

**KLASIFIKASI PENYAKIT TEH BERBASIS CITRA MENGGUNAKAN
*SUPPORT VECTOR MACHINE (SVM)***

SKRIPSI

Oleh:
AHMAD MUNAWWARUL HAQQ
NIM. 19650105



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

**KLASIFIKASI PENYAKIT TEH BERBASIS CITRA MENGGUNAKAN
*SUPPORT VECTOR MACHINE (SVM)***

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh:
AHMAD MUNAWWARUL HAQQ
NIM. 19650105

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

HALAMAN PERSETUJUAN

**KLASIFIKASI PENYAKIT TEH BERBASIS CITRA MENGGUNAKAN
SUPPORT VECTOR MACHINE (SVM)**

SKRIPSI

Oleh:

AHMAD MUNAWWARUL HAQQ
NIM. 19650105

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 19 Juni 2026

Pembimbing I,



Supriyono, M.Kom
NIP. 19841010 201903 1 012

Pembimbing II,



Ashri Shabrina Afrah, M.T
NIP. 19900430 202012 2 003

Mengetahui

Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Supriyono, M.Kom

NIP. 19841010 201903 1 012

HALAMAN PENGESAHAN

**KLASIFIKASI PENYAKIT TEH BERBASIS CITRA MENGGUNAKAN
SUPPORT VECTOR MACHINE (SVM)**

SKRIPSI

Oleh:

AHMAD MUNAWWAARUL HAQQ
NIM. 19650105

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 23 Juni 2026

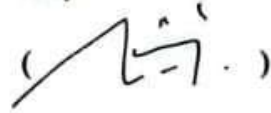
Susunan Dewan Penguji

Ketua Penguji : Prof. Dr. Ir. Muhammad Faisal, M.T
NIP. 19740510 200501 1 007

Anggota Penguji I : Dr. Ir. Yunifa Miftachul Arif, S.ST., M.T.
SMIEEE
NIP. 19830616 201101 1 004

Anggota Penguji II : Supriyono, M.Kom
NIP. 19841010 201903 1 012

Anggota Penguji III : Ashri Shabrina Afrah, M.T
NIP. 19900430 202012 2 003

()
()
()
()

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Supriyono, M.Kom
NIP. 19841010 201903 1 012

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Ahmad Munawwarul Haqq
NIM : 19650105
Fakultas / Program Studi : Sains dan Teknologi / Teknik Informatika
Judul Skripsi : Klasifikasi Penyakit Teh Berbasis Citra
Menggunakan *Support Vector Machine* (SVM)

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 23 Juni 2026

Yang membuat pernyataan,



Ahmad Munawwarul Haqq
NIM. 19650105

MOTTO

Ikatlah untamu, lalu bertawakkallah.

HALAMAN PERSEMBAHAAN

Alhamdulillah rabbil 'alamin segala puji bagi Allah *Subhanahu wa Ta'ala* yang telah melimpahkan rahmat, hidayah, serta kemudahan sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Shalawat dan salam senantiasa tercurah kepada baginda Rasulullah *Shallallahu 'Alaihi wa Sallam* yang telah menuntun umat manusia keluar dari zaman *jahiliyah* menuju cahaya Islam.

Penulis mempersembahkan karya ini kepada kedua orang tua, para dosen pengguji dan pembimbing, segenap sivitas akademika Program Studi Teknik Informatika, sahabat dan kerabat serta orang-orang yang telah membantu, mendoakan, serta menyemangati penulis hingga selesainya skripsi ini.

Penulis ucapkan "*Jazakumullah khairan katsiiraa*". Semoga ukhuwah tetap terjaga dan selalu diridhoi Allah *Subhanahu wa Ta'ala*. *Aamiin Ya Rabbal 'Alamiin*.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh.

Puji syukur penulis panjatkan atas kehadiran Allah *Subhanahu wa Ta'ala* yang telah melimpahkan rahmat dan kasih sayang-Nya, sehingga penulis mampu menyelesaikan skripsi ini dengan baik dan tepat waktu. Shalawat serta salam semoga tercurahkan kepada Nabi Muhammad *Shallallahu 'Alaihi wa Sallam*, atas syafaatnya yang telah menuntun umat manusia dari jalan kebatilan ke jalan yang benar. Semoga kita semua termasuk ke dalam golongan yang dituntun Allah *Subhanahu wa Ta'ala* serta mendapatkan pertolongan Nabi Muhammad *Shallallahu 'Alaihi wa Sallam*. Aamiin.

Selanjutnya, penulis mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu dan mendukung terselesaikannya skripsi ini. Ucapan terima kasih ini penulis sampaikan kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si., CAHRM., CRMP., selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. H. Agus Mulyono, M.Si, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Supriyono, M.Kom, selaku Ketua Jurusan Teknik Informatika.
4. Bapak Supriyono, M.Kom dan Ibu Ashri Shabrina Afrah, M.T., selaku dosen pembimbing I dan II yang telah meluangkan waktu, tenaga, dan pikiran dalam membimbing penyusunan skripsi ini hingga selesai.
5. Bapak Prof. Dr. Ir. Muhammad Faisal, M.T. dan Bapak Dr. Ir. Yunifa Miftachul Arif, S.ST., M.T., SMIEEE., selaku dosen penguji yang telah

memberikan arahan, kritik, dan saran konstruktif demi kesempurnaan skripsi ini.

6. Segenap civitas akademika Jurusan Teknik Informatika, terima kasih atas ilmu dan bimbingan yang telah diberikan selama masa perkuliahan.
7. Bapak Jumadi, S.T. dan Ibu Asnanik, S.Pd. orang tua penulis tercinta, yang senantiasa memberikan doa dan dukungan, baik secara moril maupun materi, sehingga penulis dapat menyelesaikan skripsi ini.
8. Segenap sahabat dan kerabat, yang telah menjadi teman diskusi, tempat bertukar pikiran, dan sumber semangat di setiap langkah perjuangan.
9. Semua pihak yang telah membantu dalam penyusunan skripsi ini, yang tidak dapat penulis sebutkan satu per satu. Semoga kebaikan kalian dibalas oleh Allah *Subhanahu wa Ta'ala* dengan balasan yang lebih baik.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat banyak kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan. Penulis berharap skripsi ini dapat memberikan manfaat bagi para pembaca, terutama bagi penulis sendiri, serta dapat menjadi kontribusi kecil bagi pengembangan ilmu pengetahuan, khususnya di bidang pengolahan citra dan kecerdasan buatan.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Malang, 23 Juni 2026

Penulis

DAFTAR ISI

HALAMAAN JUDUL	II
HALAMAN PERSETUJUAN	III
HALAMAN PENGESAHAN.....	IV
PERNYATAAN KEASLIAN TULISAN	V
MOTTO	VI
HALAMAN PERSEMBAHAAN.....	VII
KATA PENGANTAR.....	VIII
DAFTAR ISI.....	X
DAFTAR GAMBAR	XII
DAFTAR TABEL	XIII
DAFTAR SIMBOL	XIV
ABSTRAK	XV
ABSTRACT.....	XVI
مستخلص البحث.....	XVII
BAB I PENDAHULUAN.....	1
1.1 LATAR BELAKANG	1
1.2 PERNYATAAN MASALAH	5
1.3 BATASAN MASALAH	5
1.4 TUJUAN PENELITIAN	6
1.5 MANFAAT PENELITIAN	6
BAB II STUDI PUSTAKA	8
2.1 PENELITIAN TERKAIT	8
2.2 TANAMAN TEH	12
2.2.1 Morfologi Penyakit Teh.....	13
2.3 PENGOLAHAN CITRA DIGITAL	14
2.3.1 Prapengolahan Citra	14
2.3.1.1 <i>Center-Based Square Cropping</i>	15
2.3.2 Augmentasi Citra.....	16
2.3.3 Transformasi Ruang Warna	17
2.3.3.1 Ruang Warna HSV	17
2.3.3.2 Citra <i>Grayscale</i>	18
2.3.4 Ekstraksi dan Fusi Fitur.....	19
2.3.4.1 Ekstraksi Fitur Statistik	19
2.3.4.2 <i>Feature Fusion</i> dan Normalisasi Skala Data	20
2.4 <i>SUPPORT VECTOR MACHINE</i>	21
2.4.1 Konsep Dasar dan Formulasi Matematis SVM.....	22
2.4.2 Fungsi Kernel SVM	23
2.4.3 Klasifikasi Multikelas <i>One-vs-Rest</i> (OvR).....	25
2.5 METRIK EVALUASI PERFORMA MODEL	25
BAB III DESAIN PENELITIAN.....	27
3.1 DESAIN SISTEM	27
3.2 PENGUMPULAN DATA.....	28
3.3 DATA CITRA DAUN TEH	30

3.3.1 Prapengolahan Citra	31
3.3.2 Pembentukan Dataset	32
3.3.2.1 Dataset Tidak Seimbang	32
3.3.3 Konversi Ruang Warna HSV dan <i>Grayscale</i>	33
3.3.4 Ekstraksi Fitur	35
3.3.5 <i>Feature Fusion</i>	37
3.3.6 Normalisasi.....	38
3.4 KLASIFIKASI MENGGUNAKAN <i>SUPPORT VECTOR MACHINE</i>	39
3.4.1 Arsitektur <i>Support Vector Machine</i>	40
3.4.1.1 Parameter Fungsi Kernel	41
3.4.2 Pelatihan Model Menggunakan <i>One-vs-Rest</i>	41
3.4.3 Prediksi dan Penentuan Kelas	42
3.5 SKENARIO PENGUJIAN	43
3.5.1 Lingkungan Pengujian.....	44
3.5.2 Skenario Pengujian Klasifikasi	44
BAB IV HASIL DAN PEMBAHASAN.....	46
4.1 HASIL.....	46
4.1.1 Pengumpulan Data	46
4.1.2 Pra-pemrosesan Citra	47
4.1.3 Pembentukan Dataset	49
4.1.4 Ekstraksi Fitur dan Normalisasi	50
4.1.5 Hasil Klasifikasi SVM pada Data Tidak Seimbang	52
4.1.5.1 Analisis <i>Support Vector</i>	52
4.1.5.2 Visualisasi <i>Hyperplane</i>	53
4.1.5.3 Simulasi <i>Decision Function</i>	54
4.1.5.4 Hasil Evaluasi	54
4.1.6 Hasil Klasifikasi SVM pada Data Seimbang	57
4.1.6.1 Analisis <i>Support Vector</i>	57
4.1.6.2 Visualisasi <i>Hyperplane</i>	58
4.1.6.3 Simulasi <i>Decision Function</i>	58
4.1.6.4 Hasil Evaluasi	59
4.1.7 Perbandingan Kinerja Antar Skenario.....	61
4.1.8 Perbandingan <i>Support Vector</i>	63
4.2 PEMBAHASAN	64
4.2.1 Pengaruh Fungsi Kernel terhadap Kinerja Klasifikasi	65
4.2.2 Pengaruh <i>Hybrid Resampling</i> terhadap Kinerja Klasifikasi	66
4.2.3 Analisis Kesalahan Klasifikasi	67
4.2.4 Perbandingan dengan Penelitian Terdahulu	67
4.2.5 Integrasi Keilmuan dalam Perspektif Islam	69
BAB V KESIMPULAN DAN SARAN	74
5.1 KESIMPULAN.....	74
5.2 SARAN	75
DAFTAR PUSTAKA	
LAMPIRAN-LAMPIRAN	

DAFTAR GAMBAR

Gambar 2. 1 Konsep dasar SVM	22
Gambar 3. 1 Dasain sistem	27
Gambar 3. 2 Alur pada data citra teh	30
Gambar 3. 3 Arsitektur SVM dengan strategi OvR.....	40
Gambar 3. 4 Proses pelatihan SVM dengan strategi Ovr	42
Gambar 4. 1 Distribusi dataset awal citra daun teh	47
Gambar 4. 2 Tahapan pra-pemrosesan citra.....	48
Gambar 4. 3 Hasil prapemrosesan citra	48
Gambar 4. 4 Perbandingan Distribusi Data Latih Sebelum dan Sesudah <i>Hybrid Resampling</i>	49
Gambar 4. 5 Distribusi fitur statistik per kelas	51
Gambar 4. 6 Visualisasi <i>hyperplane</i> SVM kernel RBF pada data tidak seimbang	53
Gambar 4. 7 <i>Confusion matrix</i> kernel: a) linear, b) polynomial, c) RBF dan d) sigmoid , pada data tidak seimbang	55
Gambar 4. 8 Visualisasi <i>hyperplane</i> SVM kernel RBF pada data seimbang	58
Gambar 4. 9 <i>Confusion matrix</i> kernel: a) linear, b) polynomial, c) RBF dan d) sigmoid , pada data seimbang	60
Gambar 4. 10 Grafik perbandingan: a) akurasi, b) presisi, c) <i>recall</i> dan d) <i>F1-Score</i> , antar kernel dan scenario	63
Gambar 4. 11 Perbandingan jumlah <i>support vector</i>	63

DAFTAR TABEL

Tabel 2. 1 Ringkasan penelitian terkait	11
Tabel 2. 2 Morfologi penyakit pada daun teh	13
Tabel 3. 1 Data sampel	28
Tabel 3. 2 Parameter fungsi kernel SVM	41
Tabel 3. 3 Rancangan skenario pengujian	45
Tabel 4. 1 Distribusi dataset awal	46
Tabel 4. 2 Distribusi data latih	49
Tabel 4. 3 Contoh hasil ekstraksi fitur per kelas	50
Tabel 4. 4 Cotoh hasil normalisasi	51
Tabel 4. 5 Jumlah <i>support vector</i> pada data tidak seimbang.....	52
Tabel 4. 6 Simulasi <i>decision function</i> kernel RBF pada data tidak seimbang	54
Tabel 4. 7 Hasil evaluasi pada data tidak seimbang	55
Tabel 4. 8 Detail metrik per kelas kernel terbaik pada fata tidak seimbang	56
Tabel 4. 9 Jumlah <i>support vector</i> pada data seimbang.....	57
Tabel 4. 10 Simulasi <i>decision function</i> kernel RBF pada data seimbang	59
Tabel 4. 11 Hasil evaluasi pada data seimbang (<i>hybrid resampling</i>)	59
Tabel 4. 12 Detail metrik perkelas kernel terbaik pada data seimbang.....	61
Tabel 4. 13 Perbandingan kinerja antar skenario	62
Tabel 4. 14 Selisih metrik (seimbang dengan tidak seimbang).....	62
Tabel 4. 15 Perbandingan dengan penelitian terdahulu	68

DAFTAR SIMBOL

w	: Vektor bobot (<i>weight vector</i>) pada SVM
b	: Nilai bias (<i>intercept</i>) pada SVM
x_i	: Vektor fitur masukan ke- i
y_i	: Label kelas data ke- i
ξ_i	: Variabel kelonggaran (<i>slack variable</i>) pada <i>soft margin</i> SVM
C	: Parameter penalti yang mengontrol keseimbangan antara <i>margin</i> dan toleransi kesalahan
γ	: Parameter <i>gamma</i> fungsi kernel (pengontrol radius atau kemiringan)
d	: Derajat polinomial (<i>degree</i>) pada kernel Polynomial
r	: Koefisien independen (<i>coef0</i>) pada kernel non-linear
$K(x_i, x_j)$	: Fungsi kernel yang memetakan vektor fitur
$f_k(x)$	: Fungsi keputusan model biner ke- k
$\hat{y}$	: Label kelas hasil prediksi
H	: Komponen <i>Hue</i> (jenis warna) pada ruang warna HSV
S	: Komponen <i>Saturation</i> (kejenuhan warna) pada ruang warna HSV
V	: Komponen <i>Value</i> (kecerahan) pada ruang warna HSV
μ	: Nilai rata-rata (<i>mean</i>)
σ	: Standar deviasi (<i>standard deviation</i>)
H_{mean}	: Nilai rata-rata komponen <i>Hue</i>
S_{mean}	: Nilai rata-rata komponen <i>Saturation</i>
V_{mean}	: Nilai rata-rata komponen <i>Value</i>
μY	: Nilai rata-rata intensitas citra <i>grayscale</i>
σY	: Nilai standar deviasi intensitas citra <i>grayscale</i>
F	: Vektor fitur hasil fusi (<i>feature fusion</i>) berdimensi lima
Z	: Nilai fitur hasil normalisasi <i>Z-score</i>
R	: Nilai intensitas piksel kanal <i>Red</i> (merah)
G	: Nilai intensitas piksel kanal <i>Green</i> (hijau)
B	: Nilai intensitas piksel kanal <i>Blue</i> (biru)
Y	: Nilai intensitas piksel <i>grayscale</i>
TP	: <i>True Positive</i> (data yang diprediksi benar sebagai anggota suatu kelas)
TN	: <i>True Negative</i> (data yang diprediksi benar bukan sebagai anggota suatu kelas)
FP	: <i>False Positive</i> (data yang sebenarnya bukan anggota suatu kelas, tetapi diprediksi sebagai anggota kelas tersebut)
FN	: <i>False Negative</i> (data yang sebenarnya termasuk anggota suatu kelas, tetapi diprediksi sebagai kelas lain)

ABSTRAK

Haqq, Ahmad Munawwarul, 2026. **Klasifikasi Penyakit Teh Berbasis Citra Menggunakan Support Vector Machine (SVM)**. Skripsi. Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Supriyono, M.Kom. (II) Ashri Shabrina Afrah, M.T.

Kata kunci: *Support Vector Machine*, klasifikasi citra, penyakit daun teh, fusi fitur, *hybrid resampling*

Tanaman teh (*Camellia sinensis*) merupakan komoditas perkebunan bernilai ekonomi tinggi yang rentan terhadap serangan penyakit daun. Identifikasi penyakit secara manual memerlukan waktu lama dan rentan terhadap kesalahan akibat kemiripan visual antar kelas penyakit. Oleh karena itu, diperlukan sistem yang mampu mengklasifikasikan penyakit daun teh secara otomatis, cepat, dan akurat. Penelitian ini bertujuan menganalisis kinerja algoritma *Support Vector Machine* (SVM) dengan empat fungsi kernel, yaitu Linear, Polynomial, *Radial Basis Function* (RBF), dan Sigmoid berbasis fusi fitur statistik HSV dan *Grayscale* dalam mengklasifikasikan enam kategori kondisi daun teh (*Red Leaf Spot*, *Brown Blight*, *Bird's Eye Spot*, *Anthraco*se, *Algal Leaf*, dan *Healthy Leaf*) pada dua kondisi dataset (tidak seimbang dan hasil *hybrid resampling*). Dataset yang digunakan berjumlah 1.929 citra yang diperoleh dari repositori publik Kaggle dengan ketidakseimbangan jumlah antar kelas. Tahapan penelitian meliputi prapengolahan citra menggunakan *Center-Based Square Cropping*, pembentukan dua skenario dataset, transformasi ruang warna HSV dan *Grayscale*, ekstraksi lima fitur statistik (*Mean Hue*, *Mean Saturation*, *Mean Value*, *Mean Grayscale*, *Standard Deviation Grayscale*), fusi fitur, normalisasi *Z-score*, serta klasifikasi SVM multikelas dengan strategi *One-vs-Rest* (OvR). Hasil pengujian menunjukkan bahwa kernel RBF memberikan performa terbaik pada kedua skenario dengan akurasi 80,83%, presisi 81,25%, *recall* 80,83%, dan *F1-score* 80,67%. *Hybrid resampling* memberikan peningkatan paling signifikan pada kernel Polynomial dengan kenaikan *F1-score* sebesar 1,80%, sementara kernel RBF menunjukkan stabilitas tinggi dengan perubahan mendekati nol. Kesalahan klasifikasi tertinggi terjadi antara kelas *Brown Blight* dan *Algal Leaf* akibat kemiripan karakteristik visual. Hasil penelitian menunjukkan bahwa algoritma SVM dengan kernel RBF efektif diterapkan dalam sistem klasifikasi penyakit daun teh dan dapat dimanfaatkan sebagai dasar pengembangan sistem identifikasi penyakit tanaman berbasis citra yang akurat dan efisien secara komputasi.

ABSTRACT

Haqq, Ahmad Munawwarul, 2026. **Image-Based Tea Leaf Disease Classification Using Support Vector Machine** (SVM). Thesis. Department of Informatics Engineering, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University of Malang. Advisors: (I) Supriyono, M.Kom. (II) Ashri Shabrina Afrah, M.T.

Keywords: Support Vector Machine, image classification, tea leaf disease, feature fusion, hybrid resampling

Tea plant (*Camellia sinensis*) is a high-value plantation commodity susceptible to various leaf diseases. Manual disease identification is time-consuming and prone to errors due to visual similarities among disease classes. Therefore, a system capable of automatically, rapidly, and accurately classifying tea leaf diseases is essential. This study aims to analyze the performance of the Support Vector Machine (SVM) algorithm with four kernel functions Linear, Polynomial, Radial Basis Function (RBF), and Sigmoid based on the fusion of HSV and Grayscale statistical features in classifying six categories of tea leaf conditions (Red Leaf Spot, Brown Blight, Bird's Eye Spot, Anthracnose, Algal Leaf, and Healthy Leaf) under two dataset conditions: imbalanced and hybrid resampled. The dataset consists of 1,929 images obtained from the Kaggle public repository, with class imbalance among categories. The research stages include image preprocessing using Center-Based Square Cropping, the creation of two dataset scenarios, color space transformation to HSV and Grayscale, extraction of five statistical features (Mean Hue, Mean Saturation, Mean Value, Mean Grayscale, and Standard Deviation Grayscale), feature fusion, Z-score normalization, and multi-class SVM classification using the One-vs-Rest (OvR) strategy. The results show that the RBF kernel achieved the best performance in both scenarios with an accuracy of 80.83%, precision of 81.25%, recall of 80.83%, and F1-score of 80.67%. Hybrid resampling provided the most significant improvement on the Polynomial kernel with an F1-score increase of 1.80%, while the RBF kernel demonstrated high stability with near-zero changes. The highest classification errors occurred between the Brown Blight and Algal Leaf classes due to visual similarity. The findings indicate that the SVM algorithm with RBF kernel is effectively applicable for tea leaf disease classification and can serve as a foundation for developing an accurate and computationally efficient image-based plant disease identification system.

مستخلص البحث

الحق، أحمد منور، 2026. تصنيف أمراض الشاي على أساس الصور باستخدام آلة المتجه الداعم (SVM). البحث الجامعي. قسم الهندسة المعلوماتية، كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف الأول: سوفريونوا، الماجستير؛ المشرف الثاني: عصري صبرنا أفرح، الماجستير.

الكلمات الرئيسية: آلة متجه داعم، تصنيف صور، أمراض أوراق شاي، دمج ميزات، إعادة معاينة.

يعتبر نبات الشاي (*Camellia sinensis*) محصولاً زراعياً ذا قيمة اقتصادية عالية ومعرضاً لهجمات أمراض الأوراق. تتطلب عملية تحديد الأمراض يدوياً وقتاً طويلاً وتكون عرضة للأخطاء نتيجة التشابه البصري بين فئات الأمراض. لذلك، هناك حاجة إلى نظام قادر على تصنيف أمراض أوراق الشاي بشكل تلقائي وسريع ودقيق. هدف هذا البحث إلى تحليل أداء خوارزمية آلة المتجه الداعم (SVM) باستخدام أربع دوال نواة، وهي الخطية والمتعددة الحدود ودالة الأساس الشعاعي (RBF) ودالة سيجمويد، بناءً على دمج ميزات الإحصاءات للصور بصيغتي HSV وتدرج الرمادي في تصنيف ست فئات من حالات أوراق الشاي (بقعة الورقة الحمراء، العفن البني، بقعة عين الطائر، الأنثراكنوز، ورقة الطحلب، والورقة السليمة) في حالتين من مجموعات البيانات (غير متوازنة ونتيجة إعادة معاينة). تم استخدام مجموعة بيانات تحتوي على 1929 صورة تم الحصول عليها من المستودع العام Kaggle مع تفاوت في عدد الحالات بين الفئات. تشمل مراحل البحث معالجة الصور المسبقة باستخدام القص المربع المرتكز على المركز، وتشكيل سيناريوهين لمجموعة البيانات، وتحويل فضاء الألوان إلى HSV و Grayscale، واستخراج خمس ميزات إحصائية (متوسط التدرج اللوني، متوسط التشبع، متوسط القيمة، متوسط الرمادي، الانحراف المعياري للرمادي)، ودمج الميزات، وتطبيق Z-score، بالإضافة إلى التصنيف باستخدام SVM متعددة الفئات باستراتيجية One-vs-Rest (OvR). أظهرت نتائج الاختبار أن نواة RBF تقدم أفضل أداء في السيناريوهين مع دقة 80.83٪، وثبات 81.25٪، واسترجاع 80.83٪، ودرجة ف1 بلغت 80.67٪. قدمت إعادة المعاينة أعلى تحسن على نواة Polynomial بزيادة في درجة ف1 بمقدار 1.80٪، بينما أظهرت نواة RBF استقراراً عالياً مع تغييرات تقترب من الصفر. حدثت أعلى معدلات خطأ في التصنيف بين فئة العفن البني و ورقة الطحلب بسبب التشابه في الخصائص البصرية. توضح نتائج البحث أن خوارزمية SVM مع نواة RBF فعالة عند تطبيقها في نظام تصنيف الأمراض أوراق الشاي ويمكن الاستفادة منها كأساس لتطوير نظام تحديد أمراض النباتات المعتمد على الصور بشكل دقيق وفعال من حيث الحوسبة.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Tanaman teh (*Camellia sinensis*) merupakan komoditas bernilai ekonomi tinggi yang rentan terhadap penyakit daun, yang dapat menurunkan kualitas dan produktivitas (Ginanjari dkk., 2019). Hingga saat ini, identifikasi penyakit masih banyak dilakukan secara manual melalui pengamatan visual oleh petani atau pakar tanaman. Metode ini memerlukan waktu yang relatif lama dan rentan terhadap subjektivitas karena karakteristik bercak penyakit pada daun teh sering kali memiliki kemiripan warna dan tekstur antar kelas penyakit.

