

BAB 3

METODE PENELITIAN

Pada bab ini, akan dijelaskan mengenai metodologi penelitian yang digunakan dalam Tugas Akhir ini.

3.1 Alat Yang Digunakan

Pada Tabel 3.1 menyajikan detail alat-alat yang digunakan dalam penelitian Tugas Akhir ini, yang mencakup perangkat keras, perangkat lunak, dan data set. Laptop dengan spesifikasi RAM 12 GB, prosesor Intel(R) Core (TM) i5-8265U, dan kartu grafis NVIDIA GeForce MX230 digunakan sebagai perangkat eksekusi program untuk menjalankan berbagai tugas penelitian.

Tabel 3. 1 Alat Penelitian

| No. | Alat | Spesifikasi | Jumlah | Justifikasi |
|-----|------------------------------|---|--------|---|
| 1 | Laptop | <ul style="list-style-type: none"> • RAM 12 GB • Intel(R) Core(TM) i5-8265U CPU @ 1.60GHz (8 CPUs), ~1.8GHz • NVIDIA GeForce MX230 | 1 | Perangkat Eksekusi program |
| 2 | Operating System | Windows 11 Home Single Language 64-bit (10.0, Build 22000) (22000.co_release.210604-1628) | 1 | Sistem operasi untuk pengembangan penelitian |
| 3 | <i>Deep learning</i> IDE | Jupyter Notebook | 1 | Platform lingkungan pengembangan <i>deep learning</i> |
| 4 | <i>Library Deep learning</i> | TensorFlow | 1 | Framework pengembangan <i>deep learning</i> |

| No. | Alat | Spesifikasi | Jumlah | Justifikasi |
|-----|------------------------------------|---|--------|--|
| 5 | Python Preprocessing Library | <ul style="list-style-type: none"> • Numpy • Matplotlib | 1 | Modul pembaca data penyakit daun stroberi dan <i>Preprocessing</i> |

3.1.1 Data Set

Penelitian ini dilakukan dengan menggunakan dataset penyakit daun stroberi yang didapatkan melalui pengambilan citra secara langsung dan dataset yang bersumber dari kaggle.com. Dataset yang dibutuhkan berupa gambar daun Stroberi dengan kondisi atau gejala seperti daun sehat, *tipburn*, dan *leaf spot*.

Setiap citra memiliki label yang menandakan jenis penyakit yang terjadi pada daun stroberi tersebut. Dataset ini merupakan sumber data utama yang digunakan dalam penelitian ini untuk melatih dan menguji model klasifikasi penyakit daun stroberi menggunakan *Convolutional Neural Network (CNN)*.

Pada Tabel 3.2 menyajikan spesifikasi dataset, termasuk jumlah total citra, kelas penyakit, resolusi citra (piksel), aspek rasio, kedalaman warna (*bit*), dan format file. Dataset ini, yang mencakup citra yang diharapkan dapat meningkatkan pemahaman dan kemampuan dalam mengenali serta mengklasifikasikan penyakit pada tanaman stroberi berdasarkan citra daun yang diperoleh..

Tabel 3. 2 Spesifikasi Dataset

| Dataset | Jumlah | Kelas | Resolusi Citra (px) | Aspek Rasio | bit size | Format File |
|---------------------|--------|---|---------------------|-------------|------------------------|-------------|
| ercanavsar | 1433 | <ul style="list-style-type: none"> • Normal • Tiburn | 5664 x 4248 | 1:1 | 24 bit (True Color) | jpg |
| Asheniran- ga | 1565 | <ul style="list-style-type: none"> • Normal • Leaf Spot | 256x256 | 1:1 | 24 bit True Color) | jpg |
| Abdallaha- lidev | 1565 | <ul style="list-style-type: none"> • Normal • Leaf Spot | 256x256 | 1:1 | 24 bit True Color) | jpg |
| dev523 | 4498 | <ul style="list-style-type: none"> • Normal • Leaf Spot | 256x256 | 1:1 | 24 bit True Color) | jpg |

3.1.2 Perangkat Keras (*Hardware*)

Perangkat keras yang digunakan dalam proses penelitian ini adalah sebuah Personal Computer (PC) dengan spesifikasi sebagai berikut :

