

Klasifikasi Tanaman Calathea Menggunakan Metode SVM Berbasis Fitur HSV dan HOG

Sri Andayani¹, Nining Ariati², Ayu Elisya Natama Sianturi³

^{1,3}Fakultas Sains dan Teknologi, Universitas Katolik Misi Charitas
Jalan Bangau No.60, Palembang, Indonesia

²Ilmu Komputer dan Sains, Universitas Indo Global Mandiri
Jalan Jend. Sudirman, Palembang, Indonesia

e-mail: andayani_s@ukmc.ac.id¹, nining@uigm.ac.id², ayuelisyans@gmail.com³

Received : December, 2025

Accepted : December, 2025

Published : December, 2025

Abstract

The visual similarity among Calathea varieties—especially in leaf color, pattern, and texture, often causes misidentification by the general public and even horticulturists. This research proposes an automated classification system for five Calathea varieties (Black Lipstick, Corona, Crimson, Medallion, and Pink Jessy) using Support Vector Machine (SVM) based on combined HSV (Hue, Saturation, Value) and HOG (Histogram of Oriented Gradients) features. A dataset of 1,000 leaf images was collected under controlled conditions (white background, fixed distance, consistent lighting) and processed through preprocessing, HSV-based color segmentation, grayscale conversion, and HOG feature extraction. The SVM model with a linear kernel was evaluated using 5-fold cross-validation. The system achieved an overall average accuracy of 93.52%, with Calathea Corona showing the highest accuracy (96.50%) and precision (95.12%), while Calathea Black Lipstick recorded the lowest precision (67.50%) but the highest recall (92.00%). These results demonstrate that the fusion of color and shape features enhances classification performance, although complex leaf patterns can still pose challenges in precision. This approach offers a practical and objective solution for Calathea identification in horticultural applications.

Keywords: Calathea, support vector machine, HSV, HOG, image classification

Abstrak

Kemiripan visual antar varietas Calathea, terutama pada warna, pola, dan tekstur daun, sering menyebabkan kesalahan identifikasi oleh masyarakat umum maupun pembudidaya. Penelitian ini mengusulkan sistem klasifikasi otomatis untuk lima varietas Calathea (Black Lipstick, Corona, Crimson, Medallion, dan Pink Jessy) menggunakan Support Vector Machine (SVM) berbasis gabungan fitur HSV (Hue, Saturation, Value) dan HOG (Histogram of Oriented Gradients). Dataset terdiri atas 1.000 citra daun yang dikumpulkan dalam kondisi terkontrol (latar putih, jarak tetap, pencahayaan konsisten), kemudian diproses melalui praproses, segmentasi warna HSV, konversi ke skala keabuan, dan ekstraksi fitur HOG. Model SVM dengan kernel linear dievaluasi menggunakan validasi silang 5-fold. Sistem mencapai akurasi rata-rata keseluruhan sebesar 93,52%, dengan Calathea Corona menunjukkan akurasi tertinggi (96,50%) dan precision (95,12%), sedangkan Calathea Black Lipstick mencatat precision terendah (67,50%) namun recall tertinggi (92,00%). Temuan ini menunjukkan bahwa penggabungan fitur warna dan bentuk mampu meningkatkan kinerja klasifikasi, meskipun pola daun yang kompleks tetap menjadi tantangan. Pendekatan ini memberikan solusi objektif dan praktis untuk identifikasi Calathea dalam aplikasi hortikultura.

Kata Kunci: Calathea, support vector machine, HSV, HOG, klasifikasi citra

1. PENDAHULUAN

Perkembangan teknologi pengolahan citra digital dan *machine learning* telah mendorong berbagai penelitian dalam bidang klasifikasi citra secara signifikan [1]. Klasifikasi citra banyak diterapkan dalam berbagai bidang, seperti pertanian, kesehatan, dan industri, khususnya untuk mengidentifikasi objek berdasarkan karakteristik visual tertentu [2], [3]. Dalam sektor pertanian, khususnya pada komoditas hortikultura, teknologi ini menjadi sangat relevan untuk mengidentifikasi spesies tanaman yang memiliki karakteristik visual kompleks berupa pola warna, motif, dan tekstur daun yang beragam, yang sering kali sulit dibedakan melalui pengamatan manusia secara manual.

Salah satu tanaman hias yang memiliki nilai estetika tinggi dan sangat diminati masyarakat adalah genus *Calathea* [4]. Tanaman ini memiliki kekayaan variasi morfologi pada daunnya, namun di sisi lain, keragaman ini menciptakan tantangan dalam identifikasi. *Calathea* yang merupakan bagian dari famili Marantaceae memiliki lebih dari 569 spesies [5]. Varietas populer seperti Black Lipstick, Corona, Crimson, Medallion, dan Pink Jessy memiliki kemiripan morfologis yang sangat tinggi, terutama pada kombinasi warna ungu gelap dan merah muda yang hampir serupa. Hal ini sering menimbulkan kebingungan bagi masyarakat awam maupun pembudidaya dalam menentukan jenis varietas secara akurat [4], [5].

Identifikasi secara manual tidak hanya bersifat subjektif, tetapi juga tidak efisien untuk kebutuhan komersial, konservasi, maupun edukasi skala besar. Oleh karena itu, pendekatan berbasis *computer vision* dan *machine learning* hadir sebagai solusi strategis. Salah satu algoritma yang telah terbukti handal adalah Support Vector Machine (SVM). SVM merupakan salah satu metode dalam bidang *machine learning* yang termasuk kategori supervised learning, dimana dikenal efektif dalam menangani data berdimensi tinggi dan memiliki efisiensi memori yang baik melalui penggunaan fungsi kernel untuk mencari hyperplane optimal sebagai pemisah antar kelas dan mengenali pola dalam proses klasifikasi [6]. Kinerja model SVM sangat bergantung pada pemilihan ekstraksi fitur yang tepat. Fitur warna dan bentuk merupakan dua komponen kritis dalam pengenalan objek [7]. Ruang warna Hue Saturation Value (HSV) kerap dipilih dalam pengolahan citra karena representasinya yang

lebih selaras dengan cara manusia mempersepsikan warna dibandingkan ruang warna RGB. Komponen Hue mencerminkan rona dasar suatu warna, Saturation menggambarkan kemurnian warna, yaitu seberapa sedikit atau banyak cahaya putih bercampur dengannya, sedangkan Value menyatakan tingkat kecerahan atau intensitas cahaya yang mempengaruhi kesan tonal pada citra [8].

Dalam upaya menangkap karakteristik visual secara menyeluruh, fitur HSV dapat dikolaborasikan dengan fitur HOG guna mengekstraksi informasi warna sekaligus tekstur. Proses ekstraksi menggunakan HOG diawali dengan konversi citra menjadi skala keabuan untuk mempermudah penghitungan gradien intensitas pada tiap unit piksel. Tahap ini krusial untuk mengidentifikasi kontur objek melalui distribusi arah gradien yang kemudian dikuantisasi ke dalam orientation binus. Hasil dari proses ini adalah sebuah deskriptor histogram yang memiliki ketahanan tinggi terhadap variasi pencahayaan dan geometri objek [9], [10].

