] menggunakan EfficientNetB4 untuk klasifikasi multi-kelas penyakit daun cabai. Sementara itu, Leite et al. [6] mengkaji deteksi dan penilaian tingkat keparahan *Cercospora leaf spot* pada tanaman cabai. Penelitian tersebut memperlihatkan penerapan pendekatan berbasis pembelajaran mendalam dalam analisis citra tanaman.

Salah satu metode ekstraksi fitur yang dapat digunakan adalah *Histogram of Oriented Gradients* (HOG). Metode ini merepresentasikan karakteristik citra melalui distribusi orientasi gradien pada bagian-bagian lokal sehingga dapat menggambarkan struktur, kontur, dan pola tepi objek [7]. Dalal dan Triggs [7] memperkenalkan HOG sebagai metode representasi fitur berbasis gradien untuk pengenalan objek. Dalam bidang pertanian, Bao et al. [8] menggunakan HOG untuk mengidentifikasi spesies tanaman berdasarkan karakteristik pola daun sedangkan Ochango et al. [9] menerapkannya dalam klasifikasi penyakit daun jagung. Temuan tersebut menunjukkan bahwa HOG dapat digunakan sebagai representasi fitur dalam klasifikasi citra tanaman. Namun, HOG terutama merepresentasikan informasi gradien dan struktur citra sehingga informasi warna tidak direpresentasikan secara langsung melalui fitur HOG standar.

Penggunaan fitur warna bersama fitur bentuk dan tekstur juga telah dikaji dalam penelitian klasifikasi penyakit daun. Kainat et al. [10] menggabungkan fitur Histogram of Oriented Gradients (HOG), *Local Binary Pattern* (LBP), dan fitur warna untuk klasifikasi penyakit daun mentimun. Pada eksperimen pertama yang menggabungkan ketiga jenis fitur tersebut, metode *Fine K-Nearest Neighbor* (KNN) dan *Subspace KNN* masing-masing memperoleh akurasi sebesar 94,30%. Hasil tersebut menunjukkan bahwa kombinasi fitur gradien, tekstur, dan warna dapat digunakan sebagai representasi citra dalam klasifikasi penyakit daun. Meskipun demikian, hasil pada daun mentimun tidak dapat langsung dianggap mewakili kinerja klasifikasi penyakit daun cabai karena perbedaan dataset dan kelas yang digunakan.

Pada tahap klasifikasi, Support Vector Machine (SVM) dapat digunakan untuk mengolah vektor fitur hasil ekstraksi citra. Solihin et al. [11] membandingkan SVM, K-Nearest Neighbor (K-NN), dan *Stochastic Gradient Descent* (SGD) untuk klasifikasi penyakit daun jagung menggunakan fitur HOG. Sari *et al.* [12] menerapkan SVM dengan fitur *Gray Level Co-occurrence Matrix* (GLCM) untuk klasifikasi penyakit daun cabai dan melaporkan akurasi sebesar 88% pada lima kelas. Duka et al. [13] uga mengkaji identifikasi penyakit pada citra daun cabai menggunakan GLCM dan SVM, sedangkan Prajapati et al. [14] menggunakan SVM untuk klasifikasi penyakit tanaman padi berdasarkan karakteristik citra daun. **Studi-studi tersebut menjadi rujukan penerapan metode klasifikasi** berbasis fitur pada citra tanaman dengan karakteristik fitur dan rancangan eksperimen yang berbeda.

Kinerja Support Vector Machine (SVM) dipengaruhi oleh pemilihan hyperparameter, antara lain parameter C dan jenis kernel, serta gamma dan degree pada kernel tertentu. Penelitian ini menggunakan GridSearchCV dengan 3-fold cross-validation untuk mengevaluasi kombinasi hyperparameter berdasarkan nilai accuracy dan memilih konfigurasi model terbaik dari ruang pencarian yang telah ditentukan.

Berdasarkan penelitian terdahulu, penelitian ini bertujuan mengevaluasi kinerja model HOG-SVM yang dioptimasi menggunakan GridSearchCV dalam mengklasifikasikan citra Daun Sehat, Leaf Curl, dan Yellowish. Evaluasi tambahan dilakukan dengan menggabungkan fitur HOG dan histogram warna HSV pada SVM (HOG + HSV-SVM), serta membandingkan kinerja HOG-SVM dengan Vanilla

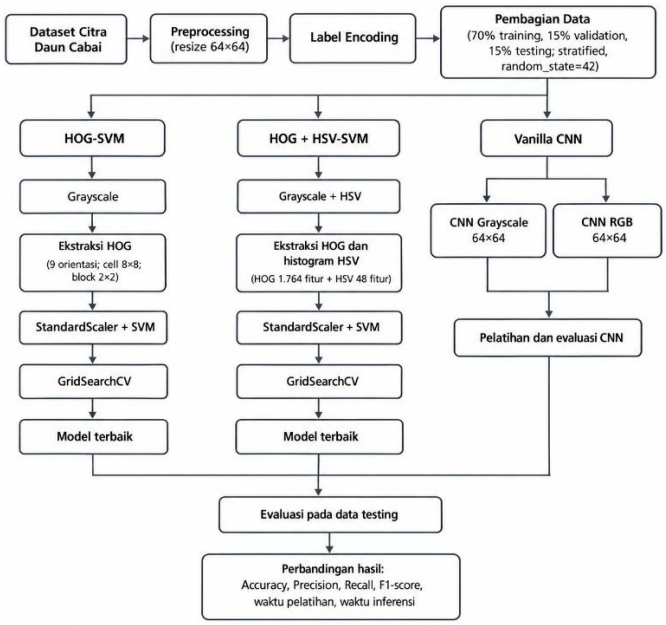
Vol. 11, No. 2, Oktober 2026, Page. 670-688
https://journal.uin-alauddin.ac.id/index.php/instek
p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

Convolutional Neural Network (CNN) menggunakan masukan grayscale dan RGB. Kinerja setiap pendekatan dinilai berdasarkan accuracy, precision, recall, dan F1-score. Hasil evaluasi digunakan untuk membandingkan kinerja model pada dataset dan pembagian data yang sama.

2. METODE PENELITIAN

2.1. Tahap Penelitian

Penelitian ini diawali dengan pengumpulan dataset citra daun cabai, kemudian dilanjutkan dengan preprocessing dan pembagian dataset menjadi data training, validation, dan testing. Pengujian dilakukan melalui tiga pendekatan, yaitu HOG-SVM, HOG + HSV-SVM, dan Vanilla CNN. Pada HOG-SVM, citra grayscale diekstraksi menggunakan fitur Histogram of Oriented Gradients (HOG), sedangkan HOG + HSV-SVM menggabungkan fitur HOG dengan histogram warna HSV. Kedua pendekatan tersebut menggunakan standardisasi fitur dan optimasi hyperparameter SVM melalui GridSearchCV dengan *3-fold cross-validation* untuk memperoleh konfigurasi model terbaik. Sementara itu, Vanilla CNN diuji menggunakan citra grayscale dan RGB. Seluruh model kemudian dievaluasi menggunakan data testing berdasarkan nilai *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*, serta waktu pelatihan dan waktu inferensi.



Gambar 1.Tahapan Penelitian Dilakukan Secara Keseluruhan

2.2. *Dataset*

Dataset yang digunakan dalam penelitian ini merupakan dataset citra daun cabai yang diperoleh dari platform *Kaggle pada akun pengguna Ratnasarii*[15]. Berdasarkan informasi pada halaman dataset, rincian mengenai kondisi akuisisi citra dan versi dataset tidak dijelaskan secara lengkap. Resolusi asli citra sebelum proses penyeragaman ukuran juga tidak dapat dipastikan berdasarkan informasi yang tersedia pada halaman tersebut. Pelabelan dalam penelitian ini mengikuti kategori kelas yang disediakan oleh dataset, tanpa mengklaim adanya verifikasi independen oleh ahli.



Gambar 2. Citra daun cabe dari *dataset*

p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

Dataset terdiri atas tiga kelas, yaitu Daun Sehat, *Leaf Curl*, dan *Yellowish*, dengan masing-masing 400 citra sehingga totalnya 1.200 citra. Gambar 2 memperlihatkan contoh citra dari setiap kelas. Secara visual, kelas Daun Sehat menunjukkan daun berwarna hijau, kelas *Leaf Curl* memperlihatkan perubahan bentuk daun, sedangkan kelas *Yellowish* menunjukkan perubahan warna daun menjadi kekuningan.

Ketiga kelas tersebut dipilih karena merupakan kategori yang tersedia dalam dataset yang digunakan. Setiap kelas terdiri atas 400 citra, sehingga jumlah data pada masing-masing kelas seimbang. Oleh karena itu, eksperimen dibatasi pada kelas Daun Sehat, *Leaf Curl*, dan *Yellowish*, dan hasil penelitian tidak dimaksudkan untuk mewakili seluruh jenis penyakit daun cabai.

Selain itu, dilakukan pemeriksaan duplikasi file citra menggunakan algoritma SHA-256 untuk mengidentifikasi file yang memiliki isi biner identik. Hasil pemeriksaan terhadap 1.200 file gambar menunjukkan terdapat 529 hash unik, 120 kelompok duplikat yang mencakup 791 file, dan tidak ditemukan file yang gagal dibaca. Hasil tersebut menunjukkan adanya sejumlah file citra dengan isi biner identik dalam dataset. Pemeriksaan SHA-256 hanya mengidentifikasi file dengan isi biner yang sama dan tidak mencakup citra yang secara visual serupa.

Seluruh citra diseragamkan menjadi ukuran 64 × 64 piksel sebelum digunakan dalam eksperimen klasifikasi.

2.3. Preprocessing Citra

Tahap *preprocessing* dilakukan untuk menyeragamkan ukuran dan format citra sebelum digunakan dalam proses ekstraksi fitur dan klasifikasi. Seluruh citra diubah ukurannya (*resize*) menjadi 64 × 64 piksel agar memiliki dimensi yang seragam.

