



**KLASIFIKASI CITRA PENYAKIT DAUN KOPI ARABIKA  
MENGUNAKAN SUPPORT VECTOR MACHINE (SVM)  
DENGAN SELEKSI FITUR INFORMATION GAIN**

**SKRIPSI**

**SYLVIANA MURNI**

**1810511016**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**

**FAKULTAS ILMU KOMPUTER**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**2022**



**KLASIFIKASI CITRA PENYAKIT DAUN KOPI ARABIKA  
MENGUNAKAN SUPPORT VECTOR MACHINE (SVM)  
DENGAN SELEKSI FITUR INFORMATION GAIN**

**SKRIPSI**

**Diajukan Sebagai Syarat Untuk Memperoleh Gelar  
Sarjana Komputer**

**SYLVIANA MURNI**

**1810511016**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI TEKNIK INFORMATIKA  
2022**

## PERNYATAAN ORISINALITAS

Tugas Akhir ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk saya nyatakan dengan benar.

Nama : Sylviana Murni

NIM : 1810511016

Tanggal : 26 Juli 2022

Bilamana dikemudian hari ditemukan ketidak sesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 26 Juli 2022

Yang Menyatakan,



(Sylviana Murni)

**PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK  
KEPENTINGAN AKADEMIS**

---

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan dibawah ini:

Nama : Sylviana Murni

NIM : 1810511016

Fakultas : Ilmu Komputer

Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non eksklusif (Non-exclusive Royalty Free Right) atas karya ilmiah saya berjudul :

**KLASIFIKASI CITRA PENYAKIT DAUN KOPI ARABIKA  
MENGUNAKAN SUPPORT VECTOR MACHINE (SVM) DENGAN  
SELEKSI FITUR INFORMATION GAIN**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 26 Juli 2022

Yang Menyatakan,



(Sylviana Murni)

## PENGESAHAN

Dengan ini dinyatakan bahwa Skripsi berikut:

Nama : Sylviana Murni  
NIM : 1810511016  
Program Studi : S1 Informatika 2018  
Judul : Klasifikasi Citra Penyakit Daun Kopi Arabika Menggunakan *Support Vector Machine* (SVM) Dengan Seleksi Fitur *Information Gain*.

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Yuni Widiastiwi, S.Kom., M.Si.

Penguji I



Desta Sandva Prasvita, M.Kom.

Penguji II



Dr. Didit Widiyanto, S.Kom., M.Si.

Pembimbing I



Catur Nugrahaeni PD, M.Kom.

Pembimbing II



Dr. Ermatita, M.Kom.

Dekan



Desta Sandva Prasvita, M.Kom.

Ketua Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 13 Juli 2022



# KLASIFIKASI CITRA PENYAKIT DAUN KOPI ARABIKA MENGUNAKAN SUPPORT VECTOR MACHINE (SVM) DENGAN SELEKSI FITUR INFORMATION GAIN

