

BAB III

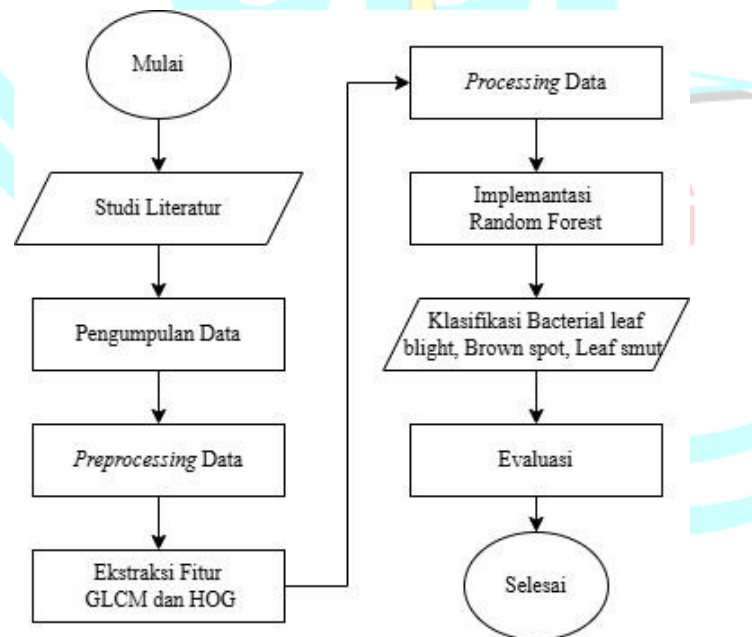
METODOLOGI

3.1 Objek Penelitian

Objek dalam penelitian ini adalah daun padi yang terindikasi mengalami penyakit dengan berbagai tingkat keparahan, yaitu *leaf smut*, *brown spot*, dan *bacterial leaf blight*. Pada kategori *leaf smut*, daun padi mulai menunjukkan bercak kecil atau perubahan warna seperti kekuningan, namun sebagian besar masih hijau. Sementara itu, pada kategori *brown spot*, daun padi mengalami perubahan warna dominan menjadi cokelat atau kuning tua dengan bercak besar. Pada kategori *bacterial leaf blight*, daun padi mengalami perubahan warna yang sangat besar, serta kemungkinan adanya jaringan daun yang rusak atau kering.

Penelitian ini dilaksanakan di Laboratorium Riset Universitas Buana Perjuangan Karawang. Adapun waktu pelaksanaan penelitian dimulai pada bulan Desember hingga penelitian selesai.

3.2 Prosedur Penelitian



Gambar 3. 1 Prosedur penelitian

Terdapat 8 (delapan) tahapan dalam prosedur penelitian ini dimulai dengan studi literatur, selanjutnya melakukan pengumpulan data, Setelah itu, dilakukan *pre-processing* data, ekstraksi fitur menggunakan metode *GLCM* dan *HOG* dan

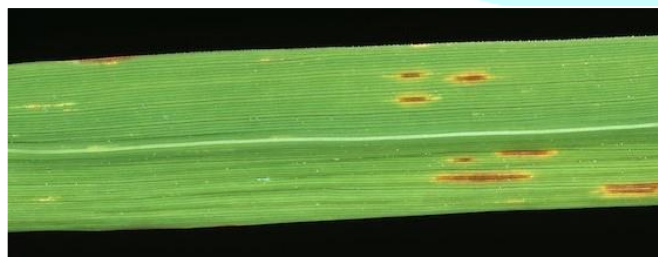
processing data, mengimplementasikan algoritma *Random Forest*. Tahap berikutnya adalah klasifikasi penyakit daun padi. Setelah dilakukan klasifikasi data dilakukan evaluasi dan digunakan dalam penelitian selanjutnya.

3.3 Studi Literatur

Pada tahapan ini melakukan observasi dan wawancara kepada Pak tani di salah satu area Pesawahan Karawang Barat dan *review* pengumpulan data dilakukan dengan mencari sumber referensi dari buku, jurnal, dan artikel yang relevan dengan topik penelitian mengenai Klasifikasi Penyakit Daun Padi Menggunakan Metode *GLCM* dan *HOG* dengan Algoritma *Random Forest*. Setelah itu, penulis akan melakukan pengumpulan serta analisis data yang dibutuhkan guna mendukung jalannya penelitian ini.

