

TESIS

**KLASIFIKASI PENYAKIT DAUN KELAPA SAWIT MENGGUNAKAN
K-NEAREST NEIGHBOR DAN SELEKSI FITUR
ALGORITMA GENETIKA**



Disusun oleh:

Nama : Enda Putri Atika
NIM : 20.51.1416
Konsentrasi : Business Intelligence

**PROGRAM STUDI S2 TEKNIK INFORMATIKA
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2022

TESIS

**KLASIFIKASI PENYAKIT DAUN KELAPA SAWIT MENGGUNAKAN
K-NEAREST NEIGHBOR DAN SELEKSI FITUR
ALGORITMA GENETIKA**

**GENETIC ALGORITHM AND K-NEAREST NEIGHBOR FOR OIL
PALM LEAF DISEASE CLASSIFICATION**



Disusun oleh:

Nama : Enda Putri Atika
NIM : 20.51.1416
Konsentrasi : Business Intelligence

**PROGRAM STUDI S2 TEKNIK INFORMATIKA
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2022

HALAMAN PENGESAHAN

**KLASIFIKASI PENYAKIT DAUN KELAPA SAWIT
MENGUNAKAN K-NEAREST NEIGHBOR DAN
SELEKSI FITUR ALGORITMA GENETIKA**

**GENETIC ALGORITHM AND K-NEAREST NEIGHBOR
FOR OIL PALM LEAF DISEASES CLASSIFICATION**

Dipersiapkan dan Disusun oleh

Enda Putri Atika

20.51.1416

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tesis
Program Studi S2 Teknik Informatika
Program Pascasarjana Universitas AMIKOM Yogyakarta
pada hari Senin, 5 September 2022

Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer

Yogyakarta, 5 September 2022

Rektor

Prof. Dr. M. Suvanto, M.M.

NIK. 190302001

HALAMAN PERSETUJUAN

KLASIFIKASI PENYAKIT DAUN KELAPA SAWIT MENGUNAKAN K-NEAREST NEIGHBOR DAN SELEKSI FITUR ALGORITMA GENETIKA

GENETIC ALGORITHM AND K-NEAREST NEIGHBOR FOR OIL PALM LEAF DISEASES CLASSIFICATION

Dipersiapkan dan Disusun oleh

Enda Putri Atika

20.51.1416

Telah Ditujikan dan Dipertahankan dalam Sidang Ujian Tesis
Program Studi S2 Teknik Informatika
Program Pascasarjana Universitas AMIKOM Yogyakarta
pada hari Senin, 5 September 2022

Pembimbing Utama

Anggota Tim Penguji

Dr. Andi Sunyoto, M.Kom.
NIK. 190302052

Alva Hendi Muhammad, S.T., M.Eng., Ph.D.
NIK. 190302493

Pembimbing Pendamping

Dr. Wing Wahyu Winarno, MAFIS, Ak.
NIK. 555195

Emha Taufiq Luthfi, S.T., M. Kom. **Dr. Andi Sunyoto, M. Kom**
NIK. 190302125 NIK. 190302052

Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer

Yogyakarta, 5 September 2022
Direktur Program Pascasarjana

Dr. Kusriani, M. Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : Enda Putri Atika
NIM : 20.51.1416
Konsentrasi : Business Intelligence

Menyatakan bahwa Tesis dengan judul berikut:
Tuliskan Judul Tesis Bahasa Indonesia

Dosen Pembimbing Utama : Dr. Andi Suryoto, M. Kom.
Dosen Pembimbing Pendamping : Emha Taufiq Luahfi, S.T., M. Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi

Yogyakarta, 5 September 2022
Yang Menyatakan,



Enda Putri Atika

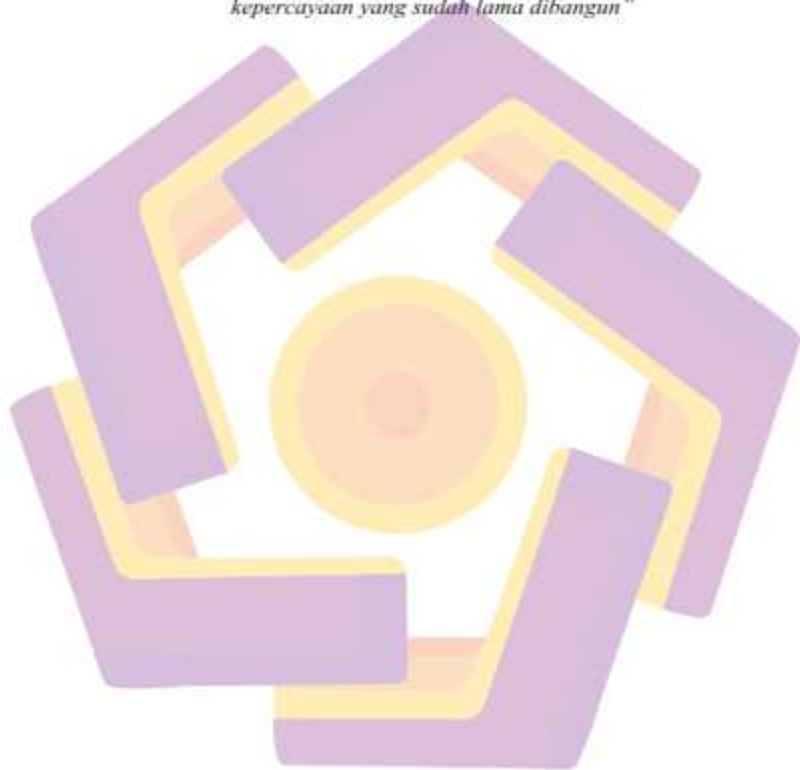
HALAMAN PERSEMBAHAN

Alhamdulillahirabbil Alamin, Puji Syukur kehadiran Allah SWT atas berkat, rahmat dan karunia-Nya serta pertolongannya sehingga saya dapat menyelesaikan tesis ini dengan baik. Saya juga menyampaikan banyak terimakasih terhadap pihak-pihak yang telah memberikan banyak kontribusi baik secara langsung maupun tidak langsung dalam penelitian maupun dalam penyusunan naskah tesis. Tesis ini dipersembahkan kepada:

1. Orang tua tersayang dan adik-adik tersayang yang sudah senantiasa memberikan dukungan, kasih sayang dan doa sehingga kakak dapat menyelesaikan skripsi ini dengan baik.
2. Bapak Dr. Andi Sunyoto, M. Kom. selaku dosen pembimbing yang sudah membimbing saya hingga dapat menyelesaikan tesis ini dan sampai memiliki publikasi ilmiah.
3. Bapak Emha Taufiq Luthfi, S.T., M. Kom. selaku dosen pembimbing kedua yang sudah banyak membantu saya dalam menyelesaikan tesis ini.
4. Bapak dan Ibu Dosen yang telah memberikan ilmunya dari saya semester 1 sampai saat ini.
5. Calon suami Muhammad Naufal Arif yang selalu support dan siap siaga untuk di ajak liburan disaat stress dengan kerjaan dan tesis.
6. Semua pihak keluarga besar universitas AMIKOM Yogyakarta yang tidak bisa saya sebutkan satu persatu.

HALAMAN MOTTO

“Teruslah menjadi orang yang jujur dimanapun, kapapun dan bagaimanapun keadaannya, karena satu kebohongan saja dapat menghancurkan kepercayaan yang sudah lama dibangun”



KATA PENGANTAR

Alhamdulillahirabbil Alamin, Puji syukur kehadiran Allah SWT atas berkat, rahmat dan karuni-Nya serta pertolongannya sehingga saya dapat menyelesaikan tesis dengan judul “Klasifikasi Penyakit Daun Kepala Sawit Menggunakan K-Nearest Neighbor Dan Seleksi Fitur Algoritma Genetika” sebagai syarat untuk menyelesaikan Program Pascasarjana (S2) Universitas AMIKOM Yogyakarta.

Dalam penyusunan tesis ini terdapat banyak hambatan dan rintangan yang dihadapi namun pada akhirnya dapat dilalui karena bantuan dan bimbingan dari berbagai pihak. Pada kesempatan ini saya mengucapkan banyak terimakasih kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Dr. Andi Sunyoto, M. Kom. dan Emha Taufiq Luthfi, S.T., M. Kom. selaku dosen pembimbing yang dengan sabar membimbing dan memberi ilmu kepada saya.
3. Bapak Dr. Kumara Ari Yuana, S.T., M.T selaku dosen mata kuliah soft computing yang telah banyak membantu saya untuk menyelesaikan tesis saya.
4. Seluruh Bapak dan Ibu Dosen Program Pascasarjana Universitas AMIKOM Yogyakarta yang telah memberikan banyak ilmu sehingga dapat menjadi bekal dalam mengerjakan tesis ini.

5. Semua pihak keluarga besar Universitas AMIKOM Yogyakarta yang tidak bisa saya sebutkan satu per satu.
6. Dan terkhusus kepada keluarga saya tercinta, Orang tua dan adik-adik saya serta pasangan saya Muhammad Naufal Arif.

Saya menyadari bahwa naskah tesis ini masih jauh dari kata sempurna, untuk itu kritik dan saran yang membangun sangat diharapkan. Akhir kata semoga tesis ini dapat bermanfaat bagi pembaca dan bagi diri sendiri.

Yogyakarta, 7 September 2022

Penulis



DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PERNYATAAN KEASLIAN TESIS.....	v
HALAMAN PERSEMBAHAN.....	vi
HALAMAN MOTTO.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiv
INTISARI.....	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah.....	3
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian.....	4
1.5. Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1. Tinjauan Pustaka.....	5
2.2. Keaslian Penelitian.....	8

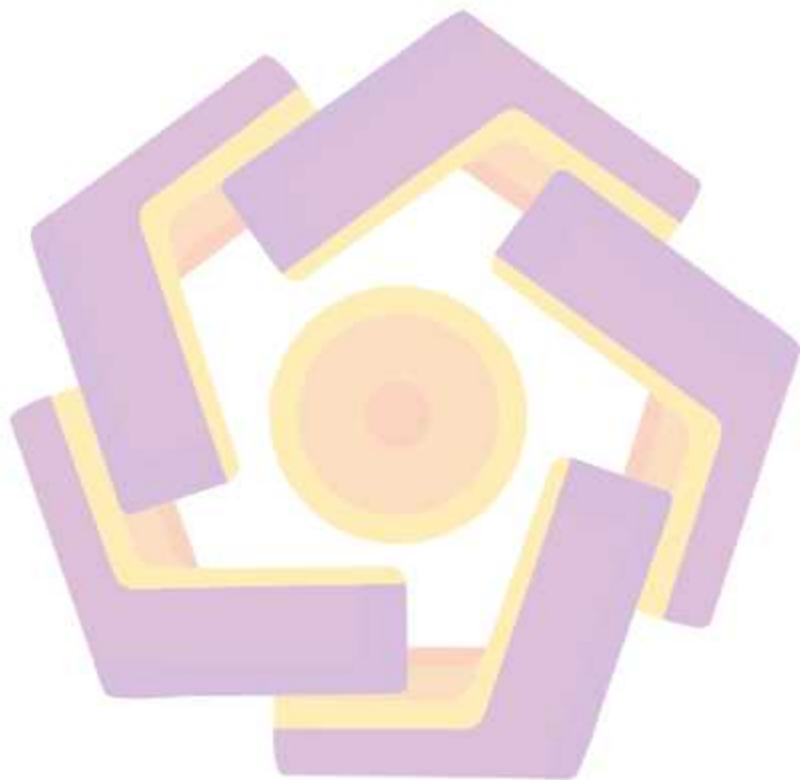
2.3. Landasan Teori.....	13
BAB III METODE PENELITIAN.....	23
3.1. Jenis, Sifat, dan Pendekatan Penelitian.....	23
3.2. Metode Pengumpulan Data.....	23
3.3. Alur Penelitian.....	24
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	31
BAB V PENUTUP.....	64
5.1. Kesimpulan.....	64
5.2. Saran.....	65
DAFTAR PUSTAKA.....	66



DAFTAR TABEL

Tabel 2. 1 Matriks literatur review dan posisi penelitian Klasifikasi Penyakit Daun Kelapa Sawit Menggunakan K-Nearest Neighbor dan Seleksi Fitur Algoritma Genetika	8
Tabel 4. 1 Urutan fitur ekstraksi	38
Tabel 4. 2 Sampel Data	41
Tabel 4. 3 Populasi awal	42
Tabel 4. 4 Sampel nilai pada individu 1	43
Tabel 4. 5 Nilai fitness dari setiap individu	43
Tabel 4. 6 Hasil perhitungan nilai fitness	45
Tabel 4. 7 Hasil perhitungan nilai fitness relatif	46
Tabel 4. 8 Hasil perhitungan nilai fitness kumulatif	47
Tabel 4. 9 Hasil penentuan calon induk crossover	49
Tabel 4. 10 Hasil penentuan induk crossover	50
Tabel 4. 11 Posisi acak dari induk crossover	50
Tabel 4. 12 Sub individu 1	51
Tabel 4. 13 Sub individu 2	51
Tabel 4. 14 Keturunan baru crossover	51
Tabel 4. 15 Hasil pencarian anak mutasi	53
Tabel 4. 16 Populasi baru	55
Tabel 4. 17 Hasil eksperimen ruang warna RGB	58
Tabel 4. 18 Hasil percobaan ruang warna LAB	58
Tabel 4. 19 Hasil percobaan ruang warna HSV	59

Tabel 4. 20 Hasil percobaan ruang warna HSI	60
Tabel 4. 21 Hasil percobaan dengan seluruh ruang warna.....	61
Tabel 4. 22 Perbandingan hasil dengan penelitian sebelumnya.....	62



DAFTAR GAMBAR

Gambar 2. 1 Daun kelapa sawit sehat	14
Gambar 2. 2 Daun kelapa sawit terinfeksi bercak daun curvularia.....	14
Gambar 2. 3 Daun kelapa sawit terinfeksi jamur.....	14
Gambar 2. 4 Komponen Algoritma Genetika.....	16
Gambar 2. 5 Contoh Kromosom.....	17
Gambar 3. 1 Diagram Proses.....	24
Gambar 3. 2 Alur proses algoritma genetika.....	27
Gambar 3. 3 Proses ekstraksi fitur.....	29
Gambar 4. 1 Resize gambar.....	31
Gambar 4. 2 Konversi saluran ruang warna.....	32
Gambar 4. 3 Memberi border pada daun untuk mendapatkan Region Of interest	34
Gambar 4. 4 Hasil forming ROI.....	34
Gambar 4. 5 Visualisasi hasil segmentasi dengan k-means clustering.....	35
Gambar 4. 6 Visualisasi hasil segmentasi warna penyakit pada daun.....	35
Gambar 4. 7 Visualisasi hasil segmentasi warna daun yang sehat.....	35
Gambar 4. 8 Visualisasi hasil segmentasi background.....	36
Gambar 4. 9 Keturunan baru hasil mutasi.....	54

INTISARI

Dunia kecerdasan buatan khususnya machine learning saat ini sedang mengalami perkembangan yang sangat pesat. Klasifikasi di dalam machine learning adalah salah satu bidang kecerdasan buatan yang sangat penting. Fitur – fitur yang digunakan dalam pengenalan pola dan pembelajaran mesin untuk klasifikasi semakin meningkat, sehingga sangat sulit untuk menemukan fitur mana yang paling relevan dalam sebuah kumpulan data yang digunakan. Banyak fitur yang dikumpulkan untuk pengklasifikasian, namun sayangnya setelah diamati kumpulan fitur tersebut tidak semuanya relevan atau bahkan berlebih untuk digunakan dalam klasifikasi. Pada penelitian sebelumnya mengenai deteksi penyakit daun kelapa sawit, tingkat akurasi yang dihasilkan sudah cukup baik, namun pada penelitian tersebut menggunakan sangat banyak fitur, sekitar 41 fitur.