Sylviana Murni

## Abstrak

Tanaman kopi arabika lebih rentan terhadap penyakit dibandingkan dengan jenis penyakit kopi lainnya. Hal tersebut menyulitkan petani kopi dalam memelihara tanaman kopi arabika sehingga diperlukan proses identifikasi penyakit sejak dini. Proses identifikasi penyakit tanaman dapat didiagnosa melalui daun. Teknik pengolahan citra digital dapat dimanfaatkan dalam mengklasifikasikan jenis penyakit daun kopi arabika melalui citra daun. Tahapan yang dilakukan yaitu pra proses, ekstraksi fitur, seleksi fitur, dan klasifikasi. Metode ekstraksi fitur yang digunakan adalah *Gray-Level Co-occurrence Matrix* (GLCM) sebagai metode ekstraksi fitur tekstur dan *Color Moment* sebagai metode ekstraksi fitur warna. Proses ekstraksi fitur GLCM dan *Color Moment* menghasilkan 29 fitur. Pemilihan fitur diperlukan untuk memperoleh fitur – fitur yang relevan yang memberikan hasil klasifikasi terbaik. Pada penelitian ini menerapkan metode seleksi fitur *Information Gain*. Algoritma klasifikasi yang digunakan yaitu *Support Vector Machine* (SVM). Total data yang digunakan pada data *imbalanced* sebanyak 1325 data dan data *balanced* sebanyak 1534 data dengan pembagian 80% data latih dan 20% data uji. Dalam menentukan jumlah fitur terbaik yang memberikan akurasi tertinggi dilakukan pengujian performa klasifikasi dengan penggunaan jumlah fitur yang berbeda pada data *imbalanced* dan data *balanced*. Perbedaan jumlah fitur diperoleh dari kombinasi atribut peringkat teratas hingga terendah hasil seleksi fitur *Information Gain*. Hasil pengujian menunjukkan, penggunaan 18 fitur dari 29 fitur pada data *imbalanced* menghasilkan akurasi tertinggi dengan nilai akurasi sebesar 68,30%, presisi sebesar 55,77%, dan *recall* sebesar 57,85%.

**Kata Kunci** : Penyakit Daun, Tanaman Kopi Arabika, *Support Vector Machine* (SVM), *Information Gain*.

# ARABICA COFFEE LEAF DISEASE IMAGE CLASSIFICATION USING SUPPORT VECTOR MACHINE (SVM) WITH INFORMATION GAIN FEATURE SELECTION

**Sylviana Murni**

## **Abstract**

*Arabica coffee plants are more susceptible to disease than other types of coffee diseases. This makes it difficult for coffee farmers to maintain arabica coffee plants so that an early disease identification process is needed. The process of identifying plant diseases can be diagnosed through the leaves. Digital image processing techniques can be used in classifying the types of Arabica coffee leaf diseases through leaf imagery. The steps taken are pre-processing, feature extraction, feature selection, and classification. The feature extraction method used is Gray-Level Co-occurrence Matrix (GLCM) as a texture feature extraction method and Color Moment as a color feature extraction method. The GLCM and Color Moment feature extraction process yielded 29 features. Feature selection is needed to obtain relevant features that give the best classification results. In this study, the Information Gain feature selection method is applied. The classification algorithm used is the Support Vector Machine (SVM). The total data used in the imbalanced data is 1325 data and the balanced data is 1534 data with the distribution of 80% training data and 20% test data. In determining the best number of features that provide the highest accuracy, classification performance testing is carried out by using a different number of features on imbalanced and balanced data. The difference in the number of features is obtained from the combination of the highest to lowest ranking attributes from the Information Gain feature selection. The test results show that the use of 18 features out of 29 features in the imbalanced data produces the highest accuracy with an accuracy value of 68.30%, precision of 55.77%, and recall of 57.85%.*

**Keywords :** *Leaf Disease, Arabica Coffee Plants, Support Vector Machine (SVM), Information Gain.*

## KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala karunia Nya, sehingga tugas akhir ini berhasil diselesaikan. Penulis ingin mengucapkan terima kasih kepada:

1. Orang tua dan keluarga yang selalu memberikan dukungan, dorongan, dan doa setiap saat kepada penulis agar dapat menyelesaikan tugas akhir ini.
2. Bapak Dr. Didit Widiyanto, S.Kom., M.Si. dan Ibu Catur Nugrahaeni PD, S.Kom, M.Kom., selaku dosen pembimbing yang berjasa dan telah memberikan bimbingan sehingga ini tugas akhir dapat terselesaikan.
3. Bapak Desta Sandya Prasvita, M.Kom., selaku Ketua Jurusan Informatika Universitas Pembangunan Nasional Veteran Jakarta.
4. Ibu Dr. Ermatita, M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.
5. Ibu Dr. Ana Feronika Cindra Irawati, SP., MP., selaku pakar penyakit tanaman yang telah membantu melakukan validasi data dalam penelitian ini.
6. Sahabat-sahabat penulis di kelas A, serta teman-teman Informatika 2018 Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.