Penyesuaian format citra dilakukan berdasarkan konfigurasi metode yang digunakan. Pada HOG-SVM dan Vanilla CNN dengan masukan grayscale, citra dikonversi dari format BGR menjadi grayscale. Konversi tersebut menghasilkan citra satu kanal yang merepresentasikan intensitas piksel dan digunakan sebagai masukan pada proses ekstraksi fitur HOG maupun CNN grayscale.

Pada HOG + HSV-SVM, citra disiapkan dalam dua representasi, yaitu grayscale untuk ekstraksi fitur HOG dan HSV untuk ekstraksi histogram warna. Kedua fitur tersebut selanjutnya digabungkan sebagai representasi citra. Sementara itu, pada Vanilla CNN dengan masukan RGB, citra dikonversi dari format BGR menjadi RGB untuk mempertahankan informasi warna dalam tiga kanal. Kedua konfigurasi masukan CNN tersebut digunakan untuk membandingkan kinerja klasifikasi berdasarkan representasi grayscale dan RGB.

Secara matematis, proses *preprocessing* untuk HOG-SVM dan Vanilla CNN grayscale dinyatakan sebagai berikut:

$$I_{gray} = Grayscale(Resize(I, 64 \times 64) \quad (1)$$

Sedangkan untuk *Vanilla CNN* dengan input RGB, proses *preprocessing* dapat dituliskan sebagai:

$$I_{RGB} = Resize(I, 64 \times 64) \quad (2)$$

Dengan I merupakan citra masukan, I_{gray} merupakan citra *grayscale* berukuran 64 × 64 piksel yang digunakan pada HOG-SVM dan *Vanilla CNN* grayscale, sedangkan I_{RGB} merupakan citra RGB berukuran 64 × 64 piksel yang digunakan pada *Vanilla CNN* RGB.

Penyeragaman citra melalui tahap *preprocessing* juga diterapkan dalam penelitian klasifikasi penyakit daun untuk menyiapkan citra sebelum digunakan pada tahap ekstraksi fitur maupun klasifikasi. Penelitian *Leaf Disease Detection and Classification* menggunakan citra daun sebagai masukan dalam membandingkan beberapa pendekatan klasifikasi berbasis pengolahan citra dan *deep learning* [16].

p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

2.4. Ekstraksi Fitur Histogram of Oriented Gradients (HOG)

Histogram of Oriented Gradients (HOG) merupakan metode ekstraksi fitur yang merepresentasikan karakteristik citra berdasarkan distribusi orientasi gradien. Metode ini digunakan untuk memperoleh informasi mengenai tepi dan pola bentuk objek melalui perubahan intensitas piksel pada citra. Proses ekstraksi HOG meliputi perhitungan gradien, penentuan magnitudo dan orientasi gradien, pembentukan histogram pada setiap *cell*, normalisasi *block*, serta penggabungan fitur menjadi sebuah vektor numerik [7],[17].

Pada penelitian ini, HOG digunakan untuk mengubah citra daun cabai menjadi representasi fitur numerik yang selanjutnya digunakan sebagai masukan pada algoritma SVM. Fitur HOG digunakan untuk menangkap karakteristik tepi dan pola bentuk yang terdapat pada citra daun. Penggunaan HOG untuk klasifikasi penyakit daun juga telah diterapkan pada penelitian sebelumnya, yang menunjukkan bahwa informasi gradien dapat digunakan sebagai representasi fitur pada citra daun [11], [18].

Ekstraksi fitur HOG dilakukan dengan menghitung gradien horizontal dan vertikal pada citra grayscale, kemudian menentukan magnitudo dan orientasi gradien. Citra dibagi menjadi cell berukuran 8 × 8 piksel dengan 9 orientasi, lalu cell digabungkan dalam block 2 × 2 dan dinormalisasi menggunakan L2-Hys. Seluruh histogram yang dihasilkan digabungkan menjadi vektor fitur. Dengan parameter tersebut, setiap citra menghasilkan 1.764 fitur HOG yang selanjutnya digunakan sebagai masukan SVM.

2.5. *Label encoding*

Dataset dalam penelitian ini memiliki tiga kelas citra daun cabai, yaitu Daun Sehat, Leaf Curl, dan Yellowish. Sebelum proses klasifikasi, setiap kelas dikonversi menjadi label numerik menggunakan LabelEncoder agar dapat diproses oleh algoritma SVM. Proses ini menghasilkan pemetaan nilai numerik untuk setiap kelas. Pemetaan tersebut disimpan oleh *LabelEncoder* sehingga label hasil prediksi model dapat dikembalikan ke nama kelas aslinya pada tahap evaluasi.

2.6. Pembagian *Dataset*

Dataset yang terdiri atas 1.200 citra dibagi menjadi tiga subset, yaitu data pelatihan (*training set*) sebesar 70%, data validasi (*validation set*) sebesar 15%, dan data pengujian (*testing set*) sebesar 15%. Pembagian dilakukan secara *stratified* untuk mempertahankan proporsi jumlah citra pada setiap kelas dalam masing-masing subset.

Pembagian data dilakukan secara *stratified* menggunakan parameter *stratify* sehingga proporsi masing-masing kelas tetap dipertahankan pada setiap subset. Nilai random_state=42 digunakan untuk memastikan proses pembagian data dapat direproduksi [19].

Pembagian *dataset* yang sama diterapkan pada seluruh eksperimen, yaitu HOG-SVM, HOG + HSV-SVM, serta Vanilla CNN dengan masukan grayscale dan RGB. Penggunaan subset yang identik bertujuan memastikan bahwa perbandingan kinerja antarmodel dilakukan pada data yang sama, sehingga perbedaan hasil dapat dianalisis berdasarkan metode dan konfigurasi input yang digunakan.

2.7. Standardisasi Fitur

Fitur HOG yang telah diperoleh selanjutnya distandardisasi menggunakan *StandardScaler*. Standardisasi dilakukan untuk menyamakan skala nilai fitur sebelum digunakan dalam proses klasifikasi SVM. Tahap ini penting karena fitur HOG terdiri atas sejumlah nilai numerik yang perlu berada pada skala yang lebih seragam agar proses pembentukan model SVM dapat dilakukan secara konsisten.

StandardScaler melakukan standardisasi dengan mengurangi setiap nilai fitur dengan nilai rata-rata, kemudian membaginya dengan standar deviasi fitur [20]. Proses standardisasi dapat dinyatakan menggunakan persamaan:

$$z = \frac{x-\mu}{\sigma} \tag{3}$$

dengan $z$merupakan nilai fitur setelah standardisasi, $x$merupakan nilai fitur sebelum standardisasi, μmerupakan nilai rata-rata fitur, dan σmerupakan standar deviasi fitur.

Pada implementasi penelitian, *StandardScaler* ditempatkan dalam *pipeline* bersama dengan SVM. Penggunaan *pipeline* membuat proses standardisasi dan klasifikasi menjadi satu rangkaian pemrosesan

p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

sehingga transformasi fitur dapat diterapkan secara konsisten pada data yang digunakan dalam pelatihan, validasi, dan pengujian. Selain itu, penggunaan *pipeline* dalam proses *GridSearchCV* memastikan bahwa proses standardisasi dilakukan sebagai bagian dari setiap proses *cross-validation*, sehingga informasi dari data evaluasi tidak digunakan dalam proses standardisasi data *training*.

2.8. Klasifikasi Menggunakan Support Vector Machine

Fitur HOG yang telah distandardisasi selanjutnya digunakan sebagai masukan untuk algoritma *Support Vector Machine* (SVM). SVM digunakan untuk menentukan batas pemisah (*hyperplane*) antar kelas berdasarkan karakteristik fitur yang diperoleh dari citra daun. Dalam proses klasifikasi, SVM mencari *hyperplane* yang dapat memisahkan data dari kelas yang berbeda dengan mempertimbangkan posisi data terhadap batas pemisah tersebut. Secara umum, persamaan *hyperplane* pada SVM dapat dituliskan sebagai:

$$f(x) = w^T x + b \tag{4}$$

dengan w merupakan vektor bobot, x merupakan vektor fitur, dan b merupakan *bias*.
SVM dapat menggunakan fungsi *kernel* untuk menangani data yang memiliki pola pemisahan yang tidak *linear*. Fungsi *kernel* memungkinkan SVM memetakan data ke ruang fitur dengan dimensi yang lebih tinggi sehingga pola pemisahan antar kelas dapat direpresentasikan dengan lebih sesuai. Dalam penelitian ini, empat jenis *kernel* dibandingkan, yaitu *Linear*, RBF (*Radial Basis Function*), *Polynomial*, dan *Sigmoid*.
Penggunaan SVM dalam klasifikasi penyakit daun telah diterapkan pada penelitian sebelumnya, termasuk pada klasifikasi penyakit daun cabai. Hasil penelitian tersebut menjadi salah satu dasar penggunaan SVM sebagai metode klasifikasi dalam penelitian ini sehingga dapat menunjukkan konfigurasi kernel yang paling sesuai. Penggunaan SVM untuk klasifikasi penyakit daun cabai juga telah dilakukan pada penelitian sebelumnya, antara lain dengan kombinasi fitur tekstur GLCM maupun fitur hasil ekstraksi lainnya [12], [13].

2.9. Optimasi *Hyperparameter* Menggunakan GridSearchCV

Performa SVM dapat dipengaruhi oleh pemilihan *hyperparameter*, seperti **C**, *kernel*, *gamma*, dan *degree*. Oleh karena itu, penelitian ini menggunakan GridSearchCV untuk mencari kombinasi *hyperparameter* yang memberikan nilai *cross-validation accuracy* tertinggi berdasarkan parameter yang telah ditentukan [21].

GridSearchCV melakukan pencarian secara sistematis terhadap seluruh kombinasi *hyperparameter* dalam *parameter grid*. Setiap kombinasi parameter kemudian dievaluasi menggunakan metode *cross-validation*. Parameter yang digunakan dalam proses optimasi ditunjukkan pada Tabel 1.