Meskipun penggunaan fitur tunggal seperti HSV atau HOG secara terpisah telah menunjukkan performa yang cukup baik dalam beberapa studi [11], [12], fitur tunggal sering kali gagal merepresentasikan seluruh karakteristik visual objek yang kompleks secara optimal, sehingga diperlukan penggabungan fitur (*feature fusion*) untuk meningkatkan akurasi dan stabilitas sistem [13], [14].

Sejumlah penelitian terdahulu telah membuktikan keberhasilan kombinasi fitur HSV, HOG, dan SVM pada berbagai objek biologis, dimana Fitur HSV lebih representatif dalam menangkap persepsi warna manusia dibandingkan RGB [7], sementara HOG efektif dalam menangkap gradien tepi dan struktur bentuk objek [15], dan metode SVM sendiri dikenal handal dalam klasifikasi data berdimensi tinggi dengan margin pemisah optimal [16].

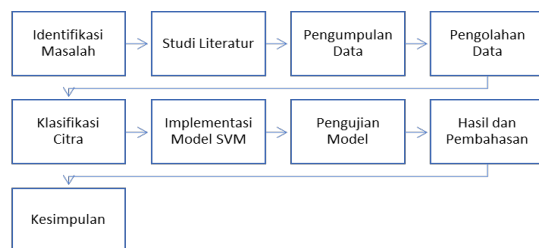
Sebagai contoh, penelitian klasifikasi bunga menggunakan pendekatan ini mampu mencapai akurasi rata-rata 86,4%, dengan performa paling optimal pada jenis bunga Daisy [9]. Studi lain pada klasifikasi sembilan genus jamur menghasilkan akurasi sebesar 82,69% [8]. Bahkan pada klasifikasi varietas terung, penggunaan SVM berbasis HSV dan HOG dengan kernel linier mampu mencatatkan akurasi tertinggi hingga 96,75% [17].

Meskipun penelitian terkait klasifikasi tanaman telah banyak dilakukan tetapi sebagian besar

studi masih berfokus pada tanaman pangan atau tanaman obat. Penerapan kombinasi fitur HSV dan HOG pada tanaman hias genus *Calathea* masih terbatas, terutama dalam menangani kemiripan warna ungu dan merah muda pada varietas yang spesifik. Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengimplementasikan metode SVM berbasis fitur HSV dan HOG guna mengklasifikasikan lima varietas *Calathea*. Diharapkan penggunaan kombinasi fitur warna (HSV) dan fitur tekstur atau bentuk (HOG) dapat menutupi kelemahan masing-masing fitur tunggal, sehingga mampu memberikan hasil klasifikasi yang lebih akurat dan objektif dalam membedakan varietas *Calathea* yang memiliki kemiripan visual tinggi.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimen berbasis computer vision dan machine learning untuk mengklasifikasikan lima varietas tanaman hias *Calathea*. Tahapan penelitian dirancang secara sistematis dimulai dari identifikasi masalah, studi literatur, pengumpulan data, pengolahan data, klasifikasi citra, implementasi model SVM, pengujian model, hasil dan pembahasan dan kesimpulan, seperti yang terlihat pada gambar 1.



Gambar 1. Tahapan Penelitian

2.1 Identifikasi Masalah

Penelitian ini diawali dengan merumuskan batasan masalah secara komprehensif melalui parameter terukur pada klasifikasi genus *Calathea*, dengan memfokuskan cakupan analisis pada lima varietas spesifik, yaitu Black Lipstick, Corona, Crimson, Medallion, dan Pink Jessy.

2.2 Studi Literatur

Studi pustaka dalam penelitian ini mencakup eksplorasi mendalam terhadap metodologi ekstraksi ciri HSV dan HOG, mekanisme klasifikasi menggunakan model SVM, serta tinjauan taksonomi mengenai tanaman hias

Calathea sebagai landasan pengembangan sistem.

2.3 Pengumpulan Data

Dataset dikumpulkan melalui pemotretan langsung menggunakan kamera ponsel Vivo V40 Lite 5G dengan resolusi 50 MP, aperture f/1.8, dan sensor Sony IMX 882, dalam kondisi terkontrol, yaitu latar belakang putih seragam, jarak kamera 60 cm dari objek (menggunakan tripod dan penggaris), serta pencahayaan konsisten. Total dataset terdiri atas 1.000 citra digital, dengan distribusi merata, yaitu 200 citra untuk masing-masing varietas *Calathea*. Semua citra disimpan dalam format JPG dengan resolusi awal 4080 × 3060 piksel, kemudian distandardisasi menjadi 800 × 600 piksel untuk keperluan komputasi.

2.4 Pengolahan Data

a. Preprocessing

Tahap awal pengolahan data melibatkan proses pra pemrosesan guna memastikan dataset siap untuk dianalisis dan dimodelkan. Langkah ini dilakukan dengan menyeragamkan dimensi seluruh citra menjadi 800 x 600 piksel untuk menjaga konsistensi skala dan resolusi spasial. Selain itu, seluruh citra ditransformasikan ke dalam format digital JPG guna menjamin keseragaman ekstensi file di seluruh dataset.

b. Segmentasi Ruang Warna HSV

Setelah fase pra pemrosesan, dilakukan segmentasi menggunakan ruang warna Hue Saturation Value (HSV). Proses identifikasi diawali dengan konversi citra dari model warna RGB ke HSV, yang dilanjutkan dengan ekstraksi kanal Value (V) untuk merepresentasikan informasi intensitas cahaya atau distribusi terang-gelap pada objek tanaman. Data intensitas ini menjadi landasan utama dalam menghasilkan citra berskala keabuan (grayscale) yang lebih representatif.

c. Transformasi Skala Keabuan (Grayscale Conversion)

Konversi ke format grayscale dilakukan setelah tahap segmentasi warna untuk mereduksi kompleksitas representasi data. Dengan mentransformasi citra berwarna menjadi skala abu-abu, pengaruh variasi warna yang tidak signifikan dapat diminimalisir, sehingga algoritma klasifikasi dapat lebih fokus pada

ekstraksi informasi struktural dan pola bentuk objek yang lebih esensial.

d. Ekstraksi Fitur HOG (Histogram of Oriented Gradients)

Ekstraksi fitur bentuk diimplementasikan melalui metode HOG, yang dimulai dengan kalkulasi besaran gradien pada tiap piksel untuk memetakan arah dan kontur tepi objek. Citra kemudian dipartisi ke dalam blok-blok kecil (sel), di mana pada setiap unit tersebut dibentuk histogram orientasi gradien berdasarkan persebaran arah gradien lokal. Sebagai tahap akhir, seluruh deskriptor histogram dari setiap blok diintegrasikan menjadi satu vektor fitur global yang merepresentasikan karakteristik visual citra secara komprehensif.