Dalam perspektif Islam, manusia diperintahkan mengkaji fenomena alam sebagai pengenalan terhadap kebesaran Allah *Subhanahu wa Ta'ala*. Allah *Subhanahu wa Ta'ala* berfirman dalam QS. Al-An'am ayat 99:

وَهُوَ الَّذِي أَنْزَلَ مِنَ السَّمَاءِ مَاءً فَأَخْرَجْنَا بِهِ نَبَاتَ كُلِّ شَيْءٍ فَأَخْرَجْنَا مِنْهُ خَضِرًا نُخْرِجُ مِنْهُ حَبًّا مُتَرَاكِبًا وَمِنَ النَّخْلِ مِنْ طَلْعِهَا قِنْوَانٌ دَانِيَةٌ وَجَنَّاتٍ مِنْ أَعْنَابٍ وَالزَّيْتُونَ وَالرُّمَّانَ مُشْتَبِهًا وَغَيْرَ مُتَشَابِهٍ انظُرُوا إِلَى ثَمَرِهِ إِذَا أَثْمَرَ وَيَنْعِهِ ۚ إِنَّ فِي ذَٰلِكُمْ لَآيَاتٍ لِّقَوْمٍ يُؤْمِنُونَ ﴿٩٩﴾

"Dialah yang menurunkan air dari langit lalu dengannya Kami menumbuhkan segala macam tumbuhan. Maka, darinya Kami mengeluarkan tanaman yang menghijau. Darinya Kami mengeluarkan butir yang bertumpuk (banyak). Dari mayang kurma (mengurai) tangkai-tangkai yang menjuntai. (Kami menumbuhkan) kebun-kebun anggur. (Kami menumbuhkan pula) zaitun dan delima yang serupa dan yang tidak serupa. Perhatikanlah buahnya pada waktu berbuah dan menjadi masak. Sesungguhnya pada yang demikian itu benar-benar terdapat tanda-tanda (kekuasaan Allah) bagi kaum yang beriman." (Q.S. Al-An'am/6: 99).

Ayat ini mengandung perintah "*unzhuru*" (perhatikanlah) yang menurut Tafsir Al-Wajiz berarti seruan untuk merenungkan keajaiban ciptaan Allah, seperti

turunnya hujan yang menumbuhkan berbagai jenis tumbuhan dengan warna, bentuk, dan rasa yang berbeda-beda (az-Zuhaili, 2019). Tafsir as-Sa'di menambahkan bahwa perintah ini mendorong manusia untuk melakukan observasi ilmiah karena dalam keanekaragaman ciptaan tersebut terdapat tanda-tanda kekuasaan, rahmat, dan kesempurnaan kodrat Allah (as-Sa'di, 2019).

Selain bernilai ibadah, penelitian ini juga merupakan bentuk tolong-menolong dalam kebaikan sebagaimana diperintahkan dalam QS. Al-Maidah ayat 2:

يَا أَيُّهَا الَّذِينَ آمَنُوا لَا تَحِلُّوا شَعَائِرَ اللَّهِ وَلَا الشَّهْرَ الْحَرَامَ وَلَا الْهَدْيَ وَلَا الْقَلَائِدَ وَلَا آمِينَ الْبَيْتِ الْحَرَامِ يَنْتَعُونَ فَضْلًا مِّن رَّبِّهِمْ وَرِضْوَانًا وَإِذَا حَلَلْتُمْ فَاصْطَادُوا وَلَا يَجْرِمَنَّكُمْ شَنَاٰنُ قَوْمٍ أَن صَدُّوكُمْ عَنِ الْمَسْجِدِ الْحَرَامِ أَن تَعْتَدُوا وَتَعَاوَنُوا عَلَى الْبِرِّ وَالتَّقْوَىٰ وَلَا تَعَاوَنُوا عَلَى الْإِثْمِ وَالْعُدْوَانِ وَاتَّقُوا اللَّهَ إِنَّ اللَّهَ شَدِيدُ الْعِقَابِ ﴿٢﴾

" Wahai orang-orang yang beriman, janganlah kamu melanggar syiar-syiar (kesucian) Allah, jangan (melanggar kehormatan) bulan-bulan haram, jangan (mengganggu) hadyu (hewan-hewan kurban) dan qalā'id (hewan-hewan kurban yang diberi tanda), dan jangan (pula mengganggu) para pengunjung Baitulharam sedangkan mereka mencari karunia dan rida Tuhannya! Apabila kamu telah bertahalul (menyelesaikan ihram), berburulah (jika mau). Janganlah sekali-kali kebencian(-mu) kepada suatu kaum, karena mereka menghalang-halangi-mu dari Masjidilharam, mendorongmu berbuat melampaui batas (kepada mereka). Tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa, dan jangan tolong-menolong dalam berbuat dosa dan permusuhan. Bertakwalah kepada Allah, sesungguhnya Allah sangat berat siksaan-Nya." (Q.S. Al-Maidah/5: 2).

Menurut Tafsir as-Sa'di, ayat ini memerintahkan kaum mukmin untuk saling membantu dalam Kebajikan yakni segala perbuatan baik yang dicintai Allah, baik yang berkaitan dengan hak Allah maupun hak manusia dan dalam takwa, yaitu sikap meninggalkan segala perbuatan yang dibenci Allah dan Rasul-Nya (as-Sa'di, 2019). Tafsir Al-Wajiz menjelaskan bahwa perintah tolong-menolong ini bersifat universal, mencakup seluruh aspek kehidupan, termasuk upaya membantu sesama dalam memecahkan masalah yang mereka hadapi (az-Zuhaili, 2019).

Deteksi dini penyakit juga memungkinkan pengendalian yang lebih tepat sasaran, mengurangi penggunaan pestisida berlebihan, dan menjaga keseimbangan ekosistem. Hal ini sejalan dengan larangan membuat kerusakan di bumi dalam QS. Al-A'raf ayat 56:

وَلَا تُفْسِدُوا فِي الْأَرْضِ بَعْدَ إِصْلَاحِهَا وَادْعُوهُ حَوْفًا وَطَمَعًا إِنَّ رَحْمَتَ اللَّهِ قَرِيبٌ مِّنَ الْمُحْسِنِينَ ﴿٥٦﴾

"Janganlah kamu berbuat kerusakan di bumi setelah diatur dengan baik. Berdoalah kepada-Nya dengan rasa takut dan penuh harap. Sesungguhnya rahmat Allah sangat dekat dengan orang-orang yang berbuat baik." (Q.S. Al-A'raf/7: 56).

Menurut Tafsir Al-Wajiz, ayat ini melarang berbuat kerusakan di bumi dengan berbuat syirik dan maksiat setelah Allah memperbaikinya melalui pengutusan para rasul dan penjelasan syariat (az-Zuhaili, 2019). Tafsir as-Sa'di menjelaskan bahwa larangan ini mencakup segala bentuk kemaksiatan yang dapat merusak akhlak, amal, dan rezeki, karena ketaatan kepada Allah justru akan memperbaiki keadaan dunia dan akhirat (as-Sa'di, 2019). Tafsir Ibnu Katsir menegaskan bahwa Allah melarang perbuatan yang memberi kemudharatan setelah adanya perbaikan, karena kerusakan tersebut membahayakan seluruh hamba (Katsir & Umar, 2017).

Perkembangan teknologi pengolahan citra digital memungkinkan identifikasi penyakit tanaman dilakukan secara otomatis melalui analisis karakteristik visual daun. Metode *deep learning* seperti *Convolutional Neural Network* (CNN) mampu memberikan akurasi tinggi, namun memerlukan data besar dan komputasi berat (Lu dkk., 2021). Sebaliknya, *K-Nearest Neighbor* (K-NN) lebih sederhana tetapi rentan terhadap *curse of dimensionality* pada data berdimensi tinggi (Zhang dkk., 2018). *Support Vector Machine* (SVM) hadir sebagai alternatif

yang mampu menangani data berdimensi tinggi dengan jumlah data terbatas (Chandra & Bedi, 2021).

Kinerja SVM sangat dipengaruhi oleh kualitas fitur dan distribusi data. Ketidakseimbangan jumlah data antar kelas dapat menyebabkan model bias pada kelas mayoritas sehingga menurunkan kemampuan klasifikasi (Mohasel & Koosha, 2025; Valkenborg dkk., 2023). Selain itu, kemiripan visual antar penyakit daun teh menuntut representasi fitur yang mampu menangkap informasi warna dan intensitas citra secara efektif. Oleh karena itu, diperlukan kombinasi fitur yang komprehensif serta strategi penyeimbangan data untuk mengurangi bias akibat distribusi kelas yang tidak merata.

Berdasarkan tinjauan penelitian terdahulu, terdapat tiga celah utama yang belum terjawab. Pertama, belum ada studi yang secara terpadu menggabungkan fusi fitur statistik HSV dan *grayscale*, perbandingan empat kernel SVM, serta pengujian pada kondisi data tidak seimbang dan hasil *hybrid resampling* untuk klasifikasi penyakit daun teh. Kedua, penelitian pada daun teh umumnya menggunakan fitur tekstur kompleks atau *deep learning*, sementara pendekatan fusi fitur statistik yang sederhana dan efisien belum banyak dieksplorasi. Ketiga, analisis proses internal SVM seperti *support vector*, *hyperplane*, dan *decision function* masih jarang dibahas secara mendalam. Penelitian ini hadir untuk mengisi celah-celah tersebut.

Berdasarkan permasalahan tersebut, penelitian ini mengimplementasikan *Support Vector Machine* (SVM) multikelas dengan strategi *One-vs-Rest* (OvR) untuk mengklasifikasikan penyakit daun teh berdasarkan fusi fitur statistik warna dan intensitas citra. Penelitian ini juga menerapkan pendekatan *hybrid resampling*

untuk menghasilkan distribusi data yang lebih seimbang serta melakukan pengujian komparatif terhadap fungsi kernel Linear, Polynomial, *Radial Basis Function* (RBF), dan Sigmoid pada dua kondisi dataset: tidak seimbang dan hasil penyeimbangan. Hasil penelitian diharapkan dapat memberikan informasi mengenai pengaruh penyeimbangan data dan pemilihan fungsi kernel terhadap performa klasifikasi, sehingga menjadi dasar pengembangan sistem identifikasi penyakit tanaman yang lebih akurat dan efisien.

1.2 Pernyataan Masalah

Bagaimana kinerja algoritma *Support Vector Machine* (SVM) menggunakan fungsi kernel Linear, Polynomial, *Radial Basis Function* (RBF), dan Sigmoid berbasis fusi fitur statistik HSV dan *Grayscale* dalam mengklasifikasikan penyakit daun teh pada kondisi dataset tidak seimbang dan dataset hasil *hybrid resampling*?

1.3 Batasan Masalah

Berikut batasan masalah yang digunakan dalam penelitian ini:

1. Dataset yang digunakan merupakan data sekunder citra daun teh dari repositori Kaggle yang mencakup enam kelas kondisi: *Red Leaf Spot*, *Brown Blight*, *Bird's Eye Spot*, *Anthrachnose*, *Algal Leaf* (*Cephaleuros virescens*), dan *Healthy Leaf*.
2. Tahap prapengolahan citra dilakukan menggunakan metode *Center-Based Square Cropping* dengan ukuran citra 225×225 piksel.

3. Fitur yang digunakan merupakan hasil fusi lima fitur statistik, yaitu *Mean Hue*, *Mean Saturation*, *Mean Value*, *Mean Grayscale*, dan *Standard Deviation Grayscale*.
4. Penanganan ketidakseimbangan data dilakukan menggunakan pendekatan *Hybrid Resampling* yang mengombinasikan teknik augmentasi citra dan *random undersampling*.
5. Proses klasifikasi multikelas dilakukan menggunakan algoritma *Support Vector Machine* (SVM) dengan strategi *One-vs-Rest* (OvR).
6. Fungsi kernel yang digunakan pada algoritma SVM meliputi kernel *Linear*, *Polynomial*, *Radial Basis Function* (RBF), dan *Sigmoid*.
7. Evaluasi kinerja model dilakukan menggunakan *confusion matrix*, *accuracy*, *precision*, *recall*, dan *F1-score*.

1.4 Tujuan Penelitian

Menganalisis kinerja algoritma *Support Vector Machine* (SVM) menggunakan fungsi kernel *Linear*, *Polynomial*, *Radial Basis Function* (RBF), dan *Sigmoid* berbasis fusi fitur statistik HSV dan *Grayscale* dalam mengklasifikasikan penyakit daun teh pada kondisi dataset tidak seimbang dan dataset hasil *hybrid resampling*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat teoritis: Memberikan kontribusi pada bidang pengolahan citra digital dan *machine learning*, khususnya terkait penerapan fusi fitur statistik

HSV dan Grayscale serta perbandingan fungsi kernel SVM dalam menangani klasifikasi multikelas pada dataset yang tidak seimbang.

2. Manfaat praktis: Menjadi landasan dalam pengembangan sistem identifikasi penyakit daun teh berbasis citra yang lebih akurat dan efisien. Selain itu, hasil penelitian dapat dijadikan referensi bagi praktisi pertanian, pakar perkebunan, maupun peneliti selanjutnya dalam menerapkan metode SVM untuk klasifikasi penyakit tanaman.

BAB II

STUDI PUSTAKA

2.1 Penelitian Terkait

Untuk memperkuat landasan penelitian ini, dipaparkan sejumlah penelitian terdahulu yang relevan dengan komponen kunci skripsi ini, yaitu metode klasifikasi CNN dan K-NN sebagai pembanding, metode SVM dengan fusi fitur, serta teknik penanganan ketidakseimbangan data. Setiap penelitian dianalisis untuk mengidentifikasi celah yang belum terjawab dan menjadi dasar orisinalitas penelitian ini.

Convolutional Neural Network (CNN) merupakan metode *deep learning* yang unggul dalam mengenali pola pada citra, namun membutuhkan data besar dan komputasi tinggi. Penelitian berjudul "*Performance enhancement of kernelized SVM with deep learning features for tea leaf disease prediction*" oleh Bhagat & Kumar (2024) mengatasi keterbatasan data pada klasifikasi penyakit daun teh dengan menerapkan proses augmentasi gambar yang meningkatkan jumlah dataset hampir empat belas kali lipat. Namun, karena jumlah data yang terbatas masih belum memadai untuk klasifikasi berbasis *deep learning* murni, penelitian ini mengusulkan teknik hibrida yang menggabungkan ekstraksi fitur *deep learning* menggunakan VGG-16 dengan *classifier* berbasis machine learning. Hasilnya, model hibrida yang menggabungkan fitur CNN dengan SVM kernel Sigmoid dan Linear mencapai akurasi tertinggi 96,67%, mengungguli pendekatan CNN murni serta *classifier* lainnya seperti KNN, XGB, dan *Random Forest*. Penelitian ini mem-

buktikan bahwa SVM tetap kompetitif pada data terbatas dengan pendekatan *hybrid feature extraction*, namun belum menguji performa pada kondisi data yang tidak seimbang. Sementara itu, penelitian berjudul "*Towards precision agriculture tea leaf disease detection using CNNs and image processing*" oleh Rahat dkk. (2025) mengembangkan model CNN kustom dengan arsitektur kompleks yang dilengkapi *residual blocks* dan *GlobalAveragePooling2D*, serta dioptimasi menggunakan *Adam optimizer*. Model ini berhasil mencapai akurasi 99% pada delapan kelas penyakit daun teh. Namun, pendekatan *deep learning* ini memerlukan sumber daya komputasi yang besar dan belum mengeksplorasi efisiensi metode SVM. Kedua penelitian ini menunjukkan bahwa pendekatan *deep learning* cenderung kurang efisien untuk skenario dengan data dan sumber daya terbatas.

K-Nearest Neighbor (K-NN) bekerja berdasarkan jarak terdekat antar data dan mudah diimplementasikan, namun sensitif terhadap data berdimensi tinggi. Penelitian berjudul "*Implementasi Metode K-NN Untuk Klasifikasi Penyakit Daun Teh Berdasarkan Fitur Tekstur GLCM*" oleh Ardhiansyah dkk. (2026) mengklasifikasikan enam penyakit daun teh menggunakan ekstraksi fitur tekstur GLCM dengan algoritma K-NN. Hasil terbaik diperoleh pada skenario pembagian data 80:20 dengan nilai $K=1$, menghasilkan akurasi sebesar 95,83%. Meskipun akurasi yang dicapai cukup tinggi, metode K-NN memiliki kelemahan mendasar yaitu sensitivitas terhadap data berdimensi tinggi (*curse of dimensionality*). Hal ini memperkuat alasan pemilihan SVM yang secara teoritis lebih tahan terhadap permasalahan data berdimensi tinggi.

Support Vector Machine (SVM) dikenal tangguh pada data berdimensi tinggi dan data terbatas. Kinerjanya sangat bergantung pada kualitas representasi fitur. Penelitian berjudul "*Deteksi Penyakit Daun Cabai Menggunakan Kombinasi GLCM Dan HSV dengan Klasifikasi SVM*" oleh Nurmadinah dkk. (2025) membandingkan fitur statistik ruang warna HSV dan RGB pada daun cabai menggunakan SVM kernel RBF. Hasilnya, fitur HSV memberikan akurasi 10% lebih tinggi dibanding RGB, dengan kombinasi GLCM dan HSV mencapai akurasi 91,30% pada parameter $C=10$ dan $\gamma=0,1$. Meski demikian, penelitian ini belum menggabungkan fitur warna HSV dengan informasi tekstur dari citra Grayscale secara spesifik pada daun teh. Lebih lanjut, penelitian berjudul "*Comparison of Tea Leaf Disease Classification Using SVM with MobileNetV2 and MobileNetV3-Small Feature Extractors*" oleh Raihan & Wijaya (2026) mengeksplorasi penggunaan MobileNet sebagai ekstraktor fitur yang dikombinasikan dengan SVM untuk klasifikasi penyakit teh. Hasilnya, *MobileNetV2-SVM* mencapai akurasi 75,3% dan *MobileNetV3-Small-SVM* mencapai 75,1%. Meskipun akurasi yang dicapai masih tergolong rendah, penelitian ini membuktikan bahwa pendekatan *hybrid* CNN-SVM dapat menjadi alternatif yang lebih efisien secara komputasi, terutama untuk *deployment* pada perangkat bergerak. Namun, penelitian ini belum menggunakan fusi fitur statistik berbasis HSV dan *Grayscale* yang lebih sederhana dan ringan secara komputasi. Penelitian berjudul "*Identifikasi dan Klasifikasi Penyakit Daun Jagung Menggunakan Support Vector Machine*" oleh Suhendra & Juliwardi (2022) menerapkan SVM dengan fusi fitur warna dan tekstur GLCM pada daun jagung,

menghasilkan akurasi 99,7% dengan fitur gabungan. Penelitian ini menguji beberapa kernel SVM dan menunjukkan bahwa kernel RBF memberikan performa terbaik. Namun, penelitian ini belum membandingkan keempat kernel SVM (Linear, Polynomial, RBF, Sigmoid) secara menyeluruh pada dataset yang sama, serta belum menguji pada kondisi data tidak seimbang.

Permasalahan ketidakseimbangan data menyebabkan model klasifikasi cenderung bias ke kelas mayoritas. Penelitian berjudul "*Application of extreme learning machine in plant disease prediction for highly imbalanced dataset*" oleh Bhatia dkk. (2020) menerapkan berbagai Teknik *resampling* seperti SMOTE, ROS, RUS, dan IMPS pada dataset penyakit tanaman yang sangat tidak seimbang. Hasilnya, teknik IMPS (*Importance Sampling*) memberikan performa terbaik dengan akurasi 89,19% dan AUC 88,57%, mengungguli SMOTE yang hanya mencapai 77,03%. Namun, teknik SMOTE tunggal diketahui rentan menghasilkan sampel *noise* pada area tumpang tindih antar kelas yang memiliki kemiripan visual tinggi. Oleh karena itu, diperlukan metode *hybrid resampling* yang menggabungkan *oversampling* dan *undersampling* untuk membersihkan *noise* dan menciptakan batas kelas yang lebih tegas.

Tabel 2. 1 Ringkasan penelitian terkait

No	Peneliti dan Tahun	Judul Penelitian	Metode dan Dataset Penelitian	Hasil	Gap
1	Bhagat & Kumar (2024)	<i>Performance enhancement of kernelized SVM with deep learning features for tea leaf disease prediction</i>	VGG-16 + SVM (Sigmoid & Linear)	Akurasi 96,67%; SVM unggul dari CNN pada data terbatas.	Belum menguji data tidak seimbang.
2	Rahat dkk. (2025)	<i>Towards precision agriculture tea leaf disease detection using</i>	Custom CNN + Adam Optimizer	Akurasi 99% pada 8 kelas penyakit teh.	Komputasi tinggi; belum

No	Peneliti dan Tahun	Judul Penelitian	Metode dan Dataset Penelitian	Hasil	Gap
		<i>CNNs and image processing</i>			mengeksplorasi efisiensi SVM.
3	Ardhiansyah dkk. (2026)	Implementasi Metode K-NN Untuk Klasifikasi Penyakit Daun Teh Berdasarkan Fitur Tekstur GLCM	K-NN + GLCM	Akurasi 95,83% dengan K=1.	K-NN sensitif data berdimensi tinggi; belum dibandingkan dengan SVM.
4	Nurmadinah dkk. (2025)	Deteksi Penyakit Daun Cabai Menggunakan Kombinasi GLCM Dan HSV dengan Klasifikasi SVM	SVM RBF + GLCM + HSV	Akurasi 91,30%; HSV unggul 10% dari RGB.	Fusi HSV + Grayscale untuk daun teh belum dieksplorasi.
5	Raihan & Wijaya (2026)	<i>Comparison of Tea Leaf Disease Classification Using SVM with MobileNetV2 and MobileNetV3-Small Feature Extractors</i>	MobileNet + SVM	Akurasi 75,3% dan 75,1%; efisien secara komputasi.	Akurasi rendah; belum menggunakan fusi fitur statistik HSV dan Grayscale.
6	Suhendra & Juliwardi (2022)	Identifikasi dan Klasifikasi Penyakit Daun Jagung Menggunakan <i>Support Vector Machine</i>	SVM + Fitur Warna & Tekstur (GLCM)	Akurasi 99,7% pada data daun jagung.	Belum membandingkan empat kernel SVM secara bersamaan.
7	Bhatia dkk. (2020)	<i>Application of extreme learning machine in plant disease prediction for highly imbalanced dataset</i>	ELM + SMOTE, ROS, RUS, IMPS	IMPS terbaik (89,19%); SMOTE rentan <i>noise</i> .	SMOTE tunggal menghasilkan <i>noise</i> ; perlu <i>hybrid resampling</i> .

2.2 Tanaman Teh

Tanaman teh (*Camellia sinensis* L.) merupakan salah satu komoditas perkebunan bernilai ekonomi tinggi yang berperan penting dalam sektor agrikultur dan agroindustri (Anjarsari dkk., 2020). Selain memiliki nilai ekonomi, daun teh juga mengandung berbagai senyawa metabolit sekunder yang bermanfaat bagi kesehatan manusia (Eskundari, 2020). Produktivitas tanaman teh sangat dipengaruhi oleh kondisi kesehatan daun sebagai organ utama tempat berlangsungnya proses fotosintesis.

Serangan patogen seperti jamur dan alga dapat menimbulkan berbagai gejala pada permukaan daun, seperti munculnya bercak, lesi, serta perubahan warna akibat kerusakan pigmen daun (Suryadi dkk., 2019). Gejala tersebut menjadi indikator penting dalam proses identifikasi penyakit. Namun, identifikasi secara visual masih rentan terhadap subjektivitas dan membutuhkan waktu yang relatif lama, terutama pada area perkebunan yang luas.

Dalam sistem pengolahan citra digital, karakteristik visual yang ditimbulkan oleh penyakit dapat direpresentasikan ke dalam bentuk fitur numerik yang menggambarkan informasi warna dan intensitas citra. Representasi tersebut selanjutnya digunakan oleh algoritma *machine learning* untuk melakukan proses klasifikasi penyakit daun teh secara otomatis.