Tabel 1. Parameter Grid Search untuk Optimasi SVM

Kernel	**C**	**Gamma**	**Degree**
Linear	0.1, 1, 10, 100	–	–
RBF	0.1, 1, 10, 100	scale, 0.01, 0.1	–
Polynomial	0.1, 1, 10, 100	scale, 0.01, 0.1	**2, 3, 4**
Sigmoid	0.1, 1, 10, 100	scale, 0.01, 0.1	–

Parameter C mengatur besarnya penalti terhadap kesalahan klasifikasi. Nilai C yang lebih kecil memberikan regularisasi yang lebih kuat, sedangkan nilai C yang lebih besar memberikan penalti yang lebih besar terhadap kesalahan klasifikasi. Pada *kernel* nonlinier, parameter *gamma* mengatur seberapa besar pengaruh suatu data terhadap data di sekitarnya. Sementara itu, *degree* digunakan pada *kernel Polynomial* untuk menentukan derajat fungsi polinomial [22].

Proses *Grid Search* dilakukan menggunakan 3-fold *cross-validation* dengan *scoring accuracy*. Pada proses ini, data *training* dibagi menjadi tiga bagian (*fold*). Dua bagian digunakan untuk proses pelatihan, sedangkan satu bagian digunakan sebagai validasi internal. Proses tersebut dilakukan secara bergantian hingga setiap bagian digunakan sebagai data validasi internal.

Untuk setiap kombinasi parameter, GridSearchCV menghitung nilai *accuracy* pada masing-masing *fold*, kemudian menghitung rata-ratanya sebagai nilai *mean cross-validation accuracy*. Nilai tersebut digunakan untuk membandingkan performa setiap kombinasi parameter dalam *parameter grid* [21].

2.10. Pemilihan Model dan Evaluasi

Setelah seluruh kombinasi *hyperparameter* diuji menggunakan 3-fold *cross-validation*, GridSearchCV memilih konfigurasi dengan nilai *cross-validation accuracy* tertinggi dari *parameter grid* yang telah ditentukan. Konfigurasi tersebut kemudian digunakan untuk membentuk model SVM yang selanjutnya dievaluasi menggunakan data *validation* dan data *testing*. Data *validation* digunakan untuk melihat kinerja model pada data yang tidak digunakan dalam proses pelatihan, sedangkan data *testing* digunakan untuk memperoleh hasil evaluasi akhir model.

2.11. Validasi Model

Model SVM dengan konfigurasi *hyperparameter* yang dipilih melalui GridSearchCV selanjutnya dievaluasi menggunakan data *validation* dan *testing*. Evaluasi dilakukan untuk mengetahui kemampuan model dalam mengklasifikasikan citra daun cabai berdasarkan kelas sebenarnya. Metrik yang digunakan dalam evaluasi meliputi *accuracy*, *precision*, *recall*, dan *F1-score*.

2.11.1. *Accuracy*

Accuracy digunakan untuk mengetahui proporsi jumlah data yang berhasil diklasifikasikan dengan benar dari seluruh data yang dievaluasi. Nilai *accuracy* dapat dihitung menggunakan persamaan berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

dengan TP (*True Positive*) merupakan jumlah data positif yang diprediksi dengan benar, TN (*True Negative*) merupakan jumlah data negatif yang diprediksi dengan benar, FP (*False Positive*) merupakan jumlah data negatif yang diprediksi sebagai positif, dan FN (*False Negative*) merupakan jumlah data positif yang diprediksi sebagai negatif.

2.11.2. *Precision*

Precision menunjukkan tingkat ketepatan model dalam memberikan prediksi positif. Dengan kata lain, metrik ini menunjukkan proporsi data yang benar-benar termasuk dalam suatu kelas dari seluruh data yang diprediksi sebagai kelas tersebut. *Precision* dapat dihitung menggunakan persamaan:

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

2.11.3. *Recall*

Recall menunjukkan kemampuan model dalam mengenali data yang benar-benar termasuk dalam suatu kelas. Nilai *recall* menggambarkan seberapa banyak data dari suatu kelas yang berhasil ditemukan oleh model. Perhitungannya dapat dinyatakan sebagai:

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

2.11.4. *F1-score*

F1-score merupakan rata-rata harmonis antara *precision* dan *recall*. Metrik ini digunakan untuk memberikan gambaran keseimbangan antara ketepatan prediksi dan kemampuan model dalam menemukan data dari suatu kelas. *F1-score* dapat dihitung menggunakan persamaan:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (8)$$

Pada penelitian ini, nilai *precision*, *recall*, dan *F1-score* dihitung menggunakan metode *weighted average*. Perhitungan tersebut mempertimbangkan jumlah sampel pada masing-masing kelas sehingga kontribusi setiap kelas disesuaikan dengan jumlah datanya.

Penggunaan *accuracy*, *precision*, *recall*, dan *F1-score* sebagai metrik evaluasi juga diterapkan pada penelitian sebelumnya yang menggunakan GridSearchCV dalam proses optimasi SVM [23], [24].

2.12. *Confusion matrix*

Confusion matrix digunakan untuk melihat hasil klasifikasi pada setiap kelas dengan membandingkan kelas sebenarnya dan kelas prediksi. Pada penelitian ini, matriks berukuran 3 × 3 karena terdiri atas kelas Daun Sehat, *Leaf Curl*, dan *Yellowish*. Nilai diagonal menunjukkan prediksi yang benar, sedangkan nilai di luar diagonal menunjukkan kesalahan klasifikasi.

2.13. Pembentukan Model *Vanilla CNN*

Sebagai metode pembanding terhadap HOG-SVM, penelitian ini menggunakan Vanilla Convolutional Neural Network (CNN) sebagai model baseline. CNN dapat mempelajari karakteristik visual citra secara langsung melalui operasi konvolusi tanpa memerlukan ekstraksi fitur HOG secara terpisah. Penggunaan CNN memungkinkan perbandingan antara pendekatan berbasis ekstraksi fitur manual dan pendekatan pembelajaran fitur otomatis.

[25], [26].

Model Vanilla CNN dibangun menggunakan TensorFlow/Keras dengan dua lapisan konvolusi. Lapisan pertama menggunakan Conv2D dengan 32 filter berukuran 3 × 3 dan fungsi aktivasi ReLU, diikuti MaxPooling2D berukuran 2 × 2. Lapisan kedua menggunakan Conv2D dengan 64 filter berukuran 3 × 3 dan fungsi aktivasi ReLU, kemudian diikuti MaxPooling2D berukuran 2 × 2. Keluaran dari kedua lapisan konvolusi diratakan menggunakan Flatten, kemudian diteruskan ke Dense layer dengan 128 neuron dan fungsi aktivasi ReLU. Selanjutnya, lapisan Dropout dengan nilai 0,5 digunakan untuk membantu mengurangi risiko overfitting. Lapisan keluaran menggunakan 3 neuron dengan fungsi aktivasi softmax untuk memprediksi kelas Daun Sehat, Leaf Curl, dan Yellowish.

Model dilatih selama 20 epoch dengan batch size 32 menggunakan optimizer Adam dengan learning rate 0,001 dan loss function sparse categorical crossentropy. Proses pelatihan tidak menggunakan early stopping maupun data augmentation. Untuk input grayscale, dimensi masukan model adalah 64 × 64 × 1, sedangkan untuk input RGB dimensinya 64 × 64 × 3.

Pelatihan dan evaluasi CNN menggunakan pembagian dataset sebagaimana dijelaskan pada Subbab 2.6. Data training digunakan untuk melatih model, data validation digunakan untuk memantau kinerja selama pelatihan, dan data testing digunakan untuk evaluasi akhir. Pengujian dilakukan pada dua konfigurasi input, yaitu grayscale dan RGB, untuk mengevaluasi kinerja model pada kedua representasi citra. Untuk mengamati variasi kinerja akibat perubahan inisialisasi acak, eksperimen tambahan dilakukan menggunakan tiga random seed, yaitu 42, 123, dan 2025. Hasil pengujian tiga seed dilaporkan dalam bentuk rata-rata dan deviasi standar.

2.14. Pembentukan Model HOG + HSV-SVM

Sebagai eksperimen tambahan, penelitian ini menggunakan kombinasi fitur *Histogram of Oriented Gradients* (HOG) dan *Hue, Saturation, Value* (HSV) dengan algoritma *Support Vector Machine* (SVM). HOG merupakan metode ekstraksi fitur yang merepresentasikan karakteristik bentuk dan tepi objek berdasarkan distribusi orientasi gradien pada bagian lokal citra [7]. Sementara itu, HSV merupakan ruang warna yang memisahkan informasi warna (*hue*), kejenuhan warna (*saturation*), dan kecerahan (*value*), sehingga dapat digunakan untuk merepresentasikan karakteristik warna citra [27].

Pada eksperimen ini, fitur HOG diekstraksi dari citra grayscale berukuran 64 × 64 piksel menggunakan konfigurasi 9 orientasi gradien, ukuran *cell* 8 × 8 piksel, ukuran *block* 2 × 2 *cell*, dan normalisasi L2-Hys. Konfigurasi tersebut menghasilkan 1.764 fitur HOG. Selanjutnya, citra dikonversi ke ruang warna HSV menggunakan OpenCV. Histogram warna dihitung pada masing-masing kanal H, S, dan V dengan 16 bin per kanal. Histogram kemudian dinormalisasi menggunakan metode L1 sehingga menghasilkan 48 fitur warna.

p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

Fitur HOG dan histogram HSV digabungkan melalui proses konkatenasi sehingga menghasilkan 1.812 fitur untuk setiap citra. Penggabungan ini bertujuan mengombinasikan informasi gradien dan bentuk daun yang diperoleh dari HOG dengan distribusi warna yang direpresentasikan oleh histogram HSV. Dengan demikian, model menggunakan informasi bentuk dan warna secara bersamaan dalam proses klasifikasi.