Penulis menyadari bahwa penulisan tugas akhir ini masih memiliki banyak kekurangan. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun sebagai bahan perbaikan untuk penelitian selanjutnya. Semoga tugas akhir ini dapat memberikan manfaat dan wawasan bagi pembaca.

Jakarta, 20 Juni 2022

Penulis  
Sylviana Murni



## DAFTAR ISI

PERNYATAAN ORISINALITAS .....	i
PERNYATAAN PERSETUJUAN PUBLIKASI .....	ii
PENGESAHAN .....	iii
Abstrak .....	iv
Abstract .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR .....	xi
DAFTAR TABEL.....	xiii
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN.....	1
1.1    Latar Belakang .....	1
1.2    Rumusan Masalah .....	3
1.3    Tujuan Penelitian.....	3
1.4    Manfaat.....	3
1.5    Ruang Lingkup Penelitian .....	4
1.6    Luaran Yang Diharapkan .....	4
1.7    Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA .....	6
2.1    Identifikasi Penyakit Tanaman Kopi Arabika .....	6
2.1.1    Pengorok daun ( <i>Leaf Miner</i> ).....	6
2.1.2    Karat daun ( <i>Leaf Rust</i> ).....	6
2.1.3    Hawar daun ( <i>Phoma</i> ).....	6
2.1.4    Bercak daun ( <i>Cercospora</i> ).....	7

2.2	Segmentasi <i>Thresholding</i> .....	7
2.3	Warna HSV .....	8
2.4	<i>Gray-level Co-Occurance Matrix (GLCM)</i> .....	9
2.5	<i>Color Moment</i> .....	11
2.6	<i>Information Gain</i> .....	12
2.7	<i>Support Vector Machine (SVM)</i> .....	14
2.8.1	<i>Support Vector Machine (SVM) Multiclass dengan One-Against-All...</i>	16
2.9	Evaluasi <i>Confusion Matrix</i> .....	17
2.10	Permasalahan Detail .....	19
2.10.1	Klasifikasi Jenis Makanan menggunakan <i>Neighbor Weighted K-Nearest Neighbor</i> dengan Seleksi Fitur <i>Information Gain</i> (2019).....	21
2.10.2	Klasifikasi Penyakit Daun Kentang Berdasarkan Fitur Tekstur dan Fitur Warna Menggunakan <i>Support Vector Machine</i> (2018).....	22
2.10.3	<i>Tomato Leaf Diseases Classification Based on Leaf Images: A Comparison between Classical Machine Learning and Deep Learning Methods</i> (2021) .....	22
2.10.4	<i>Deep Learning for Classification and Severity Estimation of Coffee Leaf Biotic Stress</i> (2019) .....	23
BAB III HASIL DAN PEMBAHASAN .....		24
3.1	Alur Penelitian.....	24
3.1.1	Identifikasi Masalah.....	25
3.1.2	Studi Pustaka .....	25
3.1.3	Akuisisi Data .....	25
3.1.4	Validasi Data .....	26
3.1.5	Pra Proses.....	26
3.1.6	Ekstraksi Fitur.....	27
3.1.7	Seleksi Fitur .....	29