2.5 Klasifikasi Citra

Metodologi klasifikasi pada tanaman *Calathea* diimplementasikan melalui dua fase fundamental, yakni fase pelatihan (training) dan fase pengujian (testing). Dalam fase pelatihan, setiap data citra yang telah melalui tahap pra pemrosesan (penyelarasan dimensi dan format) ditransformasikan dari ruang warna RGB ke HSV guna mendukung segmentasi berbasis intensitas warna, yang kemudian direduksi menjadi citra skala keabuan (grayscale). Parameter struktural dan morfologi tanaman selanjutnya diisolasi menggunakan deskriptor Histogram of Oriented Gradient (HOG). Vektor fitur yang dihasilkan dari proses ekstraksi ini menjadi basis pembelajaran bagi algoritma Support Vector Machine (SVM) untuk membangun model klasifikasi yang kokoh. Selanjutnya, pada fase pengujian, dataset baru diberikan perlakuan pra pemrosesan dan ekstraksi fitur (HSV dan HOG) yang identik dengan fase pelatihan guna menjamin konsistensi data. Vektor fitur dari data uji ini kemudian diproses oleh model SVM yang telah terbentuk untuk mengidentifikasi varietas *Calathea* berdasarkan karakteristik pola yang ditemukan. Tahap akhir dari proses ini adalah evaluasi performa menggunakan matrik penilaian tertentu untuk memvalidasi tingkat akurasi dan keandalan sistem dalam mendeteksi jenis tanaman secara presisi.

2.6 Implementasi Model SVM

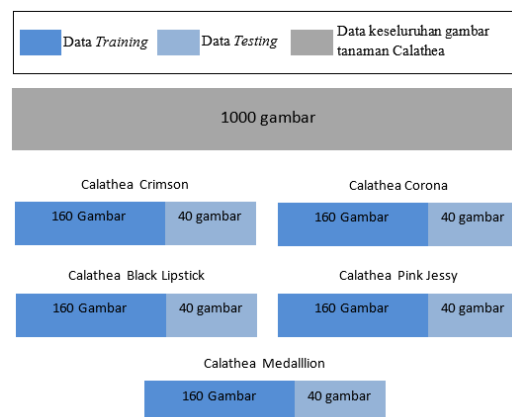
Setelah fase pelatihan model rampung, tahap implementasi dilakukan dengan mengaplikasikan model terhadap dataset

pengujian guna mengukur efektivitas kinerjanya. Prosedur pengolahan data uji diawali dengan transformasi ruang warna ke model HSV yang kemudian diikuti dengan proses ekstraksi ciri menggunakan algoritma Histogram of Oriented Gradients (HOG). Untuk menjamin validitas dan stabilitas hasil, distribusi dataset antara data latih dan data uji diatur menggunakan metode 5-fold Cross Validation dengan skema One-Time-Rejection. Proses klasifikasi akhir dieksekusi menggunakan algoritma Support Vector Machine (SVM) berbasis kernel linier untuk menghasilkan label prediksi yang akurat terhadap varietas objek yang diteliti.

2.7 Pengujian Model

Tahap pengujian melibatkan seluruh dataset untuk mengevaluasi kinerja model secara menyeluruh. Evaluasi dilakukan pada setiap varietas *Calathea* menggunakan skema 5-fold cross validation, di mana data dibagi ke dalam lima fold dan diuji secara bergantian sebagai data latih dan data uji. Rincian distribusi data pada setiap fold disajikan pada Tabel 1. Kinerja model selanjutnya dianalisis menggunakan confusion matrix untuk memperoleh nilai accuracy, precision, dan recall sebagai indikator kemampuan sistem dalam mengklasifikasikan varietas tanaman.

Tabel 1. Pembagian Jumlah *Data Training* dan *Data Testing*



2.8 Hasil dan Pembahasan

Bagian hasil dan pembahasan menyajikan analisis komprehensif terhadap data serta temuan yang diperoleh selama proses penelitian. Pada fase ini, data hasil pengujian dipaparkan secara mendetail untuk mengukur

efektivitas dan kinerja dari metodologi yang diimplementasikan. Analisis dilakukan secara kritis guna mengevaluasi korelasi antara hasil eksperimen dengan tujuan penelitian yang telah dirumuskan sebelumnya, serta memberikan interpretasi ilmiah terhadap fenomena yang muncul selama proses klasifikasi.

2.9 Kesimpulan




Sebagai tahap akhir, kesimpulan merangkum temuan-temuan krusial penelitian secara ringkas dan lugas, sekaligus memberikan jawaban eksplisit terhadap problematika penelitian serta target yang telah dicapai. Bagian ini juga menguraikan kontribusi nyata hasil penelitian terhadap pengembangan ilmu pengetahuan di bidang pengolahan citra dan memberikan rekomendasi strategis mengenai potensi implementasi praktis dari model yang telah dikembangkan untuk kebutuhan di masa depan.



3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Proses akuisisi data dilakukan melalui pengambilan gambar secara langsung untuk menyusun dataset yang terdiri dari lima varietas tanaman hias Calathea, meliputi Black Lipstick, Corona, Crimson, Medallion, dan Pink Jessy. Dataset ini menghimpun total 1.000 citra digital dengan distribusi data yang proporsional, di mana setiap varietas direpresentasikan oleh 200 sampel gambar. Rincian mengenai komposisi dan klasifikasi dataset tersebut disajikan secara komprehensif pada Tabel 2.

Tabel 2. Jenis dan Jumlah Dataset Calathea

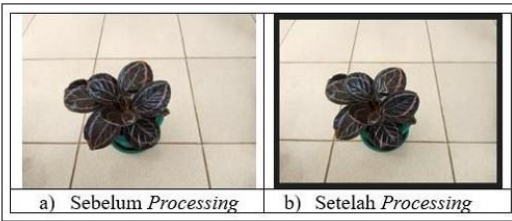
No	Jenis <i>Calathea</i>	Citra	Jumlah
1	<i>Calathea Black Lipstik</i>		200
2	<i>Calathea Corona</i>		200
3	<i>Calathea Crimson</i>		200

No	Jenis <i>Calathea</i>	Citra	Jumlah
4	<i>Calathea Medallion</i>		200
5	<i>Calathea Pink Jessy</i>		200

3.2 Pengolahan Data

a. Preprocessing

Tahapan ini dilakukan untuk menghasilkan dataset yang seragam baik dari aspek dimensi maupun format file. Citra mentah yang awalnya memiliki resolusi tinggi sebesar 4080 x 3060 piksel dengan ekstensi .IMG ditransformasikan guna mengoptimalkan efisiensi komputasi. Dalam proses ini, seluruh dataset disesuaikan ukurannya menjadi 800 x 600 piksel dan dikonversi ke format .JPG untuk menjamin konsistensi selama fase pemodelan. Perbandingan visual antara citra asli dan citra hasil pra pemrosesan diilustrasikan pada Gambar 2.



