

BAB III: METODOLOGI PENELITIAN

3.1. Objek dan Subjek Penelitian

Penelitian ini mengklasifikasi dan mendeteksi penyakit pada daun padi menggunakan model *convolutional neural network* yang dimodelkan berdasarkan arsitektur MobileNet V2. Pada proses perancangan model yang dibuat, peneliti membutuhkan referensi sebagai acuan pembuatan model. Kemudian, akan ditentukan parameter sebagai objek penelitian. Berikut data subjek dan objek dari penelitian ini.

3.1.1. Objek Penelitian

Objek yang diamati pada penelitian ini difokuskan pada 5 kelas citra pada tanaman padi yaitu sehat (*healthy*), hawar daun bakteri (*Bacterial leaf blight*), bercak coklat sempit (*Leaf Smut*), tungro, dan hawar pelepah daun (*Sheath Blight*). Gambar penyakit daun padi yang digunakan memiliki format .jpg dan .png. Dataset yang digunakan diperoleh dari beberapa sumber, yaitu dataset *Leaf Rice Disease* yang dapat di akses pada <https://www.kaggle.com/datasets/tedisetiady/leaf-rice-disease-indonesia?select=blast> , *Rice Leaf Diseases Dataset* yang dapat di akses pada <https://www.kaggle.com/datasets/vbookshelf/rice-leaf-diseases> dan pengambilan data langsung di lapangan dibantu oleh Badan Penyuluhan Pertanian (BPP) Banyumas.

3.1.2. Subjek Penelitian

Subjek penelitian ini adalah deteksi penyakit pada daun padi dengan berbagai studi literatur dengan sumber dan referensi yang jelas dan dapat dipertanggungjawabkan kebenarannya.

3.2. Alat dan Bahan Penelitian

Penelitian ini menggunakan sumber daya yang meliputi data, alat dan teknologi yang digunakan.

3.2.1. Data

1. Mendeley Data
2. Kaggle
3. Pengambilan data secara langsung yang terletak di daerah Banyumas.

3.2.2. Perangkat Lunak (*Software*)

Tabel 3.1 Tabel Perangkat Lunak yang digunakan (*Software*)

No	Nama Aplikasi	Versi	Kegunaan
1	Google Colaboratory		<i>Executable document</i> yang digunakan menulis dan menjalankan kode program
2	Tensorflow	2.8	Library pada python yang digunakan untuk machine learning yang berfokus pada <i>Deep Learning</i> seperti klasifikasi citra.
3	Python	3.7	Bahasa Pemrograman yang digunakan untuk membuat kode program

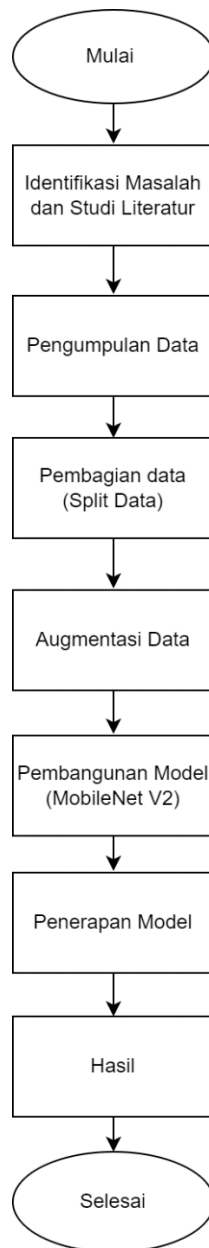
3.2.3. Perangkat Keras (*Hardware*)

Tabel 3.2 Tabel Perangkat Keras yang digunakan (*Hardware*)

No	Perangkat	Jumlah	Kegunaan
1	Acer Swift 3 Intel(R) Core (TM) i5-8250U CPU @ 1.60GHz 1.80 GHz, 12 GB RAM	1	Perangkat yang digunakan untuk proses pengerjaan tugas akhir.

3.3. Diagram Alir Penelitian

Penelitian ini dilakukan untuk mengidentifikasi penyakit pada daun padi menggunakan *transfer learning* pada algoritma *convolutional neural network* dengan arsitektur MobileNet V2. Proses penelitiannya akan digambarkan menggunakan diagram alir di bawah ini:



Gambar 3.1 Diagram Alir

Seperti yang ditunjukkan pada Gambar 3.1, proses penelitian dibagi menjadi beberapa langkah yaitu:

3.3.1. Identifikasi Masalah dan Studi Literatur

Langkah awal dalam penelitian ini setelah menentukan dan memperoleh topik penelitian adalah melakukan identifikasi masalah dan studi literatur.