2.2.1 Morfologi Penyakit Teh

Perbedaan karakteristik visual antara daun sehat dan daun yang terinfeksi menjadi fondasi utama dalam klasifikasi penyakit berbasis citra. Berdasarkan (Asterius, 2023) enam kategori kondisi daun teh yang digunakan dalam penelitian memiliki karakteristik morfologi sebagaimana dirangkum dalam Tabel 2.2 berikut:

Tabel 2. 2 Morfologi penyakit pada daun teh

No	Kategori Klasifikasi	Manifestasi Morfologi
1	Sehat (<i>Healthy</i>)	Daun berwarna hijau homogen dengan permukaan yang relatif halus dan tekstur yang seragam.
2	Hawar Coklat (<i>Brown Blight</i>)	Bercak berwarna coklat tua hingga hitam dengan bentuk tidak beraturan dan umumnya muncul pada tepi daun.
3	Bercak Mata Burung (<i>Bird's Eye Spot</i>)	Lesi berbentuk melingkar dengan pusat berwarna abu-abu yang dikelilingi zona kecokelatan.
4	Antraknosa (<i>Anthracoese</i>)	Bercak gelap yang berkembang menjadi lesi nekrotik dengan area sekitarnya mengalami klorosis.
5	Alga (<i>Algal Leaf/Cephaleuros virescens</i>)	Bercak berwarna merah karat hingga oranye dengan permukaan yang cenderung kasar.
6	Daun Merah (<i>Red Leaf Spot</i>)	Perubahan warna daun yang ditandai dengan dominasi pigmen merah akibat klorosis atau stres pada tanaman.

Setiap jenis penyakit menunjukkan manifestasi morfologis yang berbeda, mulai dari perubahan warna yang khas, bentuk bercak yang beragam, hingga tekstur permukaan daun yang bervariasi. Karakteristik-karakteristik visual ini selanjutnya diekstraksi dan direpresentasikan dalam bentuk fitur numerik yang dapat diproses oleh algoritma klasifikasi. Masing-masing kelas memiliki ciri visual yang khas, mulai dari perubahan warna, bentuk bercak, hingga tekstur permukaan daun. Perbedaan karakteristik visual inilah yang menjadi dasar dalam proses ekstraksi fitur untuk membedakan setiap kelas penyakit.

2.3 Pengolahan Citra Digital

Pengolahan citra digital (*digital image processing*) merupakan bidang ilmu yang mempelajari teknik manipulasi dan analisis citra untuk memperoleh informasi yang berguna. Dalam penelitian ini, pengolahan citra digunakan untuk mengubah citra daun teh menjadi representasi numerik yang dapat diproses oleh algoritma klasifikasi. Secara umum, tahapan yang dilakukan meliputi prapengolahan citra, penanganan ketidakseimbangan data, transformasi ruang warna, ekstraksi fitur, serta penggabungan fitur sebelum dilakukan proses klasifikasi menggunakan *Support Vector Machine* (SVM).

2.3.1 Prapengolahan Citra

Prapengolahan citra (*preprocessing*) merupakan tahapan awal yang bertujuan untuk memfokuskan area pengamatan pada objek utama dan menghasilkan ukuran citra yang seragam. Tahapan ini dilakukan untuk mengurangi

pengaruh latar belakang yang tidak relevan sehingga karakteristik visual daun dapat dipertahankan secara optimal pada proses ekstraksi fitur dan klasifikasi.

2.3.1.1 *Center-Based Square Cropping*

Metode *Center-Based Square Cropping* digunakan untuk memfokuskan area pengamatan pada objek daun sekaligus menyeragamkan dimensi spasial citra. Proses diawali dengan menentukan area objek menggunakan koordinat minimum dan maksimum kontur, sehingga diperoleh titik acuan dan dimensi *bounding box*. Koordinat sudut kiri atas serta ukuran *bounding box* dihitung menggunakan Persamaan (2.1) dan Persamaan (2.2).

$$x_{min} = \min(C_x), \quad y_{min} = \min(C_y) \quad (2.1)$$

$$W = \max(C_x) - x_{min}, H = \max(C_y) - y_{min} \quad (2.2)$$

Titik pusat *bounding box* digunakan sebagai acuan pemotongan berbentuk bujur sangkar dan dihitung menggunakan Persamaan (2.3).

$$x_c = x_{min} + \frac{W}{2}, y_c = y_{min} + \frac{H}{2} \quad (2.3)$$

Untuk mempertahankan area gejala di sekitar tepi daun, panjang sisi pemotongan (L) ditentukan berdasarkan sisi terpanjang *bounding box* dengan penambahan faktor margin (m), sebagaimana ditunjukkan pada Persamaan (2.4).

$$L = \max(w, h) \times (1 + m) \quad (2.4)$$

Selanjutnya, koordinat sudut kiri atas area pemotongan dihitung secara simetris terhadap titik pusat dengan memperhatikan batas koordinat citra, menggunakan Persamaan (2.5).

$$x'_{min} = \max\left(0, x_c - \frac{L}{2}\right), y'_{min} = \max\left(0, y_c - \frac{L}{2}\right) \quad (2.5)$$

Keterangan:

(C_x, C_y)	= himpunan koordinat kontur objek pada sumbu-x dan sumbu-y
(x_{min}, y_{min})	= koordinat minimum yang menjadi acuan sudut kiri atas <i>bounding box</i>
(W)	= lebar (<i>width</i>) <i>bounding box</i>
(H)	= tinggi (<i>height</i>) <i>bounding box</i>
(x_c, y_c)	= koordinat titik pusat (<i>centroid</i>) <i>bounding box</i>
(L)	= panjang sisi daerah pemotongan berbentuk bujur sangkar
(m)	= faktor margin yang digunakan untuk mempertahankan area di sekitar objek daun
x'_{min}, y'_{min}	= koordinat sudut kiri atas daerah pemotongan

2.3.2 Augmentasi Citra

Augmentasi data merupakan teknik manipulasi citra yang bertujuan untuk meningkatkan jumlah dan keragaman data latih tanpa mengubah informasi semantik objek. Teknik ini banyak digunakan untuk mengurangi permasalahan ketidakseimbangan jumlah sampel antar kelas (*class imbalance*) sekaligus meningkatkan kemampuan generalisasi model. Dalam penelitian ini, augmentasi citra menjadi bagian dari pendekatan *hybrid resampling* yang digunakan untuk menghasilkan distribusi data yang lebih seimbang.

Transformasi geometri yang digunakan meliputi operasi rotasi dan pembalikan (*flipping*). Rotasi dilakukan dengan memutar posisi piksel citra terhadap titik pusat sehingga koordinat piksel baru (x', y') diperoleh dari koordinat awal (x, y) melalui matriks transformasi yang ditunjukkan pada Persamaan (2.6).

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2.6)$$

Operasi pembalikan (*flipping*) terdiri atas pembalikan horizontal dan pembalikan vertikal. Pada pembalikan horizontal, koordinat piksel ditransformasikan menggunakan Persamaan (2.7).

$$x' = W - 1 - x, \quad y' = y \quad (2.7)$$

Sedangkan pada pembalikan vertikal digunakan Persamaan (2.8).

$$x' = x, \quad y' = H - 1 - y \quad (2.8)$$

Keterangan:

(x, y) = koordinat piksel sebelum transformasi

(x', y') = koordinat piksel setelah transformasi

(θ) = sudut rotasi

(W) = lebar citra

(H) = tinggi citra

2.3.3 Transformasi Ruang Warna

Transformasi ruang warna dilakukan untuk memperoleh representasi citra yang lebih sesuai dengan karakteristik fitur yang akan diekstraksi. Pada penelitian ini, citra RGB ditransformasikan ke ruang warna HSV untuk memperoleh informasi warna yang lebih stabil terhadap perubahan pencahayaan, serta ke citra *grayscale* untuk merepresentasikan informasi intensitas dan tekstur. Kedua representasi tersebut kemudian digunakan sebagai dasar dalam proses ekstraksi fitur statistik pada tahap selanjutnya.

2.3.3.1 Ruang Warna HSV

Ruang warna HSV (*Hue*, *Saturation*, *Value*) digunakan untuk merepresentasikan informasi warna secara lebih stabil terhadap perubahan intensitas pencahayaan. Komponen *Hue* (H) menyatakan jenis warna, *Saturation* (S) menunjukkan tingkat kejenuhan warna, sedangkan *Value* (V) merepresentasikan tingkat kecerahan.

Sebelum proses transformasi dilakukan, nilai piksel RGB dinormalisasi ke rentang 0–1 menggunakan Persamaan (2.9).

$$R' = \frac{R}{255}, \quad G' = \frac{G}{255}, \quad B' = \frac{B}{255} \quad (2.9)$$

Nilai maksimum, minimum, dan *chroma* didefinisikan menggunakan Persamaan (2.10) dan Persamaan (2.11).

$$C_{max} = \max(R', G', B'), \quad C_{min} = \min(R', G', B') \quad (2.10)$$

$$\Delta = C_{max} - C_{min} \quad (2.11)$$

Selanjutnya, nilai *Hue*, *Saturation*, dan *Value* dihitung menggunakan persamaan berikut:

$$H = \begin{cases} 0^\circ, & \Delta = 0 \\ 60^\circ \times \left(\frac{G' - B'}{\Delta} \bmod 6 \right), & C_{max} = R' \\ 60^\circ \times \left(\frac{B' - R'}{\Delta} + 2 \right), & C_{max} = G' \\ 60^\circ \times \left(\frac{R' - G'}{\Delta} + 4 \right), & C_{max} = B' \end{cases} \quad (2.12)$$

$$S = \begin{cases} 0, & \text{jika } C_{max} = 0 \\ \frac{\Delta}{C_{max}}, & \text{jika } C_{max} \neq 0 \end{cases} \quad (2.13)$$

$$V = C_{max} \quad (2.14)$$

Keterangan:

- (R', G', B') = nilai piksel merah, hijau, dan biru hasil normalisasi
- (C_{max}) = nilai maksimum dari ketiga kanal RGB
- (C_{min}) = nilai minimum dari ketiga kanal RGB
- (Δ) = selisih antara nilai maksimum dan minimum (*chroma*)
- (H) = komponen *Hue* yang merepresentasikan jenis warna
- (S) = komponen *Saturation* yang menunjukkan tingkat kejenuhan warna
- (V) = komponen *Value* yang menyatakan tingkat kecerahan

2.3.3.2 Citra *Grayscale*

Citra *grayscale* merupakan representasi citra berkanal tunggal yang hanya memuat informasi intensitas cahaya. Transformasi dari ruang warna RGB ke *grayscale* dilakukan menggunakan metode *luminosity* sebagaimana ditunjukkan pada Persamaan (2.15).

$$Y = 0.299R + 0.587G + 0.114B \quad (2.15)$$

Keterangan:

(Y) = intensitas piksel *grayscale*

(R, G, B) = intensitas piksel merah, hijau, dan biru

2.3.4 Ekstraksi dan Fusi Fitur

Ekstraksi dan fusi fitur merupakan tahapan yang bertujuan untuk mengubah representasi citra menjadi sekumpulan parameter numerik yang lebih ringkas dan informatif. Pada penelitian ini, fitur yang diekstraksi berasal dari karakteristik warna pada ruang HSV dan karakteristik intensitas pada citra *grayscale*. Seluruh fitur yang diperoleh kemudian digabungkan menjadi satu vektor fitur tunggal sehingga dapat digunakan sebagai masukan bagi algoritma klasifikasi.

2.3.4.1 Ekstraksi Fitur Statistik

Ekstraksi fitur statistik dilakukan untuk mereduksi dimensi citra menjadi sejumlah parameter numerik yang mampu merepresentasikan karakteristik visual daun teh. Pada penelitian ini, fitur yang digunakan terdiri atas nilai rata-rata (*mean*) dari kanal *Hue* (H), *Saturation* (S), *Value* (V), dan *grayscale*, serta nilai standar deviasi (*standard deviation*) dari citra *grayscale*. Nilai rata-rata digunakan untuk merepresentasikan karakteristik warna dominan dan intensitas rata-rata citra, sedangkan standar deviasi digunakan untuk menggambarkan tingkat variasi intensitas piksel.

Perhitungan nilai rata-rata pada setiap kanal dilakukan menggunakan Persamaan (2.16).

$$\mu = \frac{1}{N} \sum_{i=1}^N C_i \quad (2.16)$$

Pada kanal HSV, nilai μ masing-masing merepresentasikan H_{mean} , S_{mean} , dan V_{mean} , sedangkan pada citra *grayscale* nilai tersebut merepresentasikan μ_Y . Selain nilai rata-rata, karakteristik intensitas citra *grayscale* juga direpresentasikan menggunakan standar deviasi yang menggambarkan tingkat variasi intensitas piksel. Nilai standar deviasi dihitung menggunakan Persamaan (2.17).

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \mu)^2} \quad (2.17)$$

Keterangan:

- (μ) = nilai rata-rata fitur atau rata-rata intensitas *grayscale*
- (σ) = standar deviasi intensitas *grayscale*
- (C_i) = nilai intensitas piksel ke-(i) pada kanal terkait
- (Y_i) = nilai intensitas piksel *grayscale* ke-(i)
- (N) = jumlah piksel citra

2.3.4.2 *Feature Fusion* dan Normalisasi Skala Data

Feature fusion merupakan proses penggabungan beberapa fitur menjadi satu vektor fitur tunggal untuk memperoleh representasi informasi yang lebih komprehensif. Pada penelitian ini, *feature fusion* dilakukan pada tingkat fitur (*early fusion*) dengan menggabungkan tiga fitur statistik HSV dan dua fitur statistik *grayscale* sehingga diperoleh vektor fitur berdimensi lima sebagaimana ditunjukkan pada Persamaan (2.18).

$$F = [H_{mean}, S_{mean}, V_{mean}, \mu_Y, \sigma_Y] \quad (2.18)$$

Perbedaan rentang nilai antar fitur dapat memengaruhi proses optimasi *hyperplane* pada *Support Vector Machine* (SVM). Oleh karena itu, dilakukan normalisasi menggunakan metode *Z-score* untuk menyeragamkan skala setiap fitur

sehingga memiliki rata-rata mendekati nol dan standar deviasi mendekati satu. Nilai hasil normalisasi dihitung menggunakan Persamaan (2.19).

$$Z = \frac{x - \mu}{\sigma} \quad (2.19)$$

Keterangan:

Z = Nilai fitur numerik baru hasil normalisasi

x = Nilai fitur asli sebelum normalisasi

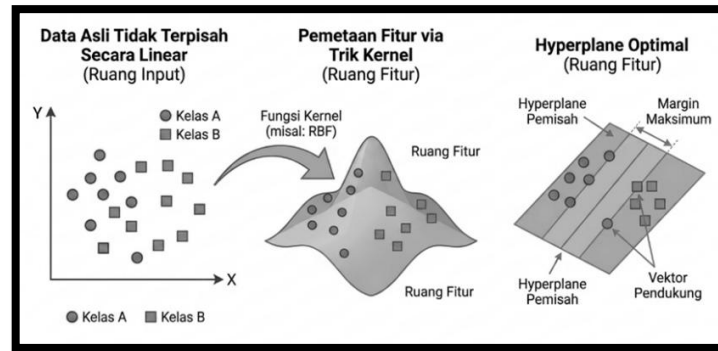
μ = Nilai rata-rata dari fitur tersebut pada seluruh dataset

σ = Nilai standar deviasi dari fitur tersebut pada seluruh dataset

2.4 Support Vector Machine

Support Vector Machine (SVM) merupakan salah satu algoritma pembelajaran terarah (*supervised learning*) yang digunakan untuk menyelesaikan permasalahan klasifikasi dengan mencari *hyperplane* optimal yang memiliki margin maksimum sehingga mampu memisahkan data antar kelas secara efektif (Fletcher, 2008; Goodfellow dkk., 2016). *Hyperplane* tersebut dibentuk oleh sejumlah titik data terdekat yang disebut sebagai *support vector*. Keberadaan *support vector* menentukan posisi dan orientasi *hyperplane* sehingga menghasilkan kemampuan generalisasi yang baik ketika model dihadapkan pada data baru yang belum pernah diamati sebelumnya (Han & Jiang, 2014). Dengan demikian, SVM tidak memerlukan seluruh data latih untuk membangun model, melainkan hanya bergantung pada *support vector* yang jumlahnya relatif kecil, menjadikannya efisien secara komputasi dan tangguh terhadap *overfitting*. Pada kasus data yang tidak dapat dipisahkan secara linear, SVM menggunakan teknik *kernel trick* untuk memetakan data ke ruang fitur berdimensi lebih tinggi tanpa perlu menghitung koordinat transformasi secara eksplisit.

2.4.1 Konsep Dasar dan Formulasi Matematis SVM



Gambar 2. 1 Konsep dasar SVM

Konsep dasar SVM dalam memisahkan dua kelompok data secara linear dapat diilustrasikan pada Gambar 2.2 Pada ilustrasi tersebut, *hyperplane* berfungsi sebagai batas pemisah antar kelas, sedangkan dua garis sejajar di sekitarnya membentuk margin. Titik-titik data yang berada paling dekat dengan *hyperplane* disebut sebagai *support vector* dan memiliki peran penting dalam menentukan posisi serta orientasi *hyperplane*. Secara matematis, *hyperplane* linear dinyatakan pada Persamaan (2.20).

$$w \cdot x + b = 0 \quad (2.20)$$

Untuk memperoleh *hyperplane* optimal, SVM meminimalkan norma vektor bobot menggunakan Persamaan (2.21) dengan kendala pada Persamaan (2.22).

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad (2.21)$$

$$y_i(w \cdot x_i + b) \geq 1, \quad \forall i = 1, 2, \dots, N \quad (2.22)$$

Pada kasus data yang tidak terpisah secara linear sempurna, SVM menerapkan prinsip *Structural Risk Minimization* (SRM) untuk meminimalkan batas kesalahan generalisasi (Vapnik, 1999). Prinsip tersebut diimplementasikan

melalui pendekatan *soft margin* dengan memperkenalkan variabel kelonggaran (ξ_i) yang memberikan toleransi terhadap sejumlah data yang berada di dalam area margin atau mengalami kesalahan klasifikasi. Keseimbangan antara maksimalisasi margin dan toleransi kesalahan dikendalikan oleh parameter penalti (C), sehingga model tetap mampu mengenali variasi pola data tanpa mengalami *overfitting* (Boser dkk., 1992; Schölkopf & Smola, 2002). Fungsi optimasi *soft margin* dinyatakan pada Persamaan (2.23).

$$\min_{w,b,\xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \quad (2.23)$$

dengan kendala:

$$y_i(w \cdot x_i + b) \geq 1, \quad \forall i = 1, 2, \dots, N \quad (2.24)$$

Keterangan:

(w) = vektor bobot (*weight vector*)

(x_i) = vektor fitur masukan

(b) = nilai bias (*intercept*)

(y_i) = label kelas data ke-(i)

(ξ_i) = variabel kelonggaran (*slack variable*)

(C) = parameter penalti yang mengontrol keseimbangan antara lebar margin dan toleransi kesalahan

(N) = jumlah sampel data

2.4.2 Fungsi Kernel SVM

Pada beberapa kasus, distribusi data tidak dapat dipisahkan secara linear pada ruang fitur asal. Untuk mengatasi kondisi tersebut, SVM memanfaatkan teknik *kernel trick* yang memungkinkan data dipetakan secara implisit ke ruang berdimensi lebih tinggi sehingga diperoleh batas keputusan yang lebih optimal (Valkenborg dkk., 2023). Pada penelitian ini digunakan empat fungsi kernel, yaitu kernel Linear, *Radial Basis Function* (RBF), Polynomial, dan Sigmoid. Masing-

masing kernel memiliki karakteristik yang berbeda dalam memetakan hubungan antar fitur sehingga dapat menghasilkan performa klasifikasi yang berbeda pula.

1. Kernel Linear, memetakan data secara linear tanpa transformasi dimensi tambahan, efektif jika hubungan antar fitur sudah cukup terpisah secara linear.
2. Kernel *Radial Basis Function* (RBF), memetakan data ke dimensi tak terhingga menggunakan fungsi basis radial (eksponensial). Kernel ini sangat fleksibel dalam menangani pola data yang non-linear.
3. Kernel Polinomial, merepresentasikan kesamaan vektor fitur melalui fungsi pangkat, yang memungkinkan pemodelan interaksi antar fitur yang lebih kompleks.
4. Kernel Sigmoid, menggunakan fungsi tangen hiperbolik yang menyerupai cara kerja jaringan saraf tiruan dalam melakukan pemetaan data.

Fungsi matematis untuk masing-masing kernel ditunjukkan pada Persamaan (2.25) hingga Persamaan (2.28).

$$K(x_i, x_j) = x_i \cdot x_j \quad (2.25)$$

$$K(x_i, x_j) = (\gamma x_i \cdot x_j + r)^d \quad (2.26)$$

$$K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2) \quad (2.27)$$

$$K(x_i, x_j) = \tanh(\gamma x_i \cdot x_j + r) \quad (2.28)$$

Keterangan:

$K(x_i, x_j)$ = Nilai hasil pemetaan fungsi kernel

x_i, x_j = Vektor fitur hasil fusi dan normalisasi

γ (*Gamma*) = Parameter pengontrol radius atau kemiringan fungsi kernel

d = Derajat polinomial (*degree*)

r = Konstanta koefisien independen (*coef0*) dalam kernel non-linear

2.4.3 Klasifikasi Multikelas *One-vs-Rest* (OvR)

Pada dasarnya, SVM dirancang untuk menyelesaikan permasalahan klasifikasi biner. Oleh karena itu, untuk menangani klasifikasi enam kelas pada penelitian ini digunakan strategi *One-vs-Rest* (OvR). Pendekatan ini membentuk sejumlah model biner sebanyak jumlah kelas, di mana setiap model dilatih untuk membedakan satu kelas terhadap seluruh kelas lainnya. Nilai fungsi keputusan untuk model ke- k dinyatakan pada Persamaan (2.29).

$$f_k(x) = w_k \cdot x + b_k \quad (2.29)$$

Prediksi kelas akhir ditentukan menggunakan fungsi *ArgMax* pada Persamaan (2.30).

$$\hat{y} = \arg \max_{k \in \{1, 2, \dots, K\}} (w_k \cdot x + b_k) \quad (2.30)$$

Keterangan:

- $\hat{y}$ = Label kelas hasil prediksi final untuk sampel citra daun teh terkait
- k = Indeks model pengklasifikasi biner (dari total $K = 6$ kategori daun teh)
- w_k = Vektor bobot (*weight vector*) yang dihasilkan oleh model biner ke- k
- b_k = Nilai bias (*intercept*) yang dihasilkan oleh model biner ke- k
- x = Vektor fitur masukan ternormalisasi (vektor berdimensi lima hasil fusi fitur)

Melalui mekanisme tersebut, kelas dengan nilai fungsi keputusan terbesar akan dipilih sebagai hasil klasifikasi akhir. Strategi *One-vs-Rest* dipilih karena implementasinya relatif sederhana dan telah banyak digunakan untuk memperluas kemampuan SVM dalam menangani permasalahan klasifikasi multikelas.

2.5 Metrik Evaluasi Performa Model

Evaluasi performa dilakukan untuk mengukur kemampuan model *Support Vector Machine* (SVM) dalam mengklasifikasikan enam kategori penyakit daun teh. Evaluasi diawali dengan membentuk *confusion matrix* berukuran 6×6 yang

memuat hubungan antara kelas sebenarnya dan kelas hasil prediksi. Dari matriks tersebut diperoleh empat parameter utama, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

TP merupakan jumlah data yang diprediksi benar sebagai anggota suatu kelas, sedangkan TN adalah jumlah data yang diprediksi benar bukan sebagai anggota kelas tersebut. FP terjadi ketika model memprediksi suatu data sebagai anggota suatu kelas padahal sebenarnya bukan, sedangkan FN terjadi ketika model memprediksi suatu data bukan sebagai anggota suatu kelas padahal sebenarnya termasuk anggota kelas tersebut.

Berdasarkan nilai TP, TN, FP, dan FN tersebut, performa model dievaluasi menggunakan metrik akurasi (*accuracy*), presisi (*precision*), *recall*, dan *F1-score*. Persamaan masing-masing metrik dinyatakan pada Persamaan (2.31) hingga Persamaan (2.34).

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.31)$$

$$\text{Presisi} = \frac{TP}{TP + FP} \quad (2.32)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2.33)$$

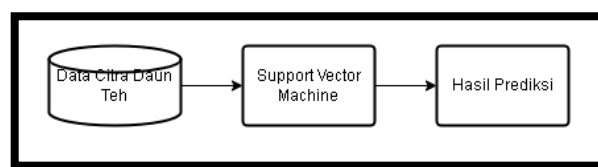
$$\text{F1-Score} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \quad (2.34)$$

BAB III

DESAIN PENELITIAN

3.1 Desain Sistem

Desain sistem yang digunakan pada penelitian ini ditunjukkan pada Gambar 3.1 berikut:



Gambar 3. 1 Dasain sistem

Secara keseluruhan, alur komputasi sistem dimulai dari input citra daun teh yang kemudian melalui tahap prapemrosesan berupa pemotongan terpusat (*center-based square cropping*) dan penyeragaman dimensi. Setelah itu, data latih diseimbangkan menggunakan skenario *Hybrid Resampling* untuk mengurangi bias kelas. Citra hasil prapemrosesan selanjutnya dikonversi ke ruang warna HSV dan *grayscale* guna diekstraksi nilai statistik visualnya, yang kemudian direpresentasikan sebagai vektor numerik berdimensi lima setelah melalui normalisasi skala.