Gambar 2. Hasil Preprocessing

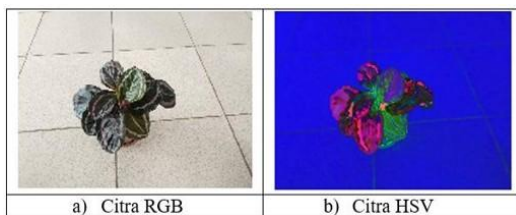
b. Segmentasi Warna HSV

Tahap segmentasi warna dilakukan dengan mentransformasikan citra tanaman Calathea dari ruang warna RGB ke model HSV guna memisahkan informasi visual menjadi komponen Hue (H) untuk warna dominan daun, Saturation (S) untuk intensitas warna, dan Value (V) untuk tingkat kecerahan. Fitur statistik yang digunakan adalah Mean untuk mempresentasikan warna dominan dan Std Dev untuk menangkap variasi warna antar area daun. Secara spesifik, ekstraksi diarahkan pada

kanal Value untuk menyederhanakan representasi citra menjadi format skala keabuan (grayscale). Seluruh prosedur segmentasi ini diimplementasikan menggunakan algoritma pengolahan data yang disusun pada perangkat lunak MATLAB R2022a.

```
% Melakukan Transformasi Ruang Warna Citra Yang Semula RGB Menjadi HSV
HSV = rgb2hsv(RGB);
subplot(3, 2, 2);
imshow(HSV);
```

Hasil yang didapatkan berupa citra RGB yang dikonversi ke dalam ruang warna HSV. Visualisasinya dapat dilihat pada Gambar 3.



Gambar 3. Hasil Citra RGB ke HSV

c. Konversi Grayscale

Tahap konversi ke grayscale dilakukan setelah nilai value (intensitas warna) didapatkan dari hasil segmentasi warna HSV. Komponen value diekstraksi untuk menghasilkan citra grayscale yang merepresentasikan tingkat kecerahan setiap piksel. Proses ini dilakukan menggunakan script pengolahan data melalui software MATLAB R2022a.

```
% Mengubah Warna HSV Menjadi Grayscale
V = HSV(:, :, 3);
```

Citra grayscale yang dihasilkan memberikan perbedaan intensitas warna secara jelas, sehingga mempermudah proses ekstraksi fitur dan klasifikasi pada tahap berikutnya. Visualisasi hasil konversi grayscale disajikan pada Gambar 4.



Gambar 4. Konversi Grayscale

d. Ekstraksi Fitur HOG

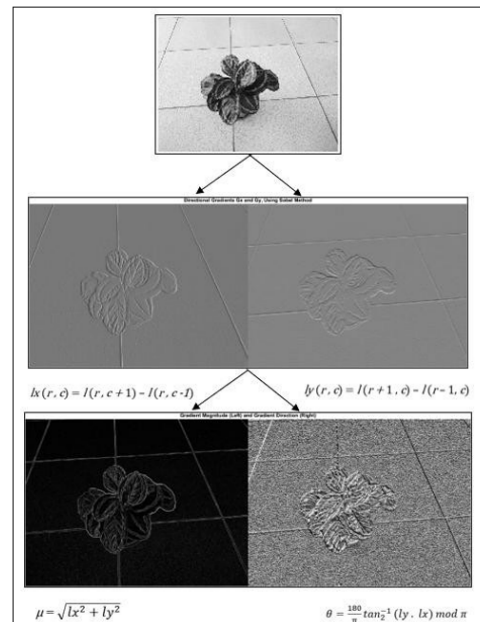
Ekstraksi fitur HOG dilakukan untuk menangkap pola urat daun, tekstur dan struktur tepi daun Calathea yang kemudian citra diubah ke skala

keabuan. Proses ini meliputi perhitungan gradien untuk mendeteksi perubahan intensitas piksel, pemetaan orientasi ke dalam bin, serta pembentukan deskriptor blok. Dari citra berukuran 800×600 piksel, turunan parsial I_x dan I_y dihitung berdasarkan posisi baris dan kolom, lalu digunakan untuk memperoleh magnitudo (μ) dan orientasi (θ). Seluruh perhitungan diimplementasikan menggunakan MATLAB R2022a.

```
%HOG
%Mengekstraksi Fitur HOG Menggunakan Citra Hasil Grayscale HSV
clc;
imageFile = imread('Datasetku/pink_jessy (1).jpg');
Gray = rgb2gray(imageFile);
imageGray = im2double(Gray);
[featureVector, hogVisualization] = extractHOGFeatures(imageGray);
figure;
imshow(imageGray);
hold on;
plot(hogVisualization);

%imshowpair(Gx,Gy,'montage')
title('Directional Gradients Gx and Gy, Using Sobel Method')
[Gmag,Gdir] = imggradient(Gx,Gy);
imshowpair(Gmag,Gdir,'montage')
title('Gradient Magnitude (Left) and Gradient Direction (Right)')
```

Visualisasi hasil perhitungan gradient horizontal dan vertikal, magnitude, serta orientasi yang didapatkan dari MATLAB R2022a dilihat pada Gambar 5.



Gambar 5. Implementasi Citra Grayscale ke HOG

Dalam implementasi HOG, citra berukuran 800×600 piksel dibagi menjadi cell berukuran 8×8 piksel, menghasilkan 100×75 cell. Setiap empat cell yang berdekatan digabung menjadi satu block berukuran 2×2 cell dengan overlap, sehingga terbentuk 99×74 block (total 7.326 block). Masing-masing cell merepresentasikan

orientasi gradien dalam 9 bin, sehingga total vektor fitur yang dihasilkan adalah $7.326 \times 4 \times 9 = 263.736$, yang digunakan sebagai masukan untuk model klasifikasi.

e. Pembentukan Model

Pembentukan model mencakup dua tahap utama: pelatihan dan pengujian. Sebelumnya, citra menjalani serangkaian praproses, yaitu penyesuaian ukuran, segmentasi warna HSV, konversi ke citra grayscale, dan ekstraksi fitur HOG untuk menghasilkan vektor fitur yang siap diklasifikasikan. Model SVM dengan kernel linear digunakan untuk membangun sistem klasifikasi. Data latih digunakan untuk menangkap ciri visual khas tiap varietas *Calathea*, sedangkan data uji (dengan alur pra proses yang identik) dipakai untuk mengevaluasi performa model. Seluruh proses diimplementasikan dalam MATLAB R2022a melalui skrip terstruktur.

```
% Memproses Data Menggunakan Fungsi Kernel Linear
t = templateSVM('KernelFunction', 'linear');
Mdl = fitcecoc(X_train_with_noise, Y(train,:), 'Learners', t);
```