Vektor fitur gabungan selanjutnya distandardisasi menggunakan StandardScaler yang ditempatkan dalam Pipeline bersama algoritma SVM. Optimasi *hyperparameter* dilakukan menggunakan GridSearchCV dengan 3-*fold cross-validation* dan metrik akurasi (*accuracy*) sebagai dasar pemilihan konfigurasi terbaik. Ruang pencarian mencakup kernel Linear, RBF, Polynomial, dan Sigmoid dengan nilai parameter C sebesar 0,1; 1; 10; dan 100. Untuk kernel RBF, Polynomial, dan Sigmoid, digunakan parameter gamma berupa scale, 0,01, dan 0,1. Khusus kernel Polynomial, pengujian juga mencakup nilai *degree* 2, 3, dan 4. Total konfigurasi yang diuji sebanyak 64 kandidat atau 192 proses *fitting*.

Model dengan konfigurasi terbaik kemudian digunakan untuk mengklasifikasikan citra ke dalam tiga kelas, yaitu Daun Sehat, *Leaf Curl*, dan *Yellowish*. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, serta *confusion matrix*. Hasil pengujian dibandingkan dengan model HOG-SVM menggunakan data pengujian yang sama untuk mengamati perubahan kinerja setelah penambahan fitur warna.

2.15. Lingkungan Eksperimen

Eksperimen klasifikasi penyakit daun cabai menggunakan metode HOG-SVM dan Vanilla CNN dilakukan menggunakan Python versi 3.12.1 pada sistem operasi Windows 11 64-bit dengan kapasitas RAM sebesar 7,68 GB. Pada metode HOG-SVM, ekstraksi fitur Histogram of Oriented Gradients (HOG) dilakukan menggunakan pustaka scikit-image versi 0.26.0. Selanjutnya, proses standardisasi fitur, klasifikasi menggunakan Support Vector Machine (SVM), dan optimasi hyperparameter melalui GridSearchCV menggunakan pustaka scikit-learn versi 1.4.2.

Sebagai pembanding, model Vanilla CNN dibangun dan dilatih menggunakan TensorFlow versi 2.18.0 dan Keras versi 3.7.0. Pustaka pendukung yang digunakan meliputi NumPy versi 1.26.4 dan SciPy versi 1.14.1 untuk pengolahan data numerik, serta OpenCV versi 4.10.0.84 untuk mendukung pemrosesan citra. Berdasarkan hasil pemeriksaan lingkungan, TensorFlow tidak mendeteksi perangkat GPU. Informasi mengenai lingkungan perangkat lunak dan perangkat keras ini disertakan untuk memperjelas konfigurasi komputasi yang digunakan serta membantu reproduksi eksperimen penelitian.

3. HASIL DAN PEMBAHASAN

3.1. Hasil Preprocessing Citra

Tahap *preprocessing* menghasilkan citra berukuran 64 × 64 piksel dalam format grayscale. Contoh hasil *preprocessing* ditampilkan pada Gambar 3 sebagai ilustrasi citra yang digunakan pada tahap ekstraksi fitur HOG.



Gambar 3. Contoh gambar *resize* dan *grayscale*

Gambar 3 menunjukkan citra daun cabai dalam tiga tahap, yaitu citra asli, hasil *resize* menjadi 64 × 64 piksel, dan hasil konversi ke format *grayscale*. Citra hasil *preprocessing* selanjutnya digunakan sebagai masukan pada tahap ekstraksi fitur HOG.

3.2. Hasil Ekstraksi Fitur HOG

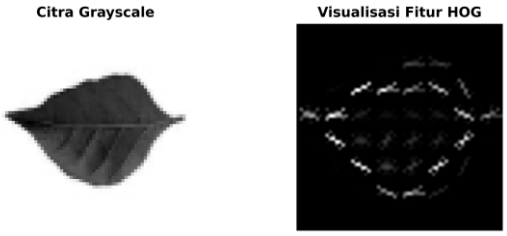
Berdasarkan konfigurasi HOG yang dijelaskan pada bagian Metode, setiap citra menghasilkan 1.764 fitur yang merepresentasikan pola gradien, tepi, dan struktur bentuk daun. Gambar 4

p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

memperlihatkan visualisasi fitur tersebut, yang selanjutnya digunakan sebagai masukan untuk klasifikasi SVM.

Tabel 2. Parameter Ekstraksi Fitur HOG

Parameter	**Nilai**
Ukuran citra	64 × 64 piksel
Orientations	9
Pixels per cell	8 × 8
Cells per block	2 × 2
Block normalization	L2-Hys
Jumlah HOG	**1.764 Fitur**



Gambar 4. Hasil Ekstraksi Fitur HOG pada Citra Daun Cabai

Gambar 4 memperlihatkan visualisasi gradien dan tepi yang direpresentasikan oleh fitur HOG. Setiap citra menghasilkan vektor berukuran 1.764 fitur yang selanjutnya digunakan pada proses klasifikasi HOG-SVM.

3.3. Pembagian *Dataset*

Pembagian *dataset* dilakukan sebagaimana dijelaskan pada Subbab 2.6. Rincian pembagian data ditampilkan pada Tabel 3 sebagai ringkasan jumlah citra pada setiap subset.

Tabel 3. Pembagian Dataset

Subset	**Persentase**	**Jumlah Data**
Training	70%	840
Validation	15%	180
Testing	15%	180
Total	**100%**	**1,200**

Tabel 3 merangkum jumlah citra pada setiap subset yang digunakan dalam eksperimen. Subset training digunakan untuk pelatihan model, validation untuk evaluasi selama pengembangan, dan testing untuk evaluasi akhir.

3.4. Hasil Optimasi *Hyperparameter* Menggunakan GridSearchCV

Hasil optimasi *hyperparameter* menggunakan GridSearchCV ditampilkan pada Tabel 4. Kernel Polynomial memperoleh rata-rata akurasi *cross-validation* tertinggi sebesar 97,38% dengan standar deviasi 0,34%, diikuti kernel RBF sebesar 96,55%, Linear sebesar 95,71%, dan Sigmoid sebesar 90,71%. Berdasarkan hasil tersebut, kernel Polynomial dipilih untuk evaluasi model selanjutnya.

Tabel 4. Hasil Pencarian Hyperparameter SVM

Kernel	**C**	**Gamma**	**Degree**	*Mean* **CV** *Accuracy*	**Std CV** *Accuracy*
Linear	0,1	–	–	95,71%	1,01%
RBF	10	scale	–	96,55%	0,34%
Polynomial	**0,1**	**0,01**	**2**	**97,38%**	**0,34%**

Sigmoid	10	scale	–	90,71%	0,29%

Beberapa hasil optimasi hyperparameter menggunakan Grid Search dengan 3-fold cross-validation menunjukkan bahwa kernel Polynomial memperoleh akurasi rata-rata tertinggi sebesar 97,38% dengan standar deviasi 0,34%. Hasil tersebut lebih tinggi dibandingkan kernel RBF sebesar 96,55%, Linear sebesar 95,71%, dan Sigmoid sebesar 90,71%. Berdasarkan hasil tersebut, konfigurasi kernel Polynomial dengan parameter C = 0,1, gamma = 0,01, dan degree = 2 dipilih sebagai model terbaik pada ruang pencarian hyperparameter yang diuji. Model selanjutnya dievaluasi menggunakan data validasi dan data pengujian.

Tabel 5. Hyperparameter Kernel Polynomial

C	**Gamma**	**Degree**	*Mean* **CV** *Accuracy*	**Std CV** *Accuracy*
0,1	0,01	2	**97,38%**	**0,34%**
0,1	0,1	2	**97,38%**	**0,34%**
1	0,01	2	**97,38%**	**0,34%**
1	0,1	2	**97,38%**	**0,34%**
10	scale	2	**97,38%**	**0,34%**
10	0,01	2	**97,38%**	**0,34%**
10	0,1	2	**97,38%**	**0,34%**
100	scale	2	**97,38%**	**0,34%**
100	0,01	2	**97,38%**	**0,34%**
100	0,1	2	**97,38%**	**0,34%**

3.5. Hasil Evaluasi Model HOG-SVM

Setelah konfigurasi *hyperparameter* diperoleh melalui GridSearchCV, model SVM dengan konfigurasi terpilih dievaluasi menggunakan data *validation* dan *testing*. Data *validation* digunakan untuk mengevaluasi kinerja model sebelum pengujian akhir, sedangkan data *testing* digunakan untuk mengukur kinerja model pada data yang tidak digunakan dalam pelatihan maupun optimasi *hyperparameter*. Hasil evaluasi keseluruhan dan berdasarkan kelas disajikan pada Tabel 6.

Tabel 6. Hasil Evaluasi Model HOG-SVM pada Data Validation dan Testing

Data/Kelas	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Validation (keseluruhan)	97,22	–	–	**–**
Testing (keseluruhan)	96,11	96,19	96,11	**96,12**
Daun Sehat	–	93,65	98,33	**95,93**
Leaf Curl	–	100,00	96,67	**98,31**
Yellowish	–	94,92	93,33	**94,12**

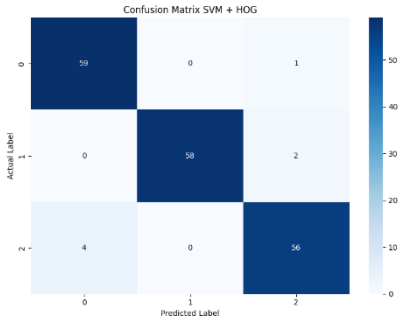
Berdasarkan Tabel 6, model HOG-SVM memperoleh accuracy sebesar 97,22% pada data validation dan 96,11% pada data testing. Dari 180 citra validation, sebanyak 175 citra diklasifikasikan dengan benar. Pada data testing, model menghasilkan precision sebesar 96,19%, recall sebesar 96,11%, dan F1-score sebesar 96,12%.

Berdasarkan hasil per kelas, Leaf Curl memperoleh precision tertinggi sebesar 100,00% dan F1-score sebesar 98,31%, sedangkan Daun Sehat memperoleh recall tertinggi sebesar 98,33%. Kelas Yellowish memiliki recall terendah, yaitu 93,33%. Distribusi kesalahan klasifikasi pada data testing selanjutnya dianalisis menggunakan confusion matrix pada Gambar 5.