Vektor fitur tersebut selanjutnya diumpankan ke dalam algoritma *Support Vector Machine* (SVM), di mana data dipetakan ke ruang fitur berdimensi tinggi melalui transformasi fungsi kernel untuk membangun batas keputusan (*hyperplane*) dengan margin optimal guna memisahkan karakteristik antar kelas penyakit. Pada tahap akhir, sistem melakukan klasifikasi multikelas menggunakan arsitektur *One-*

vs-Rest (OvR) untuk menghasilkan label keputusan final yang mengklasifikasikan citra uji ke dalam salah satu dari enam kategori kondisi daun teh.

3.2 Pengumpulan Data

Penelitian ini menggunakan data sekunder berupa citra daun teh dari repositori publik Kaggle dengan nama dataset *Plant Disease Expert* yang dipublikasikan oleh Sadman Sakib Mahi pada 31 Agustus 2024. Dataset ini dipilih karena memiliki kualitas gambar yang baik, kategori penyakit yang lengkap, serta relevansi tinggi dengan tujuan penelitian. Dari 58 kategori yang tersedia, hanya citra yang merepresentasikan kondisi daun teh yang digunakan. Contoh citra dari setiap kategori ditampilkan pada Tabel 3.1.

Tabel 3. 1 Data sampel

Daun Sehat (<i>Healthy Leaf</i>)		
		
Daun Merah (<i>Red Leaf Spot</i>)		
		
Hawar Coklat (<i>Brown Blight</i>)		
		

Bercak Mata Burung (<i>Bird's Eye Spot</i>)		
		
Antraknosa (<i>Anthracnose</i>)		
		
Alga (<i>Algal Leaf</i>)		
		

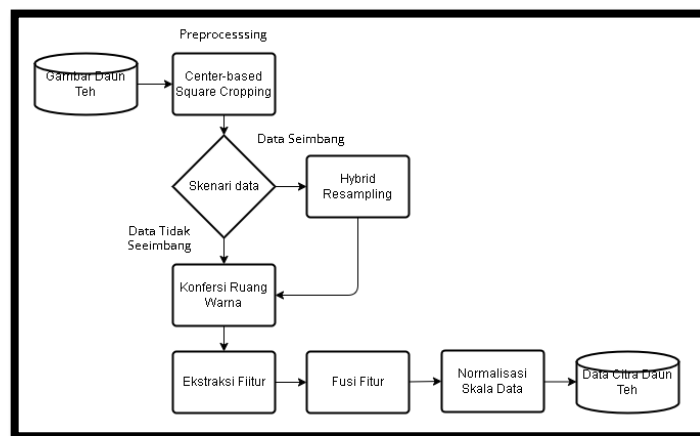
Tabel 3.1 memperlihatkan variasi karakteristik visual dari keenam kategori kondisi daun teh yang digunakan dalam penelitian ini. Perbedaan warna, bentuk bercak, dan pola kerusakan pada setiap kelas menjadi dasar utama dalam proses pengolahan citra dan klasifikasi menggunakan algoritma *Support Vector Machine* (SVM). Karakteristik visual yang khas pada masing-masing kelas—mulai dari dominasi warna merah pada *Red Leaf Spot*, bercak coklat di tepi daun pada *Brown Blight*, hingga warna hijau homogen pada *Healthy Leaf* menentukan bagaimana fitur-fitur visual diekstraksi dan direpresentasikan dalam bentuk numerik untuk dapat diproses oleh algoritma klasifikasi.

Total data berjumlah 1.929 citra yang terbagi ke dalam enam kategori, yaitu *Red Leaf Spot* (429 citra), *Brown Blight* (339 citra), *Bird's Eye Spot* (300 citra),

Anthraco (300 citra), *Algal Leaf/Cephaleuros virescens* (339 citra), dan *Healthy Leaf* (222 citra). Distribusi ini menunjukkan adanya ketidakseimbangan (*class imbalance*) yang signifikan, di mana kelas *Red Leaf Spot* memiliki jumlah terbanyak dan *Healthy Leaf* paling sedikit.

3.3 Data Citra Daun Teh

Data citra daun teh yang telah dikumpulkan kemudian diproses melalui beberapa tahapan untuk menghasilkan representasi numerik sebagai masukan algoritma SVM. Tahapan tersebut meliputi prapengolahan citra, pembentukan dataset, transformasi ruang warna, ekstraksi dan fusi fitur, serta normalisasi data. Alur pengolahan citra ditunjukkan pada Gambar 3.2.



Gambar 3. 2 Alur pada data citra teh

Gambar 3.2 menggambarkan alur sistematis pengolahan citra daun teh hingga menghasilkan vektor fitur ternormalisasi. Proses diawali dengan prapemrosesan menggunakan *Center-Based Square Cropping* untuk memfokuskan objek daun dan menyeragamkan dimensi citra. Selanjutnya, data dibagi menjadi data latih dan data uji, di mana data latih diseimbangkan melalui *Hybrid*

Resampling yang mengombinasikan augmentasi dan *random undersampling*. Citra kemudian dikonversi ke ruang warna HSV dan *grayscale* untuk ekstraksi informasi warna dan intensitas. Ekstraksi fitur statistik menghasilkan lima parameter numerik yang digabungkan melalui *feature fusion* dan dinormalisasi menggunakan metode *Z-score*. Vektor fitur hasil normalisasi inilah yang menjadi masukan bagi algoritma SVM untuk proses klasifikasi. Setiap tahapan dijelaskan lebih rinci pada sub-bab berikutnya.

3.3.1 Prapengolahan Citra

Pada tahap ini, citra akan diproses menggunakan metode *Center-Based Square Cropping* berdasarkan titik pusat (*centroid*) dari kotak pembatas (*bounding box*) objek daun yang telah dideteksi melalui kontur binarisasi citra. Sebagai representasi alur komputasi, simulasi langkah kerja pada tahapan ini adalah sebagai berikut:

1. Diasumsikan sebuah sampel citra daun teh memiliki resolusi awal 500×500 piksel. Hasil deteksi objek menghasilkan *bounding box* dengan koordinat minimum $x_{min} = 100$, $y_{min} = 80$, serta dimensi lebar $W = 200$ piksel dan tinggi $H = 150$ piksel. Koordinat titik pusat (*centroid*) objek daun dihitung sebagai berikut:

$$x_c = 100 + \frac{200}{2} = 200, y_c = 80 + \frac{150}{2} = 155 \quad (3.1)$$

2. Dengan menetapkan faktor *margin* (α) sebesar 1,5 untuk mengamankan area tepi gejala penyakit agar tidak terpotong, panjang sisi bujur sangkar pemotongan (L) ditentukan dari sisi terpanjang:

$$L = \max(200,150) \times 1,5 = 300 \text{ piksel.} \quad (3.2)$$

3. Koordinat acuan pojok kiri atas untuk area pemotongan baru ditentukan melalui fungsi pembatas:

$$x'_{min} = \max\left(0, 200 - \frac{300}{2}\right) = 50, y'_{min} = \max\left(0, 155 - \frac{300}{2}\right) = 5 \quad (3.3)$$

Sistem akan memotong citra asli dari koordinat (50,5) sebesar 300×300 piksel, lalu mengubah ukurannya (*resize*) dengan teknik interpolasi menjadi resolusi homogen 225×225 piksel.

3.3.2 Pembentukan Dataset

Setelah proses prapengolahan citra dilakukan, dataset dibagi menjadi data latih dan data uji. Sebanyak 40 citra dari setiap kelas dipilih secara acak dan digunakan sebagai data uji, sehingga diperoleh total 240 citra uji. Pemisahan data uji dilakukan sebelum proses penyeimbangan data untuk menghindari terjadinya *data leakage*. Sisa data dari masing-masing kelas selanjutnya digunakan sebagai data latih dan dibentuk menjadi dua skenario dataset, yaitu dataset tidak seimbang dan dataset seimbang menggunakan *hybrid resampling*.

3.3.2.1 Dataset Tidak Seimbang

Pada skenario pertama, data latih yang diperoleh setelah pemisahan data uji digunakan secara langsung tanpa dilakukan penyeimbangan jumlah sampel antar kelas. Dengan demikian, distribusi data latih masih mempertahankan kondisi asli dataset dan digunakan sebagai representasi kondisi awal data sebelum dilakukan proses penyeimbangan.

3.3.2.2 Dataset Seimbang

Pada skenario kedua, dilakukan penyeimbangan jumlah data latih menggunakan pendekatan *hybrid resampling* dengan target sebanyak 360 citra untuk setiap kelas. Metode ini mengombinasikan teknik *random undersampling* pada kelas mayoritas dan *oversampling* melalui augmentasi citra pada kelas minoritas. Mekanisme penyeimbangan komputasional ini dijabarkan sebagai berikut:

1. Penerapan *Random Undersampling*, diterapkan secara eksklusif pada kelas mayoritas, yaitu *Red Leaf Spot*, yang memiliki 389 data latih awal (429 total awal dikurangi 40 data uji). Sistem mengeksekusi eliminasi sebanyak 29 citra secara acak untuk mereduksi jumlah himpunan data menuju batas target 360 citra.
2. Penerapan augmentasi (*Oversampling*), dieksekusi pada lima kategori kelas minoritas guna menduplikasi jumlah sampel latih yang kurang hingga menyentuh keseimbangan target 360 citra per kelas. Proses augmentasi citra menggunakan transformasi geometri berupa rotasi acak pada rentang sudut -45° hingga $+45^\circ$, serta operasi *flipping* horizontal dan vertikal.

3.3.3 Konversi Ruang Warna HSV dan *Grayscale*

Citra daun teh hasil prapemrosesan yang berformat RGB (kanal *Red*, *Green*, *Blue*) dikonversi menjadi format ruang warna HSV (*Hue*, *Saturation*, *Value*) dan derajat keabuan (*grayscale*). Transformasi piksel demi piksel ini dilakukan untuk memisahkan informasi intensitas cahaya dari informasi warna kromatik, sehingga pola bercak penyakit dapat diekstraksi secara lebih akurat pada tahap selanjutnya.

Sebagai representasi alur komputasi pada sistem, berikut adalah simulasi perhitungan manual konversi ruang warna untuk satu titik piksel sampel yang diambil dari area bercak penyakit daun teh. Diasumsikan piksel masukan memiliki intensitas RGB awal: $R = 135, G = 62, B = 45$.

1. Konversi piksel ke *Grayscale*, nilai derajat keabuan (Y) dihitung menggunakan metode pembobotan luminositas. Algoritma mengalikan setiap intensitas piksel asli dengan koefisien sensitivitas cahaya:

$$Y = (0.299 \times R) + (0.587 \times G) + (0.114 \times B) \quad (3.4)$$

$$Y = (0.299 \times 135) + (0.587 \times 62) + (0.114 \times 45) \quad (3.5)$$

$$Y = 40.365 + 36.394 + 5.130 \quad (3.6)$$

$$Y = 81.889 \approx 81.89 \quad (3.7)$$

2. Konversi piksel ke HSV, sebelum dikonversi ke HSV, nilai RGB asli dinormalisasi terlebih dahulu ke dalam skala 0 – 1 dengan membaginya terhadap 255:

$$R' = \frac{135}{255} = 0.529, G' = \frac{62}{255} = 0.243, B' = \frac{45}{255} = 0.176 \quad (3.8)$$

Selanjutnya, sistem mencari nilai intensitas maksimum (C_{max}), minimum (C_{min}), dan selisihnya (Δ):

$$C_{max} = \max(0.529, 0.243, 0.176) = 0.529 \quad (3.9)$$

$$C_{min} = \min(0.529, 0.243, 0.176) = 0.176 \quad (3.10)$$

$$\Delta = 0.529 - 0.176 = 0.353 \quad (3.11)$$

Melalui parameter tersebut, nilai untuk masing-masing kanal *Value* (V), *Saturation* (S), dan *Hue* (H) dihitung:

$$V = C_{max} = 0.529, S = \frac{\Delta}{C_{max}} = \frac{0.353}{0.529} = 0.667 \quad (3.12)$$

Karena nilai C_{max} jatuh pada komponen R' , rumusnya adalah:

$$H = 60^\circ \times \left(\frac{G' - B'}{\Delta} \setminus \text{mod} 6 \right) \quad (3.13)$$

$$H = 60^\circ \times \left(\frac{0.243 - 0.176}{0.353} \setminus \text{mod} 6 \right) \quad (3.14)$$

$$H = 60^\circ \times (0.1898 \setminus \text{mod} 6) = 11.39^\circ \quad (3.15)$$

Melalui operasi komputasi iteratif ini, seluruh 50.625 piksel (hasil dari matriks 225×225) akan diubah menjadi representasi HSV dan *grayscale* untuk diteruskan ke tahap ekstraksi parameter statistic *mean* dan standar deviasi.

3.3.4 Ekstraksi Fitur

Setelah citra daun teh dikonversi menjadi representasi matriks HSV dan derajat keabua (*grayscale*), sistem mengeksekusi tahap ekstraksi fitur untuk mereduksi dimensi data yang besar menjadi representasi nilai statistik tunggal (vektor fitur). Penarikan fitur visual ini difokuskan pada tiga parameter rata-rata dari warna dominan (*Mean Hue, Saturation, Value*) dan dua parameter pola struktural permukaan daun (*Mean dan Standard Deviation* dari citra *grayscale*).

Sebagai representasi alur komputasi, berikut adalah simulasi perhitungan manual penarikan fitur statistik yang direpresentasikan menggunakan sampel sebagian kecil matriks berukuran 3×3 piksel:

1. Ekstraksi fitur warna (Statistik HSV), diasumsikan setelah seluruh piksel dikonversi, sistem memetakan sebaran intensitas piksel untuk kanal *Value* (*V*) ke dalam sampel matriks berikut:

$$\text{Matriks } V = \begin{bmatrix} 0.52 & 0.50 & 0.48 \\ 0.54 & 0.52 & 0.50 \\ 0.50 & 0.48 & 0.46 \end{bmatrix} \quad (3.16)$$

Sistem kemudian menarik fitur statistik berupa nilai rata-rata (*Mean*) dengan menjumlahkan seluruh elemen matriks dan membaginya dengan jumlah piksel ($N = 9$):

$$\text{a. Total Nilai } V = 0.52 + 0.50 + 0.48 + 0.54 + 0.52 + 0.50 + 0.50 + 0.48 + 0.46 = 4.50$$

$$\text{b. Fitur Mean } V = \frac{4.50}{9} = 0.500$$

Metode perhitungan nilai rata-rata ini dieksekusi secara identik oleh sistem pada matriks kanal *Hue* (*H*) dan *Saturation* (*S*), sehingga menghasilkan tiga nilai fitur warna yang merepresentasikan karakteristik visual penyakit.

2. Ekstraksi fitur tekstur (Statistik *Grayscale*), untuk menganalisis tingkat kecerahan dan variasi kekasaran tekstur bercak penyakit secara komputasional, sistem mengekstrak parameter *Mean* (μ) dan *Standard Deviation* (σ) dari citra *grayscale*. Diasumsikan sampel matriks *grayscale* (*Y*) dari area bercak daun adalah:

$$\text{Matriks } Y = \begin{bmatrix} 82 & 80 & 78 \\ 84 & 82 & 80 \\ 80 & 78 & 76 \end{bmatrix} \quad (3.17)$$

Sistem mengeksekusi ekstraksi nilai statistik tekstur dengan kalkulasi:

$$\text{a. Total Nilai } Y = 82 + 80 + 78 + 84 + 82 + 80 + 80 + 78 + 76 = 720$$

$$\text{b. Fitur Mean Grayscale } (\mu_Y) = \frac{720}{9} = 80.00$$

Selanjutnya, penyebaran deviasi untuk mengukur kekasaran tekstur dihitung menggunakan selisih kuadrat dari rata-rata:

$$\text{a. Varians } (\sigma^2) = \frac{(82-80)^2 + (80-80)^2 + \dots + (76-80)^2}{9}$$

$$\text{b. Varians } (\sigma^2) = \frac{4+0+4+16+4+0+0+4+16}{9} = \frac{52}{9} = 5.778$$

$$\text{c. Fitur Standar Deviasi } (\sigma_Y) = \sqrt{5.778} = 2.404$$

Melalui tahapan ekstraksi ini, informasi spasial citra daun teh berhasil direduksi secara matematis menjadi lima parameter fitur independen yang siap untuk digabungkan (*feature fusion*) dan dinormalisasi ke dalam skala komputasi yang seragam.

3.3.5 Feature Fusion

Setelah karakteristik warna dan tekstur berhasil diekstraksi dalam bentuk nilai statistik terpisah, sistem mengeksekusi tahap fusi fitur (*feature fusion*) yang dilanjutkan dengan normalisasi skala data. Tahapan ini sangat krusial untuk menyatukan multi-informasi visual menjadi satu kesatuan larik data tunggal, serta menghilangkan ketimpangan rentang nilai antar parameter sebelum diproses lebih lanjut oleh algoritma *Support Vector Machine* (SVM).

1. Skenario Fusi Fitur, dilakukan pada level fitur (*early fusion*) dengan cara menggabungkan tiga parameter statistik warna HSV ($H_{mean}, S_{mean}, V_{mean}$)

dan dua parameter statistik tekstur *grayscale* (μ_Y, σ_Y) ke dalam satu vektor spasial berdimensi lima (F).

2. Skenario normalisasi skala data (*Z-Score*), matriks fitur gabungan (F) memiliki perbedaan rentang nilai yang sangat signifikan—misalnya antara nilai komponen *Mean Grayscale* ($\mu_Y = 80.00$) dengan *Mean Value* ($V_{mean} = 0.500$). Jika langsung diumpankan ke model SVM, parameter dengan skala besar akan mendominasi perhitungan jarak Euclidean dan margin pembatas, sehingga mengabaikan peran parameter berskala kecil. Untuk menyetarakan kontribusi setiap fitur, diterapkan *Z-Score Normalization* (*StandardScaler*) untuk mentransformasikan seluruh nilai agar memiliki rata-rata nol ($\mu = 0$) dan standar deviasi satu ($\sigma = 1$).

Sebagai simulasi perhitungannya, merujuk pada hasil kalkulasi manual pada tahap konversi warna dan ekstraksi fitur sebelumnya, diperoleh nilai karakteristik untuk sampel citra daun teh sebagai berikut:

$$H_{mean} = 11.39, S_{mean} = 0.667, V_{mean} = 0.500, \mu_Y = 80.00, \text{ dan } \sigma_Y = 2.40 \quad (3.18)$$

Melalui operasi penyatuan vektor, diperoleh matriks fitur gabungan tunggal (F) untuk satu sampel citra tersebut:

$$F = [11.39, 0.667, 0.500, 80.00, 2.404] \quad (3.19)$$

3.3.6 Normalisasi

Proses normalisasi akan ditunjukkan pada parameter *Mean Grayscale* ($\mu_Y = 80.00$). Diasumsikan berdasarkan keseluruhan data pada fase pelatihan (*training set*), diperoleh nilai rata-rata populasi untuk fitur tersebut adalah $\mu_{dataset} = 78.00$

dengan standar deviasi populasi $\sigma_{dataset} = 2.50$. Perhitungan nilai normalisasi (Z) adalah sebagai berikut:

$$Z = \frac{x - \mu_{dataset}}{\sigma_{dataset}} \quad (3.20)$$

$$Z = \frac{80.00 - 78.00}{2.50} \quad (3.21)$$

$$Z = 0.800 \quad (3.22)$$

Melalui formula yang sama, kelima parameter di dalam vektor F akan ditransformasikan piksel demi piksel hingga menghasilkan vektor ternormalisasi baru (Z), misalnya:

$$Z = [0.500, -0.200, 0.800, 0.800, 1.200] \quad (3.23)$$

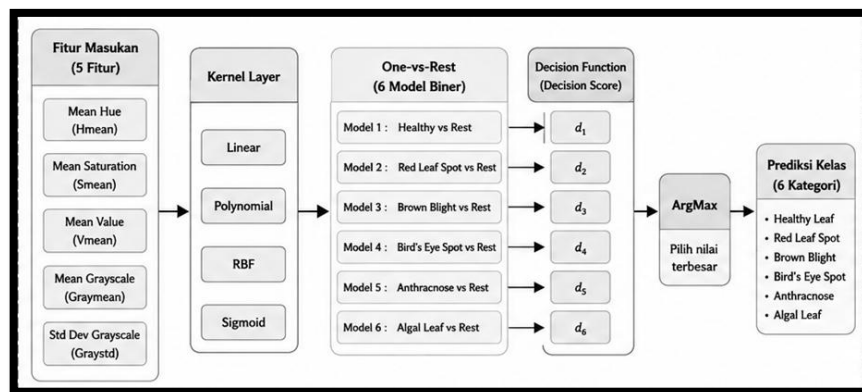
Larik data berdimensi lima yang telah setara secara skala inilah yang kemudian siap diumpankan ke dalam arsitektur klasifikasi multikelas SVM pada tahap berikutnya.

3.4 Klasifikasi Menggunakan *Support Vector Machine*

Pada penelitian ini, proses klasifikasi enam kategori kondisi daun teh dilakukan menggunakan algoritma *Support Vector Machine* (SVM) dengan pendekatan multikelas *One-vs-Rest* (OvR). Vektor fitur hasil ekstraksi dan fusi fitur yang telah dinormalisasi digunakan sebagai masukan untuk membentuk model klasifikasi dengan empat fungsi kernel, yaitu Linear, Polynomial, *Radial Basis Function* (RBF), dan Sigmoid.

3.4.1 Arsitektur *Support Vector Machine*

Arsitektur klasifikasi *Support Vector Machine* (SVM) yang digunakan pada penelitian ini dirancang untuk mengolah vektor fitur hasil ekstraksi dan menghasilkan prediksi salah satu dari enam kategori kondisi daun teh. Proses klasifikasi dilakukan menggunakan pendekatan multikelas *One-vs-Rest* (OvR) dengan memanfaatkan empat fungsi kernel yang berbeda. Alur kerja sistem klasifikasi tersebut ditunjukkan pada Gambar 3.3 berikut.



Gambar 3. 3 Arsitektur SVM dengan strategi OvR

Berdasarkan Gambar 3.3, vektor fitur hasil ekstraksi yang terdiri atas *Mean Hue*, *Mean Saturation*, *Mean Value*, *Mean Grayscale*, dan *Standard Deviation Grayscale* menjadi masukan bagi algoritma SVM. Data kemudian diproses menggunakan fungsi kernel Linear, Polynomial, RBF, dan Sigmoid untuk memetakan data ke ruang fitur yang sesuai. Klasifikasi multikelas dilakukan dengan pendekatan *One-vs-Rest* (OvR), sehingga terbentuk enam model biner yang masing-masing merepresentasikan satu kategori kondisi daun teh. Nilai fungsi keputusan dari setiap model dibandingkan untuk menentukan kelas dengan skor tertinggi sebagai hasil prediksi akhir.

3.4.1.1 Parameter Fungsi Kernel

Parameter masing-masing fungsi kernel yang digunakan dalam penelitian ini mengacu pada nilai *default* dari pustaka Scikit-Learn dengan beberapa penyesuaian sebagaimana ditunjukkan pada Tabel 3.3.

Tabel 3. 2 Parameter fungsi kernel SVM

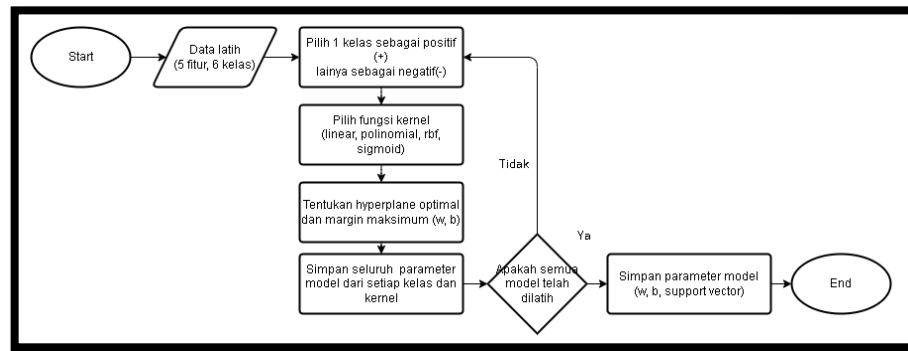
Kernel	Parameter	Nilai	Keterangan
Linear	<i>C</i>	1.0	Parameter regularisasi default
	<i>C</i>	1.0	Parameter regularisasi default
Polynomial	<i>degree</i>	3	Derajat polinomial default
	<i>coef0</i>	1.0	Koefisien independen default
RBF	<i>C</i>	1.0	Parameter regularisasi default
	<i>gamma</i>	'scale'	$\text{Gamma otomatis} = 1/(n_features * X.var())$
Sigmoid	<i>C</i>	1.0	Parameter regularisasi default
	<i>gamma</i>	'scale'	$\text{Gamma otomatis} = 1/(n_features * X.var())$

Pemilihan parameter default ini bertujuan untuk mengevaluasi kinerja dasar setiap kernel tanpa adanya optimasi *hyperparameter*, sehingga perbedaan performa yang dihasilkan lebih mencerminkan karakteristik masing-masing fungsi kernel (Schölkopf & Smola, 2002). Dengan pendekatan ini, analisis komparatif dapat berfokus pada pengaruh intrinsik fungsi kernel dan kondisi dataset terhadap kemampuan klasifikasi, bukan pada perbedaan parameter yang mungkin memengaruhi hasil secara tidak terkendali. Selain itu, penggunaan parameter *default* juga memastikan reproduibilitas penelitian karena nilai-nilai tersebut merupakan standar yang telah ditetapkan dalam pustaka Scikit-Learn.

