

## BAB 3

### METODE PENELITIAN

#### 3.1 ALAT YANG DIGUNAKAN

Dalam menyelesaikan penelitian ini, penting untuk menggunakan alat yang dapat disesuaikan dengan kebutuhan sistem model. Terbagi dalam dua komponen utama yaitu perangkat keras (*hardware*) dan perangkat lunak (*software*). Dengan spesifikasi peralatan yang sesuai untuk mendukung penelitian ini.

#### 3.2.1 DATASET

Penelitian ini bertujuan untuk mengklasifikasi penyakit pada daun padi dengan menggunakan *Convolutional Neural Network* (CNN). Data yang digunakan analisis terdiri dari data primer yang didokumentasikan menggunakan kamera dan data sekunder yang diunduh dari *website kaggle*.

**Tabel 3. 1 Dataset penyakit daun padi**

No	Klasifikasi Penyakit Daun	Jumlah Dataset
1	Blas kolar	200
3	Hawar daun bakteri	200
4	Padi sehat	200
5	Tungro	200
Jumlah		800

Tabel 3.1 menjelaskan dataset yang terdiri dari lima kelas yaitu blas kolar, hawar daun bakteri, padi sehat, dan tungro yang mana akan mengambil tiap masing-masing kelas 200 data sehingga jumlah seluruh dataset yang akan digunakan adalah 800 gambar.

#### 3.2.2 PERANGKAT KERAS (*HARDWARE*)

Spesifikasi perangkat keras yang diterapkan dalam penelitian ini adalah laptop dengan detail sebagai berikut:

1. Laptop : HP WIN-CU8H0KTSELG

2. *Processor*: Intel (R) Core(TM) i3-8130U CPU @ 2.20GHz 2.21GHz *Windows* 10 (64-bit)

3. RAM 4 GB

Spesifikasi dari *hardware* mempunyai pengaruh terhadap waktu pemrosesan atau saat melakukan *running* program. Penggunaan *PC* dengan spesifikasi yang baik dapat mempercepat pemrosesan data. Penelitian ini menggunakan *Windows 11* untuk *Operating System* (OS).

### 3.2.3 PERANGKAT LUNAK (SOFTWARE)

Bahasa pemrograman yang diterapkan dalam penelitian ini untuk mengklasifikasi penyakit daun padi menggunakan metode *Convolutional Neural Network* (CNN) adalah *Python*. Perangkat lunak yang digunakan pada penelitian ini diantaranya:

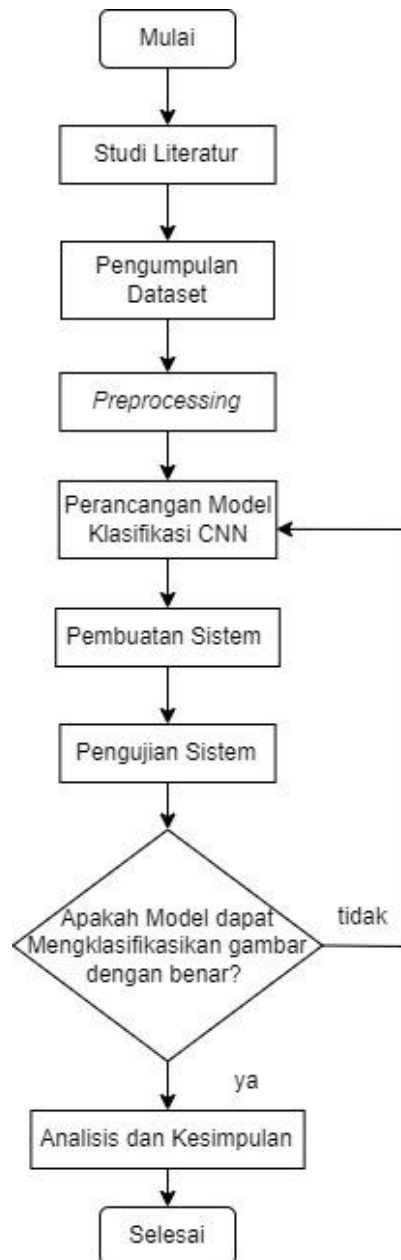
**Tabel 3. 2 Software simulasi**

No	Nama Software
1	<i>Google Colaboratory</i>
2	Bahasa Pemrograman <i>Python</i>
3	<i>Library: Tensorflow, Numpy, Matplotlib, Open CV, Keras</i>

Tabel 3.2 mencantumkan beberapa teknologi yang relevan dalam pemrograman *python*, khususnya dalam konteks pengembangan perangkat lunak, analisis data, dan pembelajaran model. Penggunaan *google colab* sebagai *platform* untuk mengeksekusi kode *python* dengan menggunakan beberapa *library* memungkinkan pengembangan dan percobaan yang efisien dalam analisis data.