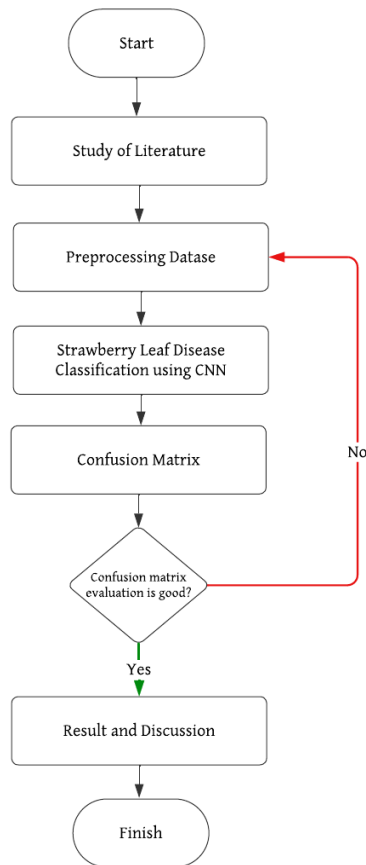
- 1) Intel core i5
- 2) Windows 10 Pro 64 – Bit
- 3) RAM 8 GB
- 4) Memory Harddisk

3.1.3 Perangkat Lunak (*Software*)

Perangkat lunak (*software*) yang digunakan dalam sistem ini adalah Google Colab dan Anaconda. Dengan menggunakan *Google Colab* dan *Anaconda*, peneliti dapat menjalankan kode pemrosesan warna dan analisis citra pada tanaman stroberi dengan efisien dan efektif. Kedua perangkat lunak ini menyediakan alat dan sumber daya yang diperlukan untuk melakukan analisis warna dan pemrosesan citra dengan mudah dan dapat diandalkan, sehingga membantu peneliti dalam memperoleh hasil yang akurat dan berkualitas tinggi dalam penelitiannya.

3.2 Diagram Alur Penelitian

Diagram alur penelitian yang tersemat pada Gambar 3.1 penelitian ini, langkah-langkah yang terstruktur, dimulai dari studi literatur hingga tahap evaluasi menggunakan *confusion matrix*, alur ini membentuk landasan sistematis bagi penelitian mengenai klasifikasi penyakit daun stroberi menggunakan *Convolutional Neural Network* (CNN). Melalui panduan ini, diharapkan dapat tercipta pemahaman yang mendalam mengenai efektivitas model dalam mengidentifikasi dan mengklasifikasikan penyakit pada tanaman daun stroberi



Gambar 3.1 Diagram Alur Penelitian

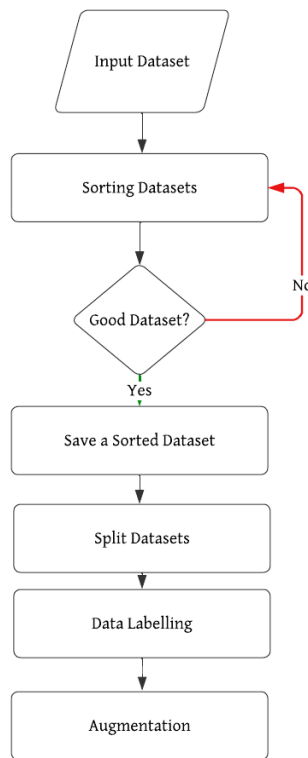
3.2.1 Perumusan Masalah

Perumusan masalah dibuat dengan mempertimbangkan beberapa hal yaitu dari segi metode, hasil penelitian, dan analisis penelitian yang akan dilakukan. Tujuan dibuatnya perumusan masalah dalam skripsi ini yaitu untuk mengetahui bagaimana sistem yang seharusnya dibuat untuk mendeteksi penyakit daun tanaman stroberi menggunakan perbandingan fitur dari dua algoritma VGG-16 dan ResNet-50.

3.2.2 Studi Literatur

Tahap studi literatur merupakan tahap untuk mencari pemahaman konsep dan pemahaman masalah *deep learning* melalui jurnal, buku ilmiah maupun penelitian sebelumnya yang selanjutnya digunakan untuk acuan penulisan dan penelitian yang akan dilakukan.