Berdasarkan hal tersebut penelitian ini bertujuan untuk mengembangkan penelitian sebelumnya supaya menghasilkan lebih sedikit fitur yang memiliki karakteristik yang lebih kuat dan relevan serta dapat meningkatkan tingkat akurasi. Pada penelitian ini untuk seleksi fitur dilakukan dengan menggunakan Algoritma Genetika dan klasifikasi penyakit daun kelapa sawit menggunakan KNN. Harapannya dalam penelitian ini Algoritma Genetika dan KNN yang digunakan dapat mengklasifikasikan penyakit daun kelapa sawit dengan menggunakan fitur yang lebih sedikit dan relevan serta menghasilkan tingkat akurasi yang lebih baik atau minimal sama dengan penelitian yang sebelumnya.

Berdasarkan percobaan yang telah dilakukan tingkat akurasi terbaik diperoleh dari skenario kedua sebesar 100%. Pasangan fitur hasil dari seleksi menggunakan algoritma genetika yang memiliki karakteristik kuat yaitu urutan fitur ke -8,13,15,40,51,52,59,61,73,77,82,90,94 urutan fitur tersebut yaitu R8, G5, G7, A8, H3-HSV, H4-HSV, S3-HSV, S5-HSV, H2-HSI, H6-HSI, S3-HSI, I3-HSI, dan I7-HSI. Rata-rata tingkat akurasi yang diperoleh setelah menggunakan fitur hasil seleksi algoritma genetika sebesar 99,2140%. Algoritma genetika mampu mereduksi sebesar 86,46%, dari total 96 fitur menjadi 13 fitur. Rata - rata tingkat akurasi yang diperoleh sebelum menggunakan fitur hasil seleksi algoritma genetika sebesar 94,3130% dan rata-rata tingkat akurasi setelah menggunakan fitur hasil seleksi algoritma genetika sebesar 97,0817%. Terjadi peningkatan akurasi sebesar 2,7687% jika menggunakan fitur hasil seleksi algoritma genetika

Kata kunci: KNN, Algoritma Genetika, Klasifikasi, Daun Kelapa Sawit, Seleksi Fitur

ABSTRACT

The world of artificial intelligence, especially machine learning, is currently experiencing very rapid development. Classification in machine learning is one of the most important fields of artificial intelligence. The features used in pattern recognition and machine learning for classification are increasing, so it is very difficult to find which features are most relevant in a data set being used. Many features are collected for classification, but unfortunately, after observing the feature set, not all of them are relevant or even excessive for use in classification. In previous studies regarding the detection of oil palm leaf disease, the accuracy level produced was quite good, but this study, used a lot of features, about 41 features.

Based on this, this study aims to develop previous research to produce fewer features that have stronger and more relevant characteristics and can increase the level of accuracy. In this study, feature selection was carried out using Genetic Algorithms and oil palm leaf disease classification using KNN. It is hoped that in this study the Genetic Algorithm and KNN used can classify oil palm leaf diseases by using fewer and more relevant features and producing a better level of accuracy or at least the same as the previous research.

Based on the experiments that have been carried out, the best level of accuracy is obtained from the second scenario of 100%. The feature pairs resulting from the selection using a genetic algorithm have strong characteristics, namely the order of features to -8,13,15,40,51,52,59,61,73,77,82,90,94 the order of the features is R8, G5, G7, A8, H3-HSV, H4-HSV, S3-HSV, S5-HSV, H2-HSI, H6-HSI, S3-HSI, I3-HSI, and I7-HSI. The average level of accuracy obtained after using the features of the genetic algorithm selection is 99.2140%. The genetic algorithm can reduce 86.46%, from a total of 96 features to 13 features. The average level of accuracy obtained before using the genetic algorithm selection feature is 94.3130% and the average accuracy level after using the genetic algorithm selection feature is 97.0817%. There is an increase in accuracy of 2.7687% when using the genetic algorithm selection result

Keyword: KNN, Genetic Algorithm, Classification, Palm Leaf, Feature Selection

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Saat ini dunia kecerdasan buatan khususnya machine learning mengalami perkembangan yang sangat pesat. Klasifikasi di dalam machine learning adalah salah satu masalah yang sangat penting. Fitur-fitur yang digunakan dalam pengenalan pola dan pembelajaran mesin untuk klasifikasi semakin meningkat, sehingga sangat sulit untuk menemukan fitur mana yang paling relevan dalam sebuah kumpulan data yang digunakan. Banyak fitur yang dikumpulkan untuk pengklasifikasian, namun sayangnya setelah diamati kumpulan fitur tersebut tidak semuanya relevan atau bahkan berlebih untuk digunakan dalam klasifikasi. Mengurangi fitur-fitur yang berlebihan dan tidak relevan dapat meningkatkan akurasi dan mengurangi overfitting (Xue et al., 2021). Oleh sebab itu, perlu dilakukan pemilihan fitur yang paling relevan supaya mendapatkan dataset yang paling baik.

Akhir-akhir ini, komputasi evolusioner dengan kemampuannya yang dapat melakukan pencarian yang kuat di semua ruang pencarian membuat teknik tersebut menarik banyak perhatian. Banyak yang telah menerapkan teknik tersebut di berbagai bidang di dunia nyata, seperti pada prognosis kanker paru-paru (Maleki et al., 2021), pemilihan sensor pada industri manufaktur (Ahn & Hur, 2020), dan intrusion detection system (IDS)(Halim et al., 2021). Beberapa penelitian tersebut menerapkan teknik komputasi evolusioner dalam memecahkan masalah pemilihan

fitur. Salah satu teknik komputasi evolusioner yang paling populer adalah Algoritma Genetika. Algoritma Genetika dapat diimplementasikan dalam masalah pemilihan fitur tanpa menyetel nilai ambang batas yang mengakibatkan ruang pencarian akan lebih kecil dibandingkan dengan PSO (Particle Swarm Optimization) (Xue et al., 2021).

Pada penelitian sebelumnya (Hamdani et al., 2021) mengenai deteksi penyakit daun kelapa sawit, tingkat akurasi yang dihasilkan sudah cukup baik, namun pada penelitian tersebut menggunakan sangat banyak fitur, sekitar 41 fitur. Adapun fitur-fitur yang digunakan yaitu RGB: R(R1, R3, R4, R6,R7, R8), G (G1, G2, G4, G5, G6, G7, G8), and B (B1, B5, B6, B7, B8), L*A*B: A (A2, A3, A4, A5, A7, A8) and B (B7, B8), HSI: H (H1-HSI, H5- HSI, H6-HSI, H7-HSI) and S (S6-HSI, S7-HSI, S8-HSI). HSV: H (H1-HSV, H3-HSV, H5-HSV, H6-HSV, H7-HSV) and S (S1-HSV, S7-HSV, S8-HSV).

Menurut jurnal (Alelyani, 2021; Wang et al., 2020) fitur yang banyak dan tidak relevan dapat menyebabkan kinerja pembelajaran mesin menjadi lebih rendah dan menghasilkan biaya serta waktu komputasi yang lebih besar. Berdasarkan hal tersebut penelitian ini bertujuan untuk mengembangkan penelitian sebelumnya (Hamdani et al., 2021) supaya menghasilkan lebih sedikit fitur yang memiliki karakteristik yang lebih kuat dan relevan serta dapat meningkatkan tingkat akurasi, karena menurut (Maleki et al., 2021) mereduksi fitur yang banyak dapat meningkatkan nilai akurasi dan mempercepat waktu komputasi.

Pada penelitian ini untuk seleksi fitur dilakukan dengan menggunakan Algoritma Genetika dan klasifikasi penyakit daun kelapa sawit menggunakan

KNN, karena berdasarkan penelitian sebelumnya (Maleki et al., 2021), menggunakan Algoritma Genetika dan KNN untuk prognosis kanker paru-paru dapat menghasilkan akurasi yang sangat tinggi sekitar 95-100%. Harapannya dalam penelitian ini Algoritma Genetika dan KNN yang digunakan dapat mengklasifikasikan penyakit daun kelapa sawit dengan menggunakan fitur yang lebih sedikit dan relevan serta menghasilkan tingkat akurasi yang lebih baik atau minimal sama dengan penelitian yang sebelumnya (Hamdani et al., 2021).

1.2. Rumusan Masalah

Adapun yang menjadi rumusan masalah adalah sebagai berikut:

- a. Fitur apa saja yang memiliki karakteristik kuat yang dihasilkan oleh seleksi fitur dari Algoritma Genetika?
- b. Berapa persen rata-rata tingkat akurasi klasifikasi yang dihasilkan setelah menggunakan fitur hasil seleksi dari algoritma genetika?
- c. Berapa persen rata-rata peningkatan akurasi sesudah menggunakan fitur hasil seleksi dari Algoritma genetika?

1.3. Batasan Masalah

Adapun batasan masalah pada penelitian ini sebagai berikut:

- a. Dataset yang digunakan terdapat 3 kelas yaitu: daun kelapa sawit sehat, terinfeksi bercak daun *curvularia* dan terinfeksi jamur.
- b. Data set yang digunakan sebanyak 1146 file gambar dengan format .jpg.
- c. Data akan diolah dengan bantuan Matlab versi R2019b.
- d. Output penelitian ini ada 3, yaitu: pasangan fitur yang dihasilkan dari seleksi algoritma genetika, persentase peningkatan akurasi klasifikasi setelah

menggunakan algoritma genetika, dan rata-rata akurasi setelah menggunakan fitur hasil seleksi algoritma genetika untuk klasifikasi.

1.4. Tujuan Penelitian

Beberapa tujuan pada penelitian ini adalah sebagai berikut:

- a. Menemukan fitur yang memiliki karakteristik yang kuat yang dapat digunakan untuk mengoptimalkan hasil klasifikasi jenis penyakit daun kelapa sawit.
- b. Untuk mengetahui tingkat akurasi yang dihasilkan dengan menggunakan fitur hasil seleksi Algoritma Genetika.
- c. Untuk mengetahui persentase peningkatan akurasi dari sebelum dan sesudah menggunakan fitur seleksi dari Algoritma Genetika.

1.5. Manfaat Penelitian

Adapun manfaat penelitian adalah sebagai berikut:

- a. Kedepannya model ini bisa dijadikan sebuah aplikasi berbasis mobile yang dapat mendeteksi jenis penyakit pada daun kelapa sawit sehingga dengan aplikasi tersebut dapat memudahkan petani dalam mendeteksi jenis penyakit pada daun kelapa sawit.
- b. Pada peneliti yang lain jurnal ini bisa dijadikan referensi jika ingin melakukan penelitian yang sejenis di bidang seleksi fitur menggunakan teknik komputasi evolusioner khususnya Algoritma Genetika.

BAB II

TINJAUAN PUSTAKA

2.1. Tinjauan Pustaka

Beberapa literatur review yang digunakan dalam tinjauan pustaka ini memiliki permasalahan yang sama yaitu mengenai pencarian solusi masalah optimasi. Masalah yang ada pada jurnal sistem deteksi intrusi yaitu dimensi data yang besar dengan fitur-fitur yang semakin meningkat namun tidak relevan membuat sistem deteksi intrusi kesulitan dalam mendeteksi serangan jaringan. Penggunaan Algoritma Genetika dalam jurnal tersebut bertujuan untuk meminimalkan jumlah fitur dengan mengurangi fitur-fitur yang tidak relevan dan terduplikasi. Pengurangan jumlah fitur tersebut ternyata memberikan hasil yang sangat baik dengan dapat menghasilkan tingkat akurasi mencapai 99.80% (Halim et al., 2021).

Penggunaan Algoritma Genetika pada bidang kesehatan untuk seleksi fitur dalam prognosis kanker paru-paru memberikan hasil yang sangat maksimal dengan akurasi mencapai 100% dalam waktu eksekusi tercepat 0.0156s. Hasil tersebut jauh lebih baik jika dibandingkan dengan penggunaan Algoritma KNN saja tanpa Algoritma Genetika yang hanya mencapai tingkat akurasi sebesar 96-99% dengan waktu eksekusi 0.0312-0.0313s (Maleki et al., 2021).

Pemilihan fitur yang mempertimbangkan kinerja, waktu eksekusi dan waktu mulai klasifikasi yang dimodelkan secara sistematis dengan menggunakan Algoritma Genetika memberikan cara yang efektif dan efisien untuk mencapai

kinerja terbaik dengan memilih jumlah sensor yang minimum yang dapat mengurangi biaya manajemen dan waktu produksi dalam bidang industri manufaktur (Ahn & Hur, 2020).

Dalam pencarian solusi baru yang menggunakan Algoritma Genetika, crossover adalah salah satu operator yang sangat penting. Operator crossover memiliki pengaruh yang kuat pada hasil optimasi. Penggunaan multi-objective binary Genetic Algorithm integrating an adaptive operator selection (MOBGA-AOS) dengan lima operator crossover yang memiliki karakteristik yang berbeda mampu menghapus sebagian besar fitur dan memastikan kesalahan terhadap klasifikasi kecil. Tidak hanya itu MOBGA-AOS juga mampu memecahkan masalah pemilihan fitur di dalam dimensi data yang lebih tinggi (Xue et al., 2021).

Algoritma Genetika juga dapat digunakan untuk pemilihan fitur dalam mendeteksi malware pada android. Fitur yang didapat dari algoritma genetika digunakan untuk klasifikasi pembelajaran mesin. Berdasarkan eksperimen dapat divalidasi bahwa algoritma genetika memberikan fitur yang paling optimal dan dapat mengurangi setengah dari fitur aslinya dan hasil akurasi yang diperoleh lebih dari 94%. Hasil tersebut lebih baik jika dibandingkan dengan sebelum menggunakan Algoritma Genetika (Fatima et al., 2019).

Penelitian (Amiri et al., 2019), membedakan bingkai darah dengan menggunakan hasil ekstraksi dari saluran warna dalam ruang warna yang berbeda. Hasil ekstraksi tersebut menghasilkan 96 fitur dari 12 saluran warna. Jumlah fitur tersebut dikurangi dengan menggunakan Algoritma Genetika kemudian hasil fitur seleksinya diklasifikasikan menggunakan jaringan syaraf MLP untuk menentukan

normal atau abnormal. Berdasarkan penelitian yang dilakukan Algoritma genetika mampu mengurangi jumlah fitur secara signifikan dan dapat meningkatkan akurasi klasifikasi.