3.1.8	Pembagian Data .....	29
3.1.9	Klasifikasi .....	29
3.1.10	Evaluasi .....	29
3.2	Perangkat Penelitian .....	30
3.2.1	Perangkat Keras ( <i>Hardware</i> ) .....	30
3.2.2	Perangkat Lunak ( <i>Software</i> ) .....	30
3.3	Jadwal Penelitian .....	30
BAB IV HASIL DAN PEMBAHASAN .....		32
4.1	Identifikasi Masalah .....	32
4.2	Akuisisi Data .....	32
4.3	Validasi Data .....	33
4.4	Pra Proses .....	33
4.4.1	Augmentasi Data .....	34
4.4.2	Segmentasi Citra .....	35
4.4.3	<i>Cropping</i> .....	37
4.4.4	<i>Resize</i> .....	38
4.4.5	<i>Grayscale</i> .....	39
4.4.6	Konversi Citra RGB ke HSV .....	40
4.4.7	Pemisahan Ruang Warna HSV .....	40
4.5	Ekstraksi Fitur .....	41
4.5.1	Gray-Level Co-occurrence Matrix (GLCM) .....	41
4.5.2	<i>Color Moment</i> .....	43
4.6	Seleksi Fitur .....	44
4.7	Pembagian Data .....	49
4.8	Klasifikasi .....	50
4.9	Evaluasi .....	52

4.9.1 Pengujian Jumlah Fitur Pada Data <i>Imbalanced</i> .....	52
4.9.2 Pengujian Jumlah Fitur Pada Data <i>Balanced</i> .....	55
BAB V PENUTUP .....	61
5.1. Kesimpulan.....	61
5.2. Saran.....	61
DAFTAR PUSTAKA .....	63
RIWAYAT HIDUP .....	66

## DAFTAR GAMBAR

Gambar 2.1 Orientasi arah piksel GLCM dengan sudut $0^\circ$ , $45^\circ$ , $90^\circ$ , $135^\circ$ (Sumber: Kadir & Susanto, 2013:555) .....	9
Gambar 2.2 (a) Kemungkinan garis pemisah dan (b) Maksimal margin garis pemisah .....	14
Gambar 3.1 Metodologi Penelitian .....	24
Gambar 3.2 Pengorok daun ( <i>Leaf Miner</i> ) .....	26
Gambar 3.3 Karat daun ( <i>Leaf Rust</i> ) .....	26
Gambar 3.4 Hawar daun ( <i>Phoma</i> ) .....	26
Gambar 3.5 Bercak daun ( <i>Cercospora</i> ) .....	26
Gambar 3.6 Sehat ( <i>Healthy</i> ) .....	26
Gambar 3.7 Alur Ekstraksi Tekstur GLCM.....	28
Gambar 3.8 Alur Ekstraksi Warna <i>Color Moment</i> .....	28
Gambar 4.1 Distribusi Data <i>Imbalanced</i> .....	34
Gambar 4.2 Distribusi Data <i>Balanced</i> .....	35
Gambar 4.3 Hasil Augmentasi Citra.....	35
Gambar 4.4 Citra Biner <i>Cercospora</i> Nilai Ambang 60 .....	36
Gambar 4.5 Citra Biner <i>Healthy</i> Nilai Ambang 60 .....	36
Gambar 4.6 Citra Biner <i>Cercospora</i> Nilai Ambang 80 .....	36
Gambar 4.7 Citra Biner <i>Healthy</i> Nilai Ambang 80 .....	36
Gambar 4.8 Citra Biner <i>Cercospora</i> Nilai Ambang 100 .....	36
Gambar 4.9 Citra Biner <i>Healthy</i> Nilai Ambang 100 .....	36
Gambar 4.10 Citra Awal .....	37
Gambar 4.11 Citra Biner Proses Morfologi.....	37
Gambar 4.12 Citra RGB Hasil Segmentasi.....	37
Gambar 4.13 Pemusatan Objek Citra.....	38
Gambar 4.14 Citra Hasil <i>Cropping</i> .....	38
Gambar 4.15 Data Citra Sebelum dan Sesudah Proses <i>Resize</i> .....	39
Gambar 4.16 Citra <i>Grayscale</i> .....	39
Gambar 4.17 Citra HSV .....	40
Gambar 4.18 Citra <i>Channel</i> Warna H.....	41