3.4 Pengumpulan Data

Tahapan ini dilakukan dengan mengumpulkan dataset dari website *Kaggle.com* serta melakukan observasi dan wawancara dengan Pak Tani di salah satu area pesawahan Karawang Barat. Data yang dikumpulkan terkait dengan Klasifikasi Penyakit Daun Padi Menggunakan Metode *GLCM* dan *HOG* dengan Algoritma *Random Forest*. Dataset yang dipakai mencakup 120 gambar citra daun padi yang terbagi ke dalam tiga kategori, yaitu *Leaf Smut*, *Brown Spot*, dan *Bacterial Leaf Blight*. Setiap kelas memiliki proporsi yang seimbang, dengan jumlah data yang didistribusikan secara merata, yakni 40 gambar per kelas. Pemilihan tiga jenis data citra daun padi yang terinfeksi penyakit bertujuan untuk mempermudah proses klasifikasi. Data yang dimanfaatkan dalam penelitian ini berasal dari hasil keputusan yang diperoleh melalui proses pengujian sistem. Contoh data yang diambil dapat dilihat pada gambar di bawah ini.



Gambar 3. 2 Daun padi bercak (leaf smut).

Pada gambar 3.2 menampilkan daun padi yang memiliki bercak pada daunnya.



Gambar 3. 3 Daun padi bercak coklat (Brown spot).

Pada gambar 3.3 menampilkan daun padi yang memiliki bercak berwarna coklat pada daunnya.



Gambar 3. 4 Hawar daun bakteri (Bacterial leaf blight).

Pada gambar 3.4 menampilkan daun padi yang memiliki hawar daun bakteri.

3.4.1 Segmentasi

Proses segmentasi citra dilakukan untuk memisahkan area bercak penyakit dari daun padi dengan tahapan dimulai dari pembacaan citra dan konversi format warna dari *BGR* ke *RGB*, kemudian ke ruang warna *HSV* untuk memudahkan deteksi berbasis rentang warna. Langkah awal adalah membuat *mask* daun menggunakan rentang warna hijau, kuning, hingga coklat sehingga area daun dapat teridentifikasi, kemudian dari area tersebut dilakukan segmentasi lanjutan untuk mengekstraksi bercak penyakit berdasarkan rentang warna coklat atau gelap yang umum menjadi indikasi gejala penyakit. Untuk meningkatkan akurasi, dilakukan operasi morfologi *opening* dan *closing* dengan kernel 3×3 guna menghilangkan *noise* kecil dan memperhalus area hasil *segmentasi*. Hasil akhir berupa citra yang hanya menampilkan bercak penyakit kemudian disimpan dalam format *PNG* pada direktori keluaran.

3.4.2 Augmentasi

Proses *augmentasi* citra dilakukan dengan menerapkan beberapa transformasi spasial tanpa mengubah warna asli citra. Transformasi yang digunakan meliputi rotasi acak sebesar 0° , 90° , 180° , atau 270° , *flipping horizontal* dan *flipping vertikal* masing-masing dengan probabilitas 50%, serta pemotongan (crop) acak pada area citra antara 80% hingga 100% dari ukuran asli yang kemudian diubah ukurannya kembali ke ukuran semula. Tujuan dari *augmentasi* ini adalah untuk meningkatkan variasi data, meminimalisasi *overfitting*, serta mensimulasikan berbagai kondisi visual pada citra daun padi agar model lebih *robust* dalam mengenali pola bercak penyakit pada berbagai sudut pandang dan komposisi gambar.

3.5 Preprocessing Data

Tahap *preprocessing* ini untuk mempersiapkan citra daun padi sebelum digunakan dalam proses ekstraksi fitur dan pelatihan model klasifikasi. Proses ini diawali dengan membaca seluruh gambar dari tiga kategori penyakit daun padi, yaitu *Bacterial Leaf Blight*, *Brown Spot*, dan *Leaf Smut* yang terdapat di dalam direktori dataset. Setiap gambar kemudian diubah ukurannya menjadi 256×256 piksel agar memiliki dimensi yang seragam dan untuk menurunkan beban komputasi. Penyeragaman ukuran ini penting agar proses ekstraksi fitur dapat dilakukan secara konsisten. Selain itu, dilakukan juga pengecekan terhadap validitas gambar untuk memastikan tidak ada file yang rusak atau kosong. Gambar yang telah melewati tahap praproses kemudian disimpan dalam bentuk *array* beserta label kelas masing-masing, dan siap digunakan pada tahap berikutnya.

3.6 Ekstraksi Fitur

Fungsi ekstraksi fitur ini melibatkan pengambilan informasi penting dari citra, seperti warna dan tekstur, untuk digunakan sebagai input model klasifikasi. Pada fitur warna, dilakukan perhitungan fitur yang digunakan mencakup histogram warna, nilai rata-rata, dan deviasi standar dari masing-masing saluran warna (Merah, Hijau, Biru). Sementara untuk fitur tekstur, diterapkan metode *Gray Level Co-occurrence Matrix* (GLCM) guna memperoleh nilai seperti *contrast*, *homogeneity*, dan *entropy* yang berguna dalam mengidentifikasi pola dalam citra. Fitur ini akan digunakan input untuk model klasifikasi.