Setelah model SVM berhasil dibangun menggunakan data fitur (X) hasil ekstraksi HOG, langkah selanjutnya adalah menguji performa model melalui proses pengujian, yang bertujuan untuk mengklasifikasikan lima jenis tanaman hias *Calathea*, yaitu Black Lipstick, Corona, Crimson, Medallion, dan Pink Jessy.

f. Uji Model

Pada pengujian model, dataset berisi 1.000 citra dibagi secara proporsional menjadi 80% data latih dan 20% data uji. Validasi dilakukan menggunakan 5-Fold Cross Validation, di mana setiap fold secara bergilir berperan sebagai data uji dan sisanya sebagai data latih. Pembagian data diotomatisasi melalui skrip berbasis MATLAB R2022a.

```
%kfold_data.m
%Membagi data testing dalam 5-Fold
%1000/5 = 200 data yang menjadi data testing
%40 citra per jenis Calathea

tic
load('Program Calathea/dataFeature_Latihan.mat');
indices5 = crossvalind('Kfold', Y, 5);

save('Program Calathea/kfold5_data1.mat', 'indices5');
toc
```

Pembagian dataset menghasilkan subset pelatihan (800 citra) dan pengujian (200 citra) yang berbeda pada setiap fold. Kinerja model dievaluasi menggunakan confusion matrix untuk menghitung metrik accuracy, precision, dan

recall berdasarkan prediksi terhadap data uji, melalui skrip pengolahan berikut:

```
%% Klasifikasi_K-Fold.m
tic
load('CodeData/MyFeature_uji.mat');
load('CodeData/kfold5_data.mat');

for i = 1:5
test = (indices5 == i);
train = ~test;

% Menambahkan Noise Ke Data Fitur Pelatihan
X_train_with_noise = X(train,:) + 1 * randn(size(X(train,:)));

% Memproses Data Menggunakan Fungsi Kernel Linear
t = templateSVM('KernelFunction', 'linear');
Mdl = fitcecoc(X_train_with_noise, Y(train,:), 'Learners', t);

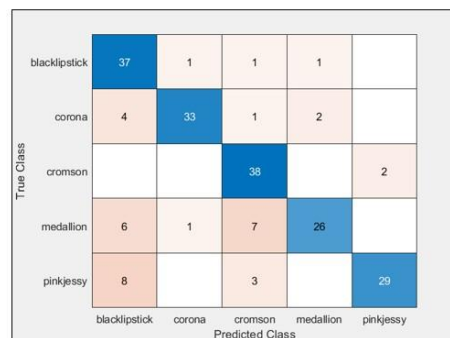
% Prediksi Hasil
Y_Predict = predict(Mdl, X(test,:));

% Menampilkan Hasil Testing Menjadi Bagan Confusion Matrix
figure;
confusionchart(Y(test,:), Y_Predict);
end
toc
```

Analisis awal menunjukkan capaian accuracy, precision, dan recall yang sempurna pada mayoritas kelas, yang secara teknis mengindikasikan gejala overfitting. Fenomena ini terjadi saat model kehilangan kemampuan generalisasi akibat penyesuaian berlebihan terhadap data latih [18]. Kurangnya variasi pada dataset serta penggunaan latar belakang putih yang sangat seragam dalam penelitian ini berisiko membatasi model dalam mengenali pola visual yang lebih luas [19]. Sebagai solusi, prinsip Structural Risk Minimization diimplementasikan melalui teknik augmentasi data dengan penyisipan noise terkontrol pada fitur data latih [20]. Pendekatan ini bertujuan memperluas keragaman data agar model lebih robust terhadap variasi data baru, yang selanjutnya divalidasi kinerjanya menggunakan confusion matrix melalui skema cross validation.

g. Confusion Matrix 5-Fold

Evaluasi kinerja klasifikasi lima varietas *Calathea* menggunakan SVM kernel linear dilakukan melalui *confusion matrix* pada perangkat lunak MATLAB R2022a.



Gambar 6. Confusion Matrix SVM Kernel Linear 1-Fold

Hasil pengujian pada satu *fold* (Gambar 6) menunjukkan bahwa model mencapai akurasi prediksi tertinggi pada varietas 'Crimson' dengan 38 data benar, sementara performa terendah tercatat pada 'Pink Jessy' dengan 29 data.

True Class	blacklipstick	corona	cromson	medallion	pinkjessy	
	35			3	2	
	6	34				
	1	1	35		3	
	5			34	1	
	7		1		32	
		Predicted Class				
		blacklipstick	corona	cromson	medallion	pinkjessy

Gambar 7. Confusion Matrix SVM Kernel Linear 2-Fold

Hasil evaluasi 2-fold menggunakan SVM kernel linear terhadap tujuh spesies *Calathea* disajikan dalam Gambar 5.7. Data menunjukkan bahwa jenis *Black Lipstick* dan *Crimson* mendominasi akurasi prediksi (35), sementara *Pink Jessy* memiliki tingkat pengenalan paling rendah (32).

True Class	blacklipstick	corona	cromson	medallion	pinkjessy	
	36			3	1	
	4	36				
	1	1	30	1	7	
	6		1	33		
	11		1	1	27	
		Predicted Class				
		blacklipstick	corona	cromson	medallion	pinkjessy

Gambar 8. Confusion Matrix SVM Kernel Linear 3-Fold

Evaluasi 3-fold pada Gambar 5.8 menunjukkan bahwa SVM kernel linear mencapai prediksi tertinggi pada varietas Black Lipstick dan Corona (masing-masing 36). Sebaliknya, performa terendah tercatat pada Pink Jessy dengan nilai 27.

True Class	blacklipstick	corona	cromson	medallion	pinkjessy	
	37				3	
	4	35		1		
	1	2	30	1	6	
	7	1		30	2	
	3		2		35	
		Predicted Class				
		blacklipstick	corona	cromson	medallion	pinkjessy

Gambar 9. Confusion Matrix SVM Kernel Linear 4-Fold

Berdasarkan confusion matrix 4-fold (Gambar 5.9), klasifikasi Black Lipstick menghasilkan akurasi tertinggi sebesar 37, sementara varietas Crimson dan Medallion menunjukkan hasil terendah dengan nilai masing-masing 30.

True Class	blacklipstick	corona	cromson	medallion	pinkjessy	
	39			1		
	2	36		2		
	2	2	33		3	
	6			34		
	5		1		34	
		Predicted Class				
		blacklipstick	corona	cromson	medallion	pinkjessy

Gambar 10. Confusion Matrix SVM Kernel Linear 5-Fold

Hasil pengujian 5-fold (Gambar 10) menunjukkan akurasi klasifikasi tertinggi pada varietas Black Lipstick dengan 39 prediksi benar, sedangkan Crimson mencatat hasil terendah sebesar 33.