3.4.2 Pelatihan Model Menggunakan *One-vs-Rest*

Pada tahap pelatihan, algoritma *Support Vector Machine* (SVM) dengan pendekatan *One-vs-Rest* (OvR) membentuk beberapa model klasifikasi biner secara independen. Setiap model dilatih dengan menjadikan satu kelas sebagai kelas

positif (+1), sedangkan seluruh kelas lainnya digabungkan sebagai kelas negatif (-1). Proses ini dilakukan secara berulang hingga seluruh kelas memiliki model klasifikasi masing-masing.



Gambar 3. 4 Proses pelatihan SVM dengan strategi Ovr

Berdasarkan Gambar 3.5, proses pelatihan diawali dengan data latih hasil ekstraksi fitur yang telah dinormalisasi. Selanjutnya, satu kelas dipilih sebagai kelas positif (+1), sedangkan lima kelas lainnya ditetapkan sebagai kelas negatif (-1). Data tersebut kemudian diproses menggunakan fungsi kernel yang digunakan pada penelitian ini, yaitu Linear, Polynomial, *Radial Basis Function* (RBF), dan Sigmoid untuk memperoleh *hyperplane* dengan margin maksimum. Parameter model yang dihasilkan disimpan, kemudian proses yang sama diulangi hingga terbentuk enam model klasifikasi biner yang merepresentasikan seluruh kategori kondisi daun teh.

3.4.3 Prediksi dan Penentuan Kelas

Setelah pelatihan selesai, setiap model SVM menghasilkan parameter untuk menghitung nilai fungsi keputusan (*decision function*) pada data uji. Pada pendekatan *One-vs-Rest* (OvR), citra uji dievaluasi oleh seluruh model biner, lalu kelas dengan nilai *decision score* tertinggi dipilih sebagai hasil akhir menggunakan

fungsi *ArgMax*. Sebagai contoh, misalkan sebuah citra uji memiliki vektor fitur hasil normalisasi sebagai berikut:

$$Z = [0.5, -0.2, 0.8, 0.8, 1.2] \quad (3.24)$$

dan parameter model biner "*Healthy vs Rest*" yang diperoleh dari proses pelatihan adalah:

$$w = [0.4, 0.1, -0.3, 0.6, 0.5], \quad b = -0.2 \quad (3.25)$$

maka nilai fungsi keputusan dihitung menggunakan hasil perkalian antara vektor bobot dan vektor fitur yang kemudian ditambahkan dengan nilai bias, sehingga diperoleh:

$$f(x) = (0.4 \times 0.5) + (0.1 \times -0.2) + (-0.3 \times 0.8) + (0.6 \times 0.8) + (0.5 \times 1.2) - 0.2 \quad (3.26)$$

$$f(x) = 0.20 - 0.02 - 0.24 + 0.48 + 0.60 - 0.2 \quad (3.27)$$

$$f(x) = 0.82 \quad (3.28)$$

Karena nilai fungsi keputusan bernilai positif ($f(x) > 0$), maka sampel tersebut diklasifikasikan sebagai anggota kelas positif pada model "*Healthy vs Rest*". Proses yang sama dilakukan pada lima model biner lainnya, kemudian kelas dengan nilai *decision score* tertinggi dipilih sebagai hasil klasifikasi akhir.

3.5 Skenario Pengujian

Skenario pengujian pada penelitian ini dirancang untuk mengevaluasi kemampuan algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan enam kategori kondisi daun teh berdasarkan vektor fitur hasil ekstraksi dan fusi fitur warna HSV serta *grayscale* yang telah dinormalisasi. Pengujian dilakukan

untuk menganalisis pengaruh penggunaan fungsi kernel dan kondisi distribusi data terhadap performa klasifikasi, sehingga dapat diperoleh model yang memiliki kemampuan generalisasi terbaik dalam mengidentifikasi penyakit daun teh.

3.5.1 Lingkungan Pengujian

Seluruh tahapan penelitian yang meliputi prapengolahan citra, pembentukan dataset, ekstraksi fitur, pelatihan model, hingga evaluasi klasifikasi diimplementasikan menggunakan bahasa pemrograman Python versi 3.12.13. Proses komputasi dijalankan pada layanan Google Colaboratory berbasis sistem operasi Linux 6.6.122+ dengan dukungan prosesor Intel Xeon @ 2,20 GHz dan RAM sebesar 13,2 GB, sedangkan perangkat lokal yang digunakan adalah laptop Lenovo V14 IIL. Implementasi sistem memanfaatkan beberapa pustaka utama, yaitu OpenCV untuk pengolahan citra, NumPy dan Pandas untuk manipulasi data numerik, TensorFlow/Keras untuk augmentasi data pada proses *hybrid resampling*, Scikit-Learn untuk normalisasi data, pembangunan model *Support Vector Machine* dengan pendekatan *One-vs-Rest* (OvR), serta perhitungan metrik evaluasi, sedangkan visualisasi hasil pengujian dilakukan menggunakan pustaka Matplotlib dan Seaborn.

3.5.2 Skenario Pengujian Klasifikasi

Untuk mengevaluasi kinerja algoritma *Support Vector Machine* (SVM), penelitian ini menggunakan delapan skenario pengujian yang diperoleh dari kombinasi empat fungsi kernel, yaitu Linear, Polynomial, *Radial Basis Function* (RBF), dan Sigmoid dengan dua kondisi dataset, yaitu dataset tidak seimbang dan

dataset hasil *hybrid resampling*. Seluruh pengujian menggunakan vektor fitur hasil fusi parameter statistik ruang warna HSV dan *grayscale* yang telah melalui proses normalisasi. Rancangan skenario pengujian yang digunakan pada penelitian ini ditunjukkan pada Tabel 3.3 berikut:

Tabel 3. 3 Rancangan skenario pengujian

No Skenario	Kernel	Jenis Data Masukan
1	RBF	Data asli (<i>imbalance</i>)
2	Linear	Data asli (<i>imbalance</i>)
3	Polynomial	Data asli (<i>imbalance</i>)
4	Sigmoid	Data asli (<i>imbalance</i>)
5	RBF	Data seimbang (<i>balance</i>)
6	Linear	Data seimbang (<i>balance</i>)
7	Polynomial	Data seimbang (<i>balance</i>)
8	Sigmoid	Data seimbang (<i>balance</i>)

Berdasarkan Tabel 3.2, setiap model dilatih menggunakan pendekatan multikelas *One-vs-Rest* (OvR) dan dievaluasi menggunakan *confusion matrix* serta metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil pengujian tersebut selanjutnya digunakan untuk menganalisis pengaruh fungsi kernel dan penyeimbangan data terhadap kemampuan model dalam mengklasifikasikan enam kategori kondisi daun teh serta menentukan kombinasi kernel dan kondisi dataset yang memberikan performa terbaik.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil

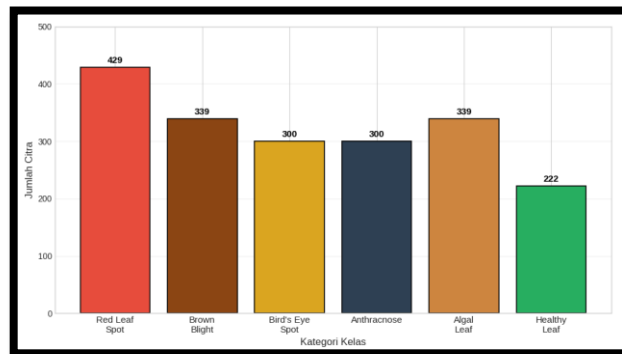
Bab ini menyajikan hasil seluruh tahapan penelitian, mulai dari pengumpulan data, pra-pemrosesan citra, pembentukan dataset, ekstraksi dan normalisasi fitur, hingga klasifikasi menggunakan SVM dengan empat fungsi kernel pada dua skenario dataset. Hasil disajikan dalam bentuk tabel, grafik, dan gambar disertai penjelasan untuk memudahkan pemahaman.

4.1.1 Pengumpulan Data

Penelitian ini menggunakan data sekunder dari repositori Kaggle bernama *Plant Disease Expert* yang dipublikasikan oleh Sadman Sakib Mahi pada 31 Agustus 2024. Dataset asli terdiri dari 58 kategori, namun hanya enam kategori kondisi daun teh yang digunakan, dengan total 1.929 citra. Distribusi data disajikan pada Tabel 4.1 dan Gambar 4.1. Berdasarkan distribusi tersebut, terlihat adanya ketidakseimbangan jumlah sampel yang signifikan antar kelas, yang menjadi salah satu fokus analisis dalam penelitian ini.

Tabel 4. 1 Distribusi dataset awal

No	Kategori	Jumlah Citra	Proporsi (%)
1	<i>Red Leaf Spot</i>	429	22.24
2	<i>Brown Blight</i>	339	17.57
3	<i>Bird's Eye Spot</i>	300	15.55
4	<i>Anthracnose</i>	300	15.55
5	<i>Algal Leaf</i>	339	17.57
6	<i>Healthy Leaf</i>	222	11.51

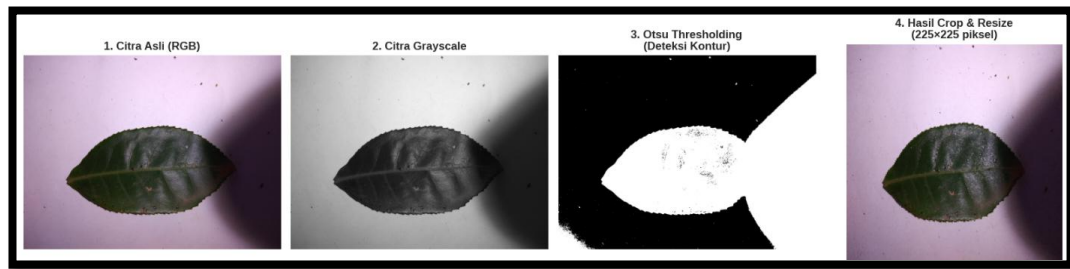


Gambar 4. 1 Distribusi dataset awal citra daun teh

Tabel 4.1 menunjukkan adanya ketidakseimbangan jumlah sampel antar kelas (*class imbalance*). Kelas *Red Leaf Spot* memiliki jumlah citra terbanyak (429 citra, 22,24%), sedangkan *Healthy Leaf* hanya memiliki 222 citra (11,51%). Perbedaan signifikan ini berpotensi menyebabkan model bias terhadap kelas mayoritas. Visualisasi kesenjangan tersebut ditunjukkan pada Gambar 4.1. Ketidakseimbangan ini menjadi dasar perlunya penyeimbangan data sebelum pelatihan model.

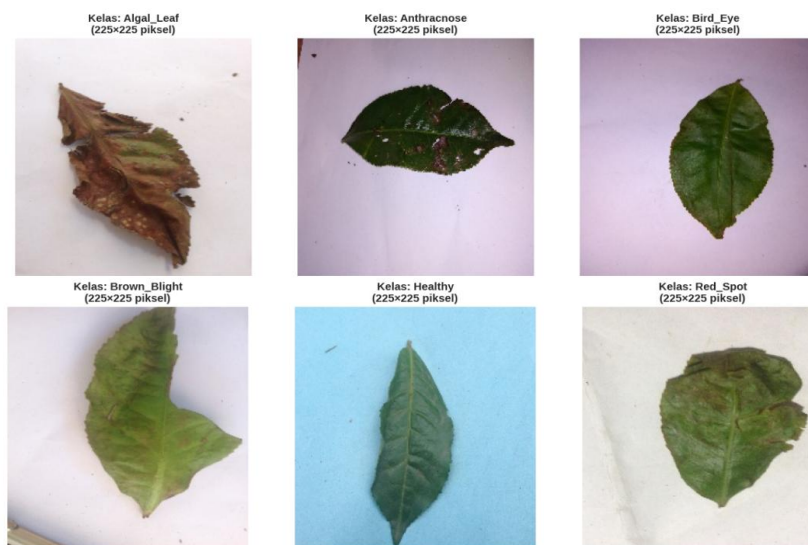
4.1.2 Pra-pemrosesan Citra

Pra-pemrosesan citra bertujuan memfokuskan area pengamatan pada objek daun dan menyeragamkan dimensi citra. Metode yang digunakan adalah *Center-Based Square Cropping* dengan faktor margin 1,5, dilanjutkan *resize* ke ukuran 225×225 piksel. Prosesnya meliputi deteksi kontur daun menggunakan *Otsu Thresholding*, penentuan bounding box, perhitungan titik pusat (*centroid*), dan pemotongan berbentuk bujur sangkar. Ilustrasi tahapan ini ditunjukkan pada Gambar 4.2.



Gambar 4. 2 Tahapan pra-pemrosesan citra

Gambar 4.2 menampilkan empat tahap utama: (1) citra asli RGB, (2) konversi ke grayscale, (3) Otsu Thresholding untuk segmentasi, dan (4) hasil akhir berupa citra terpotong dan di-resize menjadi 225×225 piksel. Hasil pra-pemrosesan seluruh kelas ditampilkan pada Gambar 4.3.



Gambar 4. 3 Hasil prapemrosesan citra

Berdasarkan Gambar 4.3, keenam kelas telah berhasil diproses dengan karakteristik visual yang berbeda. *Red Leaf Spot* didominasi warna merah, *Brown Blight* bercak coklat di tepi daun, *Bird's Eye Spot* lesi melingkar abu-abu, *Anthracnose* bercak gelap nekrotik, *Algal Leaf* bercak oranye kemerahan, dan

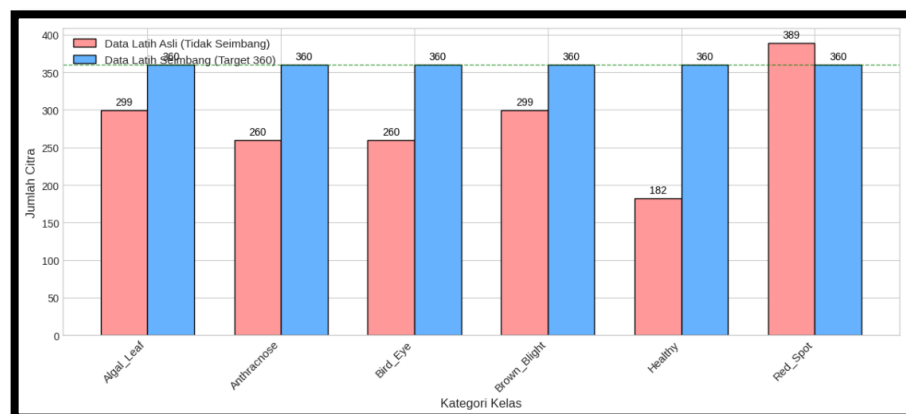
Healthy Leaf hijau homogen. Perbedaan ini menjadi dasar ekstraksi fitur dan klasifikasi.

4.1.3 Pembentukan Dataset

Setelah pra-pemrosesan, dataset dibagi menjadi data uji dan data latih. Sebanyak 40 citra per kelas (total 240 citra) dipilih secara acak sebagai data uji sebelum proses penyeimbangan untuk menghindari data leakage. Sisa data digunakan dalam dua skenario: tidak seimbang dan seimbang menggunakan hybrid resampling, seperti ditunjukkan pada Tabel 4.2.

Tabel 4. 2 Distribusi data latih

No	Kategori	Data Awal	Data Uji	Data Latih Asli	Teknik	Data Latih Seimbang
1	<i>Algal_Leaf</i>	429	40	299	<i>Undersampling</i> (-29)	360
2	<i>Anthracnose</i>	339	40	260	Augmentasi (+61)	360
3	<i>Bird_Eye</i>	300	40	260	Augmentasi (+100)	360
4	<i>Brown_Blight</i>	300	40	299	Augmentasi (+100)	360
5	<i>Healthy</i>	339	40	182	Augmentasi (+61)	360
6	<i>Red_Spot</i>	222	40	389	Augmentasi (+178)	360



Gambar 4. 4 Perbandingan Distribusi Data Latih Sebelum dan Sesudah *Hybrid Resampling*

Pada skenario tidak seimbang, jumlah data latih bervariasi antara 182–389 citra. Pada skenario seimbang, seluruh kelas diseragamkan menjadi 360 citra: kelas *Red Leaf Spot* dikurangi 29 citra melalui *random undersampling*, sedangkan lima kelas minoritas diperbanyak melalui augmentasi (rotasi -45° hingga $+45^\circ$ serta *flipping horizontal* dan *vertikal*). Perbandingan distribusi ditunjukkan pada Gambar 4.4, yang memperlihatkan seluruh kelas mencapai target 360 citra. Penyeragaman ini bertujuan mengurangi bias model terhadap kelas mayoritas dan meningkatkan kinerja pada kelas minoritas.

4.1.4 Ekstraksi Fitur dan Normalisasi

Ekstraksi fitur bertujuan mengubah citra menjadi parameter numerik untuk SVM. Lima fitur statistik diekstraksi dari setiap citra: *Mean Hue*, *Mean Saturation*, dan *Mean Value* dari ruang HSV, serta *Mean Grayscale* dan *Standard Deviation Grayscale* dari citra *grayscale*. Contoh hasil ekstraksi disajikan pada Tabel 4.3.

Tabel 4. 3 Contoh hasil ekstraksi fitur per kelas

No	Kategori	<i>Mean Hue</i>	<i>Mean Saturation</i>	<i>Mean Value</i>	<i>Mean Grayscale</i>	<i>Std Grayscale</i>
1	<i>Algal Leaf</i>	107.96	27.606	192.788	188.20	67.06
2	<i>Anthracnose</i>	122.03	41.144	167.879	154.52	54.94
3	<i>Bird Eye</i>	131.87	54.533	166.327	152.03	66.95
4	<i>Brown Blight</i>	101.80	38.151	185.072	177.45	61.14
5	<i>Healthy</i>	98.76	109.089	187.435	146.23	28.98
6	<i>Red Spot</i>	29.61	41.188	195.023	191.14	71.05

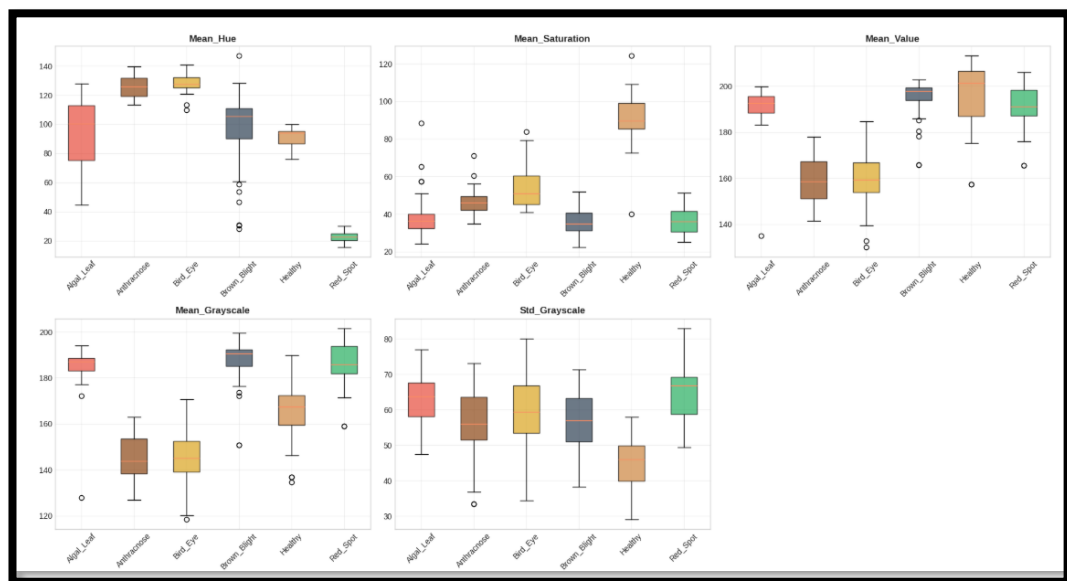
Tabel 4.3 menunjukkan perbedaan karakteristik numerik antar kelas. *Red Leaf Spot* memiliki *Mean Hue* lebih rendah karena pigmen merah, sementara *Healthy Leaf* yang hijau memiliki *Mean Grayscale* lebih terang. Perbedaan ini menjadi dasar SVM dalam membangun batas keputusan. Sebelum pelatihan, seluruh fitur dinormalisasi menggunakan *Z-Score* agar memiliki skala setara (rata-

rata mendekati 0, standar deviasi mendekati 1). Contoh hasil normalisasi ditunjukkan pada Tabel 4.4.

Tabel 4. 4 contoh hasil normalisasi

No	Fitur	Nilai Asli	Nilai Ternormalisasi
1	<i>Mean Hue</i>	107.964	0.4486
2	<i>Mean Saturation</i>	27.606	-0.9710
3	<i>Mean Value</i>	192.788	0.4762
4	<i>Mean Grayscale</i>	188.197	0.7383
5	<i>Std Grayscale</i>	67.061	0.8531

Tabel 4.4 menunjukkan transformasi fitur ke *standartscaller*, mencegah fitur dengan rentang besar mendominasi perhitungan jarak dan margin pada SVM. Sebaran nilai fitur per kelas divisualisasikan pada Gambar 4.5.



Gambar 4. 5 Distribusi fitur statistik per kelas

Gambar 4.5 menunjukkan beberapa fitur memiliki pemisahan yang baik antar kelas, terutama *Mean Hue* dan *Mean Saturation*, yang mengindikasikan informasi warna cukup diskriminatif. Sebaliknya, *Mean Value* dan *Mean Grayscale* pada beberapa kelas memiliki rentang tumpang tindih, yang berpotensi menyebabkan kesalahan klasifikasi.

4.1.5 Hasil Klasifikasi SVM pada Data Tidak Seimbang

Pada skenario pertama, model SVM dilatih menggunakan data latih tidak seimbang dengan empat fungsi kernel (Linear, Polynomial, RBF, dan Sigmoid) melalui pendekatan *One-vs-Rest* (OvR). Untuk menjelaskan proses internal SVM secara transparan, diuraikan analisis *support vector*, visualisasi *hyperplane*, simulasi *decision function*, dan hasil evaluasi.

4.1.5.1 Analisis Support Vector

Support vector adalah data latih di tepi atau dalam margin yang menentukan posisi *hyperplane*, sedangkan data lainnya tidak memengaruhinya, menjadikan SVM bersifat *sparse*. Jumlah *support vector* setiap kernel pada data tidak seimbang ditunjukkan pada Tabel 4.5.

Tabel 4. 5 Jumlah *support vector* pada data tidak seimbang

No	Kernel	<i>Red Spot</i>	<i>Brown Blight</i>	<i>Bird's Eye</i>	<i>Anthra cnose</i>	<i>Algal Leaf</i>	<i>Healthy</i>	Total
1	Linear	226	150	154	223	9	20	782
2	Polynomial	267	179	184	246	14	33	923
3	RBF	230	171	176	221	22	44	864
4	Sigmoid	225	144	175	217	30	69	860

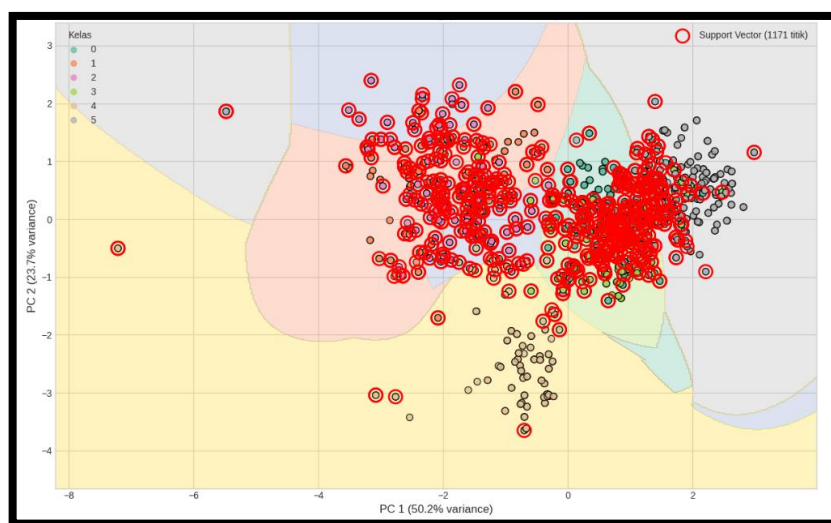
Tabel 4.5 menunjukkan kernel Polynomial menghasilkan *support vector* terbanyak (923 dari 1.689 data latih, 54,65%), mengindikasikan kompleksitas tinggi akibat derajat polinomial. Sebaliknya, kernel Linear menghasilkan paling sedikit (782, 46,30%), mencerminkan model lebih sederhana dengan *hyperplane* linear. Proporsi *support vector* mencerminkan kompleksitas batas keputusan: semakin banyak, semakin kompleks.