3.6. *Confusion matrix* HOG-SVM

Confusion matrix digunakan untuk menganalisis distribusi hasil prediksi model pada masing-masing kelas. Pada penelitian ini, *confusion matrix* terdiri atas tiga kelas, yaitu Daun Sehat, *Leaf Curl*, dan *Yellowish*. Hasil *confusion matrix* model SVM berbasis fitur *Histogram of Oriented Gradients* (HOG) ditunjukkan pada Gambar 5.

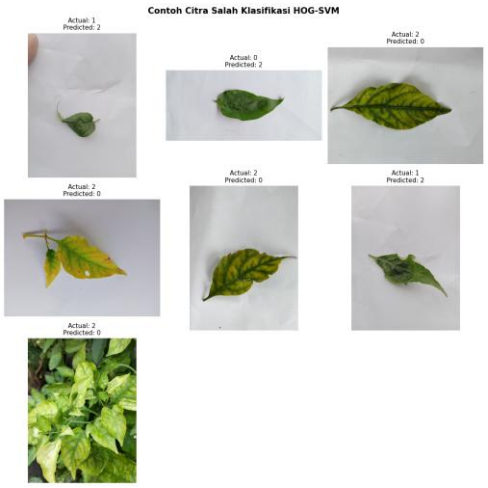
Vol. 11, No. 2, Oktober 2026, Page. 670-688
https://journal.uin-alauddin.ac.id/index.php/instek
p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI



Gambar 5. Confusion Matrix HOG-SVM

Berdasarkan Gambar 5, model berhasil mengklasifikasikan dengan benar 59 dari 60 citra Daun Sehat, 58 dari 60 citra *Leaf Curl*, dan 56 dari 60 citra *Yellowish*. Kesalahan klasifikasi meliputi 1 citra Daun Sehat yang diprediksi sebagai *Yellowish*, 2 citra *Leaf Curl* yang diprediksi sebagai *Yellowish*, serta 4 citra *Yellowish* yang diprediksi sebagai Daun Sehat. Secara keseluruhan, model mengklasifikasikan 173 dari 180 citra *testing* dengan benar.

Kelas *Yellowish* memiliki jumlah kesalahan klasifikasi terbanyak, yaitu 4 citra, dan nilai *recall* terendah sebesar 93,33%, dibandingkan dengan Daun Sehat sebesar 98,33% dan *Leaf Curl* sebesar 96,67%. Temuan ini menunjukkan bahwa model masih mengalami kesulitan dalam mengenali sebagian citra *Yellowish*. Beberapa contoh citra yang salah diklasifikasikan ditampilkan pada Gambar 6 untuk memperlihatkan karakteristik kesalahan prediksi tersebut.



Gambar 6. Contoh citra salah klasifikasi pada model HOG-SVM

Berdasarkan Gambar 6, terdapat tujuh citra yang salah diklasifikasikan. Kesalahan terbanyak terjadi pada kelas *Yellowish* yang diprediksi sebagai Daun Sehat sebanyak empat citra, diikuti dua citra *Leaf Curl* yang diprediksi sebagai *Yellowish* dan satu citra Daun Sehat yang diprediksi sebagai *Yellowish*. Pola tersebut menunjukkan bahwa kesalahan cenderung terjadi pada citra yang memiliki kemiripan bentuk, tekstur, dan pola tepi antarkelas. Pada kasus *Yellowish* diprediksi Daun Sehat, bentuk daun masih menyerupai daun sehat, sedangkan perbedaan warna menjadi karakteristik yang lebih dominan. Kondisi ini menjadi keterbatasan HOG karena fitur HOG merepresentasikan distribusi gradien intensitas, arah gradien, dan struktur tepi, bukan informasi warna secara langsung. Setelah dikonversi ke *grayscale*, perbedaan warna hijau dan kuning direpresentasikan sebagai tingkat keabuan sehingga informasi warna yang dapat membedakan kedua kelas tidak tersedia secara eksplisit. Hal ini sejalan dengan *recall* kelas *Yellowish* sebesar 93,33%.

Kesalahan pada dua citra *Leaf Curl* diprediksi *Yellowish* menunjukkan bahwa kemiripan pola permukaan dan struktur daun dapat menghasilkan distribusi gradien yang berdekatan. Sementara itu, satu citra Daun Sehat diprediksi *Yellowish* menunjukkan bahwa variasi tekstur dan intensitas permukaan

p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

juga dapat menghasilkan pola gradien yang menyerupai kelas *Yellowish*. Dengan demikian, kesalahan pada Gambar 6 tidak sepenuhnya bersifat acak, tetapi berkaitan dengan keterbatasan representasi HOG ketika karakteristik pembeda antarkelas terdapat pada warna serta ketika bentuk dan pola tepinya relatif serupa.

Keterbatasan tersebut diperkuat oleh hasil *Vanilla CNN*. CNN dengan citra *grayscale* memperoleh *F1-score* 95,54%, sedangkan konfigurasi RGB mencapai 98,33%, meningkat 2,79 poin persentase. Hasil ini menunjukkan bahwa informasi warna memberikan informasi tambahan pada *dataset* penelitian, terutama untuk membedakan kelas *Yellowish*. Dengan demikian, HOG tetap efektif dalam merepresentasikan bentuk, struktur, dan pola tepi daun, tetapi untuk gejala yang berkaitan dengan perubahan warna, representasi HOG dapat dipertimbangkan untuk dikombinasikan dengan fitur warna atau representasi citra yang mempertahankan informasi warna.

3.7. Perbandingan Kinerja HOG-SVM dan *Vanilla CNN*

Vanilla Convolutional Neural Network (CNN) digunakan sebagai metode pembanding terhadap HOG-SVM dengan menggunakan pembagian data yang sama sebagaimana dijelaskan pada Subbab 2.6. CNN diuji menggunakan input grayscale dan RGB. Selain pengujian satu eksperimen, dilakukan pengujian tambahan menggunakan tiga *random seed*, yaitu 42, 123, dan 2025, untuk mengevaluasi konsistensi kinerja model. Hasil perbandingan disajikan pada Tabel 7.

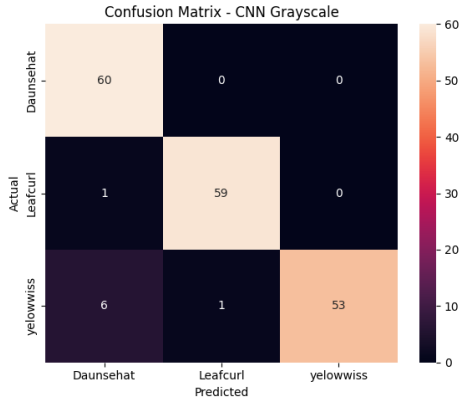
Tabel 7. Perbandingan HOG-SVM dan Vanilla CNN

Metode	Input	Jenis Pelaporan	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Training Time (s)	Inference/Image (ms)
HOG-SVM	Grayscale 64×64	Satu eksperimen	96,11	96,19	96,11	96,12	55,49	**0,426**
Vanilla CNN	Grayscale 64×64	Satu eksperimen	95,56	95,96	95,56	95,54	33,71	**1,977**
Vanilla CNN	RGB 64×64	Satu eksperimen	98,33	98,37	98,33	98,33	32,53	**1,797**
Vanilla CNN	Grayscale 64×64	Rata-rata ± SD (3 seed)	93,33 ± 3,09	93,95 ± 2,53	93,33 ± 3,09	93,18 ± 3,30	31,80 ± 2,45	**1,47 ± 0,18**
Vanilla CNN	RGB 64×64	Rata-rata ± SD (3 seed)	98,89 ± 0,96	98,90 ± 0,95	98,89 ± 0,96	98,89 ± 0,96	31,41 ± 1,22	**1,35 ± 0,00**

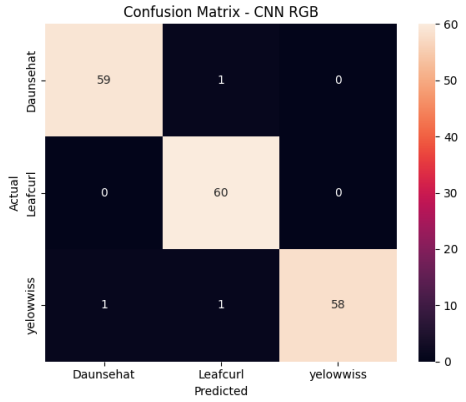
Pada pengujian satu eksperimen, HOG-SVM memperoleh accuracy sebesar 96,11% dan F1-score sebesar 96,12%, sedangkan CNN grayscale memperoleh accuracy sebesar 95,56% dan F1-score sebesar 95,54%. Selisih F1-score sebesar 0,58 poin persentase menunjukkan bahwa kinerja kedua metode relatif berdekatan. Sementara itu, CNN RGB memperoleh accuracy dan F1-score sebesar 98,33%, dengan F1-score 2,79 poin persentase lebih tinggi daripada CNN grayscale. Hasil ini mengindikasikan adanya hubungan antara penggunaan input RGB dan peningkatan kinerja CNN pada dataset yang digunakan, tetapi belum membuktikan bahwa informasi warna merupakan satu-satunya faktor yang memengaruhi hasil klasifikasi.