Berdasarkan literasi beberapa jurnal di atas, Algoritma Genetika sangat powerful dalam melakukan seleksi fitur yang dapat diterapkan diberbagai bidang. Algoritma tersebut dapat mengurangi jumlah fitur yang sangat signifikan dan menggunakan fitur yang dihasilkan dapat mempercepat waktu komputasi. Oleh sebab itu, dalam penelitian ini juga akan menggunakan Algoritma Genetika untuk mengurangi fitur yang ada pada jurnal sebelumnya (Hamdani et al., 2021). Pada penelitian sebelumnya (Hamdani et al., 2021), tingkat akurasi yang dihasilkan sudah sangat baik. Maka dari itu, pengurangan seleksi fitur di sini dimaksud untuk mendapatkan hasil akurasi yang lebih baik atau minimal sama baiknya dengan yang sebelumnya namun dengan menggunakan fitur yang lebih sedikit dan relevan supaya dapat mengurangi waktu komputasi serta mengurangi ruang penyimpanan yang lebih besar.

2.2. Keaslian Penelitian

Tabel 2. 1 Matriks literatur review dan posisi penelitian
Klasifikasi Penyakit Daun Kelapa Sawit Menggunakan K-Nearest Neighbor dan Seleksi Fitur Algoritma Genetika

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
1	Detection of oil palm leaf disease based on color Histogram and supervised classifier	Hamdani Septiarni, Anindita Sunyoto, Andi Suyanto, Suyanto Utamingrum, Fitri, Optik, 2021	Mendeteksi penyakit daun kelapa sawit untuk membedakan 2 kelas daun yang sehat dan terinfeksi dengan menggunakan ruang warna RGB, LAB, HSI dan HSV dengan membagi Histogram masing-masing warna menjadi 8 bin.	Hasil klasifikasi menunjukkan bahwa metode yang diusulkan menghasilkan kinerja yang baik yang ditunjukkan dengan nilai sensitivitas, spesifisitas, dan akurasi yang tinggi masing-masing mencapai 99,3%, 100%, dan 99,67%.	Penelitian ini berpotensi untuk dikembangkan lebih lanjut lagi untuk mencari lebih sedikit fitur dengan karakteristik yang lebih kuat dan diskriminatif dengan komputasi waktu yang lebih sedikit.	Pada jurnal tersebut algoritma seleksi yang digunakan untuk menyeleksi fitur yaitu PCA, sedangkan pada penelitian yang akan dilakukan untuk seleksi fitur menggunakan GA.
2	Efficient genetic algorithm for feature selection for early time series classification	Ahn, Gilseung Hur, Sun, Computers and Industrial Engineering, 2020	Memaksimalkan kinerja klasifikasi dan meminimalkan waktu mulai dan waktu eksekusi klasifikasi	Algoritma Genetika yang diusulkan dapat memberikan cara yang efektif dan efisien untuk memilih jumlah minimum sensor yang mencapai kinerja terbaik dengan waktu mulai paling awal dari klasifikasi.	Melakukan penelitian model dan algoritma untuk deret waktu yang multivariat	Perbedaan penelitian sebelumnya dengan yang akan dilakukan yaitu pada studi kasus atau objek penelitiannya. Pada penelitian sebelumnya penerapan GA digunakan untuk peralatan sensor di bidang manufaktur industri, untuk menemukan cara yang

Tabel 2.1 Lanjutan

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
						efektif dan efisien dalam memilih jumlah sensor yang minimum namun dapat mencapai kinerja terbaik, sedangkan pada penelitian yang akan dilakukan penerapan GA digunakan untuk pencarian fitur terbaik yang lebih sedikit dari sebelumnya namun dengan hasil akurasi yang sama dan waktu komputasi yang lebih cepat.
3	An effective genetic algorithm-based feature selection method for intrusion detection systems	Halim, Zahid Yousaf, Muhammad Nadeem Waqas, Muhammad Sulaiman, Muhammad Abbas, Ghulam Hussain, Masroor Ahmad, Iftekhar Hanif, Muhammad, Computers and Security, 2021	Merancang metode yang mempertahankan Sebagian besar informasi unik yang terkait dengan data dengan jumlah fitur minimum. Dapat mengatasi masalah pemilihan fitur dalam domain keamanan jaringan dan deteksi intrusi, karya ini memberikan kontribusi metode	Algoritma yang diusulkan memberikan hasil yang baik untuk studi kasus yang digunakan setelah dilakukan evaluasi dengan 3 dataset benchmark dan 4 metode perbandingan.	Penelitian selanjutnya dapat memanfaatkan metode pembelajaran tanpa pengawasan, seperti pengelompokan untuk membuat mesin mempelajari sendiri jenis serangan baru.	Perbedaan pada penelitian (Halim et al., 2021) dengan penelitian yang akan dilakukan yaitu terletak pada objek penelitiannya. Pada penelitian (Halim et al., 2021), GA digunakan untuk mengatasi masalah pemilihan fitur dalam keamanan jaringan dan mendeteksi intrusi. Sedangkan pada penelitian yang akan dilakukan penerapan GA digunakan untuk pencarian fitur terbaik yang lebih sedikit dari sebelumnya

Tabel 2.1 Lanjutan

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			pemilihan fitur berbasis Algoritma Genetika (GA), disebut sebagai GA-based Feature Selection (GbFS), untuk meningkatkan akurasi pengklasifikasi.			namun dengan hasil akurasi yang sama dan waktu komputasi yang lebih cepat.
4	Spectroscopic detection of rice leaf blast infection from asymptomatic to mild stages with integrated machine learning and feature selection	Tian, Long Xue, Bowen Wang, Ziyi Li, Dong Yao, Xia Cao, Qiang Zhu, Yan Cao, Weixing Cheng, Tao	Tujuan dari penelitian(Tian et al., 2021) adalah untuk menguji kelayakan dalam mendeteksi infeksi RLB(Rice Leaf Blast) dari spektrum reflektansi daun pada tahap perkembangan penyakit asimtomatik, awal dan ringan.	Hasil kesimpulan pada penelitian (Tian et al., 2021) menunjukkan bahwa spektroskopi reflektansi memiliki potensi besar dalam deteksi pra-visual infeksi RLB dan spektroskopi pencitraan udara dalam pemetaan kejadian awal dan tingkat keparahan infeksi RLB pada skala besar.	Penelitian ini masih dapat dikembangkan lagi dengan memfokuskan pada penilaian kontribusi pada masing-masing DSSF dalam mendeteksi infeksi pada RLB, atau dapat mengembangkan strategi yang lebih sederhana lagi seperti indeks spektral spesifik untuk mendeteksi tahapan infeksi yang berbeda. Penelitian deteksi penyakit masih memiliki potensi yang besar untuk dikembangkan lagi	Penelitian (Tian et al., 2021), dalam melakukan seleksi fiturnya menggunakan 3 metode yaitu Reflectance bands, Spectral Indices dan Wavelet Coefficients sedangkan untuk klasifikasinya menggunakan SVM-SFFS, K-NN-SFFS dan LDA-SFFS.

Tabel 2.1 Lanjutan

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
					dalam pemetaan kejadian di awal dan pada tingkat keparahan RLB ataupun penyakit pada tanaman serealia lainnya dalam skala besar.	
5	A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection	Maleki, Negar Zcinali, Yasser Niaki, Seyed Taghi Akhavan, Expert Systems with Applications, 2021	Menerapkan teknik k-Nearest-Neighbors dan algoritma genetika untuk pemilihan fitur yang efisien untuk mengurangi dimensi kumpulan data dan meningkatkan kecepatan pengklasifikasi yang digunakan untuk mendiagnosis stadium penyakit pada pasien.	Algoritma KNN di hibridisasi dengan algoritma genetik seleksi fitur, akurasi klasifikasi meningkat secara signifikan. Menggunakan Algoritma genetic kemudian algoritma K-NN menghasilkan tingkat akurasi 100% dalam waktu eksekusi 0.0156 detik(hanya algoritma machine learning)	Untuk peneliti selanjutnya mungkin dapat menggunakan algoritma klasifikasi pembelajaran mesin lainnya atau menggunakan fitur berbasis populasi lainnya.	Perbedaan penelitian (Maleki et al., 2021) dengan penelitian yang akan dilakukan yaitu terletak pada algoritma yang digunakan dan objek penelitiannya. Penelitian (Maleki et al., 2021) menggunakan algoritma K-NN untuk klasifikasi dan GA untuk pencarian fitur pada object penelitian kanker paru-paru. Sedangkan pada penelitian yang akan dilakukan menggunakan A-NN untuk klasifikasi dan GA untuk pencarian fitur pada objek penelitian penyakit daun kelapa sawit.

Tabel 2.1 Lanjutan

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
6	Adaptive crossover operator based multi-objective binary genetic algorithm for feature selection in classification [Formula presented]	Xue, Yu Zhu, Haokai, Liang, Jiayu Slowik, Adam, Knowledge-Based Systems, 2021	Memilih crossover yang paling efisien untuk menyelesaikan masalah pemilihan fitur yang berbeda pada algoritma genetika, terutama ketika sifat masalah pemilihan fitur tidak diketahui sebelumnya.	Hasil eksperimen mengungkapkan bahwa MOBGA-AOS mampu menghapus sejumlah besar fitur sambil memastikan kesalahan klasifikasi kecil. Selain itu, ia memperoleh keuntungan menonjol pada kumpulan data skala besar, yang menunjukkan bahwa MOBGA-AOS kompeten untuk memecahkan masalah pemilihan fitur dimensi tinggi.	Pada penelitian selanjutnya dapat menambahkan operator mutasi yang berbeda untuk membentuk kumpulan operator mutasi sehingga kombinasi yang tepat dari operator crossover dan mutasi dapat dipilih pada tahap evolusi yang berbeda untuk lebih meningkatkan kinerja algoritma.	Pada penelitian (Xue et al., 2021), fokus utama penelitian yaitu terletak pada pemilihan operator crossover untuk seleksi fitur dengan GA agar dapat mengurangi Sebagian besar fitur pada kumpulan data yang skala besar. Perbedaannya terhadap penelitian yang akan dilakukan yaitu terletak pada objek penelitiannya.

2.3. Landasan Teori

2.3.1. Dataset Daun Kelapa Sawit

Dataset yang digunakan untuk kelas daun yang terinfeksi penyakit jamur dikumpulkan langsung dari perkebunan kelapa sawit di Desa Perkebunan Bandar Selamat, Sumatera utara, sedangkan untuk dataset daun yang sehat dan daun yang terinfeksi bercak daun *curvularia* menggunakan data sekunder yang digunakan oleh penelitian sebelumnya (Hamdani et al., 2021). Pada penelitian sebelumnya (Hamdani et al., 2021), mengenai deteksi penyakit daun kelapa sawit, dataset yang digunakan merupakan data primer yang di kumpulkan dari perkebunan kelapa sawit di Muara Wahau yang ada di Kalimantan Timur.

Dataset yang digunakan dibagi menjadi 3 kelas: daun kelapa sawit sehat, terinfeksi bercak daun *curvularia* dan terinfeksi jamur. Pada dasarnya ada banyak jenis penyakit daun kelapa sawit, namun fokus penelitian ini adalah penyakit bercak daun *curvularia* dan jamur pada daun, karena penyakit itu paling sering muncul pada daun yang masih muda maupun sudah tua.

Daun kelapa sawit yang sehat berwarna hijau tanpa terdapat bercak apapun pada permukaan daunnya seperti pada Gambar 2.1. daun kelapa sawit yang terinfeksi bercak daun *curvularia* terdapat bercak daun berwarna kuning kecoklatan pada permukaan daunnya seperti yang terlihat pada Gambar 2.2. Sedangkan daun kelapa sawit yang terinfeksi jamur permukaan daunnya terdapat banyak bercak berwarna putih seperti Gambar 2.3.



Gambar 2. 1 Daun kelapa sawit sehat



Gambar 2. 2 Daun kelapa sawit terinfeksi bercak daun curvularia



Gambar 2. 3 Daun kelapa sawit terinfeksi jamur

2.3.2. Klasifikasi KNN

KNN (K-Nearest Neighbor) adalah salah satu Algoritma klasifikasi yang sangat terkenal dan banyak digunakan (Maleki et al., 2021) .

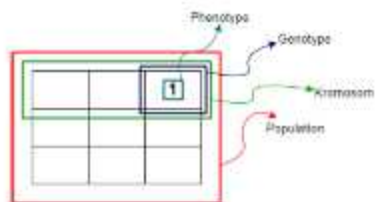
Sebuah objek dapat diklasifikasikan dengan menghitung jarak ketetanggaannya dengan menggunakan nilai K. Dimana K tersebut merupakan bilangan bulat yang bernilai positif. Jika nilai K=3, artinya objek akan masuk ke kelas yang berjarak tiga ke tetangga terdekatnya (Vaishnave & Devi, n.d.). Berdasarkan jurnal (Keramat-Jahromi et al., 2021) perhitungan KNN dengan menggunakan Euclidean dapat dilihat pada rumus (1):

$$d_0 = \sqrt{\sum_{i=1}^n (X_i - Z_i)^2} \quad (1)$$

2.3.3. Algoritma Genetika

Salah satu algoritma pada metode pencarian heuristik adalah Algoritma Genetika. Algoritma ini banyak digunakan untuk pencarian solusi pada masalah optimasi (Maleki et al., 2021). Dalam implementasinya untuk seleksi fitur, algoritma ini mampu mengurangi lebih dari setengah fitur aslinya dan dapat memberikan tingkat akurasi lebih baik dari sebelumnya dalam waktu komputasi yang lebih cepat (Amiri et al., 2019; Fatima et al., 2019).

Pada Algoritma Genetika terdapat beberapa istilah yang digunakan seperti yang ada pada Gambar 2.4. Ada populasi, kromosom, genotype dan phenotype. Populasi adalah kumpulan dari kromosom atau individu yang akan dijadikan solusi. Kromosom adalah kumpulan dari gen atau genotype, sedangkan genotype adalah elemen yang terdapat dalam kromosom, dan phenotype adalah nilai dari gen atau genotype.



Gambar 2. 4 Komponen Algoritma Genetika
Berdasarkan jurnal (Maleki et al., 2021), Algoritma Genetika memiliki 5 fase utama, yaitu:

1. Populasi Awal
2. Fungsi Fitnes
3. Seleksi
4. Crossover
5. Mutase

2.3.3.1. Populasi Awal

Dalam pengimplementasiannya, langkah pertama yang harus dilakukan dalam Algoritma Genetika adalah menggenerasi populasi awal. Generasi populasi awal dilakukan secara acak. Gen tersebut dijadikan sebagai fitur dari atribut yang ada pada data asli. Kromosom yang ada dalam populasi dapat memiliki satu solusi untuk masalah seleksi. Sebuah data yang disimpan dalam bentuk array 1 dimensi dengan sel N disebut dengan kromosom (Halim et al., 2021), seperti yang ditunjukkan oleh Gambar 2.4.