Distribusi *support vector* per kelas sangat timpang. Kelas *Algal Leaf* dan *Healthy Leaf* memiliki jumlah jauh lebih sedikit dibanding lainnya pada kernel

Linear, *Algal Leaf* hanya 9, *Healthy Leaf* 20, sedangkan *Red Leaf Spot* 226. Hal ini berkaitan dengan jumlah data latih: kelas dengan data lebih banyak cenderung menghasilkan lebih banyak *support vector*. Ketimpangan ini berpotensi menurunkan kemampuan model dalam mengenali kelas minoritas.

4.1.5.2 Visualisasi *Hyperplane*

Untuk memberikan pemahaman visual tentang pembentukan batas keputusan SVM, dilakukan visualisasi *hyperplane* kernel RBF menggunakan reduksi dimensi *Principal Component Analysis* (PCA) menjadi dua dimensi, seperti ditunjukkan pada Gambar 4.6.



Gambar 4. 6 Visualisasi *hyperplane* SVM kernel RBF pada data tidak seimbang

Gambar 4.6 memperlihatkan *hyperplane* non-linear yang dibentuk kernel RBF, dengan area berwarna berbeda sebagai wilayah keputusan tiap kelas. Titik-titik data bertanda lingkaran merah adalah *support vector* yang menentukan posisi *hyperplane*. Terlihat *support vector* terkonsentrasi di area perbatasan antar kelas, menegaskan perannya dalam membentuk batas keputusan.

4.1.5.3 Simulasi *Decision Function*

Setelah *hyperplane* terbentuk, klasifikasi dilakukan dengan menghitung *decision score* setiap model biner dalam arsitektur OvR, lalu memilih kelas dengan skor tertinggi sebagai prediksi. Simulasi untuk satu sampel per kelas menggunakan kernel RBF ditampilkan pada Tabel 4.6.

Tabel 4. 6 Simulasi *decision function* kernel RBF pada data tidak seimbang

No.	Kelas Aktual	Prediksi	Status	Skor
1	<i>Algal Leaf</i>	<i>Algal Leaf</i>	V	5.2990
2	<i>Anthraco</i> <i>nose</i>	<i>Anthraco</i> <i>nose</i>	V	5.2830
3	<i>Bird Eye</i>	<i>Bird Eye</i>	V	5.2974
4	<i>Brown Blight</i>	<i>Algal Leaf</i>	X	5.2933
5	<i>Healthy</i>	<i>Healthy</i>	V	5.2778
6	<i>Red Spot</i>	<i>Red Spot</i>	V	5.2841

Tabel 4.6 menunjukkan dari enam sampel, lima diprediksi benar dan satu salah. Kesalahan terjadi pada kelas *Brown Blight* yang diprediksi sebagai *Algal Leaf* akibat kemiripan visual kedua kelas memiliki bercak gelap pada daun sehingga fitur statistik yang diekstraksi bernilai hampir sama. Akibatnya, *decision score* untuk kedua model biner (*Algal Leaf* vs *Rest* dan *Brown Blight* vs *Rest*) sangat berdekatan, dan skor tertinggi jatuh pada *Algal Leaf* vs *Rest*. Meskipun satu kesalahan terjadi, seluruh *decision score* berada pada rentang tinggi (5,2–5,3), menunjukkan tingkat kepercayaan model yang tinggi karakteristik SVM dengan kernel RBF.

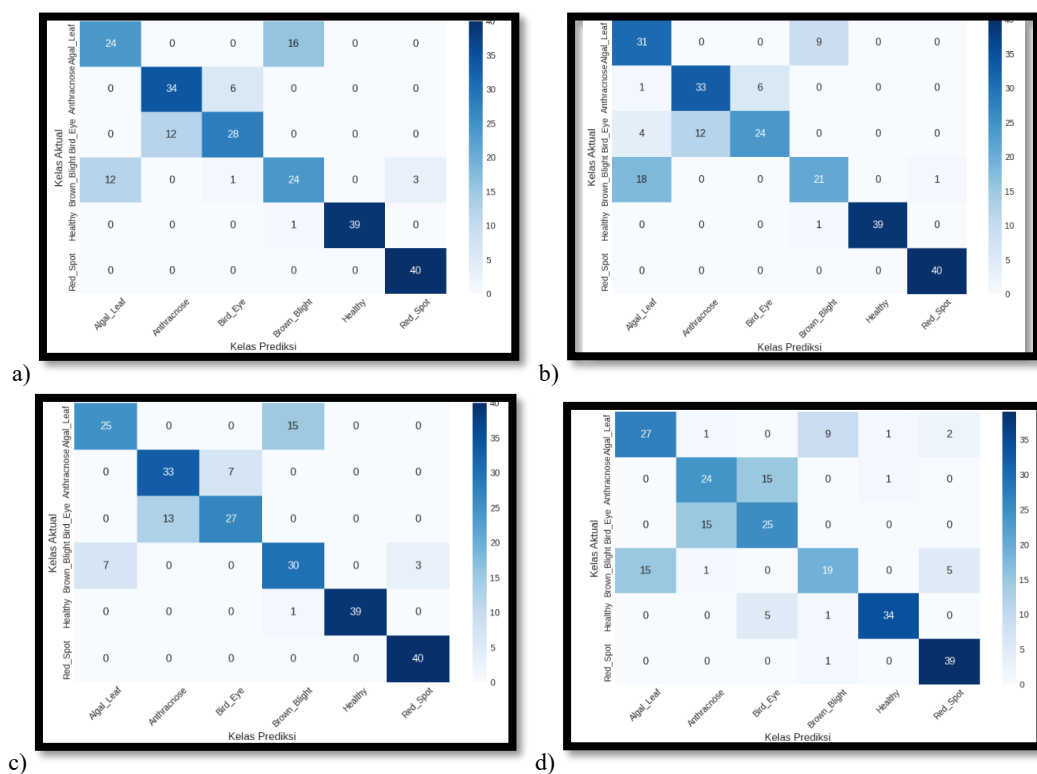
4.1.5.4 Hasil Evaluasi

Evaluasi performa model dilakukan menggunakan confusion matrix serta metrik akurasi, presisi, *recall*, dan *F1-score*. Hasil evaluasi keempat kernel pada data tidak seimbang disajikan pada Tabel 4.7.

Tabel 4. 7 Hasil evaluasi pada data tidak seimbang

No	Kernel	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)
1	Linear	78.75	78.69	78.75	78.55
2	Polynomial	78.33	79.34	78.33	78.14
3	RBF	80.83	81.25	80.83	80.67
4	Sigmoid	70.00	70.16	70.00	69.73

Tabel 4.7 menunjukkan kernel RBF sebagai performa terbaik (akurasi 80,83%, presisi 81,25%, *recall* 80,83%, *F1-score* 80,67%), sedangkan kernel Sigmoid terendah (akurasi 70,00%). Selisih akurasi RBF dan Sigmoid mencapai 10,83%, sedangkan Linear dan Polynomial relatif seimbang (selisih 0,42%) namun masih di bawah RBF. Keunggulan RBF konsisten dengan penelitian sebelumnya yang menyatakan fleksibilitasnya dalam menangani data non-linear. *Confusion matrix* setiap kernel ditunjukkan pada Gambar 4.7.



Gambar 4. 7 Confusion matrix kernel: a) linear, b) polynomial, c) RBF dan d) sigmoid , pada data tidak seimbang

Gambar 4.7 menunjukkan beberapa kelas sering tertukar, terutama kelas minoritas seperti *Healthy Leaf* yang memiliki *recall* lebih rendah. Detail metrik per kelas untuk kernel terbaik (RBF) disajikan pada Tabel 4.8.

Tabel 4. 8 Detail metrik per kelas kernel terbaik pada fata tidak seimbang

No	Kelas	TP	FP	FN	TN	Presisi (%)	Recall (%)	F1-Score (%)
1	<i>Algal Leaf</i>	25	7	15	193	78.12	62.5	69.44
2	<i>Anthraco</i>	33	13	7	187	71.74	82.5	76.74
3	<i>Bird Eye</i>	27	7	13	193	79.41	67.5	72.97
4	<i>Brown Blight</i>	30	16	10	184	65.22	75.0	69.77
5	<i>Healthy</i>	39	0	1	200	100.0	97.5	98.73
6	<i>Red Spot</i>	40	3	0	197	93.02	100.0	96.39
Rata-rata / Total		194	46	46	1154	81.25	80.83	80.67

Tabel 4.8 menunjukkan kelas mayoritas (*Red Leaf Spot*) memiliki performa terbaik (*recall* 100%, *F1-score* 96,39%), sedangkan *Healthy Leaf* juga sangat baik (presisi 100%, *recall* 97,5%) karena karakteristik visualnya yang paling berbeda (hijau homogen tanpa bercak). Sebaliknya, *Algal Leaf* memiliki *recall* terendah (62,50%) 15 dari 40 sampel gagal dikenali akibat kemiripan visual dengan *Brown Blight*, sebagaimana terlihat pada simulasi *decision function* (Tabel 4.6). *Brown Blight* sendiri memiliki presisi terendah (65,22%), menunjukkan banyak sampel kelas lain salah diprediksi sebagai *Brown Blight*.

Distribusi *False Negative* (FN) tertinggi pada *Algal Leaf* (15) dan *Bird's Eye* (13), mengindikasikan model sering gagal mendeteksi kedua kelas. Sebaliknya, *Healthy Leaf* memiliki *False Positive* (FP) = 0, menegaskan akurasi tinggi dalam membedakan daun sehat dari berpenyakit. Hasil ini menjadi justifikasi perlunya penyeimbangan data untuk meningkatkan *recall* pada kelas dengan tingkat kesalahan tinggi.

4.1.6 Hasil Klasifikasi SVM pada Data Seimbang

Pada skenario kedua, model SVM dilatih menggunakan data latih yang telah diseimbangkan melalui *hybrid resampling*. Empat fungsi kernel diuji untuk melihat pengaruh penyeimbangan data terhadap performa klasifikasi.

4.1.6.1 Analisis Support Vector

Jumlah *support vector* setiap kernel pada data seimbang ditunjukkan pada Tabel 4.9.

Tabel 4. 9 Jumlah support vector pada data seimbang

N o	Kernel	<i>Red Spot</i>	<i>Brown Blight</i>	<i>Bird's Eye</i>	<i>Anthra cnose</i>	<i>Algal Leaf</i>	<i>Health y</i>	Total
1	Linear	264	199	203	265	18	21	970
2	Polynomial	299	246	263	278	38	29	1153
3	RBF	266	221	231	259	29	40	1046
4	Sigmoid	251	194	242	252	40	86	1065

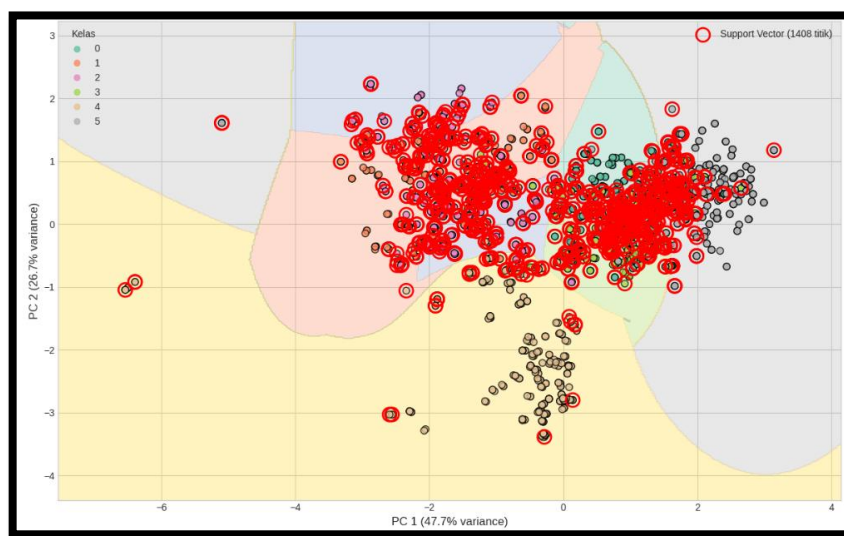
Tabel 4.9 menunjukkan distribusi *support vector* antar kelas lebih merata setelah penyeimbangan data, dengan peningkatan total pada seluruh kernel: Linear (+188 menjadi 970), Polynomial (+230 menjadi 1.153), RBF (+182 menjadi 1.046), dan Sigmoid (+205 menjadi 1.065). Peningkatan ini menunjukkan penambahan data augmentasi pada kelas minoritas menghasilkan lebih banyak titik data yang menjadi *support vector*.

Peningkatan paling signifikan terjadi pada kernel Polynomial (1.153 dari 2.160 data latih, 53,38%), wajar karena derajat 3 memerlukan lebih banyak *support vector* untuk *hyperplane* kompleks. Meskipun total meningkat, proporsi *support vector* terhadap total data latih justru menurun pada kernel RBF turun dari 51,15% menjadi 48,43%. Penurunan ini mengindikasikan peningkatan efisiensi model,

karena persentase data yang menjadi *support vector* lebih kecil, berpotensi meningkatkan kemampuan generalisasi.

4.1.6.2 Visualisasi *Hyperplane*

Visualisasi *hyperplane* kernel RBF pada data seimbang ditunjukkan pada Gambar 4.8.



Gambar 4. 8 Visualisasi *hyperplane* SVM kernel RBF pada data seimbang

Gambar 4.8 menunjukkan sebaran data lebih merata setelah penyeimbangan, dengan bentuk *hyperplane* yang berbeda dari skenario tidak seimbang. Hal ini menegaskan bahwa distribusi data memengaruhi bentuk batas keputusan yang dihasilkan SVM.

4.1.6.3 Simulasi *Decision Function*

Simulasi *decision function* kernel RBF pada data seimbang ditunjukkan pada Tabel 4.10.

Tabel 4. 10 Simulasi *decision function* kernel RBF pada data seimbang

No	Kelas Aktual	Prediksi	Status	Skor
1	<i>Algal Leaf</i>	<i>Algal Leaf</i>	V	5.2996
2	<i>Anthraco</i>	<i>Anthraco</i>	V	5.2824
3	<i>Bird Eye</i>	<i>Bird Eye</i>	V	5.2986
4	<i>Brown Blight</i>	<i>Algal Leaf</i>	X	5.2926
5	<i>Healthy</i>	<i>Healthy</i>	V	5.2786
6	<i>Red Spot</i>	<i>Red Spot</i>	v	5.2835

Tabel 4.10 menunjukkan jumlah prediksi benar tetap 5 dari 6 sampel, dengan kesalahan yang sama seperti skenario tidak seimbang *Brown Blight* tetap diprediksi sebagai *Algal Leaf*. Hal ini mengindikasikan *hybrid resampling* belum cukup mengatasi kemiripan visual antara kedua kelas.

Dibandingkan Tabel 4.6, skor *Brown Blight* sedikit menurun (5,2933 menjadi 5,2926), sedangkan skor *Algal Leaf* sedikit meningkat (5,2990 menjadi 5,2996). Perubahan sangat kecil ini menunjukkan karakteristik fundamental kedua kelas yang mirip tetap sulit dipisahkan meskipun data telah diseimbangkan. Diperlukan fitur lebih diskriminatif, seperti tekstur, untuk membedakan keduanya secara lebih akurat.

4.1.6.4 Hasil Evaluasi

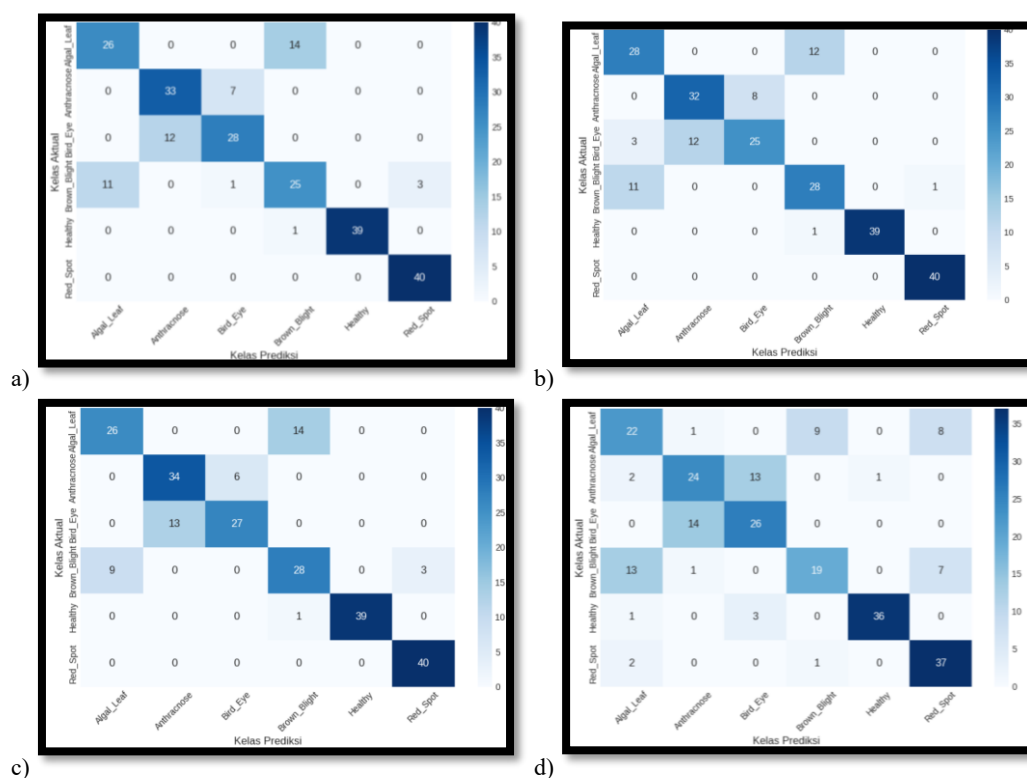
Hasil evaluasi keempat kernel pada data seimbang disajikan pada Tabel 4.11.

Tabel 4. 11 Hasil evaluasi pada data seimbang (*hybrid resampling*)

No	Kernel	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)
1	Linear	79.58	79.48	79.58	79.41
2	Polynomial	80.00	80.17	80.00	79.94
3	RBF	80.83	81.10	80.83	80.68
4	Sigmoid	68.33	68.48	68.33	67.90

Tabel 4.11 menunjukkan kernel RBF tetap sebagai performa terbaik (akurasi 80,83%, presisi 81,10%, *recall* 80,83%, *F1-score* 80,68%). Dibanding skenario

tidak seimbang (Tabel 4.7), Linear meningkat (78,75% menjadi 79,58%), Polynomial meningkat (78,33% menjadi 80,00%), RBF stabil (80,83%), sedangkan Sigmoid menurun (70,00% menjadi 68,33%). Stabilitas RBF menegaskan ketahanannya terhadap perubahan distribusi data, menjadikannya andal untuk dataset penyakit daun teh. Peningkatan Linear dan Polynomial menunjukkan manfaat penyeimbangan data bagi kernel sederhana, sedangkan penurunan Sigmoid menegaskan ketidakcocokannya untuk dataset ini. *Confusion matrix* setiap kernel pada data seimbang ditunjukkan pada Gambar 4.9.



Gambar 4. 9 Confusion matrix kernel: a) linear, b) polynomial, c) RBF dan d) sigmoid , pada data seimbang

Detail metrik per kelas untuk kernel terbaik (RBF) pada data seimbang disajikan pada Tabel 4.12.

Tabel 4. 12 Detail metrik perkelas kernel terbaik pada data seimbang

No	Kelas	TP	FP	FN	TN	Presisi (%)	Recall (%)	F1-Score (%)
1	<i>Algal Leaf</i>	26	9	14	191	74.29	65.00	69.33
2	<i>Anthracnose</i>	34	13	6	187	72.34	85.00	78.16
3	<i>Bird Eye</i>	27	6	13	194	81.82	67.50	73.97
4	<i>Brown Blight</i>	28	15	12	185	65.12	70.00	67.47
5	<i>Healthy</i>	39	0	1	200	100.00	97.50	98.73
6	<i>Red Spot</i>	40	3	0	197	93.02	100.00	96.39
Rata-rata / Total		194	46	46	1154	81.10	80.83	80.68

Dibanding Tabel 4.8 (skenario tidak seimbang), total TP, FP, FN, TN tidak berubah (TP=194, FP=46, FN=46, TN=1.154), namun terjadi pergeseran distribusi kesalahan. *Algal Leaf* meningkat *recall*-nya (62,50% menjadi 65,00%) dan *Anthracnose* juga meningkat (82,50% menjadi 85,00%), menunjukkan perbaikan deteksi setelah penyeimbangan. Sebaliknya, *Brown Blight* justru menurun (75,00% menjadi 70,00%), mengindikasikan beberapa sampelnya lebih sulit dikenali.

Kelas *Healthy Leaf* dan *Red Leaf Spot* tetap stabil dengan performa sangat baik (presisi/*recall* mendekati 100%) pada kedua skenario, menunjukkan karakteristik visualnya sangat berbeda dari kelas lain. Secara keseluruhan, *hybrid resampling* berdampak positif pada *Algal Leaf* dan *Anthracnose*, tetapi belum efektif untuk *Brown Blight*. Diperlukan augmentasi lebih spesifik atau penambahan fitur lain untuk meningkatkan kemampuan model membedakan kelas dengan kemiripan visual tinggi.

4.1.7 Perbandingan Kinerja Antar Skenario.

Perbandingan kinerja antara skenario data tidak seimbang dan data seimbang disajikan pada Tabel 4.13 untuk melihat pengaruh penyeimbangan data dan pemilihan kernel secara menyeluruh.

Tabel 4. 13 Perbandingan kinerja antar skenario

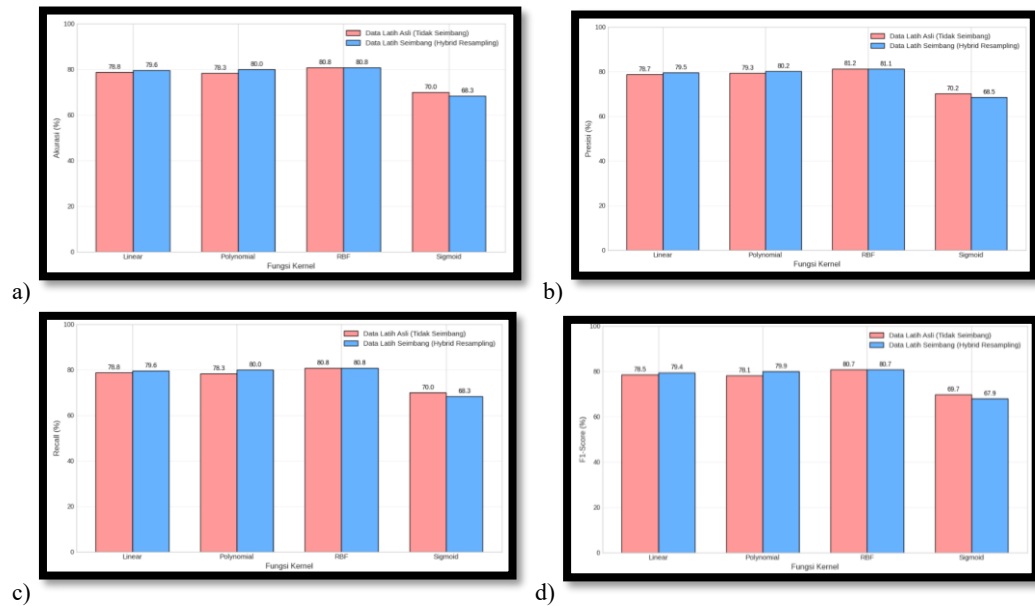
No	Kernel	Akurasi (Imb)	Akurasi (Bal)	Presisi (Imb)	Presisi (Bal)	Recall (Imb)	Recall (Bal)	F1-Score (Imb)	F1-Score (Bal)
1	Linear	78.75	79.58	78.69	79.48	78.75	79.58	78.55	79.41
2	Polynomial	78.33	80.00	79.34	80.17	78.33	80.00	78.14	79.94
3	RBF	80.83	80.83	81.25	81.10	80.83	80.83	80.67	80.68
4	Sigmoid	70.00	68.33	70.16	68.48	70.00	68.33	69.73	67.90

Tabel 4.13 menunjukkan *hybrid resampling* umumnya memberikan pengaruh positif. Selisih metrik antar skenario disajikan pada Tabel 4.14.