Secara keseluruhan, evaluasi melalui cross-validation 1 hingga 5-fold menunjukkan konsistensi performa pada varietas Black Lipstick yang sering mencapai nilai prediksi tertinggi. Sebaliknya, fluktuasi hasil terendah terlihat pada varietas Pink Jessy dan Crimson di beberapa tahapan pengujian.

h. Hasil Uji Model Jenis Calathea

Bagian ini memaparkan evaluasi metrik accuracy, precision, dan recall untuk setiap varietas *Calathea* yang diperoleh melalui pengujian 5-fold cross validation menggunakan metode SVM.

Tabel 3. Hasil Accuracy, Precision dan Recall Pada 1-Fold

Jenis <i>Calathea</i>	Parameter	Persentase
Black Lipstick	Accuracy	89,50%
	Precision	67,27%
	Recall	92,50%
Corona	Accuracy	95,50%
	Precision	94,29%
	Recall	82,50%
Crimson	Accuracy	93,00%

	<i>Precision</i>	76,00%
	<i>Recall</i>	95,00%
Jenis <i>Calathea</i>	Parameter	Persentase
<i>Medallion</i>	<i>Accuracy</i>	91,50%
	<i>Precision</i>	89,66%
	<i>Recall</i>	65,00%
<i>Pink Jessy</i>	<i>Accuracy</i>	93,50%
	<i>Precision</i>	93,55%
	<i>Recall</i>	72,50%

Data pada Tabel 3 menunjukkan bahwa varietas Corona mendominasi performa dengan nilai accuracy (95,50%) dan precision (94,29%) tertinggi. Sementara itu, nilai recall tertinggi dicapai oleh varietas Crimson sebesar 95,00%.

Tabel 4. Hasil Accuracy, Precision dan Recall Pada 2-Fold

Jenis <i>Calathea</i>	Parameter	Persentase
<i>Black Lipstick</i>	<i>Accuracy</i>	88,00%
	<i>Precision</i>	64,81%
	<i>Recall</i>	87,50%
<i>Corona</i>	<i>Accuracy</i>	96,50%
	<i>Precision</i>	97,14%
	<i>Recall</i>	85,00%
<i>Crimson</i>	<i>Accuracy</i>	97,00%
	<i>Precision</i>	97,22%
	<i>Recall</i>	87,50%
<i>Medallion</i>	<i>Accuracy</i>	95,50%
	<i>Precision</i>	91,89%
	<i>Recall</i>	85,00%
<i>Pink Jessy</i>	<i>Accuracy</i>	93,00%
	<i>Precision</i>	84,21%
	<i>Recall</i>	80,00%

Merujuk pada Tabel 4, varietas Crimson mencatatkan nilai accuracy (97,00%) dan precision (97,22%) tertinggi. Adapun nilai recall optimal diperoleh secara bersamaan oleh varietas Black Lipstick dan Crimson dengan nilai masing-masing 87,50%.

Tabel 5. Hasil Accuracy, Precision dan Recall Pada 3-Fold

Jenis <i>Calathea</i>	Parameter	Persentase
<i>Black Lipstick</i>	<i>Accuracy</i>	87,00%
	<i>Precision</i>	62,07%
	<i>Recall</i>	90,00%
	<i>Accuracy</i>	97,50%

<i>Corona</i>	<i>Precision</i>	97,30%
	<i>Recall</i>	90,00%
Jenis <i>Calathea</i>	Parameter	Persentase
<i>Crimson</i>	<i>Accuracy</i>	94,00%
	<i>Precision</i>	93,75%
	<i>Recall</i>	75,00%
<i>Medallion</i>	<i>Accuracy</i>	94,00%
	<i>Precision</i>	86,84%
	<i>Recall</i>	82,50%
<i>Pink Jessy</i>	<i>Accuracy</i>	89,50%
	<i>Precision</i>	77,14%
	<i>Recall</i>	67,50%

Berdasarkan Tabel 5, varietas Corona menunjukkan performa unggul dengan nilai accuracy 97,50% dan precision 97,30%. Sementara itu, skor recall tertinggi (90,00%) dicapai secara bersamaan oleh jenis Black Lipstick dan Corona.

Tabel 6. Hasil Accuracy, Precision dan Recall Pada 4-Fold

Jenis <i>Calathea</i>	Parameter	Persentase
<i>Black Lipstick</i>	<i>Accuracy</i>	91,00%
	<i>Precision</i>	71,15%
	<i>Recall</i>	92,50%
<i>Corona</i>	<i>Accuracy</i>	96,00%
	<i>Precision</i>	92,11%
	<i>Recall</i>	87,50%
<i>Crimson</i>	<i>Accuracy</i>	94,00%
	<i>Precision</i>	93,75%
	<i>Recall</i>	75,00%
<i>Medallion</i>	<i>Accuracy</i>	94,00%
	<i>Precision</i>	93,75%
	<i>Recall</i>	75,00%
<i>Pink Jessy</i>	<i>Accuracy</i>	92,00%
	<i>Precision</i>	76,09%
	<i>Recall</i>	87,50%

Data pada Tabel 6 menunjukkan bahwa accuracy tertinggi diraih oleh Corona sebesar 96,00%. Nilai precision optimal dicatatkan oleh Crimson dan Medallion (masing-masing

93,75%), sedangkan recall terbaik diperoleh varietas Black Lipstick dengan nilai 92,50%.

Tabel 7. Hasil Accuracy, Precision dan Recall Pada 5-Fold

Jenis <i>Calathea</i>	Parameter	Persentase
<i>Black Lipstick</i>	<i>Accuracy</i>	92,00%
	<i>Precision</i>	72,22%
	<i>Recall</i>	97,50%
<i>Corona</i>	<i>Accuracy</i>	97,00%
	<i>Precision</i>	94,74%
	<i>Recall</i>	90,00%
<i>Crimson</i>	<i>Accuracy</i>	96,00%
	<i>Precision</i>	97,06%
	<i>Recall</i>	82,50%
<i>Medallion</i>	<i>Accuracy</i>	95,50%
	<i>Precision</i>	91,89%
	<i>Recall</i>	85,00%
<i>Pink Jessy</i>	<i>Accuracy</i>	95,50%
	<i>Precision</i>	91,89%
	<i>Recall</i>	85,00%

Merujuk pada Tabel 7, tingkat *accuracy* tertinggi kembali ditunjukkan oleh *Corona* (97,00%). Adapun nilai *precision* tertinggi dicapai oleh *Crimson* sebesar 97,06%, sementara *Black Lipstick* mencatatkan nilai *recall* tertinggi sebesar 97,50%.

i. Hasil Rata-Rata Uji Model Jenis *Calathea* 5-Fold

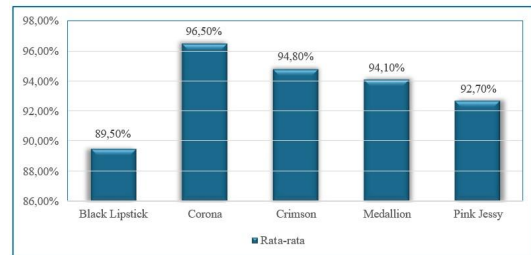
Bagian ini memaparkan rerata metrik *accuracy*, *precision*, dan *recall* untuk seluruh jenis *Calathea* yang diteliti. Rangkuman rerata nilai *accuracy* untuk setiap varietas disajikan secara mendalam pada Tabel 8.