Gambar 2. 5 Contoh Kromosom

2.3.3.2. Fungsi Fitness

Fungsi fitness memfasilitasi insstrument supaya dapat mengetahui beberapa fitur yang mempunyai kesamaan rendah diantara mereka. Hal tersebut memungkinkan fungsi fitness untuk mengumpulkan fitur-fitur yang dapat mewakili sebagian besar informasi dengan keragaman yang lebih tinggi yang ada dalam dataset (Halim et al., 2021). Kromosom yang memiliki nilai kebugaran tinggi mempunyai peluang unggul untuk membawa individu baru ke dalam populasi (Maleki et al., 2021), formula untuk menghitung nilai fitness menurut jurnal (Zhang & Lin, 2019) dapat dilihat pada rumus (2):

$$Fitness = \frac{k - Rank + 1}{k} \quad (2)$$

2.3.3.3. Seleksi

Seleksi dilakukan untuk mencari orang tua agar menghasilkan keturunan. Salah satu strategi yang bisa digunakan untuk seleksi ini adalah roda roulette. Strategi ini saat diimplementasikan secara paralel dapat membuat waktu eksekusi menjadi berkurang dan tidak membutuhkan penyortiran seperti metode lainnya.

Seleksi untuk pemilihan orang tua berbanding lurus dengan hasil fitness yang didapat dari fungsi fitness. Peluang terpilih akan semakin tinggi jika memiliki nilai fitness yang tinggi. Pemilihan dalam roda roulette menggunakan prinsip pencarian linier dimana slot roda akan di timbang. Pada strategi roda roulette, kromosom yang memiliki nilai fitness yang lebih tinggi mencakup lebih banyak area. Sehingga yang dipilih memiliki probabilitas yang lebih tinggi (Halim et al., 2021)

2.3.3.4. Crossover dan Mutasi

Membawa kromosom baru ke dalam populasi secara turun temurun disebut dengan operasi crossover dan mutasi (Maleki et al., 2021). Membentuk dua kromosom baru untuk dijadikan keturunan dari dua individu yang telah dipilih sebagai orang tua dengan menukar bagian genom mereka adalah operasi dari crossover (Maleki et al., 2021).

Operasi mutasi sendiri dilakukan untuk mencegah penggabungan sebelum waktunya ke nilai terbaik terdekat dengan mengamati secara acak fokus baru di ruang pencarian, hal itu dilakukan dengan membalik bit tanpa pola tertentu yang memiliki beberapa kemungkinan (Maleki et al., 2021).

2.3.4. RGB

RGB adalah sebuah ruang warna yang memiliki 3 saluran warna, yaitu: Red, Green dan Blue. Ketiga warna tersebut merupakan golongan warna-warna dasar atau primer yang ditemukan oleh Young-Helmholtz (Utama & Light, 2016). Jika ketiga warna tersebut digabungkan maka akan menghasilkan warna-warna yang baru. Nilai intensitas untuk masing-masing saluran warna RGB yaitu 256.

Ruang warna RGB adalah model ruang warna yang paling banyak digunakan dalam merepresentasikan dan menampilkan gambar (Khekade, 2015). Prinsip model ruang warna RGB pada dasarnya sama dengan mata manusia yang sensitif terhadap cahaya yang tampak (Utama & Light, 2016). Salah satu keunggulan yang ada pada ruang warna RGB adalah dapat digunakan untuk area-area sensitif terhadap warna, karena ruang warna RGB dapat memetakan banyak warna (Utama & Light, 2016).

2.3.5. HSI

Ruang warna HIS terdiri dari Hue, Saturation dan Intensitas warna (Chen et al., 2021). Secara analogi HIS memiliki koordinat sumbu ruang warna yang berbentuk silinder. Dalam ruang warna HIS, satu warna dapat diwakilkan sebagai titik yang memiliki tiga koordinat yaitu hue, saturasi dan intensitas. Hue adalah jenis warna yang memiliki sudut tegak lurus terhadap sumbu vertikal di dalam

sebuah koordinat kutub. Saturasi bertindak sebagai wakil dari radial jarak ke sumbu dengan skala abu-abu vertikal. Sedangkan intensitas merupakan sebuah jarak ke dasar sepanjang sumbu vertikal (Chen et al., 2021).

Hue memiliki sudut yang berada pada rentang nilai kisaran $0-360^\circ$, 0° adalah warna merah, $120-240^\circ$ warna hijau dan biru, dan diakhiri 360° warna merah. Nilai H berfokus pada Panjang gelombang spektral warna. Nilai S berfungsi untuk mengukur tingkat penipisan cahaya murni oleh cahaya putih yang berada dalam rentang nilai kisaran $0-1$, semakin tinggi nilai S menunjukkan warna lebih kaya dan mencolok. I berada dalam rentang nilai kisaran $0-1$. Nilai 0 digunakan untuk bagian bawah dan nilai 1 digunakan untuk bagian atas sumbu vertikal untuk menunjukkan kecerahan relatif. Nilai H tidak akan terdefinisi jika S bernilai 0 dan S tidak akan terdefinisi jika I bernilai 0 (Chen et al., 2021).

Ruang warna His lebih unggul jika dibandingkan dengan ruang warna RGB karena dapat menunjukkan secara langsung kecerahan dan chrominance. Formula yang digunakan jurnal (Hamdani et al., 2021), untuk perhitungan HIS dapat dilihat pada rumus (7), (8), (9),(10).

$$H = \begin{cases} \theta, & \text{jika } B \leq G \\ 360 - \theta, & \text{jika } B > G \end{cases} \quad (3)$$

Dimana,

$$\theta = \text{Cos}^{-1} \left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{[x(R-G)^2+(R-B)(G-B)]^{1/2}} \right\} \quad (4)$$

$$S = 1 - \frac{[\min(R,G,B)]}{I} \quad (5)$$

$$I = \frac{1}{3}(R + G + B) \quad (6)$$

2.3.6. HSV

Fitur HSV memberikan invariant dan penerangan yang terjadi karena adanya perbedaan kondisi pada pencahayaan. HSV terdiri dari Hue, Saturation dan Value (Ali et al., 2017). Hue merupakan jenis warna, saturasi mewakili jumlah warna dan value adalah jumlah cahaya (Kartika & Herumurti, 2017).

- H adalah sejenis warna yang memiliki rentang derajat kisaran 0-360°
- S adalah kejemihan warna yang memiliki rentang kisaran 0% - 100%.
- V adalah kecerahan warna yang memiliki rentang kisaran 0% - 100%.

Rumus (3),(4),(5),(6) adalah formula yang digunakan jumul (Hamdani et al., 2021) untuk menghitung HSV.

$$H = \{\theta, B \leq G; 360 - \theta, B > G\} \quad (7)$$

Dimana,

$$\theta = \text{Cos}^{-1} \left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{[x(R-G)^2+(R-B)(G-B)]^{1/2}} \right\} \quad (8)$$

$$S = \begin{cases} 0, \max(R, G, B) = 0; \\ 1 - \frac{\min(R, G, B)}{\max(R, G, B)}, \text{ jika tidak} \end{cases} \quad (9)$$

$$V = \max f_0(R, G, B) \quad (10)$$

2.3.7. LAB

Ruang warna LAB atau biasa dikenal dengan sebutan CIELAB merupakan ruang warna yang sangat lengkap yang ditetapkan oleh Komisi Internasional tentang iluminasi warna (French Commission Internationale de l'eclairage dikenal dengan CIE), ruang warna LAB memiliki rentang warna yang lebar dan dapat mengekspresikan keseluruhan warna yang bisa dirasakan oleh mata manusia dan mampu mengatasi masalah distribusi warna yang ada dalam ruang warna RGB, LAB juga sangat sering digunakan sebagai referensi ruang warna (Dai et al., 2018; Rulaningtyas et al., 2015). Menurut jurnal (Hamdani et al., 2021)h, formula yang dapat digunakan untuk menghitung ruang warna LAB seperti yang ditunjukkan oleh rumus (11), (12), dan (13).

$$L = 0.2126R + 0.7152G + 0.0722 \quad (11)$$

$$a = 1.4749 (0.2213R - 0.339G + 0.1177B) + 128 \quad (12)$$

$$b = 0.6245 (0.1949R + 0.6057G - 0.8006B) + 128 \quad (13)$$

BAB III

METODE PENELITIAN

3.1. Jenis, Sifat, dan Pendekatan Penelitian

Penelitian yang dilakukan menggunakan pendekatan kuantitatif. Penelitian ini termasuk ke dalam jenis penelitian eksperimen. Menurut Sudaryono (2017:85) penelitian eksperimen adalah metode yang dapat menguji hipotesis mengenai hubungan sebab akibat.

3.2. Metode Pengumpulan Data

Sumber data yang digunakan dalam penelitian ini ada 2. Sumber data yang pertama yaitu data sekunder diantaranya dataset kelas daun kelapa sawit yang sehat dan kelas daun kelapa sawit yang terinfeksi bercak daun *curvularia*. Dataset tersebut adalah data yang sama yang digunakan oleh penelitian sebelumnya (Hamdani et al., 2021). Sedangkan untuk sumber data yang kedua yaitu data primer, meliputi kelas daun kelapa sawit yang terinfeksi jamur.

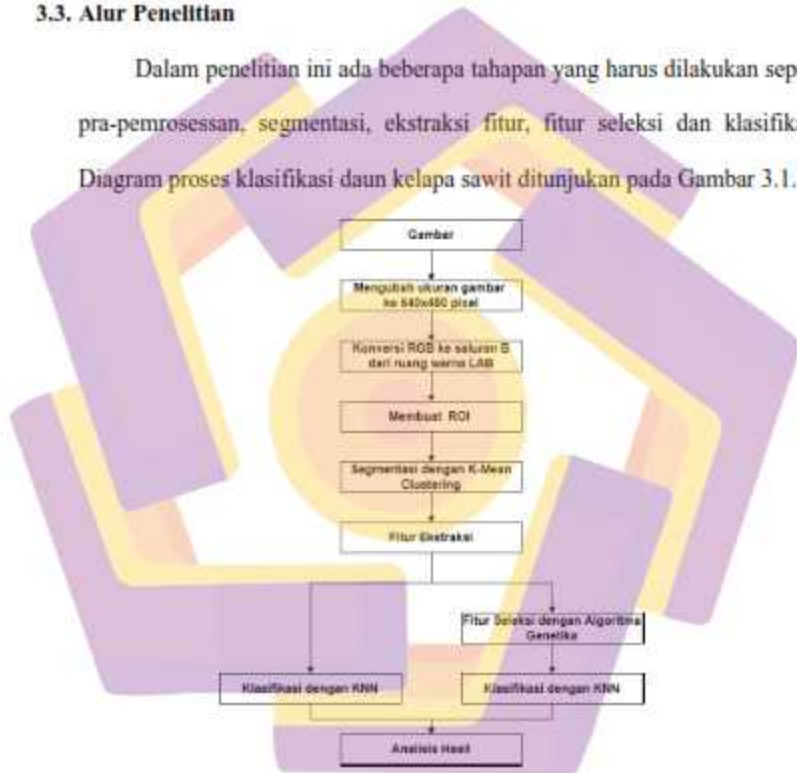
Dataset untuk kelas daun penyakit yang terinfeksi jamur diambil dengan menggunakan kamera belakang handphone realme 3 pro yang memiliki resolusi 16 MP sensor sony IMX 519 + 5 MP untuk depth sensor. Teknik pengambilan gambar mengikuti jurnal sebelumnya (Hamdani et al., 2021), supaya gambar yang dihasilkan nantinya tidak terdapat perbedaan yang sangat signifikan.

Objek daun yang akan di foto diletakan di atas kertas berwarna putih untuk backgroundnya. objek daun tersebut di foto dengan jarak +- 20 cm dari

kamera. Satu gambar daun di foto sebanyak 5 kali dari berbagai sudut yang berbeda, untuk melatih sistem dalam mengenali sebuah objek dari berbagai sudut pandang bukan hanya dari satu sisi saja.

3.3. Alur Penelitian

Dalam penelitian ini ada beberapa tahapan yang harus dilakukan seperti pra-pemrosesan, segmentasi, ekstraksi fitur, fitur seleksi dan klasifikasi. Diagram proses klasifikasi daun kelapa sawit ditunjukkan pada Gambar 3.1.



Gambar 3. 1 Diagram Proses

3.3.1. Dataset

Dataset primer yang digunakan dalam penelitian ini diambil dari perkebunan kelapa sawit di wilayah Perkebunan Bandar Selamat, Sumatera Utara, Indonesia. Dataset tersebut berupa gambar

daun kelapa sawit dengan format .jpg sebanyak 1146 gambar yang dibagi ke dalam 3 kelas, yaitu: kelas daun kelapa sawit sehat sebanyak 382 gambar, kelas daun kelapa sawit yang terinfeksi jamur sebanyak 382 gambar dan kelas daun kelapa sawit terinfeksi bercak daun *curvularia* sebanyak 382. Data kelas pertama adalah daun kelapa sawit sehat, data kelas kedua adalah daun kelapa sawit terinfeksi jamur dan data kelas ketiga adalah daun kelapa sawit terinfeksi bercak daun *curvularia*.

3.3.2. Pra-Pemrosesan

Pra-pemrosesan yang dilakukan dalam penelitian ini yaitu *resize* gambar dari yang ukurannya bervariasi di ubah menjadi 640 x 480 pixel (Hamdani et al., 2021) dan *conversi* warna dari RGB ke LAB. Namun pada ruang warna LAB hanya saluran warna B saja yang digunakan karena pada saluran B objek daun dan latar belakang sangat jelas dibedakan dibandingkan dengan saluran lain, *konversi* ruang warna juga dilakukan dengan tujuan untuk mengurangi waktu *komputasi* (Hamdani et al., 2021).

3.3.3. Segmentasi

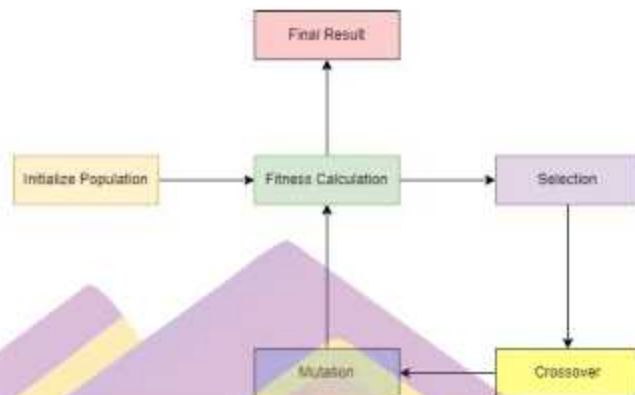
Pada penelitian ini *segmentasi* dilakukan dengan tujuan untuk membentuk ROI (Region of Interest) atau biasa disebut dengan sub-citra. Untuk mendapatkan ROI pada gambar dilakukan *pengaturan* area daun sebagai dasar penentuan batas citra ROI.

Segmentasi dilakukan pada gambar yang sudah dikonversi ke saluran B dalam ruang warna LAB.

Proses segmentasi untuk membedakan objek daun dengan latar belakangnya menggunakan algoritma K-Means Clustering. Jumlah K pada K-Means clustering ditentukan berdasarkan jumlah kelas yang ada. Pada penelitian ini ada 3 kelas yang digunakan karena pada gambar terbagi menjadi 3 area, yang pertama area daun yang masih sehat atau berwarna hijau, yang kedua area daun yang terinfeksi penyakit dan yang ketiga bentuk katar belakang daun secara keseluruhan. Thresholding yang digunakan dalam penelitian ini yaitu dilatasi atau erosi. Thresholding tersebut digunakan untuk mengatasi masalah yang ada pada klasifikasi.