1. Identifikasi masalah dilakukan bertujuan untuk menegaskan batasan-batasan permasalahan yang dibahas, sehingga penelitian yang dilakukan tidak keluar dari tujuan yang sudah ditentukan. Identifikasi masalah dilakukan melalui analisis terhadap model-model yang telah dibuat sebelumnya melalui beberapa penelitian terkait yang relevan, jelas dan dapat dipertanggungjawabkan kebenarannya.
2. Studi literatur yang dilakukan dalam penelitian ini yaitu menghimpun informasi yang relevan dengan topik atau masalah yang akan diteliti dari buku, laporan penelitian, jurnal ilmiah, dan sumber-sumber tertulis lainnya. Tujuan melakukan studi literatur yaitu agar peneliti dapat memanfaatkan semua informasi dan pemikiran-pemikiran yang relevan dengan penelitian. Sehingga diharapkan dapat membantu dalam melakukan analisis, mempraktekan teknik dan strategi dalam pelaksanaan penelitian.

3.3.2. Pengumpulan Data

Penelitian ini menggunakan 5 kelas data citra tanaman padi, yaitu sehat (*healthy*), hawar daun bakteri (*Bacterial leaf blight*), bercak coklat sempit (*Leaf Smut*), tungro, dan hawar pelepah daun (*Sheath Blight*). Pengumpulan data yang digunakan diambil dari tiga sumber, yaitu mendeley data, kaggle dan pengambilan data langsung dari lapangan.

3.3.2.1. Mendeley Data

Data yang bersumber dari Mendeley data diambil dari dataset *Dhan-Shomadhan: A Dataset of Rice Leaf Disease Classification for Bangladeshi Local Rice* yang dapat diakses pada <https://data.mendeley.com/datasets/znsxdctwtt/1> .

3.3.2.2. Kaggle

Data yang bersumber dari website kaggle diambil dari beberapa dataset berikut.

1. Dataset *Leaf Rice Disease* yang dapat di akses pada <https://www.kaggle.com/datasets/tedisetiady/leaf-rice-disease-indonesia?select=blast>.

2. *Rice Leaf Diseases Dataset* yang dapat diakses pada <https://www.kaggle.com/datasets/vbookshelf/rice-leaf-diseases>.
3. *Rice Diseases Image Dataset* yang dapat diakses di <https://www.kaggle.com/datasets/minhhuy2810/rice-diseases-image-dataset>.

3.3.2.3. *Pengambilan Data Langsung*

Pengambilan data secara langsung dilakukan melalui arahan dari Badan Penyuluhan Pertanian (BPP) Banyumas yang terletak di desa Danaraja, Banyumas. Pengambilan data dilakukan di sawah yang terletak di beberapa desa di Banyumas yang memiliki sawah dengan usia padi sekitar 1- 2 bulan.

3.3.3. *Pembagian Data (Split Data)*

Tahap ini dataset yang telah dikumpulkan akan dibagi menjadi data *Training* dan data *Testing*. Pada penelitian ini data akan dibagi menjadi 80% data *Training*, 10% data validasi dan 10% data *Testing*. Setiap bagian data berisi 5 kelas citra penyakit pada daun padi.

1. *Data Training*

Data yang digunakan untuk *Training* atau pelatihan dinamakan sebagai *Training set*. *Training set* merupakan bagian dari dataset yang dilatih untuk membuat sebuah model. Dalam penelitian ini proses *Training* dilakukan dengan menggunakan 80% data *Training* dan 10% data validasi.

2. *Data Testing*

Data yang digunakan untuk proses *Testing* atau pengujian dinamakan sebagai *Testing set*. Dalam penelitian ini digunakan 10% data untuk data *Testing*.

3.3.4. *Augmentasi Data*

Tahap selanjutnya adalah augmentasi data dimana sejumlah teknik augmentasi diterapkan menggunakan *ImageDataGenerator* yang tersedia di keras. Proses

augmentasi yang dilakukan adalah proses *rescale 1./255*, *rotation_range=20*, *horizontal_flip*, *shear_range=0.2*, *zoom_range=0.2*, *fill_mode=nearest*.

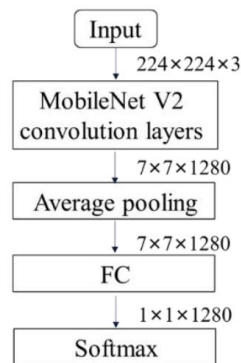
3.3.5. Pembangunan Model (MobileNet V2)

Sebelum Model dibangun, citra daun padi yang telah dikumpulkan di *resize* menjadi 224 x 224 piksel. Resolusi yang didukung oleh MobileNetV2 adalah 128, 160, 192, dan 224 px [52]. Resolusi 224 x 224 dipilih karena resolusi citra yang terlalu kecil merupakan faktor model CNN tidak mempelajari *features* dengan baik. Beberapa arsitektur CNN yang terkenal menyarankan ukuran citra (shape) yaitu 224x224x3 piksel (3 adalah channel RGB) [55].