Tabel 8. Hasil Rata-Rata Accuracy Pada Jenis Tanaman Hias *Calathea*

Hasil Rata-rata Accuracy Jenis Tanaman Hias <i>Calathea</i>						
Jenis <i>Calathea</i>	Pengujian Kernel Linier					Rata-rata
	1-Fold	2-Fold	3-Fold	4-Fold	5-Fold	
<i>Black Lipstick</i>	89,50%	88,00%	87,00%	91,00%	92,00%	89,50%
<i>Corona</i>	95,50%	96,50%	97,50%	96,00%	97,00%	96,50%
<i>Crimson</i>	93,00%	97,00%	94,00%	94,00%	96,00%	94,80%
<i>Medallion</i>	91,50%	95,50%	94,00%	94,00%	95,50%	94,10%

<i>Pink Jessy</i>	93,50%	93,00%	89,50%	92,00%	95,50%	92,70%
Rata-Rata Accuracy Seluruh Jenis Tanaman Hias <i>Calathea</i>						93,52%

Tabel 8 menunjukkan bahwa varietas *Corona* meraih rerata akurasi tertinggi sebesar 96,50%, sementara *Black Lipstick* mencatat nilai terendah sebesar 89,50%. Secara kolektif, model mencapai rerata akurasi sistem sebesar 93,52%, dengan distribusi visual setiap jenis tanaman ditampilkan pada Gambar 11.



Gambar 11. Grafik Hasil Rata-Rata Accuracy Jenis Tanaman Hias *Calathea*

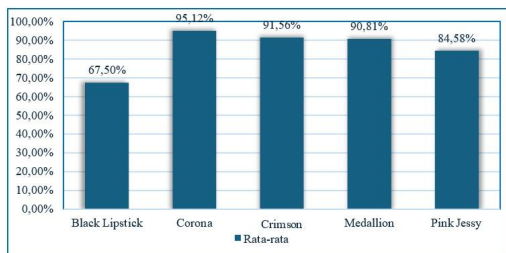
Selanjutnya hasil pengujian berupa rata-rata nilai *precision* disajikan pada Tabel 9.

Tabel 9. Hasil Rata-Rata Precision Pada Jenis Tanaman Hias *Calathea*

Hasil Rata-rata Precision Jenis Tanaman Hias <i>Calathea</i>						
Jenis <i>Calathea</i>	Pengujian Kernel Linier					Rata-rata
	1-Fold	2-Fold	3-Fold	4-Fold	5-Fold	
<i>Black Lipstick</i>	67,27%	64,81%	62,07%	71,15%	72,22%	67,50%
<i>Corona</i>	94,29%	97,14%	97,30%	92,11%	94,74%	95,12%
<i>Crimson</i>	76,00%	97,22%	93,75%	93,75%	97,06%	91,56%
<i>Medallion</i>	89,66%	91,89%	86,84%	93,75%	91,89%	90,81%
<i>Pink Jessy</i>	93,55%	84,21%	77,14%	76,09%	91,89%	84,58%
Rata-Rata Precision Seluruh Jenis Tanaman Hias <i>Calathea</i>						85,91%

Berdasarkan Tabel 9, *Calathea Corona* mencatat *precision* tertinggi (95,12%), sedangkan *Calathea Black Lipstick* terendah (67,50%). Secara keseluruhan, rata-rata *precision* mencapai 85,91%, menunjukkan bahwa model umumnya mampu mengklasifikasikan dengan akurat, meski variasi pola dan warna daun, terutama pada *Black Lipstick* yang coraknya kompleks menyebabkan lebih banyak kesalahan

prediksi. Hasil ini divisualisasikan dalam Gambar 12.



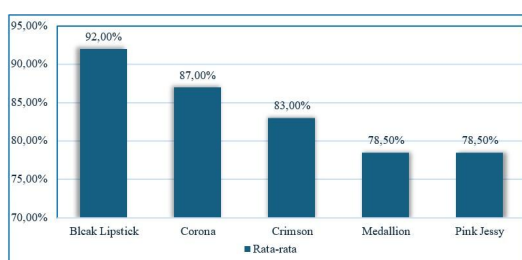
Gambar 12. Grafik Hasil Rata-Rata Precision Jenis Tanaman Hias Calathea

Sementara itu hasil pengujian berupa rata-rata nilai recall disajikan pada Tabel 10.

Tabel 10. Hasil Rata-Rata Recall Pada Jenis Tanaman Hias Calathea

Hasil Rata-rata Recall Jenis Tanaman Hias Calathea						
Jenis Calathea	Pengujian Kernel Linier					Rata-rata
	1-Fold	2-Fold	3-Fold	4-Fold	5-Fold	
Black Lipstick	92,50%	87,50%	90,00%	92,50%	97,50%	92,00%
Corona	82,50%	85,00%	90,00%	87,50%	90,00%	87,00%
Crimson	95,00%	87,50%	75,00%	75,00%	82,50%	83,00%
Medallion	65,00%	85,00%	82,50%	75,00%	85,00%	78,50%
Pink Jessy	72,50%	80,00%	67,50%	87,50%	85,00%	78,50%
Rata-Rata Recall Seluruh Jenis Tanaman Hias Calathea						83,80%

Berdasarkan Tabel 10, Calathea Black Lipstick mencatat recall tertinggi (92,00%), sedangkan Medallion dan Pink Jessy terendah (78,50% masing-masing). Rata-rata recall keseluruhan sebesar 83,80%, menunjukkan model cukup mampu mengenali sampel positif, meski dipengaruhi kompleksitas pola daun—Black Lipstick lebih mudah dikenali, sementara Medallion dan Pink Jessy cenderung menghasilkan prediksi terlewat karena motifnya yang rumit. Hasil ini divisualisasikan dalam Gambar 13.



Gambar 13. Grafik Hasil Rata-Rata Recall Jenis Tanaman Hias Calathea

Berdasarkan hasil pengujian, Calathea Corona menunjukkan kinerja terbaik dalam *accuracy* (96,50%) dan *precision* (95,12%), sedangkan Calathea Black Lipstick memiliki *accuracy* terendah (89,50%) dan *precision* paling rendah (67,50%). Namun, Black Lipstick unggul dalam *recall* (92,00%), sementara Medallion dan Pink Jessy mencatat *recall* terendah (78,50% masing-masing). Temuan ini mengindikasikan bahwa kompleksitas pola dan warna daun memengaruhi keseimbangan antara kemampuan model mengenali sampel positif (*recall*) dan menghindari kesalahan prediksi (*precision* dan *accuracy*).