3.3.4. Ekstraksi Fitur

Warna dapat mendeskripsikan dengan baik untuk mengetahui sebuah penampakan penyakit pada daun yang diteliti dari citra karena adanya perbedaan warna antara daun yang sehat dan daun yang terinfeksi penyakit. Pada penelitian (Hamdani et al., 2021), warna mampu untuk mendeteksi penyakit pada daun.



Gambar 3. 2 Alur proses algoritma genetika

Proses ekstraksi fitur yang dilakukan pada penelitian seperti yang terdapat pada Gambar 3.2, langkah pertama yang dilakukan adalah menginisialisasikan atau menggenerate populasi awal yang akan di jadikan calon individu dengan menggunakan nilai yang random. Setelah itu dilakukan perhitungan nilai fitness, kromosom atau individu yang menghasilkan nilai paling besar akan dijadikan calon individu.

Setelah memiliki nilai fitness langkah selanjutnya yang harus dilakukan yaitu menyeleksi individu-individu tersebut dengan metode seleksi roda roulette. Setelah individu tersebut diseleksi langkah selanjutnya dilakukan operasi crossover dan mutasi untuk mendapatkan calon individu. Individu yang terpilih nantinya yang memiliki nilai evaluasi paling kecil.

Ekstraksi fitur yang dilakukan pada penelitian ini menggunakan 4 ruang warna, yaitu: RGB, LAB, HSV dan HSI. Sebelum dilakukan ekstraksi fitur, perlu konversi warna terlebih dahulu dari ruang warna RGB ke tiga ruang warna yang digunakan, seperti RGB ke HSV, RGB ke LAB, RGB ke HSI.

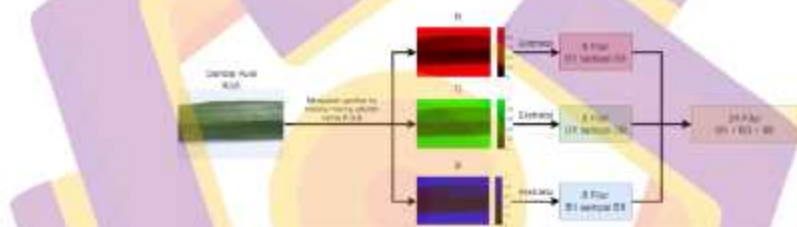
Ekstraksi yang dilakukan menghasilkan 12 Histogram warna. Masing-masing saluran warna terdiri dari 256 nilai intensitas. Untuk meminimalisir jumlah fitur yang terlalu banyak, dilakukan pemecahan Histogram ke dalam 8 bin karena berdasarkan percobaan yang telah dilakukan pada penelitian (Hamdani et al., 2021), pemecahan Histogram ke 8 bin memberikan hasil yang paling baik. Histogram 8 bin dapat menghasilkan sekitar 96 fitur.

Pada Gambar 3.2 proses ekstraksi fitur menggunakan ruang warna RGB sehingga gambar awal yang digunakan adalah gambar awal dari RGB, kemudian gambar awal tersebut diubah ke masing-masing saluran warna pada ruang warna RGB yaitu: Red, Green dan Blue, setelah itu masing-masing saluran warna tersebut diekstrak.

Satu saluran warna R pada dasarnya memiliki 256 nilai intensitas yang artinya satu saluran warna memiliki jumlah fitur sebanyak 256 saat di ekstrak, karena jumlah fiturnya sangat banyak maka pada penelitian ini 256 nilai atau fitur tersebut di split ke 8 binary, hanya urutan 1 sampai 8 saja yang digunakan oleh sebab itu

pada penelitian ini masing-masing saluran warna yang sudah di ekstrak memiliki 8 nilai intensitas atau 8 fitur.

Seperti yang terdapat pada Gambar 3.2, saluran warna R menghasilkan 8 fitur yaitu R1 sampai R8, saluran warna G menghasilkan 8 fitur yaitu G1 sampai G8 dan saluran warna B menghasilkan 8 fitur yaitu: B1 sampai B8 sehingga jika dijumlahkan, satu ruang warna RGB menghasilkan total fitur sebanyak 24 fitur.



Gambar 3. 3 Proses ekstraksi fitur

Pada ruang warna yang lainnya seperti HSV, HIS dan LAB, proses ekstraksi yang dilakukan sama seperti yang ada pada Gambar 3.2, jika yang di ekstraksi ruang warna HSV maka gambar awal yang digunakan adalah gambar awal HSV, begitu juga dengan ruang warna HIS dan LAB.

Setelah itu gambar awal HSV tadi di ubah ke masing – masing saluran warna HSV seperti hue, saturation dan value, begitu juga untuk ruang warna HIS dan LAB gambar awal diubah ke masing-masing saluran warna pada ruang warna tersebut. Kemudian

masing – masing saluran warna tersebut di ekstrak dan setiap saluran warna yang diekstrak menghasilkan 8 fitur, karena pada penelitian ini menggunakan 12 saluran warna dari 4 ruang warna yang ada, maka jika dijumlahkan, total keseluruhan fitur yang digunakan ada 96 fitur.

3.3.5. Seleksi Fitur

Algoritma yang digunakan dalam pemilihan fitur adalah algoritma genetika. Total fitur yang akan diseleksi sebanyak 96 fitur, seleksi fitur dilakukan untuk menemukan fitur warna terbaik yang memiliki peranan penting untuk dapat membedakan daun yang sehat dengan daun yang terinfeksi penyakit.

3.3.6. Klasifikasi

Pada penelitian ini klasifikasi dilakukan dengan menggunakan algoritma KNN (K-Nearest Neighbors). Algoritma ini termasuk ke dalam jenis algoritma supervised learning yang sangat terkenal dan banyak digunakan karena memberikan hasil yang sangat baik untuk masalah klasifikasi. Berdasarkan jurnal (Maleki et al., 2021), kombinasi algoritma KNN untuk klasifikasi dan Algoritma Genetika untuk seleksi fitur memberikan hasil yang sangat baik. Dalam prinsip kerjanya algoritma KNN menggunakan jarak terdekat data dengan k-tetangga untuk mengklasifikasikan data.

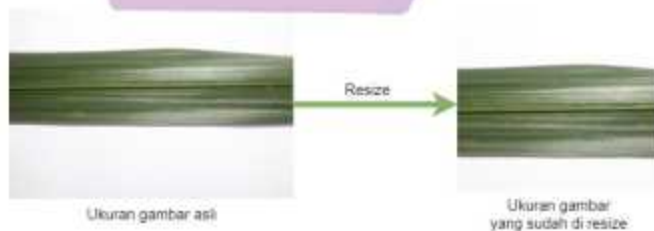
BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

4.1. Pra-Pemrosesan

4.1.1 Mengubah ukuran gambar

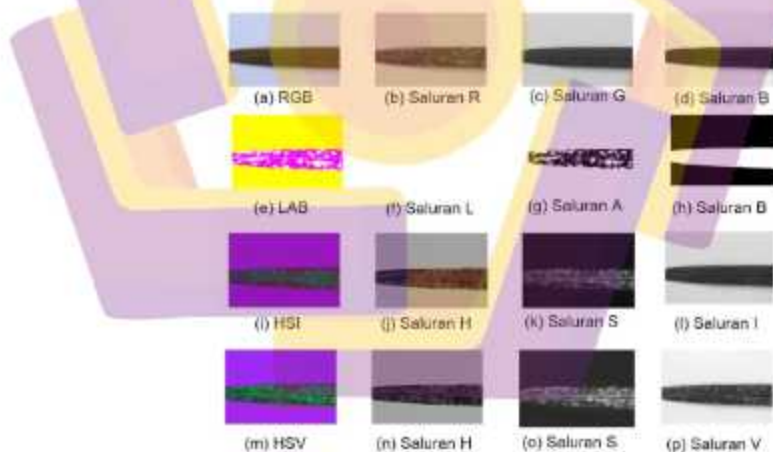
Gambar pada dataset yang digunakan memiliki ukuran yang sangat bervariasi, maka dari itu sebelum dilakukan pemrosesan lebih lanjut perlu dilakukan resize gambar supaya gambar memiliki ukuran yang sama. Gambar akan di resize ke ukuran 640 x 480 pixel seperti pada jurnal (Hamdani et al., 2021). Ukuran gambar asli dapat dilihat seperti yang di tunjukan oleh Gambar 4.1. Gambar yang telah di resize akan memiliki ukuran yang lebih kecil dibandingkan dengan gambar aslinya yang lebih ukurannya lebih lebar seperti yang di tunjukan oleh Gambar. 4.1. Resize gambar ini juga dilakukan supaya tidak memerlukan waktu komputasi yang sangat lama dan ruang memori yang besar karena disebabkan oleh ukuran gambar yang besar. Berdasarkan eksperimen yang telah dilakukan semakin besar gambar yang digunakan untuk diproses maka akan semakin besar memori dan waktu komputasi yang di butuhkan.



Gambar 4. 1 Resize gambar

4.1.2 Konversi RGB ke saluran B dari ruang warna LAB

Konversi warna dari RGB ke LAB dilakukan untuk membedakan luas area daun dengan area latar belakang sebelum dilakukannya proses segmentasi dengan K-Means Clustering. Pada tahap ini dilakukan percobaan visualisasi objek daun dengan menggunakan seluruh saluran ruang warna yang digunakan. Tahap pertama konversi gambar asli ke ruang warna RGB, kemudian gambar asli dikonversi ke ruang warna LAB, tahap ke tiga gambar asli dikonversi ke HSI dan tahap ke empat gambar asli dikonversi ke HSV. Masing-masing saluran ruang warna memberikan hasil yang berbeda-beda seperti yang ditunjukkan oleh Gambar. 4.2.



Gambar 4. 2 Konversi saluran ruang warna

Dari hasil eksperimen yang dilakukan keempat ruang warna yang digunakan seperti yang dapat dilihat pada Gambar. 4.2, saluran b dari ruang warna LAB yang sangat baik dalam membedakan

antara luas area daun dengan latar belakang dibandingkan dengan ruang warna yang lain karena pada saluran b dari ruang warna LAB antara daun dan latar belakang sangat terlihat jelas perbedaannya, dimana objek yang membentuk daun akan berwarna putih dan yang latar belakang berwarna hitam tanpa ada warna lain untuk bercak-bercak daunnya, sedangkan untuk ruang warna RGB, HSV dan HSI masing-masing saluran masih terdapat warna bercak-bercak bagian penyakitnya dan tidak terlihat jelas perbedaan antara area dan dan latar belakang karena memiliki warna yang tidak terlalu kontras perbedaannya.

4.2. Membentuk ROI dan segmentasi dengan K-Means Clustering

Sebelum segmentasi dilakukan, objek gambar terlebih dahulu dilakukan forming untuk mengambil bagian ROI, supaya gambar yang akan diproses bisa lebih fokus ke objek daunnya. Objek gambar terlebih dahulu diberi border atau garis pembatas di sekitar tepi daunnya seperti Gambar 4.3, kemudian objek akan dipotong fokus ke objek daunnya sesuai border yang telah dibuat, dan hasilnya dapat dilihat pada Gambar 4.4. Gambar yang dihasilkan akan lebih fokus ke objek daunnya dan mengurangi sebagian besar dari latar belakangnya.



Gambar 4. 3 Memberi border pada daun untuk mendapatkan Region Of interest

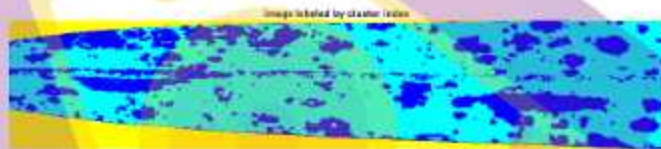


Gambar 4. 4 Hasil forming ROI

Segmentasi yang dilakukan pada penelitian ini menggunakan algoritma *K-Means Clustering* dalam mengklaster warna yang ada pada daun kelapa sawit. Banyaknya kelas yang digunakan untuk segmentasi gambar ada 3, yaitu: kelas warna penyakit pada daun, kelas warna daun yang sehat dan latar belakang. Hasil dari proses segmentasi yang telah dilakukan terdapat 3 kelas warna, yaitu warna biru tua untuk menandai bagian warna penyakit pada daun, warna biru muda untuk menandai warna daun yang sehat, dan warna kuning untuk menandai latar belakang seperti yang ditunjukkan pada Gambar 4.5.

Setelah objek di segmentasi dengan *k-means clustering* seperti Gambar 4.5, selanjutnya objek akan dikembalikan ke warna aslinya yaitu RGB. Hasil

segmentasi untuk kelas warna penyakit pada daun hanya terlihat bercak-bercak coklat penyakitnya saja seperti yang ditunjukkan oleh Gambar 4.6. Sedangkan untuk kelas warna daun yang sehat maka hanya akan terlihat bagian-bagian daun yang masih berwarna hijau seperti yang ditunjukkan oleh Gambar 4.7, dan untuk kelas warna latar belakang yang terlihat hanya warna putih saja sebagai backgroundnya karena pada objek gambarnya background yang digunakan berwarna putih, bagian hitam pada gambar adalah bagian area daunnya yang dapat dilihat pada Gambar 4.8. Dari hasil segmentasi ini terlihat sangat jelas masing-masing bagian daun, baik yang sehat penyakit ataupun latar belakang.



Gambar 4. 5 Visualisasi hasil segmentasi dengan k-means clustering



Gambar 4. 6 Visualisasi hasil segmentasi warna penyakit pada daun



Gambar 4. 7 Visualisasi hasil segmentasi warna daun yang sehat



Gambar 4. 8 Visualisasi hasil segmentasi background

4.3. Fitur Ekstraksi

Ekstraksi fitur dilakukan dengan menggunakan keempat ruang warna yaitu RGB, HSI, HSV dan LAB. Masing – masing ruang warna akan diektrak dan di split atau dipecah menjadi 8 bin dari 256 bin yang ada supaya tidak menghasilkan terlalu banyak fitur karena berdasarkan jurnal (Hamdani et al., 2021), fitur yang di split menjadi 8 bin memberikan hasil yang terbaik.

Sampel fitur hasil ekstraksi dari masing-masing ruang warna dapat dilihat pada Tabel 4.1. Pada tabel tersebut terdapat total 96 fitur beserta urutannya dari satu data gambar, urutan 1 sampai 24 fitur dari ruang warna RGB yang diberi label (R1, R2, R3, R4, R5, R6, R7, R8, G1, G2, G3, G4, G5, G6, G7, G8, B1, B2, B3, B4, B5, B6, B7 dan B8), 25 sampai 48 fitur dari ruang warna LAB yang diberi label (L1, L2, L3, L4, L5, L6, L7, L8 L9, A1, A2, A3, A4, A5, A6, A7, A8, B1, B2, B3, B4, B5, B6, B7 dan B8), 49 sampai 72 fitur dari ruang warna HSV yang di beri label (H1, H2, H3, H4, H5, H6, H7, H8, S1, S2, S3, S4, S5, S6, S7, S8, V1, V2, V3, V4, V5, V6, V7 , V8), dan 73 sampai 96 fitur dari ruang warna HSI yang diberi label (H1, H2, H3, H4, H5, H6, H7, H8, S1, S2, S3, S4, S5, S6, S7, S8, I1, I2, I3, I4, I5, I6, I7 dan I8) . Masing-masing ruang warna, menghasilkan 24 fitur yang berbeda-beda. Pada penelitian ini jumlah dataset yang digunakan sebanyak 1146 gambar oleh sebab jika seluruh fitur

dari masing-masing ruang warna digabungkan maka akan terbentuk matriks 1146×96 .