## 3.2 ALUR PENELITIAN

Penelitian ini dilakukan melalui beberapa langkah. Langkah yang dilakukan dalam penelitian ini, yang dapat dipahami secara menyeluruh melalui *flowchart* pada gambar 3.1.



**Gambar 3. 1 Diagram alur penelitian**

### 3.2.1 Studi Literatur

Pada langkah pertama dalam penelitian ini adalah melakukan studi literatur yang dilakukan dengan mencari informasi dan dataset yang dibutuhkan. Dalam penelitian ini studi literatur digunakan sebagai panduan dalam menyusun penelitian ini. Studi literatur dalam penelitian ini dilakukan dengan mencari dan membaca jurnal, buku-buku, maupun artikel dari beberapa *website* mengenai materi yang berhubungan dengan pengklasifikasian citra. Beberapa materi yang dipelajari untuk

mendukung penelitian ini diantaranya mengenai klasifikasi citra, *deep learning*, *Convolutional Neural Network (CNN)*, *training*, *validation*, *accuracy* dan *loss*.

### 3.2.2 Pengumpulan Data

Selama proses pengumpulan dataset, dilakukan pengumpulan berupa citra berupa penyakit pada daun tanaman padi. Pengumpulan dataset diperoleh dari kombinasi antara dari *website kaggle* dan mengumpulkan dataset melalui pengambilan gambar secara langsung sehingga menjadi keunikan dari penelitian ini karena pada penelitian ini menggunakan citra penyakit daun padi, maka dilakukan pengambilan gambar langsung dari sawah. Pada dataset tersebut terbagi menjadi empat kelas diantaranya yaitu kondisi daun dengan penyakit blas kolar, hawar daun bakteri, padi sehat dan tungro.

Data citra penyakit daun padi tersebut diperoleh dengan pengambilan gambar secara langsung menggunakan ponsel (*smartphone*). Pada langkah selanjutnya, kegiatan yang dilakukan adalah mempersiapkan dataset yang digunakan sebelum proses pengolahan dimulai. Data dalam penelitian ini disusun dengan mengelompokkan gambar-gambar ke dalam folder berdasarkan kategori yang sama. Artinya pada penelitian ini data dikelompokkan menjadi empat folder yaitu folder padi dengan kondisi daun blas kolar, hawar daun bakteri, padi sehat dan tungro.

Data citra penyakit pada daun padi yang telah dikelompokkan ke dalam empat folder berdasarkan kelasnya siap untuk melanjutkan ke tahap berikutnya. Pada tahap *preprocessing* ini, data citra dipisahkan menjadi data *training* (data latih) dan data *testing* untuk menguji data. Dataset ini displit menggunakan *script code* yang diatur besar persentasenya. Presentase yang digunakan adalah 70% : 30% dimana 70% untuk *training* dan 30% untuk *testing*.

### 3.2.3 Preprocessing

Istilah “*garbage in - garbage out*” yang artinya bahwa jika kualitas data input kurang baik, maka *output* yang dihasilkan juga akan kurang baik. Oleh karena itu, setelah tahap pengumpulan data, diperlukan tahap *preprocessing*. Karena citra input yang memiliki variasi yang signifikan, maka dilakukan *pre-processing* ini untuk

menyeragamkan gambar yang akan memasuki tahap analisis. Proses ini juga bertujuan untuk meningkatkan kejelasan fitur atau ciri dari citra [52]. Proses *preprocessing* ini dilakukan dengan tujuan untuk menyiapkan citra sebelum dilakukan proses lebih lanjut. Prapemrosesan atau *preprocessing* ini juga digunakan untuk mendeteksi pola, menormalisasi, menonjolkan fitur, dan menghilangkan *noise* [53]. Jenis *preprocessing* yang dilakukan yaitu *resize* yang digunakan untuk mengganti atau mengubah ukuran citra. *Resize* dilakukan untuk mempermudah dan mempercepat proses perhitungan.