3.6.1 Tahapan Ekstraksi Fitur

1. Ekstraksi Fitur Warna:

Hitung histogram warna, rata-rata, dan deviasi standar dari setiap saluran warna (RGB).

a. Konversi ke Ruang Warna RGB

Baca citra daun padi menggunakan OpenCV atau PIL.

Pastikan citra dalam format RGB.

b. Hitung Histogram Warna

Pisahkan citra ke dalam tiga saluran warna: Red (R), Green (G), dan Blue (B).

Hitung histogram intensitas piksel untuk setiap saluran warna.

c. Hitung Rata-Rata Warna

Hitung nilai rata-rata intensitas piksel untuk setiap saluran (R, G, B).

d. Hitung Deviasi Standar Warna

Hitung deviasi standar intensitas piksel untuk setiap saluran warna.

e. Simpan Fitur Warna

Gabungkan nilai histogram, rata-rata, dan deviasi standar ke dalam vektor fitur warna.

2. Ekstraksi Fitur Tekstur:

Gunakan metode seperti *Gray Level Co-occurrence Matrix* (GLCM) untuk memperoleh nilai *contrast*, *homogeneity*, dan *entropy*.

a. Konversi Gambar ke Grayscale

Karena *GLCM* hanya bekerja pada citra grayscale, ubah semua gambar ke skala abu-abu.

b. Normalisasi Gambar (Opsional, jika diperlukan)

Normalisasi dilakukan untuk menyamakan skala intensitas piksel agar lebih seragam.

c. Perhitungan *GLCM*

Buat *GLCM* dengan menentukan parameter berikut:

Sudut hubungan antar piksel (misal: 0° , 45° , 90° , 135°).

Jarak antar piksel (misal: 1, 3, 5).

Jumlah level intensitas grayscale (misalnya, 256 untuk citra 8-bit).

d. Ekstraksi Fitur Tekstur

Dari *GLCM* yang dihasilkan, hitung fitur tekstur utama:

Contrast → Membedakan intensitas piksel pada citra.

Homogeneity → Mengukur kehalusan tekstur.

Entropy → Mengukur tingkat ketidakaturan tekstur.

e. Simpan Fitur ke dalam Dataset

Nilai *contrast*, *homogeneity*, dan *entropy* disimpan sebagai fitur dalam dataset untuk proses klasifikasi lebih lanjut.

3. Ekstraksi fitur bentuk

Gunakan *Histogram of Oriented Gradients* (HOG) untuk ekstraksi fitur bentuk.

a. Konversi Gambar ke Grayscale

Sama seperti *GLCM*, *HOG* juga bekerja lebih baik pada citra grayscale untuk mengurangi kompleksitas perhitungan.

b. Normalisasi Kontras (Opsional, jika diperlukan)

Meningkatkan visibilitas tepi pada gambar untuk hasil yang lebih baik

c. Perhitungan Gradien Gambar

Hitung gradien intensitas piksel dalam arah *horizontal* (dx) dan *vertikal* (dy) menggunakan operator Sobel.

Gradien ini akan digunakan untuk menentukan arah tepi objek dalam gambar.

d. Pembentukan Sel dan Blok

Bagi gambar menjadi sel kecil (misalnya 8×8 piksel).

Gabungkan beberapa sel menjadi blok (misalnya 2×2 sel).

e. Pembuatan *Histogram Gradien*

Setiap sel menghasilkan histogram berdasarkan arah gradiennya.

Jumlah bin histogram ditentukan (misalnya 9 bin untuk $0^\circ - 180^\circ$ dalam 20° per bin).

f. Normalisasi Fitur *HOG*

Fitur dalam blok dinormalisasi untuk mengurangi pengaruh perubahan pencahayaan.

g. Simpan Fitur ke dalam Dataset

Vektor hasil ekstraksi fitur *HOG* disimpan dalam dataset untuk digunakan dalam proses klasifikasi dengan *Random Forest*.

4. Penyusunan Fitur:

Setelah fitur warna, tekstur, dan bentuk diekstraksi, langkah selanjutnya adalah menggabungkan semua fitur tersebut ke dalam satu dataset yang akan digunakan dalam proses analisis dan pemodelan klasifikasi. Proses penyusunan fitur dilakukan melalui tahapan berikut:

a. Penggabungan Fitur

Semua fitur yang telah diekstraksi dari citra daun padi, yaitu fitur warna, fitur tekstur, dan fitur bentuk, digabungkan ke dalam satu vektor fitur. Penggabungan ini memastikan bahwa setiap citra direpresentasikan dengan informasi yang komprehensif.

b. Format Data

Dataset yang dihasilkan berbentuk tabel, di mana setiap baris mewakili satu citra daun padi dan setiap kolom berisi nilai fitur yang telah dihitung. Fitur warna mencakup histogram intensitas warna, nilai rata-rata, dan standar deviasi dari setiap saluran warna. Fitur tekstur terdiri dari nilai contrast, homogeneity, dan entropy hasil perhitungan GLCM. Sementara itu, fitur bentuk diperoleh dari hasil ekstraksi Histogram of Oriented Gradients (HOG).