Gambar 4.19 Citra <i>Channel</i> Warna S .....	41
Gambar 4.20 Citra <i>Channel</i> Warna V .....	41
Gambar 4.21 Grafik Hubungan Jumlah Fitur Terhadap Akurasi Untuk Nilai C ( <i>Cost</i> ) =100 .....	59
Gambar 4.22 Grafik Hubungan Jumlah Fitur Terhadap Akurasi Untuk Nilai C ( <i>Cost</i> ) =1000 .....	59

## DAFTAR TABEL

Tabel 2.1 Rentang Nilai HSV .....	8
Tabel 2.2 Contoh Masalah Klasifikasi Menggunakan Pendekatan <i>One-Against-All</i> .....	17
Tabel 2.3 Multiclass Confusion Matrix .....	17
Tabel 2.4 Penelitian Terdahulu .....	19
Tabel 3.1 Jadwal Penelitian .....	31
Tabel 4.1 Jumlah Data Citra Daun Kopi Arabika.....	33
Tabel 4.2 Hasil Ekstraksi Fitur GLCM Arah Sudut 0 Derajat.....	42
Tabel 4.3 Hasil Ekstraksi Fitur GLCM Arah Sudut 45 Derajat.....	42
Tabel 4.4 Hasil Ekstraksi Fitur GLCM Arah Sudut 90 Derajat.....	42
Tabel 4.5 Hasil Ekstraksi Fitur GLCM Arah Sudut 135 Derajat.....	43
Tabel 4.6 Hasil Ekstraksi Fitur Warna <i>Color Moment Channel</i> Warna H .....	43
Tabel 4.7 Hasil Ekstraksi Fitur Warna <i>Color Moment Channel</i> Warna S.....	44
Tabel 4.8 Hasil Ekstraksi Fitur Warna <i>Color Moment Channel</i> Warna V .....	44
Tabel 4.9 Sampel Data Fitur Contrast_0.....	45
Tabel 4.10 Tabel Data Kelompok .....	46
Tabel 4. 11 Hasil Seleksi Fitur Information Gain Data <i>Imbalanced</i> .....	47
Tabel 4.12 Hasil Seleksi Fitur Information Gain Data <i>Balanced</i> .....	48
Tabel 4.13 Pembagian Data <i>Imbalanced</i> .....	50
Tabel 4.14 Pembagian Data <i>Balanced</i> .....	50
Tabel 4.15 Hasil Pengujian Jumlah Fitur Pada Data <i>Imbalanced</i> Untuk Nilai C=100.....	52
Tabel 4. 16 Hasil Pengujian Jumlah Fitur Pada Data <i>Imbalanced</i> Untuk Nilai C=1000.....	53
Tabel 4.17 Hasil Pengujian Jumlah Fitur Pada Data <i>Balanced</i> Untuk Nilai C=100 .....	55
Tabel 4.18 Hasil Pengujian Jumlah Fitur Pada Data <i>Balanced</i> Untuk Nilai C=1000.....	56
Tabel 4.19 Hasil Pengujian Jumlah Fitur Terbaik .....	58

## DAFTAR LAMPIRAN

Lampiran 1 Bukti Pelaksanaan Validasi Data.....	68
Lampiran 2 Citra Penyakit Daun Kopi Arabika.....	69
Lampiran 3 Citra Pra Proses .....	72
Lampiran 4 <i>Source Code</i> Pra Proses.....	74
Lampiran 5 <i>Source Code</i> GLCM dan <i>Color Moment</i> .....	76
Lampiran 6 <i>Source Code</i> Proses Data Citra .....	80
Lampiran 7 <i>Source Code</i> Seleksi Fitur <i>Information Gain</i> .....	82
Lampiran 8 Hasil Ekstraksi Fitur GLCM dan <i>Color Moment</i> Data .....	86
Lampiran 9 Hasil Seleksi Fitur <i>Information Gain</i> .....	87
Lampiran 10 Tabel Daftar Kombinasi Fitur .....	88
Lampiran 11 Uji Turnitin.....	95