### 3.2.4 Perancangan Model Klasifikasi CNN

Pada penelitian ini akan membahas terkait klasifikasi penyakit daun padi menggunakan metode *Convolutional Neural Network* (CNN) yang mampu mendeteksi fitur secara otomatis dari data yang diberikan. Klasifikasi gambar atau *image classification* menjadi tahap penting dalam penelitian ini. Fungsinya adalah untuk mendeteksi citra penyakit daun padi secara akurat. CNN dipilih untuk klasifikasi citra karena CNN dirancang khusus untuk mengolah dan memproses data dua dimensi, seperti data citra. Dalam prosesnya, CNN menggunakan proses konvolusi pada gambar, maka komputer akan memperoleh informasi melalui perkalian bagian gambar dengan filter. Pada tahapan ini menjadi penting karena digunakan untuk mencapai hasil akurasi yang tinggi pada model yang diuji. Model klasifikasi ini dirancang dengan memanfaatkan beberapa lapisan untuk melatih dan mengenali data tersebut. Data *training* kemudian digunakan untuk proses *training* CNN, di mana seluruh dataset akan melewati proses *training* pada jaringan saraf hingga dikembalikan ke awal untuk satu kali putaran (*epoch*).

Parameter pada penelitian ini mencakup optimasi, jumlah *epoch* yang dilakukan, jumlah lapisan konvolusi, ukuran *batch*, *learning rate*, ukuran kernel yang digunakan, dan *step* per *epoch*. Parameter-parameter ini disebut sebagai *hyperparameter* karena peneliti dapat menentukan nilainya sendiri. Besar *weight* akan selalu disesuaikan dengan *learning rate* (kecepatan pembelajaran) selama *epoch* masih berjalan. Selama proses *training* data, dataset secara keseluruhan akan dipelajari dan dijalankan setiap *batch*-nya. Setelah melalui proses *training* data, akan dihasilkan model yang terbaik untuk proses klasifikasi.

### 3.2.5 Pengujian Sistem

Pengujian model dilakukan untuk mengevaluasi program yang telah dikembangkan. Pada pengujian ini dimulai dengan *training* data untuk melatih model dalam mengenali *input* berupa citra penyakit pada daun padi. Tujuannya adalah agar ketika diberikan citra baru, model dapat mengklasifikasikannya dengan benar ke dalam kelas yang sesuai. Pengujian dilakukan dengan melakukan uji coba klasifikasi menggunakan data testing. Dimana kelas hasil uji coba tersebut akan dibandingkan dengan kelas aktualnya, menggunakan *confusion matrix*.

Beberapa parameter yang menjadi tolak ukur dari kinerja model CNN meliputi nilai akurasi, presisi, *recall*, dan *F1-Score*. Model algoritma yang telah dibuat akan dievaluasi untuk melihat seberapa baik kinerja dan akan diukur dengan tingkat akurasi dari hasil prediksi terhadap data uji. Pada rumus 2.6 cara untuk menghitung nilai akurasi:

$$\text{Akurasi} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\% \quad (2.6)$$

Untuk memperjelas rumus di atas, dapat dilihat melalui tabel 3. 3 *Confusion matrix* sebagai berikut:

**Tabel 3. 3 *Confusion matrix* kelas blas kolar**

		Kelas Prediksi	
		Blas Kolar	Bukan Blas Kolar
Kelas Aktual	Blas Kolar	TP	FN
	Bukan Blas Kolar	FP	TN

Keterangan:

a. *True Positive* (TP)

Dapat dikatakan *true positive* jika pada data positif hasil yang diprediksi didapatkan adalah penyakit blas kolar dan hasil klasifikasi pada kelas sebenarnya juga blas kolar. Karena hasil keduanya sama-sama blas kolar atau *true* maka disebut *true positive*.

b. *False Positive* (FP)

Jika pada hasil yang diprediksi didapatkan adalah blas kolar dengan hasil klasifikasi pada kelas sebenarnya adalah bukan blas kolar, maka dinamakan *False Positive*.

c. *False Negative* (FN)

Jika pada hasil diprediksi didapatkan adalah bukan blas kolar dengan hasil klasifikasi pada kelas sebenarnya adalah blas kolar, maka dinamakan *False Negative*.

d. *True Negative* (TN)

Pada kondisi ini kelas prediksi jika didapatkan adalah bukan blas kolar tetapi terjadi pada kelas sebenarnya adalah bukan blas kolar maka dinamakan *True Negative*. Dapat dilihat melalui tabel 3. 4 merupakan *confusion matrix* dari kelas hawar daun bakteri.