3.2.3 Preprocessing Data



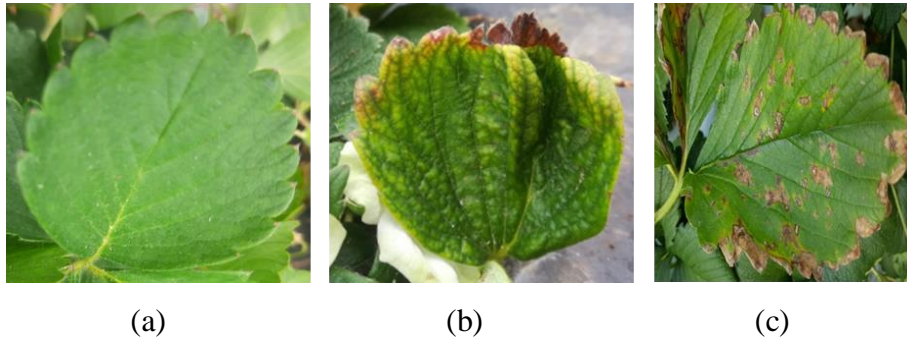
Gambar 3. 2 Diagram Alir *Preprocessing* Dataset

Gambar 3.2 yang merupakan diagram alur preprocessing dataset, memberikan gambaran rinci tentang langkah-langkah yang diambil dari input dataset hingga pembentukan dataset yang siap untuk pelatihan model. Proses dimulai dengan penyortiran dataset, di mana dataset diarahkan ke good dataset jika memenuhi kriteria tertentu; jika ya, dataset tersebut disimpan sebagai dataset yang telah diurutkan. Selanjutnya, dataset yang telah diurutkan dibagi menjadi bagian-bagian yang sesuai, dilanjutkan dengan tahap pemberian label data dan augmentasi. Namun, jika suatu dataset tidak memenuhi kriteria (no), maka proses kembali ke tahap penyortiran dataset untuk peninjauan ulang. Dengan gambaran ini, diagram alur tersebut memberikan pemahaman mendalam mengenai proses *preprocessing dataset* yang dilakukan sebelum mengintegrasikan data ke dalam model klasifikasi menggunakan *Convolutional Neural Network* (CNN).

3.2.3.1 *Collecting Dataset*

Pada proses pengumpulan dataset, indikator utama yang menjadi dataset dalam penelitian ini adalah gambar daun Stroberi dengan kondisi permukaan daun

memiliki gejala; daun sehat, *tipburn*, dan *leaf spot*. Dataset tersebut bersumber dari *kaggle.com* dan *datadryad.com* [11] [12]. Pada Gambar 3.3 menunjukkan visualisasi dari daun Stroberi dengan label. Setiap gambar memberikan representasi visual yang penting untuk pemahaman dan analisis terkait kondisi kesehatan tanaman stroberi dalam konteks penelitian ini.



Gambar 3. 3 Tampilan (a) *Healthy Leaf*, (b) *Tipburn* (c) *Leaf Spot* [11] [12]

3.2.3.2 Split Dataset

Dataset daun stroberi tersebut selanjutnya akan di *input* ke folder dataset yang dikategorikan; *training* dataset dan *validation* dataset. Proses *input* data yang dikelompokkan menjadi 2 bagian, yaitu 80% dataset di *input* ke folder *training* dataset dan 20% ke *validation* dataset. Pembagian data tersebut membantu menghindari kesalahan pengukuran performa model.

3.2.3.3 Data Labelling

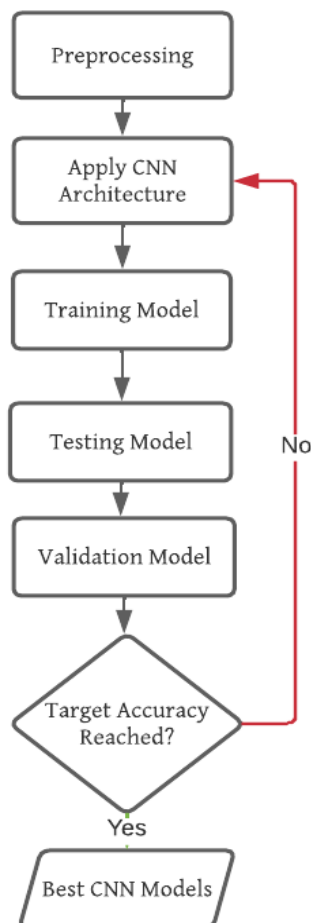
Proses pelabelan data dilakukan pada citra (gambar) yang dihasilkan oleh *image generator*. Setiap gambar diberikan label yang sesuai dengan klasifikasi yang telah disusun sebelumnya. Hal ini memungkinkan pengidentifikasi kelas atau kategori pada setiap data gambar. Setelah pelabelan, dilakukan verifikasi dan koreksi terhadap hasil pelabelan. Hal ini penting untuk memastikan keakuratan dan konsistensi dalam proses pelabelan. Jika terdapat kesalahan pelabelan, perlu dilakukan koreksi agar dataset yang digunakan memiliki label yang benar dan konsisten. Terakhir, dataset dapat dipisahkan menjadi set pelatihan, set validasi, dan set pengujian sesuai dengan kebutuhan. Pemisahan ini membantu dalam menghindari kontaminasi dan memastikan evaluasi yang obyektif pada tahap selanjutnya dalam pengembangan model *machine learning* atau *deep learning*.