Tahap pembangunan model dilakukan dengan menggunakan metode *transfer learning*. Metode *transfer learning* yang digunakan adalah MobileNet V2 yang sudah dilatih sebelumnya oleh ImageNet sebagai *feature extractor*. ImageNet adalah sekumpulan data yang besar yang terdiri dari 1,4 juta gambar dan 1000 kelas [71]. Dalam arsitektur MobileNet V2 dasar data akan diolah melalui empat tahap, pertama Input citra yang masuk pertama kali akan diolah pada proses *depthwise separable convolution* yang bertujuan untuk mengurangi komputasi/agar lebih sedikit parameter, sehingga menghasilkan ukuran model yang lebih kecil. Proses *depthwise separable convolution* terbagi menjadi tiga lapis konvolusi yaitu lapisan *depthwise convolution*, *1x1 convolution linearity*, dan *1x1 convolution non linearity*. Konvolusi 1x1 atau disebut *pointwise convolutional* yang digunakan untuk menggabungkan seluruh hasil konvolusi *depthwise convolution*. *Depthwise separable convolution* merupakan pengganti *drop in* untuk lapisan konvolusi standar. Konvolusi 1x1 *linearity* menggunakan ReLU sebagai fungsi aktivasinya, fungsi paling umum ReLU adalah $\psi(x) = x1_{x \geq 0}$ [72].

Kemudian menambahkan beberapa parameter lainnya, seperti Average Pooling yang digunakan untuk menghitung nilai rata-rata dari setiap baris matriks pada hasil konvolusi sebelumnya [73], *dropout* untuk mencegah model *overfitting* [74], dense layer untuk menggunakan fungsi aktivasi ReLU dan *softmax*. Fungsi aktivasi ReLU digunakan untuk melakukan normalisasi batch ketika nilainya negatif maka akan

diubah menjadi nol dan ketika tidak negatif nilainya akan tetap [75] , sedangkan fungsi aktivasi *softmax* digunakan sebagai output layer yang akan mengklasifikasi gambar ke dalam banyak kelas [76].



Gambar 3.2 arsitektur MobileNet V2

3.3.6. Penerapan Model

Tahap penerapan model terbagi menjadi dua tahapan yaitu tahap pelatihan dan pengujian.

3.3.6.1. Pelatihan

Proses pelatihan atau *Training* merupakan proses yang dilakukan untuk melatih arsitektur MobileNetV2 yang telah dibangun. Tahap pelatihan dilakukan dengan menggunakan data *Training* atau *Training set*. Pada penelitian ini, *Training* model dilakukan dengan menggunakan 3 skenario parameter *epoch* yaitu dengan menggunakan 10 *epoch*, 20 *epoch* dan 30 *epoch*. Digunakan 3 skenario tersebut adalah untuk mengukur performa model berdasarkan nilai epoch yang berbeda-beda.

3.3.6.2. Pengujian

Tahap pengujian atau *Testing* dilakukan bertujuan untuk menguji model yang digunakan dalam penelitian sudah dapat memprediksi dengan baik atau belum. Dalam penelitian ini tahap pengujian dilakukan dalam dua langkah, yaitu prediksi dan evaluasi. Langkah prediksi dilakukan menggunakan *Testing set* yang telah dibagi sebelumnya yaitu 10%. Kemudian dilakukan evaluasi model, pada penelitian

ini dilakukan dengan menggunakan *confusion matrix*. Evaluasi model dilakukan dengan memonitor jumlah empat kombinasi nilai prediksi yaitu *true positive*, *true negative*, *false positive*, dan *false negative*. Kemudian dari data tersebut dapat dihitung nilai akurasi, precision, dan recall dari model. Pada penelitian ini empat kombinasi nilai prediksi dan nilai aktual yang berbeda adalah sebagai berikut:

1. True Positive (TP)

True positive merupakan data positif yang diprediksi benar. Pada penelitian ini, daun padi yang terjangkit penyakit (class 1) dan dari model yang dibuat memprediksi terjangkit penyakit (class 1).

2. True Negative (TN)

True negative merupakan data negatif yang diprediksi benar. Pada penelitian ini, daun padi yang sehat/ tidak terjangkit penyakit (class 2) dan dari model yang dibuat memprediksi daun padi sehat/tidak terjangkit penyakit (class 2).

3. False Positive (FP) - Type 1 Error

False positive merupakan data negatif namun diprediksi sebagai data positif. Pada penelitian ini, daun padi yang sehat/ tidak terjangkit penyakit (class 2) tetapi dari model yang dibuat memprediksi daun padi terjangkit penyakit (class 1).

4. False Negative - Type 2 Error

False negative merupakan data positif namun diprediksi sebagai data negatif. Pada penelitian ini, daun padi yang terjangkit penyakit (class 1) tetapi dari model yang dibuat memprediksi daun padi sehat/ tidak terjangkit penyakit (class 2).

3.3.7. Hasil

Hasil dari penelitian ini berupa hasil dari evaluasi model yang dilakukan pada proses pengujian. Pada tahap ini, hasil pengujian akan dilihat apakah model sudah memiliki performa yang bagus untuk mengidentifikasi penyakit pada daun padi.