4. KESIMPULAN

Penerapan algoritma Support Vector Machine (SVM) dengan kernel linear terbukti efektif dalam mengklasifikasikan jenis tanaman hias Calathea, menghasilkan rata-rata akurasi sistem sebesar 93,52%. Varietas Calathea Corona menjadi objek yang paling konsisten dikenali oleh model dengan tingkat akurasi rata-rata 96,50%. Hal ini menunjukkan bahwa fitur visual pada Corona memiliki karakteristik yang kuat dan unik bagi hyperplane SVM. Nilai precision terendah ditemukan pada varietas Black Lipstick (67,50%), yang mengindikasikan bahwa kompleksitas corak dan kemiripan warna daun antar spesies tertentu masih menjadi faktor utama terjadinya kesalahan prediksi (falsafah klasifikasi). Penambahan noise terkontrol dan penggunaan k-fold cross validation berhasil menjaga stabilitas model dan meminimalkan risiko overfitting pada dataset dengan latar belakang seragam, sehingga model memiliki kemampuan generalisasi yang baik terhadap data baru.

Saran untuk penelitian berikutnya adalah tahap praproses citra dapat ditingkatkan dengan metode segmentasi daun yang lebih akurat, guna memisahkan objek daun dari latar belakang secara optimal sehingga fitur yang diekstraksi lebih representatif dan menggunakan dataset dengan jumlah citra yang lebih besar serta variasi kondisi pengambilan gambar seperti perbedaan pencahayaan, sudut pandang dan latar belakang.

DAFTAR PUSTAKA

- [1] M. Y. F., *Dasar Pengolahan Citra Digital*. 2020.
- [2] S. Andayani, L. Kusneti, U. Katolik, M. Charitas, and A. History, "Using the Support Vector Machine method with the HOG feature for classification of orchid types," *AITI J. Teknol. Inf.*, vol. 21, no. 1, pp. 82–95, 2024.
- [3] R. Dijaya, *Buku Ajar Pengolahan Citra Digital*. 2023.
- [4] K. P. D. J. HORTIKULTURA, *Tanaman Hias Berdaun Indah*. KEMENTERIAN PERTANIAN DIREKTORAT JENDERAL HORTIKULTURA DIREKTORAT BUAH DAN FLORIKULTURA, 2021.
- [5] N. R. Hanik, F. A. Cahyanti, and T. Wahyuni, "Identification of Pests and Diseases of Calathea Ornamental Plants in Ngledoksari Village , Karanganyar," *J. Biol. Trop.*, vol. 24, pp. 641–650, 2024.
- [6] M. Fansyuri, "Analisis Support Vector Machine (SVM) dalam Pengenalan Citra Bunga Menggunakan Fitur Warna dan Bentuk," *OKTAL J. Ilmu Komput. dan Sci.*, vol. 1, no. 09, pp. 1579–1590, 2022.
- [7] H. Durairaj and P. Sapient, "Interactive Color Image Segmentation using HSV Color Space," *Sci. Technol. J. .*, vol. 7, no. 1, 2020, doi: 10.22232/stj.2019.07.01.05.
- [8] D. Udjulawa and T. I. Sariyo, "Klasifikasi Jenis Jamur Menggunakan SVM dengan Fitur HSV dan HOG," *PETIR J. Pengkaj. dan Penerapan Tek. Inform.*, vol. 15, no. 1, pp. 113–120, 2022.
- [9] H. Leidiyana and J. Warta, "Implementasi Metode SVM untuk Klasifikasi Bunga dengan Ekstraksi Fitur Histogram of Gradient (HOG)," *J. Inf. Inf. Secur.*, vol. 3, no. 1, pp. 89–98, 2022.
- [10] Rial Prasthio, Yohannes, and S. Devella, "Penggunaan Fitur HOG Dan HSV Untuk Klasifikasi Citra Sel Darah Putih," *J. Algoritm.*, vol. 2, no. 2, pp. 120–132, 2022.
- [11] A. Y. Kuswari, M. A. Rahman, D. Alamsyah, and S. V. Machine, "Klasifikasi Jenis Bunga Menggunakan Support Vector Machine dengan Fitur HSV dan HOG," no. x, 2018.
- [12] Marini and Hariyanto, "Implementasi Metode HOG (Histogram Of Oriented Gradient) Untuk Mendeteksi Pejalan Kaki Pada Citra Video," *Innov. J. Soc. Sci. Res.*, vol. 4, no. 4, pp. 13964–13970, 2024.
- [13] N. Rachmat, Yohannes, and A. Mahendra, "Klasifikasi Jenis Ikan Laut Menggunakan Metode SVM Dengan Fitur HOG Dan HSV," *J. Tek. Inform. dan Sist. Inf.*, vol. 8, no. 4, pp. 2235–2247, 2021.
- [14] S. Singh, D. Singh, and V. Yadav, "Face Recognition Using HOG Feature Extraction and SVM Classifier," *Int. J. Emerg. Trends Eng. Res.*, no. 9, pp. 6437–6440, 2020.
- [15] S. Rajaa, "Facial Expression Recognition System Based on SVM and HOG Techniques," *Int. J. Image Process.*, vol. 15, no. 2, pp. 14–21, 2021.
- [16] Y. G. Lestari and H. Irsyad, "Penggunaan Metode SVM Dengan Fitur HSV HOG Dalam Mengklasifikasi Jenis Ikan Guppy," *J. Algoritm.*, vol. 4, no. 1, pp. 21–31, 2023, doi: 10.35957/algoritme.xxxx.
- [17] A. B. Kurniawan et al., "KLASIFIKASI JENIS TERUNG MENGGUNAKAN METODE SVM," in *MDP STUDENT CONFERENCE (MSC)*, 2023, pp. 185–191.
- [18] I. Badruddin, T. R. Nugrahanto, O. D. Pangesti, T. Agustin, and J. Tengah, "Pendekatan pengurangan overfitting pada mobilenet untuk klasifikasi citra sampah," in *SEMINAR NASIONAL AMIKOM SURAKARTA (SEMNAS)*, 2024, pp. 226–234.
- [19] T. J. Bradshaw, "A Guide to Cross-Validation for Artificial Intelligence in," *Radiol. Artif. Intell.*, no. 12, 2023.
- [20] M. R. Saputra and H. Irsyad, "Klasifikasi Tingkat Kemanisan Alpukat Berdasarkan Fitur Hue Saturation Value (HSV) dengan Menggunakan Support Vector Machine (SVM)," *Algoritme*, vol. 2, no. 2, pp. 113–119, 2022.