Proses ekstraksi yang dilakukan di masing-masing ruang warna sama seperti yang ada pada Gambar 3.2. Gambar awal diubah terlebih dahulu ke masing-masing saluran warna kemudian setiap saluran warna tersebut di ekstrak dan menghasilkan 8 fitur yang nantinya akan digunakan pada proses seleksi fitur dan klasifikasi.



Tabel 4. 1 Urutan fitur ekstraksi

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
R1	R2	R3	R4	R5	R6	R7	R8	G1	G2	G3	G4	G5	G6	G7	G8	B1	B2	B3	B4	B5	B6	B7	B8
21021	72	57	27	0	0	0	0	21021	72	57	27	0	0	0	0	21021	72	57	27	0	0	0	0
25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
L1	L2	L3	L4	L5	L6	L7	L8	A1	A2	A3	A4	A5	A6	A7	A8	B1	B2	B3	B4	B5	B6	B7	B8
20915	2	47	0	0	1	5	207	21110	0	10	28	9	7	1	12	21121	33	0	9	9	3	0	2
49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
H1	H2	H3	H4	H5	H6	H7	H8	S1	S2	S3	S4	S5	S6	S7	S8	V1	V2	V3	V4	V5	V6	V7	V8
21116	20	3	0	4	4	30	0	21157	10	0	0	1	0	0	9	21021	72	57	27	0	0	0	0
73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96
H1	H2	H3	H4	H5	H6	H7	H8	S1	S2	S3	S4	S5	S6	S7	S8	I1	I2	I3	I4	I5	I6	I7	I8
21116	20	3	0	4	4	30	0	21164	3	1	0	0	0	0	9	21021	72	57	27	0	0	0	0

4.4. Fitur Seleksi

Algoritma yang digunakan dalam seleksi fitur yaitu Algoritma Genetika. Total fitur yang akan di seleksi sebanyak 24 dan 96 fitur, untuk skenario pertama total fitur yang akan diseleksi yaitu 24 fitur dari masing-masing ruang warna, fitur-fitur tersebut akan di seleksi untuk di cari fitur terbaiknya yang relevan dan memiliki karakteristik kuat kemudian hasil fitur yang di seleksi akan diklasifikasikan, sedangkan untuk skenario kedua total fitur yang akan diseleksi yaitu 96 fitur, fitur – fitur tersebut kemudian diseleksi dengan Algoritma Genetika untuk mendapatkan fitur terbaik.

Seleksi fitur bertujuan untuk mengurangi waktu komputasi, biaya serta ruang penyimpanan karena semakin besar jumlah fitur yang digunakan akan membutuhkan waktu komputasi yang lebih lama, biaya yang lebih mahal dan ruang penyimpanan yang lebih besar. Berikut ini adalah pseudocode Algoritma Genetika untuk mendapatkan pasangan fitur terbaik.

START

Membentuk kromosom

Mendefinisikan fungsi fitness

Membentuk populasi awal (n kromosom)

Menentukan nilai crossover, mutasi dan banyaknya generasi

For $i=1$ to banyaknya generasi

 For $j = 1$ to n

 Seleksi

Crossover

Mutasi

End for

Gabung semua kromosom

Pilih n kromosom terbaik (populasi baru)

End for

STOP

Berikut ini adalah contoh perhitungan Algoritma Genetika dengan menggunakan sampel data yang ada pada penelitian.

Parameter-parameter yang digunakan yaitu:

Nilai Mutasi = 0.01

Nilai Crossover = 0.6

Jumlah Populasi Awal = 4

Maksimum Generasi = 1

Seleksi = Roda roulette

Crossover = PMX

Mutasi = Reciprocal exchange

Tabel 4. 2 Sampel Data

Urutan Fitur	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Data 1	1	456 59	0	0	323 22	657 8	136 51	7 5 5	0	29	274 89	0	192 17	147 42	463 68	145 3	44 10	0	471 96	0	325 69	0	423	403 6
Data 2	58 12	666 7	454 94	722 00	128 25	22	0	0	0	235 56	607 2	721 72	0	411 87	33	983 43	18 95 7	0	163 91	931 3	0	0	16	0
Data 3	18	104 5	0	392 4	0	269 11	119 58	0	94 95	493 05	351 86	544	0	0	516 9	63	35 21 6	0	401 44	175 00	137 44	440 89	0	188 5
Data 4	27 2	435 2	0	0	0	488 30	158 92	2 0	0	468 72	142 2	181 37	280 43	2	512 1	0	78 69	0	804	318 43	455 67	0	135 30	302 69

- Langkah pertama yang dilakukan adalah membuat populasi awal untuk calon solusi

Jumlah individu yang digunakan untuk populasi awal sebanyak 4 individu. Setiap individu atau kromosom terdiri dari 24 gen.

Nilai setiap gen untuk populasi awal di inisialisasi oleh bilangan bulat dari 1 sampai 24. Untuk membangun 1 set kromosom diisi dengan bilangan acak dari 1 sampai 24, seperti yang ada pada Tabel 4.3.

Tabel 4. 3 Populasi awal

Kromosom	Inisialisasi populasi																							
Individu 1	7	4	8	9	3	5	2	6	1	15	13	17	11	18	12	14	10	16	19	24	20	23	22	21
Individu 2	1	5	3	2	4	6	9	7	8	10	13	11	15	12	14	18	19	16	17	20	23	22	21	24
Individu 3	7	6	9	10	8	3	2	1	5	4	20	22	24	23	21	15	13	16	12	11	18	19	17	14
Individu 4	22	21	24	23	16	19	18	15	17	12	14	11	13	7	10	9	5	8	6	3	4	1	2	20

- Langkah kedua yaitu mencari nilai evaluasi dengan menggunakan rumus (14).

$$fk(x) = \text{Nilai gen 1} + \text{nilai gen 2} + \dots + \text{nilai gen}_n \quad (14)$$

- Contoh perhitungan nilai evaluasi untuk individu 1:

Nilai untuk setiap individu diperoleh dari sampel data yang ada pada Tabel 4.2, sampel dataset yang digunakan dalam penelitian ini menggunakan ruang warna RGB dari data ke 1 sampai data ke 4. Pada Tabel 4.4, pemberian nilai sesuai dengan urutan yang ada pada Tabel 4.2, contohnya pada individu 1 kromosom pertama urutan ke 7 maka nilai yang diisi adalah nilai urutan ke 7.

Tabel 4. 4 Sampel nilai pada individu 1

Individu 1	7	4	8	9	3	5	2	6	1	15	13	17	11	18	12	14	10	16	19	24	20	23	22	21
Nilai	1	45659	0	0	32322	6578	13651	755	0	29	27489	0	19217	14742	46368	1453	4410	0	47196	0	32569	0	423	4036

Nilai Evaluasi Individu 1

$$= 1 + 45659 + 0 + 0 + 32322 + 6578 + 13651 + 755 + 0 + 29 + 27489 + 19217 + 14742 + 46368 + 1453 + 4410 + 0 + 47196 + 0 + 32569 + 0 + 423 + 4036 = 296898$$

Perhitungan dengan cara yang sama dilakukan pada individu ke 2 dan seterusnya, hasil perhitungan diperoleh seperti yang tertera pada Tabel 4.5.

Tabel 4. 5 Nilai fitness dari setiap individu

Kromosom	Nilai Evaluasi
Individu 1	296898
Individu 2	429060
Individu 3	296196
Individu 4	298845

- Langkah ketiga yang harus dilakukan yaitu mencari nilai fitness menggunakan rumus (15).

$$\text{Nilai fitness} = \sum_{k=1}^{10} fk(x) - fk(x) \quad (15)$$

Contoh perhitungan nilai fitness untuk individu 1 adalah sebagai berikut:

- Menjumlahkan nilai evaluasi dari individu 1 sampai individu 4

$$\text{Total nilai evaluasi} = 296898 + 429060 + 296196 + 298845 = 1320999$$

- Mencari nilai fitness untuk setiap individu seperti di bawah ini menggunakan rumus (15).

$$\text{Nilai Fitness Individu 1} = \text{Total nilai evaluasi} - \text{nilai evaluasi individu}$$

$$\text{Nilai Fitness Individu 1} = 1320999 - 296898 = 1024101$$

Untuk perhitungan nilai fitness pada individu 2 dan seterusnya menggunakan cara yang sama seperti mencari nilai fitness individu 1. Hasil perhitungan nilai fitness dari masing-masing individu dapat dilihat pada Tabel 4.6.

Tabel 4. 6 Hasil perhitungan nilai fitness

Kromosom	Nilai Fitness
Individu 1	1024101
Individu 2	891939
Individu 3	1024803
Individu 4	1022154
Total Fitness	3962997

- Langkah keempat yaitu mencari nilai fitness relatif dan nilai fitness kumulatif.

Mencari nilai fitness relatif dapat menggunakan rumus (16) dan rumus (17).

$$Fitness\ Relatif = \frac{Nilai\ Evaluasi}{Total\ Fitness} \quad (16)$$

$$Fitness\ Kumulatif = \text{jumlah fitness relatif dari individu sebelumnya} \quad (17)$$

Contoh perhitungan fitness relatif untuk individu 1 sebagai berikut:

$$Fitness\ relatif\ individu\ 1 = \frac{296898}{1024101} = 0.25841579$$

Untuk menghitung fitness relatif individu 2 dan seterusnya menggunakan cara yang sama seperti mencari nilai fitness individu

1. Hasil perhitungan nilai fitness relatif semua individu dapat dilihat pada Tabel 4.7.

Tabel 4. 7 Hasil perhitungan nilai fitness relatif

Kromosom	Nilai Fitness Relatif
Individu 1	0.25841579
Individu 2	0.225066787
Individu 3	0.258592929
Individu 4	0.257924495

Contoh perhitungan fitness kumulatif adalah sebagai berikut:

Fitness kumulatif individu 1 = 0.25841579

Khusus nilai fitness kumulatif individu 1 sama dengan fitness relatif individu satu, karena belum ada individu lain selain individu 1 untuk di jumlahkan.

Fitness kumulatif individu 2= nilai fitness relatif individu 1 + nilai fitness relatif individu 2

Fitness kumulatif individu 2 = 0.25841579 + 0.225066787 = 0.483482576

Fitness kumulatif individu 3= nilai fitness relatif individu 1 + nilai fitness relatif individu 2 + nilai fitness individu 3

Fitness kumulatif individu 4 = nilai fitness relatif individu 1 + nilai fitness relatif individu 2 + nilai fitness individu 3 + nilai fitness individu 4

Fitness kumulatif individu 5 = nilai fitness relatif individu 1 + nilai fitness relatif individu 2 + nilai fitness individu 3 + nilai fitness individu 4 + nilai fitness individu 5

Hasil perhitungan nilai fitness Kumulatif dapat dilihat pada Tabel 4.8.

Tabel 4. 8 Hasil perhitungan nilai fitness kumulatif

Kromosom	Nilai Fitness Kumulatif
Individu 1	0.25841579
Individu 2	0.483482576
Individu 3	0.742075505
Individu 4	1

- Langkah kelima yaitu melakukan operasi crossover.

Individu yang dipilih untuk calon induk crossover yang nilai bilangan acaknya kurang dari nilai fitness kumulatifnya. Bilangan nilai acak ditentukan secara random dari range nilai 0 sampai 1.

Contohnya:

- Bilangan acak pertama nilainya lebih besar dari nilai fitness kumulatif individu 1 maka individu 1 tidak bisa di pilih jadi calon induk crossover. Bilangan acak pertama juga nilainya lebih besar dari nilai fitness kumulatif individu 2, maka individu 2 juga tidak bisa di pilih jadi calon induk. Tetapi bilangan acak pertama kurang dari individu 3, maka individu 3 dapat di jadikan calon induk crossover yang pertama.
- Bilangan acak kedua nilainya lebih kecil dari nilai fitness kumulatif individu ke 2, maka individu 2 dapat di jadikan calon induk crossover yang kedua.
- Bilangan acak ketiga nilainya lebih kecil dari nilai fitness kumulatif ke 3, maka individu 3 dapat di jadikan calon induk yang ketiga
- Bilangan acak keempat nilainya lebih kecil dari nilai fitness kumulatif keempat, maka individu 4 dapat di jadikan calon induk crossover yang keempat.

Hasil penentuan calon induk crossover seperti yang sudah dijelaskan di atas dapat dilihat pada Tabel 4.9.

Tabel 4. 9 Hasil penentuan calon induk crossover

Kromosom	Nilai Fitnes Kumulatif	Bilangan acak	Calon Induk Crossover
Individu 1	0.25841579	0.529002	Individu 3
Individu 2	0.483482576	0.333667	Individu 2
Individu 3	0.742075505	0.639029	Individu 3
Individu 4	1	0.855140	Individu 4

- Langkah keenam yaitu menentukan induk crossover.

Parameter nilai crossover yang digunakan pada contoh perhitungan ini adalah 0,6, maka individu yang dapat dijadikan induk dari crossover adalah individu yang memiliki nilai acak < 0.6 . Hasil penentuan induk crossover individu yang dapat dijadikan induk adalah individu 3 dan individu 4 karena individu tersebut memiliki bilangan acak < 0.6 , seperti yang dapat dilihat pada Tabel 4.10.

Tabel 4. 10 Hasil penentuan induk crossover

Calon Induk Crossover	Bilangan acak	Induk Crossover
Individu 3	0.387645	Individu 3
Individu 2	0.987654	
Individu 3	0.743684	
Individu 4	0.294708	Individu 4

Langkah ketujuh yaitu melakukan operasi crossover dari 2 individu yang telah dipilih pada langkah sebelumnya, yaitu individu 3 dan individu 4. Selanjutnya untuk melanjutkan melakukan operasi crossover supaya mendapatkan keturunan baru, memilih 2 posisi secara acak, pada perhitungan ini posisi yang di pilih yaitu 12 dan 22 untuk menentukan sud individu 1 dan sub individu 2, seperti yang dapat dilihat pada Tabel 4.11.