Tabel 4. 14 Selisih metrik (seimbang dengan tidak seimbang)

No	Kernel	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)
1	Linear	0.83	0.79	0.83	0.86
2	Polynomial	1.67	0.83	1.67	1.80
3	RBF	0.00	-0.15	0.00	0.01
4	Sigmoid	-1.67	-1.68	-1.67	-1.83

Tabel 4.14 menunjukkan peningkatan terbesar pada kernel Polynomial (*F1-Score* +1,80%, akurasi +1,67%), karena karakteristik fungsi pangkatnya yang lebih diuntungkan dengan bertambahnya data augmentasi untuk mempelajari interaksi fitur kompleks. Kernel RBF sangat stabil (selisih mendekati nol), menegaskan ketahanannya terhadap perubahan distribusi data. Kernel Linear juga meningkat (*F1-Score* +0,86%), sedangkan Sigmoid menurun pada seluruh metrik (*F1-Score* -1,83%), semakin menegaskan ketidakcocokannya untuk dataset ini. Secara keseluruhan, hasil ini mengonfirmasi bahwa pemilihan kernel yang tepat dan strategi penyeimbangan data merupakan dua faktor kunci yang saling memengaruhi dalam membangun model klasifikasi SVM yang optimal. Perbandingan visual metrik evaluasi ditunjukkan pada Gambar 4.10.

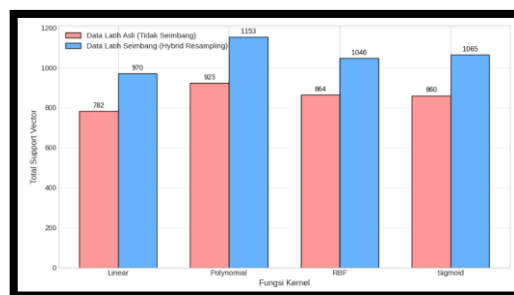


Gambar 4. 10 Grafik perbandingan: a) akurasi, b) presisi, c) *recall* dan d) *F1-Score*, antar kernel dan skenario

Gambar 4.10 mengonfirmasi kernel RBF sebagai performa terbaik konsisten pada kedua skenario, sedangkan Sigmoid terendah. *Recall* mengalami peningkatan paling signifikan setelah penyeimbangan, sesuai tujuan *hybrid resampling* untuk meningkatkan deteksi kelas minoritas.

4.1.8 Perbandingan *Support Vector*

Jumlah *support vector* mencerminkan kompleksitas model. Perbandingan total *support vector* setiap kernel pada kedua skenario ditunjukkan pada Gambar 4.11.



Gambar 4. 11 Perbandingan jumlah *support vector*

Gambar 4.11 menunjukkan kernel Polynomial menghasilkan *support vector* terbanyak pada kedua skenario (923 tidak seimbang, 1.153 seimbang) karena fungsi pangkat derajat 3 memetakan data ke ruang fitur lebih tinggi, membangun *hyperplane* paling kompleks. Sebaliknya, kernel Linear menghasilkan paling sedikit (782 tidak seimbang, 970 seimbang), menegaskan *hyperplane* paling sederhana (garis linear).

Setelah penyeimbangan, total *support vector* meningkat pada seluruh kernel: Linear +188 (782 menjadi 970), Polynomial +230 (923 menjadi 1.153), RBF +182 (864 menjadi 1.046), Sigmoid +205 (860 menjadi 1.065). Peningkatan ini wajar karena data latih bertambah dari 1.689 menjadi 2.160, sehingga lebih banyak titik berpotensi menjadi *support vector*.

Meskipun jumlah absolut meningkat, proporsi *support vector* terhadap total data latih justru menurun pada sebagian besar kernel: Linear (46,30% menjadi 44,91%), Polynomial (54,65% menjadi 53,38%), RBF (51,15% menjadi 48,43%). Penurunan proporsi ini mengindikasikan peningkatan efisiensi model persentase data yang menjadi *support vector* lebih kecil, berpotensi meningkatkan generalisasi. Sigmoid juga menurun (50,92% menjadi 49,31%). Temuan ini menegaskan *hybrid resampling* tidak hanya menyeimbangkan data tetapi juga membantu membangun *hyperplane* lebih efisien.

4.2 Pembahasan

Pembahasan ini menganalisis temuan sub-bab 4.1 secara lebih mendalam, mencakup pengaruh fungsi kernel terhadap kinerja klasifikasi, efektivitas *hybrid resampling* dalam menangani ketidakseimbangan data, pola kesalahan klasifikasi,

perbandingan dengan penelitian terdahulu, serta implikasi dan keterbatasan penelitian. Setiap temuan dikaitkan dengan teori relevan dan hasil penelitian sebelumnya untuk memberikan pemahaman yang komprehensif.

4.2.1 Pengaruh Fungsi Kernel terhadap Kinerja Klasifikasi

Berdasarkan Tabel 4.7 dan 4.11, pemilihan fungsi kernel berpengaruh signifikan terhadap performa SVM. Kernel RBF konsisten menjadi yang terbaik pada kedua skenario (akurasi 80,83%), sejalan dengan Suhendra & Juliwardi (2022) yang menyatakan keunggulan RBF pada data non-linear.

Keunggulan RBF berasal dari pemetaan data ke ruang dimensi tak terhingga melalui fungsi eksponensial $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$, membentuk *hyperplane* fleksibel untuk memisahkan kelas dengan fitur tumpang tindih sesuai dengan karakteristik penyakit daun teh yang memiliki kemiripan visual. Meskipun demikian, RBF bukan penghasil *support vector* terbanyak; Polynomial justru menghasilkan lebih banyak (923 dan 1.153 dibandingkan 864 dan 1.046 pada RBF). Hal ini menunjukkan kompleksitas model tidak selalu sebanding dengan performa RBF lebih efisien karena mencapai akurasi tertinggi dengan *support vector* lebih sedikit.

Kernel Linear kompetitif (78,75% tidak seimbang, 79,58% seimbang) meskipun hanya membentuk *hyperplane* linear, mengindikasikan data memiliki keterpisahan linear yang cukup baik. Keunggulannya pada jumlah *support vector* paling sedikit, berarti model lebih sederhana dan komputasi lebih cepat. Sebaliknya, kernel Sigmoid konsisten terendah (70,00% tidak seimbang, 68,33% seimbang) dan justru menurun setelah penyeimbangan. Hal ini disebabkan fungsi *tangen*

hiperbolik yang tidak selalu positif semi-definit, sehingga pemetaan ke ruang fitur tidak optimal (Schölkopf & Smola, 2002). Temuan ini mengonfirmasi ketidakcocokan kernel Sigmoid untuk dataset penyakit daun teh.

4.2.2 Pengaruh *Hybrid Resampling* terhadap Kinerja Klasifikasi

Penerapan *hybrid resampling* mengombinasikan *random undersampling* pada kelas mayoritas (*Red Leaf Spot*) dan augmentasi pada lima kelas minoritas memberikan pengaruh bervariasi terhadap performa. Berdasarkan Tabel 4.14, peningkatan terbesar terjadi pada kernel Polynomial (*F1-Score* +1,80%, akurasi +1,67%), diikuti Linear (*F1-Score* +0,86%), sedangkan RBF stabil dan Sigmoid justru menurun. Peningkatan pada Linear dan Polynomial menunjukkan penyeimbangan data bermanfaat bagi kernel sederhana, karena augmentasi memperkaya variasi sampel kelas minoritas sehingga generalisasi meningkat sejalan dengan Bhatia dkk. (2020) penanganan ketidakseimbangan data dapat meningkatkan *recall* kelas minoritas.

Namun, Tabel 4.12 menunjukkan tidak semua kelas minoritas meningkat. *Algal Leaf* (62,50% menjadi 65,00%) dan *Anthracnose* (82,50% menjadi 85,00%) meningkat, tetapi *Brown Blight* justru menurun (75,00% menjadi 70,00%). Penurunan ini mengindikasikan augmentasi rotasi dan *flipping* belum cukup membedakan *Brown Blight* dari *Algal Leaf* yang memiliki kemiripan visual tinggi. Stabilitas kernel RBF pada kedua skenario (akurasi tetap 80,83%) menegaskan ketahanannya terhadap perubahan distribusi data, menjadikannya pilihan andal untuk implementasi sistem identifikasi penyakit daun teh meskipun data tidak seimbang.

4.2.3 Analisis Kesalahan Klasifikasi

Berdasarkan *confusion matrix* (Tabel 4.8 dan 4.12), pasangan *Brown Blight* dan *Algal Leaf* paling sering tertukar. Pada data tidak seimbang, *Brown Blight* memiliki FN 10 (25% data uji) dan FP 16; pada data seimbang kondisinya memburuk dengan FN meningkat menjadi 12 (30%). Simulasi *decision function* (Tabel 4.6 dan 4.10) mengonfirmasi skor kedua kelas sangat berdekatan. Kesalahan ini disebabkan kemiripan visual: *Brown Blight* (bercak cokelat tua di tepi daun) dan *Algal Leaf* (bercak oranye kemerahan yang bisa tampak kecokelatan). Fitur statistik yang digunakan, terutama *Mean Hue* dan *Mean Saturation*, menghasilkan nilai tumpang tindih untuk kedua kelas, sebagaimana terlihat pada *boxplot* Gambar 4.5.

Kelas *Bird's Eye Spot* juga memiliki *recall* rendah (67,50% pada kedua skenario), dengan 13 sampel gagal dikenali menunjukkan model kesulitan membedakan lesi melingkar kecil dari bercak penyakit lain. Hal ini mengindikasikan lima fitur statistik yang digunakan belum cukup menangkap perbedaan tekstur halus antar kelas. Di sisi positif, *Healthy Leaf* dan *Red Leaf Spot* menunjukkan performa sangat baik (*recall* mendekati 100%). *Healthy Leaf* dengan presisi 100% menegaskan tidak ada daun berpenyakit yang salah diklasifikasi sebagai sehat hal krusial karena kesalahan tersebut dapat berakibat fatal (penyakit tidak tertangani).

4.2.4 Perbandingan dengan Penelitian Terdahulu

Untuk memposisikan hasil penelitian ini dalam konteks yang lebih luas, dilakukan perbandingan dengan beberapa penelitian relevan pada Tabel 4.15.

Tabel 4. 15 Perbandingan dengan penelitian terdahulu

No	Peneliti	Metode	Dataset	Akurasi Terbaik
1	Bhagat & Kumar (2024)	VGG-16 + SVM Sigmoid & Linear	Daun teh (4 kelas)	96,67%
2	Rahat dkk. (2025)	<i>Custom CNN</i>	Daun teh (8 kelas)	99,00%
3	Ardhiansyah dkk. (2026)	K-NN + GLCM	Daun teh (6 kelas)	95,83%
4	Nurmadinah dkk. (2025)	SVM RBF + GLCM + HSV	Daun cabai (5 kelas)	91,30%
5	Raihan & Wijaya (2026)	MobileNetV2 + SVM	Daun teh (7 kelas)	75,30%
6	Suhendra & Juliwardi (2022)	SVM + Warna + GLCM	Daun jagung (2 kelas)	99,70%
7	Penelitian Ini	SVM RBF + Fusi HSV & Grayscale + Hybrid Resampling	Daun teh (6 kelas)	80,83%

Tabel 4.15 menunjukkan akurasi penelitian ini (80,83%) lebih tinggi dari Raihan & Wijaya (75,3%) namun lebih rendah dari lainnya. Perbedaan ini disebabkan beberapa faktor. Pertama, penelitian ini hanya menggunakan lima fitur statistik sederhana, sementara Ardhiansyah dkk. (2026) pada dataset yang sama mencapai 95,83% dengan fitur tekstur GLCM yang lebih kaya (*contrast, correlation, energy, homogeneity*) menunjukkan penambahan fitur tekstur dapat meningkatkan akurasi signifikan pada data dengan kemiripan visual tinggi.

Kedua, Bhagat & Kumar (2024) dan Rahat dkk. (2025) menggunakan *deep learning* (VGG-16 dan CNN kustom) yang mengekstraksi fitur otomatis, namun memerlukan komputasi besar dan data latih banyak— berbeda dengan penelitian ini yang mengutamakan kesederhanaan dan efisiensi. Ketiga, meskipun akurasi lebih rendah dari Ardhiansyah dkk., penelitian ini memberikan kontribusi melalui analisis mendalam proses internal SVM (*support vector, hyperplane, decision function*) yang jarang dibahas, serta menguji secara eksplisit pengaruh *hybrid*

resampling pada enam kelas penyakit teh dengan ketidakseimbangan data signifikan.

4.2.5 Integrasi Keilmuan dalam Perspektif Islam

Penelitian klasifikasi penyakit daun teh berbasis citra menggunakan *Support Vector Machine* (SVM) ini pada hakikatnya merupakan bentuk implementasi dari perintah Allah *Subhanahu wa Ta'ala* untuk mengkaji dan merenungkan ciptaan-Nya. Dalam perspektif Islam, manusia diperintahkan untuk mengamati fenomena alam sebagai sarana mengenali kebesaran dan kekuasaan Allah. Hal ini sejalan dengan firman Allah dalam QS. Al-An'am ayat 99:

وَهُوَ الَّذِي أَنْزَلَ مِنَ السَّمَاءِ مَاءً فَأَخْرَجْنَا بِهِ نَبَاتَ كُلِّ شَيْءٍ فَأَخْرَجْنَا مِنْهُ خَضِرًا مَخْرُجًا مِنْهُ حَبًّا مُتَرَاكِبًا وَمِنَ النَّخْلِ مِنْ طَلْعِهَا قِنْوَانٌ دَانِيَةٌ وَجَنَّاتٍ مِنْ أَعْنَابٍ وَالزَّيْتُونَ وَالرُّمَّانَ مُشْتَبِهًا وَغَيْرَ مُتَشَابِهٍ انظُرُوا إِلَى ثَمَرِهِ إِذَا أَثْمَرَ وَيَنْعِهِ ۚ إِنَّ فِي ذَلِكَ لَآيَاتٍ لِقَوْمٍ يُؤْمِنُونَ ﴿٩٩﴾

"Dialah yang menurunkan air dari langit lalu dengannya Kami menumbuhkan segala macam tumbuhan. Maka, darinya Kami mengeluarkan tanaman yang menghijau. Darinya Kami mengeluarkan butir yang bertumpuk (banyak). Dari mayang kurma (mengurai) tangkai-tangkai yang menjuntai. (Kami menumbuhkan) kebun-kebun anggur. (Kami menumbuhkan pula) zaitun dan delima yang serupa dan yang tidak serupa. Perhatikanlah buahnya pada waktu berbuah dan menjadi masak. Sesungguhnya pada yang demikian itu benar-benar terdapat tanda-tanda (kekuasaan Allah) bagi kaum yang beriman." (Q.S. Al-An'am/6: 99).

Menurut Tafsir Al-Wajiz, ayat ini menjelaskan bahwa Allah adalah Dzāt yang menurunkan hujan dari awan, lalu mengeluarkan macam-macam tumbuhan yang berbeda di bumi. Allah menumbuhkan kebun-kebun anggur, zaitun, dan delima yang ukuran dan warnanya hampir serupa, namun rasanya tidak serupa (az-Zuhaili, 2019). Tafsir as-Sa'di menambahkan bahwa perintah "*unzhuru*" (perhatikanlah) dalam ayat ini adalah seruan untuk melakukan observasi dan perenungan mendalam terhadap keajaiban ciptaan Allah, karena hal itu

mengandung pelajaran-pelajaran dan tanda-tanda kebesaran Allah sebagai bukti rahmat-Nya, keluasan karunia dan kemurahan-Nya, serta kesempurnaan kodrat-Nya (as-Sa'di, 2019).

Dalam konteks penelitian ini, implementasi ekstraksi fitur visual melalui ruang warna HSV dan statistik tekstur *grayscale* menjadi wujud nyata dari metode observasi modern untuk menangkap tanda-tanda kebesaran Allah yang termanifestasi dalam keanekaragaman warna dan tekstur daun teh. Mesin komputasi dilatih untuk menganalisis anomali spektrum warna dan tingkat kekasaran tekstur dengan presisi tinggi, membedah detail nilai matriks piksel yang sering kali luput dari pandangan mata telanjang manusia. Dengan demikian, proses klasifikasi penyakit daun teh ini merupakan bentuk tadabbur (perenungan) terhadap keanekaragaman ciptaan Allah.

Penelitian ini juga merupakan bentuk tolong-menolong dalam kebaikan sebagaimana diperintahkan dalam QS. Al-Maidah ayat 2:

يَا أَيُّهَا الَّذِينَ آمَنُوا لَا تَحْلُوا شَعَائِرَ اللَّهِ وَلَا الشَّهْرَ الْحَرَامَ وَلَا الْهَدْيَ وَلَا الْقَلَائِدَ وَلَا آمِينَ الْبَيْتِ الْحَرَامَ يَبْتَغُونَ فَضْلًا مِّن رَّبِّهِمْ وَرِضْوَانًا وَإِذَا حَلَلْتُمْ فَاصْطَادُوا وَلَا يَجْرِمَنَّكُمْ شَنَاٰنُ قَوْمٍ أَن صَدُّوكُمْ عَنِ الْمَسْجِدِ الْحَرَامِ أَن تَعْتَدُوا وَتَعَاوَنُوا عَلَى الْبِرِّ وَالتَّقْوَىٰ وَلَا تَعَاوَنُوا عَلَى الْإِثْمِ وَالْعُدْوَانِ وَاتَّقُوا اللَّهَ إِنَّ اللَّهَ شَدِيدُ الْعِقَابِ ﴿٢﴾

"Wahai orang-orang yang beriman, janganlah kamu melanggar syiar-syiar (kesucian) Allah, jangan (melanggar kehormatan) bulan-bulan haram, jangan (mengganggu) hadyu (hewan-hewan kurban) dan qalā'id (hewan-hewan kurban yang diberi tanda), dan jangan (pula mengganggu) para pengunjung Baitulharam sedangkan mereka mencari karunia dan rida Tuhannya! Apabila kamu telah bertahalul (menyelesaikan ihram), berburulah (jika mau). Janganlah sekali-kali kebencian(-mu) kepada suatu kaum, karena mereka menghalang-halangimu dari Masjidilharam, mendorongmu berbuat melampaui batas (kepada mereka). Tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa, dan jangan tolong-menolong dalam berbuat dosa dan permusuhan. Bertakwalah kepada Allah, sesungguhnya Allah sangat berat siksaan-Nya." (Q.S. Al-Maidah/5: 2).

Menurut Tafsir as-Sa'di, ayat ini memerintahkan kaum mukmin untuk saling membantu dalam kebaikan dan takwa. Kebajikan mencakup segala perbuatan baik yang dicintai Allah, baik yang berkaitan dengan hak Allah maupun hak manusia, sedangkan takwa adalah sikap meninggalkan segala perbuatan yang dibenci Allah dan Rasul-Nya (as-Sa'di, 2019). Tafsir Al-Wajiz juga menjelaskan bahwa perintah tolong-menolong ini mencakup seluruh aspek kehidupan, termasuk upaya membantu sesama dalam memecahkan masalah yang mereka hadapi (az-Zuhaili, 2019).

Sistem identifikasi penyakit yang dikembangkan dalam penelitian ini bertujuan membantu petani teh mendeteksi penyakit secara cepat dan akurat, sehingga dapat mengurangi kerugian ekonomi dan meningkatkan produktivitas. Hal ini merupakan implementasi nyata dari perintah tolong-menolong dalam kebaikan, karena dengan menyediakan alat bantu deteksi penyakit yang andal, penulis turut membantu meringankan beban petani dan berkontribusi pada kesejahteraan mereka. Nilai-nilai *itqan* (bekerja dengan tekun, teliti, dan profesional) juga terefleksikan secara kuat melalui proses optimalisasi arsitektur komputasi yang dilakukan dalam penelitian ini.

Lebih jauh, deteksi dini penyakit memungkinkan pengendalian yang lebih tepat sasaran, mengurangi penggunaan pestisida berlebihan, dan menjaga keseimbangan ekosistem. Hal ini sejalan dengan larangan membuat kerusakan di bumi dalam QS. Al-A'raf ayat 56:

وَلَا تُفْسِدُوا فِي الْأَرْضِ بَعْدَ إِصْلَاحِهَا وَادْعُوهُ حَوْفًا وَقَطْمًا إِنَّ رَحْمَتَ اللَّهِ قَرِيبٌ مِّنَ الْمُحْسِنِينَ ﴿٥٦﴾

"Janganlah kamu berbuat kerusakan di bumi setelah diatur dengan baik. Berdoalah kepada-Nya dengan rasa takut dan penuh harap. Sesungguhnya rahmat Allah sangat dekat dengan orang-orang yang berbuat baik."(Q.S. Al-A'raf/7: 56).

Menurut Tafsir Al-Wajiz, ayat ini melarang berbuat kerusakan di bumi dengan berbuat syirik dan maksiat setelah bumi diperbaiki dengan pengutusan para rasul, penurunan kitab, dan dijelaskannya syariat (az-Zuhaili, 2019). Tafsir as-Sa'di menjelaskan bahwa larangan berbuat kerusakan di muka bumi mencakup segala bentuk kemaksiatan yang dapat merusak akhlak, amal, dan rizki. Sebaliknya, ketaatan akan memperbaiki keadaan dunia dan akhirat (as-Sa'di, 2019). Tafsir Ibnu Katsir juga menegaskan bahwa Allah melarang berbuat kerusakan di bumi dan hal-hal yang memberi kemudharatan setelah adanya perbaikan, karena kerusakan tersebut membahayakan seluruh hamba (Katsir & Umar, 2017).

Penelitian ini sejalan dengan nilai tersebut karena deteksi dini penyakit memungkinkan pengendalian yang tepat sasaran, sehingga petani dapat menghindari penggunaan pestisida secara berlebihan yang dapat mencemari tanah dan air, serta menjaga keseimbangan ekosistem. Dengan demikian, penelitian ini tidak hanya bermanfaat bagi peningkatan produktivitas pertanian, tetapi juga merupakan bagian dari upaya menjaga kelestarian alam sebagaimana diperintahkan dalam Al-Qur'an.

Pada akhirnya, perancangan sistem cerdas dalam ranah patologi tanaman ini menegaskan kembali kedudukan manusia sebagai *khalifah fil ardh*, yang memikul amanah mulia untuk memelihara kelestarian ekosistem dan mengawal ketahanan pangan. Pengembangan sistem pengklasifikasi ini tidak hanya berdiri sebagai pencapaian teknologi murni, melainkan menjelma sebagai sebuah ikhtiar akademis

bernilai ibadah guna meminimalkan potensi kerugian para petani, sekaligus merawat keberkahan rezeki yang ditumbuhkan oleh Allah *Subhanahu wa Ta'ala* dari hasil bumi. Dengan demikian, penelitian ini mengintegrasikan tiga dimensi hubungan kepada Allah (ibadah melalui pengamatan ciptaan-Nya), kepada manusia (tolong-menolong kepada petani), dan kepada alam (pelestarian lingkungan) dalam satu kerangka keilmuan yang utuh.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian dan pembahasan, kesimpulan yang diperoleh adalah sebagai berikut:

1. Performa kernel terbaik: Kernel RBF secara konsisten unggul pada kedua skenario dengan akurasi 80,83%, presisi 81,25%, *recall* 80,83%, dan *F1-score* 80,67%. Sebaliknya, kernel Sigmoid menjadi yang terendah dengan akurasi 70,00% (data tidak seimbang) dan 68,33% (data seimbang).
2. Pengaruh hybrid resampling: Peningkatan paling signifikan terjadi pada kernel Polynomial (*F1-score* +1,80%), sementara kernel RBF stabil dengan perubahan sangat kecil ($\Delta F1\text{-score} = 0,01\%$). Stabilitas RBF menjadikannya pilihan andal untuk klasifikasi penyakit daun teh.
3. Analisis kesalahan klasifikasi: Kesalahan tertinggi terjadi antara kelas *Brown Blight* dan *Algal Leaf* akibat kemiripan visual. Setelah *hybrid resampling*, *recall Algal Leaf* meningkat (62,50% menjadi 65,00%), namun *recall Brown Blight* justru menurun (75,00% menjadi 70,00%). Sebaliknya, *Healthy Leaf* dan *Red Leaf Spot* menunjukkan performa sangat baik (*recall* > 97%).
4. Efisiensi model: Kernel Linear menghasilkan *support vector* paling sedikit (782 dan 970), mencerminkan model paling sederhana dan efisien secara

komputasi. Proporsi *support vector* terhadap total data latih menurun setelah penyeimbangan, mengindikasikan peningkatan efisiensi model.