Dari sisi komputasi, waktu pelatihan HOG-SVM sebesar 55,49 detik lebih lama daripada CNN grayscale dan RGB, masing-masing sebesar 33,71 dan 32,53 detik. Namun, perbandingan ini perlu mempertimbangkan perbedaan prosedur pelatihan karena waktu HOG-SVM mencakup optimasi GridSearchCV, sedangkan CNN dilatih selama 20 epoch pada satu konfigurasi model. Waktu inferensi HOG-SVM sebesar 0,426 ms per citra juga lebih rendah daripada CNN grayscale (1,977 ms) dan RGB (1,797 ms). Nilai tersebut tidak mencakup ekstraksi fitur HOG dari citra mentah sehingga tidak merepresentasikan keseluruhan waktu pemrosesan hingga prediksi dihasilkan.

p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI



Gambar 7. Confusion Matrix Vanilla CNN dengan Input Grayscale



Gambar 8. Confusion Matrix Vanilla CNN dengan Input RGB

Confusion matrix pada Gambar 7 menunjukkan bahwa CNN grayscale mengklasifikasikan dengan benar 172 dari 180 citra pengujian. Terdapat delapan kesalahan, yaitu satu citra Leaf Curl diprediksi sebagai Daun Sehat, enam citra Yellowish diprediksi sebagai Daun Sehat, dan satu citra Yellowish diprediksi sebagai Leaf Curl. Pada konfigurasi RGB (Gambar 8), sebanyak 177 dari 180 citra diklasifikasikan dengan benar, dengan tiga kesalahan: satu citra Daun Sehat diprediksi sebagai Leaf Curl, satu citra Yellowish sebagai Daun Sehat, dan satu citra Yellowish sebagai Leaf Curl. Kesalahan pada kedua konfigurasi menunjukkan bahwa kelas Yellowish masih dapat tertukar dengan kelas lain. Jumlah kesalahan yang lebih sedikit pada CNN RGB konsisten dengan kemungkinan bahwa informasi warna membantu membedakan kelas dalam dataset ini, tetapi temuan tersebut belum dapat digeneralisasikan ke dataset dengan kondisi akuisisi berbeda.

Pengujian tiga *random seed* menunjukkan bahwa CNN grayscale memperoleh rata-rata accuracy sebesar 93,33 ± 3,09% dan F1-score sebesar 93,18 ± 3,30%. CNN RGB memperoleh rata-rata accuracy sebesar 98,89 ± 0,96% dan F1-score sebesar 98,89 ± 0,96%. Pada tiga seed yang diuji, CNN RGB menunjukkan rata-rata kinerja lebih tinggi dan variasi hasil lebih kecil dibandingkan CNN grayscale. Hasil ini memperkuat pola yang terlihat pada satu eksperimen, tetapi kesimpulannya tetap terbatas pada dataset, konfigurasi model, dan jumlah seed yang digunakan.

Secara keseluruhan, CNN RGB menghasilkan kinerja klasifikasi tertinggi di antara konfigurasi yang dibandingkan pada penelitian ini, sedangkan HOG-SVM memiliki waktu inferensi per citra terendah setelah fitur HOG tersedia. HOG-SVM dan CNN grayscale menunjukkan kinerja yang relatif berdekatan pada pengujian satu eksperimen. Oleh karena itu, pemilihan metode perlu mempertimbangkan kinerja klasifikasi serta cakupan proses komputasi yang diukur, tanpa menetapkan satu metode sebagai pemenang mutlak.

3.8. Evaluasi Model HOG + HSV-SVM

Eksperimen HOG + HSV-SVM menghasilkan konfigurasi terbaik menggunakan kernel Linear dengan nilai C = 0,1 dan rata-rata akurasi cross-validation sebesar 97,86%. Pada data validation, model memperoleh accuracy sebesar 97,78%, precision 97,83%, recall 97,78%, dan F1-score 97,79%. Sementara itu, hasil pengujian pada data testing dan perbandingan kinerja serta waktu komputasi dengan model lainnya disajikan pada Tabel 9.

Tabel 9. Kinerja dan Waktu Komputasi Model

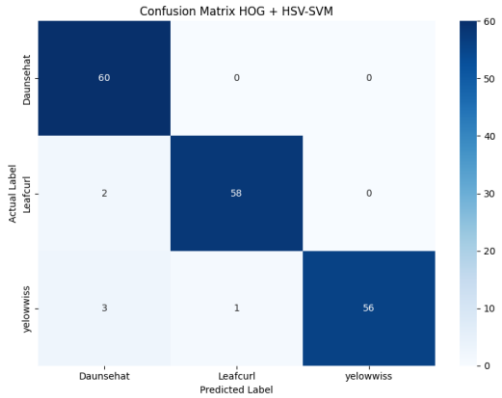
Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Training (s)	Inferensi (ms/citra)
HOG-SVM	96,11	96,19	96,11	96,12	55,487928	**0,426262**
HOG + HSV-SVM	96,67	96,87	96,67	96,68	52,892265	**0,146603**
CNN Grayscale	95,56	95,96	95,56	95,54	33,713100	**1,977230**
CNN RGB	98,33	98,37	98,33	98,33	32,530400	**1,797265**

Berdasarkan Tabel 9, HOG + HSV-SVM memperoleh accuracy sebesar 96,67%, precision 96,87%, recall 96,67%, dan F1-score 96,68% pada data testing. Dibandingkan dengan HOG-SVM, penambahan fitur HSV meningkatkan accuracy dan recall sebesar 0,56 poin persentase, precision sebesar 0,68 poin persentase, serta F1-score sebesar 0,56 poin persentase. Peningkatan tersebut menunjukkan adanya perbaikan kinerja pada eksperimen ini, meskipun selisih hasilnya relatif kecil.

Jika dibandingkan dengan Vanilla CNN, HOG + HSV-SVM memperoleh hasil lebih tinggi daripada CNN grayscale pada accuracy dan F1-score, tetapi masih berada di bawah CNN RGB. CNN RGB menghasilkan accuracy tertinggi sebesar 98,33% dan F1-score sebesar 98,33%. Hasil ini menunjukkan bahwa CNN RGB memberikan kinerja terbaik pada pengujian yang dilakukan, sedangkan HOG + HSV-SVM menunjukkan peningkatan dibandingkan HOG-SVM awal.

Dari sisi waktu komputasi, HOG + HSV-SVM mencatat waktu training sebesar 52,892265 detik dan inferensi sebesar 0,146603 ms per citra. Waktu inferensi tersebut lebih rendah dibandingkan HOG-SVM, CNN grayscale, dan CNN RGB. Namun, perbandingan waktu perlu diperhatikan karena pengukuran inferensi model berbasis HOG belum mencakup ekstraksi fitur, sedangkan waktu training SVM mencakup GridSearchCV dan waktu training CNN mencakup 20 epoch. Oleh karena itu, angka tersebut belum dapat digunakan untuk menyimpulkan perbandingan efisiensi keseluruhan kedua pendekatan secara langsung.

Untuk mengetahui distribusi kesalahan klasifikasi pada setiap kelas, confusion matrix model HOG + HSV-SVM ditampilkan pada Gambar 9.



Gambar 9. Confusion Matrix HOG + HSV-SVM

Berdasarkan Gambar 9, seluruh 60 citra Daun Sehat berhasil diklasifikasikan dengan benar. Pada kelas *Leaf Curl*, sebanyak 58 citra diklasifikasikan dengan benar, sedangkan 2 citra lainnya diprediksi sebagai Daun Sehat. Sementara itu, pada kelas *Yellowish*, sebanyak 56 citra diklasifikasikan dengan benar, 3 citra diprediksi sebagai Daun Sehat, dan 1 citra diprediksi sebagai *Leaf Curl*.

Hasil tersebut menunjukkan bahwa recall kelas Daun Sehat mencapai 100%, *Leaf Curl* sebesar 96,67%, dan *Yellowish* sebesar 93,33%. Kesalahan klasifikasi paling banyak ditemukan pada kelas *Yellowish*. Hal ini mengindikasikan bahwa karakteristik citra pada kelas tersebut masih memiliki kemiripan dengan kelas lain berdasarkan fitur yang digunakan. Meskipun penggabungan HOG dan HSV menghasilkan peningkatan kinerja dibandingkan HOG-SVM, hasil ini belum cukup untuk menyimpulkan bahwa penambahan fitur warna selalu meningkatkan akurasi pada dataset atau kondisi citra yang berbeda.

3.9. Pembahasan

Hasil penelitian menunjukkan bahwa model Histogram of Oriented Gradients–Support Vector Machine (HOG-SVM) yang dioptimasi menggunakan GridSearchCV mampu mengklasifikasikan citra daun cabai ke dalam tiga kelas, yaitu Daun Sehat, *Leaf Curl*, dan *Yellowish*. Konfigurasi terbaik menggunakan kernel *Polynomial* dengan (C=0,1), (\gamma=0,01), dan *degree* = 2, dengan *mean cross-validation accuracy* sebesar 97,38% dan standar deviasi 0,34%. Kernel *Polynomial* memperoleh rata-rata akurasi validasi silang tertinggi, sedangkan kernel *Sigmoid* menghasilkan 90,71%. Hasil ini menunjukkan bahwa pemilihan kernel dan konfigurasi *hyperparameter* berkaitan dengan kinerja model pada ruang pencarian yang digunakan. Temuan tersebut sejalan dengan Ardiansyah et al. [23] dan

Siregar [24], yang menerapkan GridSearchCV untuk optimasi SVM pada permasalahan klasifikasi lainnya. Namun, hasil ini tidak membuktikan bahwa kernel *Polynomial* selalu lebih unggul pada seluruh dataset penyakit daun cabai.

Kernel Polynomial memungkinkan SVM merepresentasikan hubungan nonlinier antarfitur HOG. Pada penelitian ini, konfigurasi terpilih adalah C = 0,1, gamma = 0,01, dan degree = 2, dengan mean cross-validation accuracy sebesar 97,38% dan standar deviasi 0,34%. Kesamaan hasil pada beberapa kombinasi hyperparameter perlu dipahami berdasarkan interaksi parameter kernel, bukan sebagai bukti bahwa setiap parameter secara individual menghasilkan kinerja yang berbeda. Dengan coef0 = 0, perubahan gamma memengaruhi skala fungsi kernel Polynomial dan dapat berinteraksi dengan C dalam menentukan batas keputusan. Apabila beberapa kandidat menghasilkan skor tertinggi yang sama, pemilihan C = 0,1 oleh GridSearchCV mengikuti urutan kandidat dalam ruang pencarian. Oleh karena itu, hasil ini tidak dapat digunakan untuk menyimpulkan bahwa C = 0,1 merupakan nilai yang paling optimal secara universal. Kinerja yang diperoleh tetap berlaku dalam batas dataset, pembagian data, dan ruang pencarian hyperparameter yang digunakan. Dalal dan Triggs [7] menjelaskan bahwa normalisasi blok pada HOG membantu meningkatkan ketahanan representasi gradien terhadap variasi pencahayaan. Namun, HOG tetap merepresentasikan gradien intensitas sehingga kemiripan struktur antarobjek dapat menjadi tantangan klasifikasi. Hal ini berkaitan dengan penelitian Solihin et al. [11] mengenai SVM berbasis fitur HOG untuk klasifikasi daun jagung serta Sari et al. [12] yang menerapkan SVM dengan fitur tekstur GLCM pada daun cabai. Penelitian tersebut menjadi konteks pendukung penggunaan metode berbasis ekstraksi fitur, tetapi tidak membuktikan bahwa satu metode selalu lebih unggul daripada metode lainnya.