Tabel 4. 11 Posisi acak dari induk crossover

Induk Crossover	Set Kromosom																							
	Individu 3	7	6	9	10	8	3	2	1	5	4	20	22	24	23	21	15	13	16	12	11	18	19	17
Individu 4	22	21	24	23	16	19	18	15	17	12	14	11	13	7	10	9	5	8	6	3	4	1	2	20

Tabel 4. 12 Sub individu 1

Set kromosom sub individu 1														
22	24	23	21	15	13	16	12	11	18	19				

Tabel 4. 13 Sub individu 2

Set kromosom individu 2										
11	13	7	10	9	5	8	6	3	4	1

Sub individu 1 dan sub individu 2 ditentukan dengan mengambil data dari posisi acak pertama ke posisi acak kedua, hasil sub individu 1 dan sub individu 2 dapat dilihat pada Tabel 12 dan Tabel 13. Selanjutnya sub individu tersebut di silangkan, untuk menghasilkan keturunan pertama crossover, sub individu 2 digabungkan ke individu 3. Dan untuk menghasilkan keturunan kedua crossover sub individu ke 1 di gabungkan dengan individu 4. Hasil dari operasi crossover yang telah dilakukan menghasilkan 2 keturunan baru dari crossover, anak 1 dan anak 2, seperti yang dapat dilihat pada Tabel 14.

Tabel 4. 14 Keturunan baru crossover

Keturunan Baru Crossover	Pasangan kromosom baru																							
	Anak 1	7	6	9	10	8	3	2	1	5	4	20	11	13	7	10	9	5	8	6	3	4	1	17
Anak 2	22	21	24	23	16	19	18	15	17	12	14	22	24	23	21	15	13	16	12	11	18	19	2	20

- Langkah kedelapan melakukan proses mutasi untuk mencari anak mutasi.

Pencarian anak mutasi dilakukan dengan cara, menentukan bilangan acak pertama untuk mencari calon induk mutasi, dan bilangan acak kedua untuk menentukan induk mutasi. Bilangan acak pertama dan bilangan acak kedua di ambil dari nilai antara 0-1.

- Individu yang dapat dijadikan calon induk mutasi adalah yang bilangan acaknya tidak lebih dari nilai fitness kumulatif individunya.

Contoh:

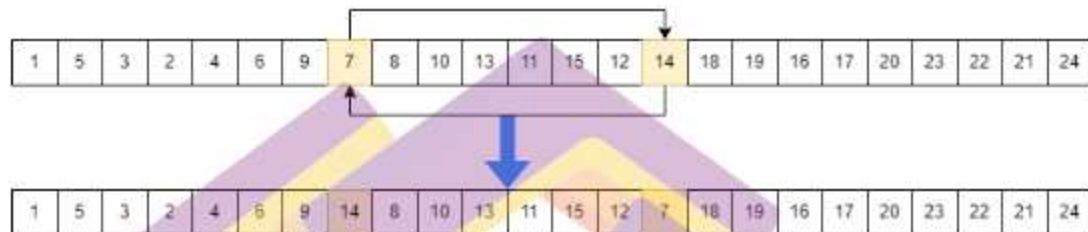
- Bilangan acak pertama $>$ individu 1, maka individu 1 tidak dapat di jadikan calon induk yang pertama.
Bilangan acak pertama $>$ individu 2, maka individu 2 juga tidak dapat dijadikan calon induk pertama.
Bilangan acak pertama $<$ individu 3, maka individu 3 dapat dijadikan calon induk pertama
- Bilangan acak kedua $>$ individu 1, individu 2 dan individu 3, namun bilangan acak 2 $<$ individu 4, maka yang dapat di jadikan calon induk kedua adalah individu 4.
- Bilangan acak ketiga $<$ individu 1, individu 2, individu 3 dan individu 4, namun dari semua itu individu yang paling mendekati bilangan acak ketiga adalah individu 1, maka individu 1 yang dapat dijadikan calon induk ketiga.

- Bilangan acak keempat > individu 1, namun bilangan acak keempat < individu 2, maka individu 2 dapat dijadikan calon induk keempat.
- Individu yang dapat dijadikan induk mutasi adalah calon induk mutasi yang memiliki bilangan acak kedua < 0.01, karena 0.01 adalah parameter untuk operasi mutasi. Berdasarkan ketentuan yang ada, individu yang dapat dijadikan induk mutasi adalah calon individu 2, karena hanya calon individu 2 yang memiliki nilai bilangan acak kedua < 0.01. Hasil penentuan calon induk mutasi dan induk mutasi seperti yang sudah dijelaskan sebelumnya dapat dilihat pada Tabel 11.

Tabel 4. 15 Hasil pencarian anak mutasi

Kromosom	Nilai Fitnes Kumulatif	Bilangan acak I	Calon Induk Mutasi	Bilangan acak II	Induk Mutasi
Individu 1	0.2584158	0.634824	Individu 3	0.342648	
Individu 2	0.4834826	0.873276	Individu 4	0.5836278	
Individu 3	0.7420755	0.218398	Individu 1	0.123739	
Individu 4	1	0.427838	Individu 2	0.0102236	Individu 2

- Kemudian pilih 2 posisi acak untuk induk mutasi individu 2, setelah itu tukarkan kedua posisi tersebut untuk mendapatkan keturunan baru anak mutasi, seperti yang terdapat pada Gambar 4.9.



Gambar 4. 9 Keturunan baru hasil mutasi

- Langkah kesembilan, membuat populasi baru dari anak keturunan crossover dan mutasi, individu 5, individu 6 dan individu 7 yang ada pada Tabel 16, adalah individu baru hasil dari proses crossover dan mutasi. Setelah populasi baru dibentuk, selanjutnya mencari nilai evaluasi, menggunakan cara yang sama seperti langkah kedua, dengan menjumlahkan seluruh nilai setiap individu. Set individu terbaik yang dihasilkan GA adalah yang memiliki nilai evaluasi paling kecil, karena nilai evaluasi yang kecil nantinya akan menghasilkan nilai fitness yang besar. Berdasarkan data yang ada pada Tabel 16, nilai fitness terkecil ada pada individu 5 dengan nilai fitness terkecil sebesar 277113.

Tabel 4. 16 Populasi baru

Calon Solusi	Set Kromosom																								Nilai Evaluasi
	7	4	8	9	3	5	2	6	1	15	13	17	11	18	12	14	10	16	19	24	20	23	22	21	
Individu 1	7	4	8	9	3	5	2	6	1	15	13	17	11	18	12	14	10	16	19	24	20	23	22	21	
Nilai	1	45 65 9	0	0	32 32 2	65 78 1	13 65 1	7 5 5	0	29	48 9	0	19 21 7	14 74 2	46 36 8	14 53 10	44 10	0	47 19 6	0	32 56 9	0	42 3	40 36	296898
Individu 2	1	5	3	2	4	6	0	7	8	10	13	11	15	12	14	18	19	16	17	20	23	22	21	24	
Nilai	58 12	66 67	45 49 4	72 20 0	12 82 5	22	0	0	0	23	60 6	72 17 2	0	41 18 7	33	98 34 3	18 95 7	0	16 39 1	93 13	0	0	16	0	429060
Individu 3	7	6	9	10	8	3	2	1	5	4	20	22	24	23	21	15	13	16	12	11	18	19	17	14	
Nilai	18	10 45	0	39 24	0	26 91 1	11 95 8	0	94 95	49 30 5	35 18 6	54 4	0	0	51 69	63	35 21 6	0	40 14 4	17 50 0	13 74 4	44 08 9	0	18 85	296196
Individu 4	22	21	24	23	16	19	18	15	17	12	14	11	13	7	10	9	5	8	6	3	4	1	2	20	
Nilai	27 2	43 52	0	0	0	48 83 0	15 89 2	2	0	46 87 2	14 22	18 13 7	28 04 3	2	51 21	0	78 69	0	80 4	31 84 3	45 56 7	0	13 53 0	30 26 9	298845
Individu 5	7	6	9	10	8	3	2	1	5	4	20	11	13	7	10	9	5	8	6	3	4	1	17	14	
Nilai	18	10 45	0	39 24	0	26 91 1	11 95 8	0	94 95	49 30 5	35 18 6	18 13 7	28 04 3	2	51 21	0	78 69	0	80 4	31 84 3	45 56 7	0	0	18 85	277113
Individu 6	22	21	24	23	16	19	18	15	17	12	14	22	24	23	21	15	13	16	12	11	18	19	2	20	

Tabel 4.16 Lanjutan

Nilai	27 2	43 52	0	0	0	48 83 0	15 89 2	2 0	0	46 87 2	14 22	54 4	0	0	51 69	63	35 21 6	0	40 14 4	17 50 0	13 74 4	44 08 0	13 53 0	30 26 9	317928
Individu 7	1	5	3	2	4	6	9	1 4	8	10	13	11	15	12	7	18	19	1 6	17	20	23	22	21	24	
Nilai	58 12	66 67	45 49 4	72 20 0	12 82 5	22	0	3 3	0	23 55 6	60 72	72 17 2	0	41 18 7	0	98 34 3	18 95 7	0	16 39 1	93 13	0	0	16	0	429060

Berdasarkan data yang ada pada Tabel 16, nilai fitness terkecil ada pada individu 5 dengan nilai fitness terkecil sebesar 277113, oleh sebab itu pasangan fitur terbaik adalah fitur dengan nomor urut ke - **7, 6, 9, 10, 8, 3, 2, 1, 5, 4, 20, 11, 13, 17, 14**, urutan fitur tersebut adalah **R7, R6, G1, G2, R8, R3, R2, R1, R5, R4, B4, G3, G5, B1, G6**, karena terdapat beberapa nomor fitur yang sama, maka nomor fitur tersebut cukup ditulis 1 kali saja. Dari total 24 fitur yang digunakan, selama 1 kali percobaan Algoritma Genetika mampu menyeleksi fitur tersebut menjadi 15 fitur terbaik yang memiliki nilai fitness paling besar. Fitur – fitur tersebut yang kemudian digunakan untuk dataset klasifikasi.

4.5. Klasifikasi

Eksperimen yang akan dilakukan menggunakan algoritma KNN untuk klasifikasinya. Nilai K yang digunakan dalam percobaan adalah $K=4$ dengan 100 kali iterasi. Perbandingan antara data training dan data testing yang digunakan sebesar 8:2, 80% untuk data training dan 20% untuk data testing. Pada tahapan klasifikasi juga dilakukan dengan 2 skenario. Skenario pertama dilakukan klasifikasi dengan menggunakan seluruh fitur tanpa dilakukan fitur seleksi terlebih dahulu, sedangkan pada skenario kedua klasifikasi dilakukan dengan menggunakan fitur-fitur hasil seleksi yang diperoleh dari tahapan sebelumnya yaitu fitur seleksi dengan Algoritma Genetika. Masing-masing skenario akan dilakukan percobaan sebanyak 5 kali.

Hasil dari eksperimen yang telah dilakukan untuk skenario pertama pada ruang warna RGB dapat dilihat pada Tabel 4.17. Berdasarkan hasil yang ada pada Tabel 4.17, tingkat akurasi yang didapatkan dengan menggunakan algoritma genetika jauh lebih baik jika dibandingkan tanpa menggunakan algoritma genetika. Tingkat akurasi tertinggi untuk ruang warna RGB sebesar **98.69%** dengan pasangan fitur **2,4,5,7,13,15,18,19,21,22** pasangan fitur tersebut adalah R2,R4, R5, R7, G5, G7, B2 B3, B5, B6,. Berdasarkan hasil tersebut ruang warna RGB dapat mereduksi fitur sebesar **58.3%**. Sebelum menggunakan Algoritma Genetika rata-rata tingkat akurasi sebesar **94.7598%** namun setelah menggunakan Algoritma Genetika rata-rata tingkat akurasi sebesar **97.9040%**. Terdapat peningkatan akurasi sebesar **3.14%** setelah menggunakan Algoritma Genetika.

Tabel 4. 17 Hasil eksperimen ruang warna RGB

Percobaan	Akurasi Tanpa GA(%)	Akurasi Dengan GA	
		Akurasi (%)	Nomor Feature
1	95.6332	97.8166	5,10,11,18,19,21,22
2	93.4498	97.3799	2,3,5,6,13,15,17,18,19,22
3	95.6332	98.69	2,4,5,7,13,15,18,19,21,22
4	94.7598	97.3799	2,6,8,17,18,21,24
5	94.3231	98.2533	2,5,6,11,13,19,23
Rata-rata	94.7598	97.9040	

Eksperimen ruang warna LAB dilakukan percobaan sebanyak 5 kali. Hasil percobaan tersebut dapat dilihat pada Tabel. 4.18. Berdasarkan hasil yang ada pada Tabel. 4.18, tingkat akurasi tertinggi pada ruang warna LAB sebesar **94.7598%** dengan pasangan fitur **1,10,13,16,22,23** pasangan fitur tersebut adalah **L1, A2, A5, A8, B6, B7**. Ruang warna LAB mampu mereduksi fitur sebesar **75%** dari 24 fitur yang digunakan. Rata – rata tingkat akurasi sebelum menggunakan Algoritma Genetika sebesar **88.9083%**, sedangkan rata-rata tingkat akurasi setelah menggunakan Algoritma Genetika sebesar **93.1878%**. Tingkat akurasi setelah menggunakan algoritma genetika rata-rata meningkat sebesar **4.28%**.

Tabel 4. 18 Hasil percobaan ruang warna LAB

Percobaan	Akurasi Tanpa GA (%)	Akurasi Dengan GA	
		Akurasi (%)	Nomor Feature
1	88.6463	94.7598	1,10,13,16,22,23
2	89.9563	91.7031	1,8,9,10,16,22,24
3	88.2096	93.0131	1,10,12,14,16,17
4	89.5197	92.1397	1,9,11,14,16,17,20,22
5	88.2096	94.3231	1,10,11,12,14,16,17,21
Rata-rata	88.9083	93.1878	

Data hasil percobaan pada ruang warna HSV dapat dilihat pada Tabel. 4.19 Berdasarkan hasil yang ada pada Tabel. 4.19, tingkat akurasi tertinggi sebesar **99.1266%** dengan pasangan fitur **1,2,3,12,14,17,19,22,23**, pasangan fitur tersebut adalah **H1, H2, H3, S4, S6, V1, V3, V6, V7**. Pada ruang warna HSV algoritma genetika mampu mereduksi fitur sebesar **62.5%** dari 24 fitur yang digunakan. Pada ruang warna HSV rata-rata tingkat akurasi sebelum menggunakan Algoritma Genetika sebesar **95.7205%**, sedangkan setelah menggunakan Algoritma Genetika rata-rata tingkat akurasi **97.72293%**. Tingkat akurasi setelah menggunakan algoritma genetika rata-rata meningkat sebesar **2.87%** dari sebelumnya.

Tabel 4. 19 Hasil percobaan ruang warna HSV

Percobaan	Akurasi Tanpa GA (%)	Akurasi Dengan GA	
		Akurasi (%)	Nomor Feature
1	95.1965	96.9432	1,3,5,11,12,18,20,21,23,24
2	96.0699	99.1266	1,2,3,12,14,17,19,22,23
3	94.7598	98.2533	1,3,10,11,14,17,18,22,23
4	95.1965	96.0699	1,2,6,18,19,20,22,24
5	97.3799	98.2533	1,2,6,9,12,16,20,22,24
Rata-rata	95.7205	97.72293	

Hasil percobaan yang dilakukan pada ruang warna HSI dapat dilihat pada Tabel 4.20. Berdasarkan data hasil percobaan yang ada pada Tabel 4.20, nilai akurasi tertinggi sebesar **98.2533%** dengan pasangan feature **1,3,6,10,13,20,24**, pasangan fitur tersebut adalah **H1, H3, S2, S5, 14, 18**. Pada ruang warna HSI algoritma genetika mampu mereduksi **70.83%** dari 24 fitur yang digunakan. Rata-rata tingkat akurasi yang diperoleh sebelum menggunakan Algoritma Genetika sebesar **95.1092%** sedangkan rata-rata tingkat akurasi setelah menggunakan

Algoritma Genetika sebesar **97.3799%**. Terdapat peningkatan akurasi sebelum dan sesudah menggunakan Algoritma Genetika, tingkat akurasi yang diperoleh setelah menggunakan algoritma genetika rata-rata meningkat **2.27%** dari sebelumnya.