5. Output yang Dihasilkan :

Proses penyusunan fitur menghasilkan dataset yang siap digunakan dalam tahap analisis dan pemodelan klasifikasi. Dataset ini berisi nilai fitur warna, tekstur, dan bentuk dari masing-masing citra daun padi yang telah melalui tahap pemrosesan. Output akhirnya berupa file data yang dapat dimanfaatkan dalam proses pelatihan model machine learning untuk mengelompokkan jenis penyakit pada daun padi berdasarkan karakteristik visualnya.

3.7 Processing Data

Tahap pemrosesan data mencakup penyusunan dataset fitur dan pembagian dataset untuk memastikan data siap digunakan dalam analisis dan pengembangan model. Penyusunan dataset dilakukan dengan menggabungkan hasil ekstraksi fitur, seperti tekstur dan bentuk, ke dalam tabel dataset yang terstruktur. Setelah itu,

dataset data tersebut dipisahkan menjadi data pelatihan dan data pengujian dengan proporsi tertentu guna memastikan evaluasi model secara maksimal. Pada penelitian ini, algoritma yang diterapkan adalah *Random Forest*, yaitu metode ensemble yang berbasis pada pohon keputusan, yang digunakan untuk mendukung proses klasifikasi dan analisis lebih mendalam.

3.8 Implementasi Algoritma

Tahap implementasi algoritma klasifikasi dalam penelitian ini klasifikasi menggunakan Algoritma *Random forest*.

3.8.1 *Random forest*

Random Forest adalah algoritma klasifikasi yang menggunakan kombinasi dari beberapa pohon keputusan (decision tree) untuk menghasilkan prediksi yang lebih stabil dan akurat. Algoritma ini bekerja dengan membangun sejumlah pohon keputusan dari berbagai subset data yang berbeda. kemudian menentukan hasil akhir berdasarkan suara mayoritas dari seluruh pohon yang telah dibangun.

Langkah – Langkah algoritma *Random forest*:

1. Mulai.
2. Siapkan data:
 - a. Data training: digunakan untuk melatih model.
 X_{train} : fitur (tekstur, warna, dan pola bercak daun).
 Y_{train} : label (jenis penyakit: sehat, penyakit X, penyakit Y).
 - b. Data testing: digunakan untuk menguji hasil prediksi.
 X_{test} : fitur yang akan diprediksi labelnya.
3. Lakukan prediksi untuk data uji:
 - a. Setiap pohon keputusan dalam Random Forest memberikan prediksi kelas untuk setiap data dalam X_{test} .
4. Gunakan voting mayoritas:
 - a. Kelas akhir ditentukan berdasarkan mayoritas prediksi dari semua pohon keputusan.
5. Simpan hasil prediksi untuk setiap data uji
6. Selesai.

3.8.2 Pengujian Model

Pada tahap ini, dilakukan pengujian terhadap model *Random Forest* yang telah dilatih untuk mengklasifikasikan penyakit pada daun padi. Pengujian dilakukan menggunakan teknik *GridSearchCV*, yaitu metode pencarian *hiperparameter* terbaik melalui proses *cross-validation*. Dalam teknik ini, dataset dibagi ke dalam beberapa *fold*, di mana masing-masing *fold* secara bergiliran berperan sebagai data pengujian, sementara *fold-fold* yang tersisa digunakan sebagai data pelatihan. Pendekatan ini bertujuan untuk mengevaluasi performa model secara menyeluruh, menghindari *overfitting*, serta memastikan bahwa model yang dikembangkan memiliki kemampuan generalisasi yang baik terhadap data yang belum pernah dilihat sebelumnya.

3.9 Evaluasi

Evaluasi model ini dilakukan dengan menggunakan *confusion matrix* untuk mengukur kinerja dari model random forest dalam klasifikasi penyakit daun padi. *Confusion matrix* memberikan gambaran yang lebih detail tentang bagaimana model mengklasifikasikan data ke dalam setiap kategori penyakit pada daun padi, sehingga dapat dianalisis tingkat keberhasilan serta kesalahan prediksi yang terjadi. Dengan menggunakan *confusion matrix*, dapat diketahui jumlah prediksi yang benar dan salah dalam setiap kelas, yang kemudian digunakan untuk menghitung metrik evaluasi lainnya seperti akurasi, presisi, recall, dan F1-score.