Evaluasi menghasilkan *accuracy* sebesar 97,22% pada data validasi dan 96,11% pada data *testing*, dengan *precision* 96,19%, *recall* 96,11%, dan *F1-score* 96,12%. Selisih akurasi validasi dan pengujian sebesar 1,11 poin persentase menunjukkan bahwa kedua hasil relatif berdekatan. Kinerja ini sebanding dengan pendekatan berbasis fitur yang dilaporkan dalam penelitian klasifikasi penyakit daun cabai menggunakan GLCM-SVM oleh Sari et al. [12] dan identifikasi penyakit daun cabai oleh Duka dan Ado [13]. Meskipun demikian, perbandingan akurasi perlu mempertimbangkan perbedaan dataset, pembagian data, jumlah kelas, dan prosedur evaluasi. Dengan demikian, hasil ini menunjukkan kinerja HOG-SVM yang baik pada dataset yang digunakan, tetapi belum memastikan generalisasinya pada dataset eksternal.

Berdasarkan *confusion matrix* pada Gambar 6, kesalahan klasifikasi terbanyak berasal dari kelas *Yellowish* yang diprediksi sebagai Daun Sehat. Sebanyak empat citra *Yellowish* salah diklasifikasikan sebagai Daun Sehat, sedangkan satu citra Daun Sehat dan dua citra *Leaf Curl* salah diprediksi sebagai *Yellowish*. Hal ini tercermin pada *recall* kelas *Yellowish* sebesar 93,33%, terendah di antara ketiga kelas. Salah satu kemungkinan penyebabnya adalah kemiripan bentuk daun dan struktur tepi antara kelas *Yellowish* dan Daun Sehat. Karena HOG merepresentasikan gradien intensitas, bukan warna secara langsung, konversi ke *grayscale* menghilangkan representasi warna hijau dan kuning yang eksplisit. Kondisi ini berpotensi menyulitkan klasifikasi ketika bentuk dan struktur tepi antarkelas serupa, meskipun hilangnya informasi warna belum dapat dipastikan sebagai satu-satunya penyebab kesalahan.

Perbandingan dengan *Vanilla Convolutional Neural Network* (CNN) menunjukkan bahwa CNN *grayscale* memperoleh *accuracy* 95,56% dan *F1-score* 95,54%, sedangkan CNN RGB memperoleh *accuracy* dan *F1-score* sebesar 98,33%. Dengan demikian, *F1-score* CNN RGB lebih tinggi 2,79 poin persentase daripada CNN *grayscale*. Sementara itu, *F1-score* HOG-SVM sebesar 96,12%, berbeda 0,58 poin persentase dari CNN *grayscale*. Hasil ini menunjukkan bahwa kinerja HOG-SVM dan CNN *grayscale* relatif berdekatan, sedangkan CNN RGB memperoleh skor tertinggi di antara ketiga konfigurasi tersebut. Penggunaan input RGB berpotensi memberikan informasi yang bermanfaat bagi klasifikasi, tetapi perbedaan skor belum membuktikan bahwa warna merupakan satu-satunya faktor penyebab peningkatan karena representasi fitur dan mekanisme pembelajaran kedua pendekatan juga berbeda.

Sebagai evaluasi tambahan, pengujian *Vanilla CNN* dilakukan menggunakan tiga *random seed* untuk mengamati variasi kinerja model. Hasil menunjukkan bahwa CNN RGB memperoleh rata-rata *accuracy* sebesar 98,89 ± 0,96% dan *F1-score* sebesar 98,89 ± 0,96%, sedangkan CNN grayscale memperoleh *accuracy* sebesar 93,33 ± 3,09% dan *F1-score* sebesar 93,18 ± 3,30%. Standar deviasi yang lebih rendah pada CNN RGB menunjukkan bahwa variasi hasilnya lebih kecil dibandingkan CNN grayscale pada ketiga *seed* yang diuji. Hasil ini memperkuat temuan bahwa input RGB memiliki kinerja

rata-rata lebih tinggi dan lebih konsisten pada eksperimen ini. Namun, hasil tersebut terbatas pada tiga *seed* dan konfigurasi model yang digunakan.

Eksperimen HOG + HSV-SVM menunjukkan bahwa kombinasi fitur HOG dan histogram warna HSV memperoleh *accuracy* 96,67%, *precision* 96,87%, *recall* 96,67%, dan *F1-score* 96,68% pada data *testing*. Dibandingkan HOG-SVM, penambahan fitur HSV meningkatkan *accuracy* dan *F1-score* masing-masing sebesar 0,56 poin persentase. Hasil ini menunjukkan adanya peningkatan kinerja pada dataset dan konfigurasi eksperimen yang digunakan, tetapi belum menjamin hasil serupa pada dataset lain. Berdasarkan *confusion matrix*, seluruh citra Daun Sehat diklasifikasikan dengan benar. Namun, dua citra *Leaf Curl* diprediksi sebagai Daun Sehat, sedangkan tiga citra *Yellowish* diprediksi sebagai Daun Sehat dan satu citra sebagai *Leaf Curl*. Dengan demikian, penambahan fitur warna belum sepenuhnya mengatasi kesalahan klasifikasi pada kelas dengan karakteristik visual serupa.

Dari sisi komputasi, model HOG-SVM dan HOG + HSV-SVM menunjukkan waktu inferensi per citra yang lebih rendah dibandingkan CNN grayscale dan CNN RGB pada pengukuran yang dilakukan. Namun, hasil tersebut perlu diinterpretasikan secara hati-hati karena pengukuran inferensi kedua model SVM belum mencakup proses prapemrosesan dan ekstraksi fitur dari citra mentah. Dengan demikian, waktu yang dilaporkan belum merepresentasikan keseluruhan proses mulai dari masukan citra hingga menghasilkan prediksi. Pada aspek pelatihan, waktu yang dibutuhkan model SVM mencakup proses optimasi hyperparameter menggunakan GridSearchCV, sedangkan waktu pelatihan CNN dihitung berdasarkan 20 epoch pada satu konfigurasi model. Perbedaan cakupan pengukuran tersebut menyebabkan waktu pelatihan kedua pendekatan tidak dapat dibandingkan secara langsung. Oleh karena itu, evaluasi efisiensi komputasi secara menyeluruh masih memerlukan pengukuran dengan tahapan pemrosesan yang setara, termasuk prapemrosesan, ekstraksi fitur, proses prediksi, serta pengujian pada perangkat target dan pengukuran kebutuhan memori.

Secara keseluruhan, hasil penelitian menunjukkan bahwa HOG-SVM yang dioptimasi menggunakan GridSearchCV memperoleh *mean cross-validation accuracy* sebesar 97,38% dan *accuracy* pengujian sebesar 96,11%. Eksperimen HOG + HSV-SVM memperlihatkan peningkatan kinerja setelah penambahan fitur warna, sedangkan CNN RGB menghasilkan skor tertinggi dalam perbandingan ini. Temuan tersebut menegaskan pentingnya pemilihan representasi fitur dalam klasifikasi penyakit daun cabai. Penelitian selanjutnya dapat mengevaluasi kombinasi fitur bentuk, tekstur, dan warna secara sistematis pada variasi pencahayaan, latar belakang, dan tingkat keparahan gejala. Pendekatan hibrida yang menggabungkan CNN sebagai ekstraktor fitur dan SVM sebagai klasifikator juga dapat dipertimbangkan [8], [25].

Penelitian ini menggunakan satu dataset yang terdiri atas 1.200 citra daun cabai dari platform Kaggle[15], dengan tiga kelas yang tersedia: Daun Sehat, *Leaf Curl*, dan *Yellowish*. Informasi mengenai kondisi akuisisi citra, resolusi asli, dan versi dataset tidak dijelaskan secara lengkap pada sumber yang digunakan. Selain itu, label mengikuti kategori yang disediakan oleh dataset dan tidak dinyatakan telah diverifikasi secara independen oleh ahli. Oleh karena itu, hasil penelitian tidak dapat langsung digeneralisasikan pada citra yang diambil dalam kondisi lapangan yang berbeda atau pada jenis penyakit daun cabai lainnya. Ruang pencarian *hyperparameter* GridSearchCV juga dibatasi pada kombinasi nilai yang telah ditentukan, sehingga konfigurasi di luar ruang tersebut belum dievaluasi. Pengujian pada dataset eksternal, variasi kondisi pengambilan citra, dan cakupan kelas penyakit yang lebih luas diperlukan untuk mengetahui ketahanan model terhadap variasi data yang lebih beragam.

Secara keseluruhan, optimasi *hyperparameter* menggunakan GridSearchCV menghasilkan model HOG-SVM dengan *mean cross-validation accuracy* sebesar 97,38% dan *accuracy* pengujian sebesar 96,11%. Perbandingan dengan CNN *grayscale*, CNN RGB, dan HOG + HSV-SVM menunjukkan bahwa setiap pendekatan memiliki hasil yang berbeda pada dataset yang digunakan. Kontribusi penelitian ini mencakup evaluasi sistematis HOG-SVM yang dioptimasi, pengujian tambahan fitur warna HSV, serta perbandingan dengan CNN pada input *grayscale* dan RGB. Analisis kesalahan juga mengidentifikasi kelas *Yellowish* sebagai salah satu kelas yang masih menghadapi kesalahan klasifikasi. Temuan tersebut memberikan dasar untuk mengevaluasi representasi warna dan struktur citra lebih lanjut tanpa mengasumsikan bahwa satu karakteristik visual merupakan satu-satunya penyebab kesalahan model.