**Tabel 3. 4 *Confusion matrix* kelas hawar daun bakteri**

		Kelas Prediksi	
		Hawar daun bakteri	Bukan Hawar daun bakteri
Kelas Aktual	Hawar daun bakteri	TP	FN
	Bukan Hawar daun bakteri	FP	TN

Keterangan:

a. *True Positive* (TP)

Dapat dikatakan *true positive* jika pada data positif hasil yang diprediksi didapatkan adalah penyakit hawar daun bakteri dan hasil klasifikasi pada kelas sebenarnya juga hawar daun bakteri. Karena hasil keduanya sama-sama hawar daun bakteri atau *true* maka disebut *true positive*.

b. *False Positive* (FP)

Jika pada hasil yang diprediksi didapatkan adalah hawar daun bakteri dengan hasil klasifikasi pada kelas sebenarnya adalah bukan hawar daun bakteri, maka dinamakan *False Positive*.

c. *False Negative* (FN)

Jika pada hasil diprediksi didapatkan adalah bukan hawar daun bakteri dengan hasil klasifikasi pada kelas sebenarnya adalah hawar daun bakteri, maka dinamakan *False Negative*.

d. *True Negative* (TN)

Pada kondisi ini kelas prediksi jika didapatkan adalah bukan hawar daun bakteri tetapi terjadi pada kelas sebenarnya adalah bukan hawar daun bakteri maka dinamakan *True Negative*. Terlihat pada tabel 3.5 merupakan *confusion matrix* dari kelas padi sehat.

**Tabel 3. 5 *Confusion matrix* kelas padi sehat**

		Kelas Prediksi	
		Padi Sehat (P)	Tidak Sehat
Kelas Aktual	Padi Sehat	TP	FN
	Tidak Sehat	FP	TN

Keterangan:

a. *True Positive* (TP)

Dapat dikatakan *true positive* jika pada data positif hasil yang diprediksi didapatkan adalah padi sehat dan hasil klasifikasi pada kelas sebenarnya juga padi sehat. Karena hasil keduanya padi sehat atau *true* maka disebut *true positive*.

b. *False Positive* (FP)

Jika pada hasil yang diprediksi didapatkan adalah padi sehat dengan hasil klasifikasi pada kelas sebenarnya adalah tidak sehat, maka dinamakan *False Positive*.

c. *False Negative* (FN)

Jika pada hasil diprediksi didapatkan adalah tidak sehat dengan hasil klasifikasi pada kelas sebenarnya adalah kelas padi sehat, maka dinamakan *False Negative*.



d. *True Negative* (TN)

Pada kondisi ini kelas prediksi jika didapatkan adalah tidak sehat tetapi terjadi pada kelas sebenarnya adalah tidak sehat maka dinamakan *True Negative*. Dapat dilihat melalui tabel 3. 6 merupakan *confusion matrix* dari kelas tungro.

**Tabel 3. 6 *Confusion matrix* kelas tungro**

		Kelas Prediksi	
		Tungro	Bukan Tungro
Kelas Aktual	Tungro	TP	FN
	Bukan Tungro	FP	TN

Keterangan:

a. *True Positive* (TP)

Dapat dikatakan *true positive* jika pada data positif hasil yang diprediksi didapatkan adalah penyakit tungro dan hasil klasifikasi pada kelas sebenarnya juga tungro. Karena hasil keduanya sama-sama tungro atau *true* maka disebut *true positive*.

b. *False Positive* (FP)

Jika pada hasil yang diprediksi didapatkan adalah tungro dengan hasil klasifikasi pada kelas sebenarnya adalah bukan tungro, maka dinamakan *False Positive*.

c. *False Negative* (FN)

Jika pada hasil diprediksi didapatkan adalah bukan tungro dengan hasil klasifikasi pada kelas sebenarnya adalah penyakit tungro, maka dinamakan *False Negative*.

d. *True Negative* (TN)

Pada kondisi ini kelas prediksi jika didapatkan adalah bukan tungro tetapi terjadi pada kelas sebenarnya adalah bukan tungro, maka dinamakan kondisi *True Negative*.

Langkah berikutnya adalah melakukan validasi data untuk menguji apakah model mampu mengenali citra baru yang diinputkan. Pada tahap akhir adalah *testing* atau pengujian. Pada tahap ini, proses dilakukan dua kali yaitu *testing* per

*image* dan *testing* per folder. Pada *testing* per *image* ini, digunakan untuk melihat kelas prediksi dari setiap citra, yang kemudian akan dibandingkan dengan *testing* per folder untuk melihat apakah kelas citra penyakit pada daun padi yang dideteksi akan sama atau berbeda. Program yang sudah dibuat menggunakan bahasa *python* diuji pada tahap pengujian model. Pengujian model ini dilakukan *accuracy*, *precision*, *recall* dan *fi-score* dari model.

### **3.2.6 Analisis dan Kesimpulan**

Pada analisis masing-masing hasil pengujian model akan dibandingkan untuk menentukan performa terbaik dari masing-masing model. Tahap ini akan menghasilkan model dengan performa terbaik dan terlihat berapa akurasinya. Penarikan kesimpulan dan analisis dilakukan berdasarkan tahap sebelumnya yaitu tahap pengujian. Kesimpulan mencakup hal-hal yang berkaitan dengan tujuan dilakukannya penelitian ini.