Tabel 4. 20 Hasil percobaan ruang warna HSI

Percobaan	Akurasi Tanpa GA (%)	Akurasi Dengan GA	
		Akurasi (%)	Nomor Feature
1	93.4498	96.9432	1,2,6,11,12,13,19,21,24
2	96.0699	96.9432	1,3,4,6,9,10,18,21,24
3	94.7598	98.2533	1,2,3,7,12,17,18,23
4	95.1965	96.5066	1,2,5,6,10,12,13,14,15,16,18,23,24
5	96.0699	98.2533	1,3,6,10,13,20,24
Rata-rata	95.1092	97.3799	

Berdasarkan percobaan yang telah dilakukan pada skenario pertama, tingkat akurasi terbaik diperoleh oleh ruang warna HSV dengan tingkat akurasi sebesar **99.1266%** dengan pasangan fitur **1,2,3,12,14,17,19,22,23**, pasangan fitur tersebut adalah **H1, H2, H3, S4, S6, V1, V3, V6, V7** dan fitur yang direduksi sebesar **62.5%**. Pada skenario kedua, percobaan dilakukan sebanyak 5 kali. Skenario kedua dilakukan dengan menggabungkan seluruh fitur yang ada, mulai dari ruang warna RGB, LAB, HSV dan HSI.

Total fitur yang akan diseleksi untuk skenario kedua sebanyak 96 fitur. Data hasil percobaan skenario kedua dapat dilihat pada Tabel 4.21. Berdasarkan data yang ada pada Tabel 4.21, tingkat akurasi tertinggi sebesar 100% dengan pasangan fitur **8,13,15,40,51,52,59,61,73,77,82,90,94**, pasangan fitur tersebut adalah **R8, G5, G7, A8, H3-HSV, H4-HSV, S3-HSV, S5-HSV, H2-HSI, H6-HSI, S3-HSI, I3-HSI, dan I7-HSI**. Dari total 96 fitur yang ada, algoritma genetika mampu

mereduksi sebesar **86.46%** fitur dengan tingkat akurasi rata-rata meningkat **2.15%** setelah menggunakan algoritma genetika.

Tabel 4. 21 Hasil percobaan dengan seluruh ruang warna

Percobaan	Akurasi Tanpa GA (%)	Akurasi Dengan GA(%)	
		Akurasi (%)	Nomor Feature
1	95.1965	98.2533	6,16,18,19,42,49,50,59,61,82,84,94,96
2	97.5633	99.5633	10,40,50,53,64,82,93,96
3	98.69	99.5633	7,11,32,33,55,61,77,90
4	97.3799	98.69	8,37,40,50,51,54,57,62,82,89
5	96.5066	100	8,13,15,40,51,52,59,61,73,77,82,90,94
Rata-rata	97.0673	99.2140	

Berdasarkan data yang dihasilkan oleh skenario pertama dan kedua, rata-rata tingkat akurasi meningkat sebesar 2.7687% setelah menggunakan algoritma genetika. Tingkat akurasi tertinggi yang dihasilkan oleh skenario kedua sebesar 100%, dan banyaknya fitur yang direduksi sebesar 86,46% dari total 96 fitur menjadi 13 fitur yang digunakan untuk menghasilkan tingkat akurasi sebesar 100%.

Rata-rata tingkat akurasi yang dihasilkan setelah menggunakan algoritma genetika untuk skenario pertama sebesar 96.5502% dan untuk skenario kedua sebesar 99.2140%. Sedangkan tingkat akurasi rata-rata untuk skenario pertama sebelum menggunakan algoritma genetika sebesar 93.6246% Dan skenario kedua sebesar 97.0673%. Berdasarkan data tersebut percobaan dengan menggunakan scenario kedua memberikan hasil yang lebih baik dari pada scenario yang pertama.

Tabel 4. 22 Perbandingan hasil dengan penelitian sebelumnya

Referensi	Dataset	Jumlah fitur awal	Algoritma fitur seleksi	Algoritma klasifikasi	Accuration	Jumlah fitur setelah seleksi
(Hamdani et al., 2021)	Daun kelapa sawit sehat dan daun kelapa sawit terinfeksi penyakit curvularia	96 fitur	correlation-based feature selection (CFS), correlation, gain ratio, and PCA	ANN	100 %	41
Penelitian saat ini	Daun kelapa sawit sehat dan daun kelapa sawit terinfeksi penyakit curvularia dan terinfeksi jamur	96 fitur	Algoritma Genetika	KNN	100 %	8

Berdasarkan data yang ada pada Tabel 4.22, sebagian besar data yang digunakan adalah data yang sama kecuali data kelas penyakit daun kelapa sawit terinfeksi jamur, dari data yang digunakan hasil yang diperoleh dalam penelitian ini jauh lebih baik jika dibandingkan dengan penelitian sebelumnya (Hamdani et al., 2021).

Pada penelitian sebelumnya (Hamdani et al., 2021) tingkat akurasi tertinggi yang dihasilkan sebesar 100%, menggunakan banyak algoritma seleksi fitur, namun jumlah fitur yang digunakan jauh lebih banyak, sekitar 41 fitur. Sedangkan pada penelitian ini jumlah kelas yang digunakan lebih banyak dari

sebelumnya, sehingga menghasilkan proses yang lebih kompleks, namun tingkat akurasi yang dihasilkan masih sama baiknya yaitu sebesar 100% dengan jumlah fitur yang digunakan hanya 13 fitur dan satu algoritma seleksi fitur.

Cara Algoritma Genetika dalam menyeleksi fitur hingga menghasilkan tingkat akurasi yang sama baiknya adalah dengan mencari nilai evaluasi yang paling minimum dari setiap generasi untuk menghasilkan generasi baru, karena nilai evaluasi yang kecil akan menghasilkan nilai fitness yang besar, dan individu yang memiliki nilai fitness yang besar memiliki peluang yang besar pula untuk menjadi individu generasi baru.

Hasil 100% yang diperoleh sama dengan penelitian sebelumnya (Hamdani et al., 2021), hal tersebut terjadi karena data yang digunakan overfitting. Data – data yang digunakan kurang bervariasi, data tersebut dengan sengaja dipilih yang terbaik, antara data kelas satu dengan kelas yang lainnya terdapat perbedaan yang sangat jelas, misalnya daun yang sehat sengaja dipilih yang warna berwarna hijau tanpa ada bercak, sedangkan data daun yang sakit yang dipilih, satu daun dipenuhi dengan bercak-bercak penyakit, sehingga hasil akurasi yang diperoleh juga sangat baik.

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan percobaan dan analisa yang telah dilakukan dapat disimpulkan bahwa:

1. Penelitian ini jauh lebih baik dari penelitian sebelumnya (Hamdani et al., 2021), karena pada penelitian ini menggunakan kelas yang jauh lebih banyak. Jika sebelumnya menggunakan 2 kelas, maka pada penelitian ini menggunakan 3 kelas, sehingga proses yang dihasilkan jauh lebih kompleks namun tingkat akurasi yang diperoleh sama baiknya yaitu 100% dan fitur yang digunakan dari hasil seleksi lebih sedikit. Jika penelitian sebelumnya harus menggunakan 41 fitur untuk menghasilkan tingkat akurasi sebesar 100%, pada penelitian ini hanya menggunakan 13 fitur untuk mencapai tingkat akurasi 100%.
2. Tingkat akurasi terbaik diperoleh dari skenario kedua sebesar 100%. Pasangan fitur hasil dari seleksi menggunakan algoritma genetika yang memiliki karakteristik kuat yaitu urutan fitur ke - **8,13,15,40,51,52,59,61,73,77,82,90,94** urutan fitur tersebut yaitu **R8, G5, G7, A8, H3-HSV, H4-HSV, S3-HSV, S5-HSV, H2-HSI, H6-HSI, S3-HSI, I3-HSI, dan 17-HSI**.
3. Rata-rata tingkat akurasi yang diperoleh setelah menggunakan fitur hasil seleksi algoritma genetika sebesar **99.2140%**.

4. Algoritma genetika mampu mereduksi sebesar 86,46%, dari total 96 fitur menjadi 13 fitur.
5. Rata-rata tingkat akurasi yang diperoleh sebelum menggunakan fitur hasil seleksi algoritma genetika sebesar 94.3130% dan rata-rata tingkat akurasi setelah menggunakan fitur hasil seleksi algoritma genetika sebesar 97.0817%. Terjadi peningkatan akurasi sebesar 2.7687% jika menggunakan fitur hasil seleksi algoritma genetika.
6. Dari kedua skenario yang telah dilakukan dan dianalisa, tingkat akurasi yang dihasilkan akan lebih baik jika fitur diseleksi terlebih dahulu sebelum dilakukan klasifikasi.

5.2. Saran

Pada penelitian yang akan datang bisa meneliti lagi menggunakan model yang sama tetapi dengan data yang berbeda, dan menganalisisnya tidak hanya menggunakan akurasi saja tetapi juga bisa menggunakan seperti F1 score, precision, dan recall, untuk mengetahui model tersebut memang bagus untuk diterapkan disemua data atau hanya di data tertentu saja karena overfitting.

DAFTAR PUSTAKA

- Ahn, G., & Hur, S. (2020). Efficient genetic algorithm for feature selection for early time series classification. *Computers and Industrial Engineering*, 142(February), 106345. <https://doi.org/10.1016/j.cie.2020.106345>
- Alelyani, S. (2021). STabel bagging feature selection on medical data. *Journal of Big Data*, 8(1). <https://doi.org/10.1186/s40537-020-00385-8>
- Ali, H., Lali, M. I., Nawaz, M. Z., Sharif, M., & Saleem, B. A. (2017). Symptom based automated detection of citrus diseases using color histogram and textural descriptors. *Computers and Electronics in Agriculture*, 138, 92–104. <https://doi.org/10.1016/j.compag.2017.04.008>
- Amiri, Z., Hassanpour, H., & Beghdadi, A. (2019). Feature Selection for Bleeding Detection in Capsule Endoscopy Images using Genetic Algorithm. *5th Iranian Conference on Signal Processing and Intelligent Systems, ICSPIS 2019, December*, 18–19. <https://doi.org/10.1109/ICSPIS48872.2019.9066008>
- Chen, B., Shi, S., Sun, J., Chen, B., Guo, K., Du, L., Yang, J., Xu, Q., Song, S., & Gong, W. (2021). Using HSI Color Space to Improve the Multispectral Lidar Classification Error Caused by Measurement Geometry. *IEEE Transactions on Geoscience and Remote Sensing*, 59(4), 3567–3579. <https://doi.org/10.1109/TGRS.2020.3006577>
- Dai, F., Fan, B., & Peng, Y. (2018). Space and Bilateral Filtering. *2018 Chinese Control And Decision Conference (CCDC)*, 5945–5948.
- Fatima, A., Maurya, R., Dutta, M. K., Burget, R., & Masek, J. (2019). Android malware detection using genetic algorithm based optimized feature selection and machine learning. *2019 42nd International Conference on Telecommunications and Signal Processing, TSP 2019*, 220–2223. <https://doi.org/10.1109/TSP.2019.8769039>
- Halim, Z., Yousaf, M. N., Waqas, M., Sulaiman, M., Abbas, G., Hussain, M., Ahmad, I., & Hanif, M. (2021). An effective genetic algorithm-based feature selection method for intrusion detection systems. *Computers and Security*, 110, 102448. <https://doi.org/10.1016/j.cose.2021.102448>
- Hamdani, H., Septiarini, A., Sunyoto, A., Suyanto, S., & Utaminigrum, F. (2021). Detection of oil palm leaf disease based on color histogram and supervised classifier. *Optik*, 245(August), 167753. <https://doi.org/10.1016/j.ijleo.2021.167753>
- Kartika, D. S. Y., & Herumurti, D. (2017). Koi fish classification based on HSV color space. *Proceedings of 2016 International Conference on Information*

and Communication Technology and Systems, ICTS 2016, 96–100.
<https://doi.org/10.1109/ICTS.2016.7910280>

- Keramat-Jahromi, M., Mohtasebi, S. S., Mousazadeh, H., Ghasemi-Varnamkhandi, M., & Rahimi-Movassagh, M. (2021). Real-time moisture ratio study of drying date fruit chips based on on-line image attributes using kNN and random forest regression methods. *Measurement: Journal of the International Measurement Confederation*, 172(October 2020), 108899. <https://doi.org/10.1016/j.measurement.2020.108899>
- Khekade, A. (2015). *Shadow detection based on RGB and YIQ Color models in Color aerial images*. *Ablaze*, 144–147.
- Maleki, N., Zeinali, Y., & Niaki, S. T. A. (2021). A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection. *Expert Systems with Applications*, 164(September 2020), 113981. <https://doi.org/10.1016/j.eswa.2020.113981>
- Rulaningtyas, R., B. Suksmono, A., L. R. Mengko, T., & Putri Saptawati, G. A. (2015). Segmentasi Citra Berwarna dengan Menggunakan Metode Clustering Berbasis Patch untuk Identifikasi Mycobacterium Tuberculosis. *Jurnal Biosains Pascasarjana*, 17(1), 19. <https://doi.org/10.20473/jbp.v17i1.2015.19-25>
- Tian, L., Xue, B., Wang, Z., Li, D., Yao, X., Cao, Q., Zhu, Y., Cao, W., & Cheng, T. (2021). Spectroscopic detection of rice leaf blast infection from asymptomatic to mild stages with integrated machine learning and feature selection. *Remote Sensing of Environment*, 257(August 2020), 112350. <https://doi.org/10.1016/j.rse.2021.112350>
- Utama, D. Q., & Light, A. (2016). *Primary Color Cluster*.
- Vaishnav, M. P., & Devi, K. S. (n.d.). *Diseases using KNN classifier*. 1–5.
- Wang, D., Chen, H., Li, T., Wan, J., & Huang, Y. (2020). A novel quantum grasshopper optimization algorithm for feature selection. *International Journal of Approximate Reasoning*, 127, 33–53. <https://doi.org/10.1016/j.ijar.2020.08.010>
- Xue, Y., Zhu, H., Liang, J., & Slowik, A. (2021). Adaptive crossover operator based multi-objective binary genetic algorithm for feature selection in classification[Formula presented]. *Knowledge-Based Systems*, 227, 107218. <https://doi.org/10.1016/j.knsys.2021.107218>
- Zhang, T., & Lin, J. (2019). A genetic algorithm design based on self-organizing dynamic network. *Chinese Control Conference, CCC, 2019-July*, 1039–1044. <https://doi.org/10.23919/ChiCC.2019.8865566>