5.2 Saran

Berdasarkan hasil dan keterbatasan penelitian, saran untuk pengembangan selanjutnya adalah:

1. Penambahan fitur tekstur GLCM: Fitur yang digunakan hanya terbatas pada lima fitur statistik. Penambahan *Gray Level Co-occurrence Matrix* (GLCM) dengan parameter *contrast*, *correlation*, *energy*, dan *homogeneity* direkomendasikan untuk meningkatkan kemampuan model membedakan kelas dengan kemiripan visual tinggi, terutama *Brown Blight* dan *Algal Leaf*. Penelitian Ardiansyah dkk. (2026) menunjukkan GLCM mampu meningkatkan akurasi hingga 95,83% pada dataset yang sama.
2. Optimasi *hyperparameter* dengan *GridSearchCV*: Parameter kernel SVM menggunakan nilai default *Scikit-Learn* ($C=1,0$, $\gamma='scale'$). Penggunaan *GridSearchCV* atau *RandomizedSearchCV* untuk mencari kombinasi C dan γ optimal berpotensi meningkatkan performa, terutama pada kernel RBF dan Polynomial.
3. Pengembangan aplikasi mobile *real-time*: Hasil penelitian dapat dikembangkan menjadi sistem deteksi penyakit daun teh berbasis aplikasi *mobile* untuk membantu petani di lapangan. Dengan kernel Linear yang memiliki *support vector* paling sedikit, model dapat dijalankan pada perangkat dengan sumber daya komputasi terbatas.

4. Uji coba di lapangan: Pengujian pada dataset dengan variasi pencahayaan, sudut pengambilan gambar, dan latar belakang berbeda diperlukan untuk mengukur generalisasi model sebelum implementasi nyata.
5. Eksplorasi metode *ensemble*: Untuk meningkatkan akurasi tanpa mengorbankan efisiensi, penelitian selanjutnya dapat mengeksplorasi *Voting Classifier* yang menggabungkan prediksi dari beberapa kernel SVM, atau arsitektur *deep learning* ringan seperti *MobileNet* yang dikombinasikan dengan SVM.

DAFTAR PUSTAKA

- as-Sa'di, A. R. (2019). Tafsir as-Sa'di (Penerjemah, Trans.). Darul Haq.
- az-Zuhaili, W. (2019). Tafsir Al-Wajiz (Penerjemah, Trans.). Gema Insani.
- Anjarsari, I. R. D., Ariyanti, M., & Rosniawaty, S. (2020). Studi ekofisiologis tanaman teh guna meningkatkan pertumbuhan, hasil, dan kualitas teh. *Kultivasi*, 19(3), 1181–1188. <https://doi.org/10.24198/kultivasi.v19i3.26623>
- Ardhiansyah, R. A., Anggraeny, F. T., & Maulana, H. (2026). Implementasi Metode K-NN Untuk Klasifikasi Penyakit Daun Teh Berdasarkan Fitur Tekstur GLCM. *CICES (Cyberpreneurship Innovative and Creative Exact and Social Science)*. <https://doi.org/10.33050/35m6gw66>
- Asterius, P. (2023). *Klasifikasi Penyakit pada Tanaman Teh Menggunakan Algoritma C4.5* [Universitas Medan Area]. <https://repositori.uma.ac.id/handle/123456789/22684>
- Bhagat, M., & Kumar, D. (2024). Performance enhancement of kernelized SVM with deep learning features for tea leaf disease prediction. *Multimedia Tools and Applications*, 83(13), 39117–39134. <https://doi.org/10.1007/s11042-023-17172-1>
- Bhatia, A., Chug, A., & Prakash Singh, A. (2020). Application of extreme learning machine in plant disease prediction for highly imbalanced dataset. *Journal of Statistics and Management Systems*, 23(6), 1059–1068. <https://doi.org/10.1080/09720510.2020.1799504>
- Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on computational learning theory* (pp. 144–152). ACM Press. <https://doi.org/10.1145/130385.130401>
- Chandra, M. A., & Bedi, S. S. (2021). Survey on SVM and their application in image classification. *International Journal of Information Technology (Singapore)*, 13(5). <https://doi.org/10.1007/s41870-017-0080-1>
- Eskundari, D. R. (2020). Budidaya Jaringan Tanaman Teh di Indonesia. *BIOMA: JURNAL BIOLOGI MAKASSAR*, 5(2), 121–130. <http://journal.unhas.ac.id/index.php/bioma>
- Fletcher, T. (2008). *Support Vector Machines Explained*. www.cs.ucl.ac.uk/staff/T.Fletcher/
- Ginanjari, B., Budiman, M. A., & Trimono, L. (2019). Usahatani Tanaman Teh Rakyat (Camellia Sinensis). *Jurnal Ilmiah Mahasiswa AGROINFO GALUH*, 6(1), 169–182. <https://doi.org/http://dx.doi.org/10.25157/jimag.v6i1.1512>

- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*.
- Han, H., & Jiang, X. (2014). Overcome Support Vector Machine Diagnosis Overfitting. *Cancer Informatics*, 13s1, CIN.S13875. <https://doi.org/10.4137/cin.s13875>
- Katsir, I. (2017). Tafsir Al-Qur'an Al-'Azhim (I. b. Umar, Ed.; 1st ed., Vol. 7). Al-Maktabah Al-Islamiyyah.
- Lu, J., Tan, L., & Jiang, H. (2021). Review on convolutional neural network (CNN) applied to plant leaf disease classification. Dalam *Agriculture (Switzerland)* (Vol. 11, Nomor 8). MDPI AG. <https://doi.org/10.3390/agriculture11080707>
- Mohasel, S. M., & Koosha, H. (2025). Robust Support Vector Machines for Imbalanced and Noisy Data via Benders Decomposition. *ARXIV PREPRINT*. <http://arxiv.org/abs/2503.14873>
- Nurmadinah, Wajidi, F., & Arifin, N. (2025). Deteksi Penyakit Daun Cabai Menggunakan Kombinasi GLCM Dan HSV dengan Klasifikasi SVM. *INSECT*, 11(02). <https://data.mendeley.com/datasets/w9mr3vf56s/1>
- Rahat, I. S., Ghosh, H., Dara, S., & Kant, S. (2025). Towards precision agriculture tea leaf disease detection using CNNs and image processing. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-02378-0>
- Raihan, M. D., & Wijaya, N. (2026). Comparison of Tea Leaf Disease Classification Using SVM with MobileNetV2 and MobileNetV3-Small Feature Extractors. *Green Intelligent Systems and Applications*, 6(1). <https://doi.org/10.53623/gisa.v6i1.1158>
- Schölkopf, B., & Smola, A. J. (2002). Learning with kernels: Support vector machines, regularization, optimization, and beyond. MIT Press.
- Suhendra, R., & Juliwardi, I. (2022). Identifikasi dan Klasifikasi Penyakit Daun Jagung Menggunakan Support Vector Machine. *Jurnal Teknologi Informasi*, 1(1), 29–35. <http://jurnal.utu.ac.id/JTI>
- Suryadi, A., Johan, D., & Febriyanti, E. L. (2019). Sistem Pakar Pendiaknosaan Hama Dan Penyakit Pada tanaman teh menggunakan metode Case-Based Reasoning Berbasis Android. *Jurnal Sistem Informasi dan Manajemen*, 7(2), 47–54. <https://ejournal.stmikgici.ac.id/>
- Valkenborg, D., Rousseau, A. J., Geubbelmans, M., & Burzykowski, T. (2023). Support vector machines. Dalam *American Journal of Orthodontics and Dentofacial Orthopedics* (Vol. 164, Nomor 5). Elsevier Inc. <https://doi.org/10.1016/j.ajodo.2023.08.003>

Zhang, S., Li, X., Zong, M., Zhu, X., & Wang, R. (2018). Efficient kNN classification with different numbers of nearest neighbors. *IEEE Transactions on Neural Networks and Learning Systems*, 29(5), 1774–1785. <https://doi.org/10.1109/TNNLS.2017.2673241>

LAMPIRAN-LAMPIRAN

Lampiran 1 *Confusion matrix* linear dengan data tidak seimbang

No	Input	Output	TP	TN	FP	FN
1	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
2	<i>Algal Leaf</i>	<i>Brown Blight</i>	0	4	1	1
3	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
4	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
5	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
6	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
7	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
8	<i>Algal Leaf</i>	<i>Brown Blight</i>	0	4	1	1
9	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
10	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
11	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
12	<i>Anthracnose</i>	<i>Bird Eye</i>	0	4	1	1
13	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
14	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
15	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
16	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
17	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
18	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
19	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
20	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
21	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
22	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
23	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
24	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
25	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
26	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
27	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
28	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
29	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
30	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
31	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
32	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
33	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
34	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
35	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
36	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
37	<i>Brown Blight</i>	<i>Red Spot</i>	0	4	1	1
38	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
39	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
40	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
41	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
42	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
43	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
44	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
45	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0

No	Input	Output	TP	TN	FP	FN
46	Healthy	Healthy	1	5	0	0
47	Healthy	Healthy	1	5	0	0
48	Healthy	Healthy	1	5	0	0
49	Healthy	Healthy	1	5	0	0
50	Healthy	Healthy	1	5	0	0
51	Red Spot	Red Spot	1	5	0	0
52	Red Spot	Red Spot	1	5	0	0
53	Red Spot	Red Spot	1	5	0	0
54	Red Spot	Red Spot	1	5	0	0
55	Red Spot	Red Spot	1	5	0	0
56	Red Spot	Red Spot	1	5	0	0
57	Red Spot	Red Spot	1	5	0	0
58	Red Spot	Red Spot	1	5	0	0
59	Red Spot	Red Spot	1	5	0	0
60	Red Spot	Red Spot	1	5	0	0

Lampiran 2 *Confusion matrix* linear dengan seimbang

No	Input	Output	TP	TN	FP	FN
1	Algal Leaf	Algal Leaf	1	5	0	0
2	Algal Leaf	Brown Blight	0	4	1	1
3	Algal Leaf	Algal Leaf	1	5	0	0
4	Algal Leaf	Algal Leaf	1	5	0	0
5	Algal Leaf	Algal Leaf	1	5	0	0
6	Algal Leaf	Algal Leaf	1	5	0	0
7	Algal Leaf	Algal Leaf	1	5	0	0
8	Algal Leaf	Brown Blight	0	4	1	1
9	Algal Leaf	Algal Leaf	1	5	0	0
10	Algal Leaf	Algal Leaf	1	5	0	0
11	Anthracnose	Anthracnose	1	5	0	0
12	Anthracnose	Bird Eye	0	4	1	1
13	Anthracnose	Anthracnose	1	5	0	0
14	Anthracnose	Anthracnose	1	5	0	0
15	Anthracnose	Anthracnose	1	5	0	0
16	Anthracnose	Anthracnose	1	5	0	0
17	Anthracnose	Anthracnose	1	5	0	0
18	Anthracnose	Anthracnose	1	5	0	0
19	Anthracnose	Anthracnose	1	5	0	0
20	Anthracnose	Anthracnose	1	5	0	0
21	Bird Eye	Bird Eye	1	5	0	0
22	Bird Eye	Anthracnose	0	4	1	1
23	Bird Eye	Bird Eye	1	5	0	0
24	Bird Eye	Anthracnose	0	4	1	1
25	Bird Eye	Bird Eye	1	5	0	0
26	Bird Eye	Bird Eye	1	5	0	0
27	Bird Eye	Bird Eye	1	5	0	0
28	Bird Eye	Anthracnose	0	4	1	1
29	Bird Eye	Bird Eye	1	5	0	0
30	Bird Eye	Anthracnose	0	4	1	1
31	Brown Blight	Algal Leaf	0	4	1	1

No	Input	Output	TP	TN	FP	FN
32	Brown Blight	Algal Leaf	0	4	1	1
33	Brown Blight	Algal Leaf	0	4	1	1
34	Brown Blight	Brown Blight	1	5	0	0
35	Brown Blight	Brown Blight	1	5	0	0
36	Brown Blight	Algal Leaf	0	4	1	1
37	Brown Blight	Red Spot	0	4	1	1
38	Brown Blight	Algal Leaf	0	4	1	1
39	Brown Blight	Brown Blight	1	5	0	0
40	Brown Blight	Brown Blight	1	5	0	0
41	Healthy	Healthy	1	5	0	0
42	Healthy	Healthy	1	5	0	0
43	Healthy	Healthy	1	5	0	0
44	Healthy	Healthy	1	5	0	0
45	Healthy	Healthy	1	5	0	0
46	Healthy	Healthy	1	5	0	0
47	Healthy	Healthy	1	5	0	0
48	Healthy	Healthy	1	5	0	0
49	Healthy	Healthy	1	5	0	0
50	Healthy	Healthy	1	5	0	0
51	Red Spot	Red Spot	1	5	0	0
52	Red Spot	Red Spot	1	5	0	0
53	Red Spot	Red Spot	1	5	0	0
54	Red Spot	Red Spot	1	5	0	0
55	Red Spot	Red Spot	1	5	0	0
56	Red Spot	Red Spot	1	5	0	0
57	Red Spot	Red Spot	1	5	0	0
58	Red Spot	Red Spot	1	5	0	0
59	Red Spot	Red Spot	1	5	0	0
60	Red Spot	Red Spot	1	5	0	0

Lampiran 3 *Confusion matrix* polinomial dengan sata tidak seimbang

No	Input	Output	TP	TN	FP	FN
1	Algal Leaf	Algal Leaf	1	5	0	0
2	Algal Leaf	Brown Blight	0	4	1	1
3	Algal Leaf	Algal Leaf	1	5	0	0
4	Algal Leaf	Algal Leaf	1	5	0	0
5	Algal Leaf	Algal Leaf	1	5	0	0
6	Algal Leaf	Algal Leaf	1	5	0	0
7	Algal Leaf	Algal Leaf	1	5	0	0
8	Algal Leaf	Brown Blight	0	4	1	1
9	Algal Leaf	Algal Leaf	1	5	0	0
10	Algal Leaf	Algal Leaf	1	5	0	0
11	Anthracnose	Anthracnose	1	5	0	0
12	Anthracnose	Bird Eye	0	4	1	1
13	Anthracnose	Anthracnose	1	5	0	0
14	Anthracnose	Anthracnose	1	5	0	0
15	Anthracnose	Anthracnose	1	5	0	0
16	Anthracnose	Anthracnose	1	5	0	0
17	Anthracnose	Anthracnose	1	5	0	0

No	Input	Output	TP	TN	FP	FN
18	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
19	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
20	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
21	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
22	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
23	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
24	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
25	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
26	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
27	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
28	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
29	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
30	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
31	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
32	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
33	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
34	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
35	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
36	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
37	<i>Brown Blight</i>	<i>Red Spot</i>	0	4	1	1
38	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
39	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
40	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
41	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
42	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
43	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
44	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
45	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
46	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
47	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
48	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
49	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
50	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
51	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
52	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
53	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
54	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
55	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
56	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
57	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
58	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
59	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
60	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0

Lampiran 4 Confusion matrix polinomial dengan seimbang

No	Input	Output	TP	TN	FP	FN
1	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
2	<i>Algal Leaf</i>	<i>Brown Blight</i>	0	4	1	1
3	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0

No	Input	Output	TP	TN	FP	FN
4	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
5	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
6	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
7	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
8	<i>Algal Leaf</i>	<i>Brown Blight</i>	0	4	1	1
9	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
10	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
11	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
12	<i>Anthracnose</i>	<i>Bird Eye</i>	0	4	1	1
13	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
14	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
15	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
16	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
17	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
18	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
19	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
20	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
21	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
22	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
23	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
24	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
25	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
26	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
27	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
28	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
29	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
30	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
31	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
32	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
33	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
34	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
35	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
36	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
37	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
38	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
39	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
40	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
41	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
42	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
43	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
44	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
45	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
46	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
47	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
48	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
49	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
50	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
51	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
52	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
53	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
54	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
55	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0

No	Input	Output	TP	TN	FP	FN
56	Red Spot	Red Spot	1	5	0	0
57	Red Spot	Red Spot	1	5	0	0
58	Red Spot	Red Spot	1	5	0	0
59	Red Spot	Red Spot	1	5	0	0
60	Red Spot	Red Spot	1	5	0	0

Lampiran 5 Confusion matrix RBF dengan data tidak seimbang

No	Input	Output	TP	TN	FP	FN
1	Algal Leaf	Algal Leaf	1	5	0	0
2	Algal Leaf	Brown Blight	0	4	1	1
3	Algal Leaf	Algal Leaf	1	5	0	0
4	Algal Leaf	Algal Leaf	1	5	0	0
5	Algal Leaf	Algal Leaf	1	5	0	0
6	Algal Leaf	Algal Leaf	1	5	0	0
7	Algal Leaf	Algal Leaf	1	5	0	0
8	Algal Leaf	Brown Blight	0	4	1	1
9	Algal Leaf	Algal Leaf	1	5	0	0
10	Algal Leaf	Algal Leaf	1	5	0	0
11	Anthracnose	Anthracnose	1	5	0	0
12	Anthracnose	Bird Eye	0	4	1	1
13	Anthracnose	Anthracnose	1	5	0	0
14	Anthracnose	Anthracnose	1	5	0	0
15	Anthracnose	Anthracnose	1	5	0	0
16	Anthracnose	Anthracnose	1	5	0	0
17	Anthracnose	Anthracnose	1	5	0	0
18	Anthracnose	Anthracnose	1	5	0	0
19	Anthracnose	Anthracnose	1	5	0	0
20	Anthracnose	Anthracnose	1	5	0	0
21	Bird Eye	Bird Eye	1	5	0	0
22	Bird Eye	Anthracnose	0	4	1	1
23	Bird Eye	Bird Eye	1	5	0	0
24	Bird Eye	Anthracnose	0	4	1	1
25	Bird Eye	Bird Eye	1	5	0	0
26	Bird Eye	Bird Eye	1	5	0	0
27	Bird Eye	Bird Eye	1	5	0	0
28	Bird Eye	Anthracnose	0	4	1	1
29	Bird Eye	Bird Eye	1	5	0	0
30	Bird Eye	Anthracnose	0	4	1	1
31	Brown Blight	Algal Leaf	0	4	1	1
32	Brown Blight	Brown Blight	1	5	0	0
33	Brown Blight	Brown Blight	1	5	0	0
34	Brown Blight	Brown Blight	1	5	0	0
35	Brown Blight	Brown Blight	1	5	0	0
36	Brown Blight	Brown Blight	1	5	0	0
37	Brown Blight	Red Spot	0	4	1	1
38	Brown Blight	Algal Leaf	0	4	1	1
39	Brown Blight	Brown Blight	1	5	0	0
40	Brown Blight	Brown Blight	1	5	0	0
41	Healthy	Healthy	1	5	0	0

No	Input	Output	TP	TN	FP	FN
42	Healthy	Healthy	1	5	0	0
43	Healthy	Healthy	1	5	0	0
44	Healthy	Healthy	1	5	0	0
45	Healthy	Healthy	1	5	0	0
46	Healthy	Healthy	1	5	0	0
47	Healthy	Healthy	1	5	0	0
48	Healthy	Healthy	1	5	0	0
49	Healthy	Healthy	1	5	0	0
50	Healthy	Healthy	1	5	0	0
51	Red Spot	Red Spot	1	5	0	0
52	Red Spot	Red Spot	1	5	0	0
53	Red Spot	Red Spot	1	5	0	0
54	Red Spot	Red Spot	1	5	0	0
55	Red Spot	Red Spot	1	5	0	0
56	Red Spot	Red Spot	1	5	0	0
57	Red Spot	Red Spot	1	5	0	0
58	Red Spot	Red Spot	1	5	0	0
59	Red Spot	Red Spot	1	5	0	0
60	Red Spot	Red Spot	1	5	0	0

Lampiran 6 *Confusion matrix* RBF dengan seimbang

No	Input	Output	TP	TN	FP	FN
1	Algal Leaf	Algal Leaf	1	5	0	0
2	Algal Leaf	Brown Blight	0	4	1	1
3	Algal Leaf	Algal Leaf	1	5	0	0
4	Algal Leaf	Algal Leaf	1	5	0	0
5	Algal Leaf	Algal Leaf	1	5	0	0
6	Algal Leaf	Algal Leaf	1	5	0	0
7	Algal Leaf	Algal Leaf	1	5	0	0
8	Algal Leaf	Brown Blight	0	4	1	1
9	Algal Leaf	Algal Leaf	1	5	0	0
10	Algal Leaf	Algal Leaf	1	5	0	0
11	Anthracnose	Anthracnose	1	5	0	0
12	Anthracnose	Bird Eye	0	4	1	1
13	Anthracnose	Anthracnose	1	5	0	0
14	Anthracnose	Anthracnose	1	5	0	0
15	Anthracnose	Anthracnose	1	5	0	0
16	Anthracnose	Anthracnose	1	5	0	0
17	Anthracnose	Anthracnose	1	5	0	0
18	Anthracnose	Anthracnose	1	5	0	0
19	Anthracnose	Anthracnose	1	5	0	0
20	Anthracnose	Anthracnose	1	5	0	0
21	Bird Eye	Bird Eye	1	5	0	0
22	Bird Eye	Anthracnose	0	4	1	1
23	Bird Eye	Bird Eye	1	5	0	0
24	Bird Eye	Anthracnose	0	4	1	1
25	Bird Eye	Bird Eye	1	5	0	0
26	Bird Eye	Bird Eye	1	5	0	0
27	Bird Eye	Bird Eye	1	5	0	0

No	Input	Output	TP	TN	FP	FN
28	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
29	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
30	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
31	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
32	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
33	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
34	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
35	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
36	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
37	<i>Brown Blight</i>	<i>Red Spot</i>	0	4	1	1
38	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
39	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
40	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
41	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
42	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
43	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
44	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
45	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
46	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
47	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
48	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
49	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
50	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
51	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
52	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
53	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
54	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
55	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
56	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
57	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
58	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
59	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
60	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0

Lampiran 7 *Confusion matrix* Sigmoid dengan data tidakseimbang

No	Input	Output	TP	TN	FP	FN
1	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
2	<i>Algal Leaf</i>	<i>Brown Blight</i>	0	4	1	1
3	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
4	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
5	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
6	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
7	<i>Algal Leaf</i>	<i>Healthy</i>	0	4	1	1
8	<i>Algal Leaf</i>	<i>Brown Blight</i>	0	4	1	1
9	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
10	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
11	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
12	<i>Anthracnose</i>	<i>Bird Eye</i>	0	4	1	1
13	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0

No	Input	Output	TP	TN	FP	FN
14	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
15	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
16	<i>Anthracnose</i>	<i>Bird Eye</i>	0	4	1	1
17	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
18	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
19	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
20	<i>Anthracnose</i>	<i>Bird Eye</i>	0	4	1	1
21	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
22	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
23	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
24	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
25	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
26	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
27	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
28	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
29	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
30	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
31	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
32	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
33	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
34	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
35	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
36	<i>Brown Blight</i>	<i>Red Spot</i>	0	4	1	1
37	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
38	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
39	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
40	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
41	<i>Healthy</i>	<i>Bird Eye</i>	0	4	1	1
42	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
43	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
44	<i>Healthy</i>	<i>Bird Eye</i>	0	4	1	1
45	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
46	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
47	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
48	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
49	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
50	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
51	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
52	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
53	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
54	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
55	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
56	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
57	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
58	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
59	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
60	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0

Lampiran 8 *Confusion matrix* sigmoid dengan seimbang

No	Input	Output	TP	TN	FP	FN
1	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
2	<i>Algal Leaf</i>	<i>Brown Blight</i>	0	4	1	1
3	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
4	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
5	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
6	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
7	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
8	<i>Algal Leaf</i>	<i>Brown Blight</i>	0	4	1	1
9	<i>Algal Leaf</i>	<i>Algal Leaf</i>	1	5	0	0
10	<i>Algal Leaf</i>	<i>Brown Blight</i>	0	4	1	1
11	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
12	<i>Anthracnose</i>	<i>Bird Eye</i>	0	4	1	1
13	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
14	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
15	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
16	<i>Anthracnose</i>	<i>Algal Leaf</i>	0	4	1	1
17	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
18	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
19	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
20	<i>Anthracnose</i>	<i>Anthracnose</i>	1	5	0	0
21	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
22	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
23	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
24	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
25	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
26	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
27	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
28	<i>Bird Eye</i>	<i>Bird Eye</i>	1	5	0	0
29	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
30	<i>Bird Eye</i>	<i>Anthracnose</i>	0	4	1	1
31	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
32	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
33	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
34	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
35	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
36	<i>Brown Blight</i>	<i>Red Spot</i>	0	4	1	1
37	<i>Brown Blight</i>	<i>Red Spot</i>	0	4	1	1
38	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
39	<i>Brown Blight</i>	<i>Brown Blight</i>	1	5	0	0
40	<i>Brown Blight</i>	<i>Algal Leaf</i>	0	4	1	1
41	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
42	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
43	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
44	<i>Healthy</i>	<i>Bird Eye</i>	0	4	1	1
45	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
46	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
47	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
48	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
49	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0
50	<i>Healthy</i>	<i>Healthy</i>	1	5	0	0

No	<i>Input</i>	<i>Output</i>	TP	TN	FP	FN
51	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
52	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
53	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
54	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
55	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
56	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
57	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
58	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
59	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0
60	<i>Red Spot</i>	<i>Red Spot</i>	1	5	0	0

Catatan: Ditampilkan 10 dari 40 data uji per kelas sebagai representasi.