4. KESIMPULAN

Penelitian ini mengevaluasi klasifikasi citra daun cabai menggunakan Histogram of Oriented Gradients–Support Vector Machine (HOG-SVM) dengan optimasi hyperparameter menggunakan GridSearchCV. Konfigurasi terbaik pada ruang pencarian yang diuji menggunakan kernel Polynomial dengan C = 0,1, gamma = 0,01, dan degree = 2, yang menghasilkan mean cross-validation accuracy sebesar 97,38%. Pada data testing, model memperoleh accuracy sebesar 96,11% dan F1-score sebesar 96,12%. Hasil tersebut menunjukkan bahwa HOG-SVM mampu mengklasifikasikan citra Daun Sehat, Leaf Curl, dan Yellowish dengan kinerja yang baik pada dataset penelitian.

Eksperimen tambahan menunjukkan bahwa penggabungan fitur HOG dan histogram warna HSV menghasilkan accuracy sebesar 96,67% dan F1-score sebesar 96,68%, sedikit lebih tinggi dibandingkan HOG-SVM tanpa fitur warna. Sementara itu, perbandingan dengan Vanilla CNN menunjukkan bahwa HOG-SVM memiliki kinerja yang relatif sebanding dengan CNN grayscale pada pengujian satu eksperimen, sedangkan CNN RGB memperoleh F1-score tertinggi sebesar 98,33%. Hasil pengujian tiga random seed juga menunjukkan bahwa CNN RGB memperoleh rata-rata F1-score sebesar 98,89 ± 0,96%, lebih tinggi dibandingkan CNN grayscale sebesar 93,18 ± 3,30%. Temuan ini menunjukkan bahwa representasi citra yang mempertahankan informasi warna berpotensi memberikan informasi tambahan dalam klasifikasi penyakit daun cabai, meskipun pengaruhnya masih perlu dikaji pada kondisi data yang lebih beragam.

Penelitian ini memiliki keterbatasan karena menggunakan satu dataset dengan tiga kategori citra yang labelnya mengikuti informasi dari sumber dataset dan belum diverifikasi secara independen oleh ahli. Selain itu, informasi mengenai kondisi akuisisi dan resolusi asli citra tidak tersedia secara lengkap. Penelitian selanjutnya disarankan untuk melakukan pengujian pada dataset eksternal, memperluas variasi kelas dan kondisi pengambilan citra, serta mengevaluasi kombinasi fitur bentuk, tekstur, dan warna dengan prosedur eksperimen yang konsisten. Pengukuran waktu komputasi juga perlu mencakup seluruh tahapan pemrosesan citra agar perbandingan efisiensi antarmodel dapat dilakukan secara lebih menyeluruh.

DAFTAR PUSTAKA

[1] R. Rosalina and A. Wijaya, "Pendeteksian Penyakit pada Daun Cabai dengan Menggunakan Metode Deep Learning," *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 3, pp. 452–461, 2020, doi: 10.28932/jutisi.v6i3.2857.

[2] M. L. Sa'idi and L. N. Huda, "Deteksi penyakit pada tanaman cabai menggunakan teknologi pengolahan citra dan machine learning," *J. Informatics Comput. Res.*, vol. 1, no. 1, pp. 1–5, 2024.

[3] L. C. Ngugi, M. Abelwahab, and M. Abo-Zahhad, "Recent advances in image processing techniques for automated leaf pest and disease recognition – A review," *Inf. Process. Agric.*, vol. 8, no. 1, pp. 27–51, 2021, doi: 10.1016/j.inpa.2020.04.004.

[4] Y. Fu, L. Guo, and F. Huang, "A lightweight CNN model for pepper leaf disease recognition in a human palm background," *Heliyon*, vol. 10, no. 12, p. e33447, 2024, doi: 10.1016/j.heliyon.2024.e33447.

[5] V. K. Pratap and N. Suresh Kumar, "High-precision multiclass classification of chili leaf disease through customized EffecientNetB4 from chili leaf images," *Smart Agric. Technol.*, vol. 5, no. April, p. 100295, 2023, doi: 10.1016/j.atech.2023.100295.

[6] D. V. Leite, A. V. de Brito, G. G. Faccioli, and G. Haddad Souza Vieira, "Deep Learning Models for Detection and Severity Assessment of Cercospora Leaf Spot (Cercospora capsici) in Chili Peppers Under Natural Conditions," *Plants*, vol. 14, no. 13, 2025, doi: 10.3390/plants14132011.

[7] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, IEEE, 2005, pp. 886–893. doi: 10.1109/CVPR.2005.177.

[8] T. Q. Bao, N. T. T. Kiet, T. Q. Dinh, and H. X. Hiep, "Plant species identification from leaf patterns using histogram of oriented gradients feature space and convolution neural networks," *J. Inf. Telecommun.*, vol. 4, no. 2, pp. 140–150, 2020, doi: 10.1080/24751839.2019.1666625.

[9] V. M. Ochango, G. M. Wambugu, and J. G. Ndia, "Feature Extraction using Histogram of Oriented Gradients for Image Classification in Maize Leaf Diseases," *Int. J. Comput. Inf.*

Technol., vol. 11, no. 2, pp. 54–63, 2022.

[10] J. Kainat, S. S. Ullah, F. S. Alharithi, R. Alroobaea, S. Hussain, and S. Nazir, "Blended Features Classification of Leaf-Based Cucumber Disease Using Image Processing Techniques," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/9736179.

[11] F. Solihin, M. Syarief, E. M. S. Rochman, and A. Rachmad, "Comparison of Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), and Stochastic Gradient Descent (SGD) for Classifying Corn Leaf Disease based on Histogram of Oriented Gradients (HOG) Feature Extraction," *Elinvo (Electronics, Informatics, Vocat. Educ.*, vol. 8, no. 1, pp. 121–129, 2023, doi: 10.21831/elinvo.v8i1.55759.

[12] Y. Sari, A. R. Baskara, and R. Wahyuni, "Classification of Chili Leaf Disease Using the Gray Level Co-occurrence Matrix (GLCM) and the Support Vector Machine (SVM) Methods," *2021 6th Int. Conf. Informatics Comput. ICIC 2021*, 2021, doi: 10.1109/ICIC54025.2021.9632920.

[13] I. O. Duka and H. A. Ado, "Identifikasi Penyakit Tanaman Citra Daun Cabe Menggunakan Gray Level Co-Occurrencce Matrix Dan Support Vector Machine," *Neptunus J. Ilmu Komput. Dan Teknol. Inf.*, vol. 2, no. 2, pp. 41–52, 2024, doi: 10.61132/neptunus.v2i2.86.

[14] H. B. Prajapati, J. P. Shah, and V. B. Dabhi, "Detection and Classification of Rice Plant Diseases," *Intell. Decis. Technol.*, vol. 11, no. 3, pp. 357–373, 2017, doi: 10.3233/IDT-170301.

[15] Ratnasarii, "penyakitdauncabai," Kaggle. [Online]. Available: https://www.kaggle.com/datasets/ratnasarii/penyakitdauncabai

[16] B. V. Nikith, N. K. S. Keerthan, M. S. Praneeth, and T. Amrita, "Leaf Disease Detection and Classification," *Procedia Comput. Sci.*, vol. 218, pp. 291–300, 2023, doi: 10.1016/j.procs.2023.01.011.

[17] S. Developers, "Histogram of Oriented Gradients." Accessed: Oct. 26, 2026. [Online]. Available: https://scikit-image.org/docs/stable/auto_examples/features_detection/plot_hog.html

[18] P. Bose, S. Dutta, V. Goyal, and S. K. Bandyopadhyay, "Leaf Diseases Detection of Medicinal Plants Based on Support Vector Machine Classification Algorithm," *J. Pharm. Res. Int.*, vol. 33, no. 42A, pp. 111–119, 2021, doi: 10.9734/JPRI/2021/v33i42A32391.

[19] Scikit-learn Developers, "train_test_split." Accessed: Sep. 16, 2026. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html

[20] Scikit-learn Developers, "StandardScaler." Accessed: Sep. 16, 2026. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

[21] S. Developers, "GridSearchCV." Accessed: Sep. 16, 2026. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html

[22] S. Developers, "Support Vector Machines." Accessed: Sep. 16, 2026. [Online]. Available: https://scikit-learn.org/stable/modules/svm.html

[23] D. P. Ardiansyah, R., Pamungkas, M. A. D. Widyadara, and A. Santika, "Analysis of the Effect of GridSearchCV Hyperparameter Tuning Strategy with Variations in Scoring, Test Size, and Random State on Water Quality Classification Performance Using SVM RBF Kernel," *Multitek Indones. J. Ilm.*, vol. 20, no. 1, pp. 104–115, 2026, doi: 10.24269/mtkind.v20i1.13841.

[24] R. K. Siregar, "Klasifikasi Status Penerimaan Pinjaman Nasabah Menggunakan Support Vector Machine dengan Optimasi Hyperparameter Tuning GridSearchCV.," *Kohesi J. Sains dan Teknol.*, vol. 10, no. 12, pp. 91–100, 2026, doi: 10.2238/sqnyp204.

[25] I. Pacal, I. Kunduracioglu, M. Hakki, and A. Muhammet, "A systematic review of deep learning techniques for plant diseases," *Artif. Intell. Rev.*, vol. 57, no. 11, 2024, doi: 10.1007/s10462-024-10944-7.

[26] J. Yao, S. O. N. N. Tran, S. Garg, and S. Sawyer, "Deep Learning for Plant Identification and Disease Classification from Leaf Images : Multi-prediction," *ACM Comput. Surv.*, vol. 56, no. 6, pp. 1–37, 2024, doi: 10.1145/3639816.

[27] Moukthika, "Color Spaces in OpenCV," 29 April 2025. Accessed: Oct. 26, 2026. [Online]. Available: https://opencv.org/blog/color-spaces-in-opencv