Dengan dataset yang terlabel dengan baik, model *machine learning* atau *deep learning* dapat dilatih dan dievaluasi dengan lebih akurat dan efektif.

3.2.3.4 Augmentation Dataset

Dataset *training* akan diaugmentasi agar menghindari overfitting pada pelatihan klasifikasi. Hal tersebut bertujuan agar model dapat mengenali variasi dasar pada citra obyek daun Stroberi. Transformasi yang digunakan dalam augmentasi dataset adalah rotasi, refleksi secara horizontal dan vertikal, *zoom*, *scale*, *width* dan *height shift*, serta *shear*. Hasil data *training* yang sudah diaugmentasi dan data yang telah tervalidasi disimpan pada *storage* data.

3.2.4 Klasifikasi Perancangan Penyakit Daun Menggunakan CNN



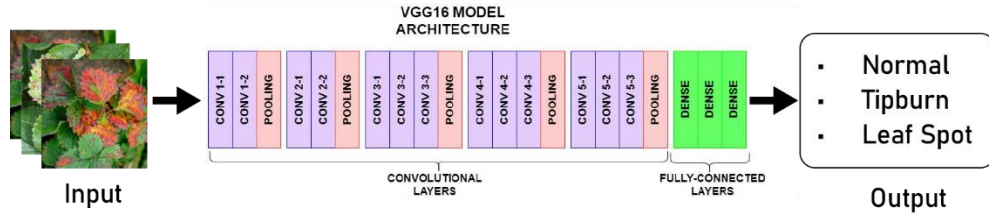
Gambar 3. 4 Klasifikasi Penyakit Daun Stroberi Menggunakan CNN

Pada Gambar 3.4 menunjukkan alur pengklasifikasian penyakit daun stroberi menggunakan *Convolutional Neural Network* (CNN) yang dilakukan dengan beberapa tahap yakni:

- 1) *Preprocessing Data*: Data yang digunakan untuk pelatihan model *deep learning* biasanya perlu dipreproses agar lebih mudah dikelola dan digunakan. Ini bisa termasuk konversi format data, normalisasi nilai, augmentasi data, dan lain-lain. Tujuan dari *preprocessing* data adalah untuk meningkatkan kualitas data dan akurasi model.
- 2) Implementasi Arsitektur CNN: Selanjutnya, Arsitektur CNN digunakan dengan menggunakan *library framework tensorflow* dan Keras [25]. Tahapan ini juga melibatkan penentuan jumlah *layer*, jumlah filter pada tiap *layer*, fungsi aktivasi, dan lain sebagainya.
- 3) Pelatihan Model: Setelah pembuatan model, model dilatih dengan menggunakan data pelatihan. Pada tahap ini, model mempelajari pola-pola dalam data dan menghasilkan model yang dapat digunakan untuk melakukan prediksi. Pelatihan model dilakukan dengan menggunakan metode backpropagation dan optimisasi pada *loss function*.
- 4) *Testing Model*: Setelah model berhasil dilatih, tahap terakhir adalah pengujian model pada data uji. Pada tahap ini, model diuji untuk melihat apakah dapat mengklasifikasikan data uji dengan akurat atau tidak.
- 5) *Confusion Matrix*: Tahap *confusion matrix* membantu mengevaluasi kinerja model dan memahami seberapa baik model dapat mengklasifikasikan data dengan benar. Ini penting untuk mengukur keefektifan model dan meningkatkan kinerjanya jika perlu.
- 6) Best CNN Models: Jika model telah dilatih dan diuji dengan baik, model dapat diekspor ke dalam format yang dapat digunakan oleh aplikasi. Salah satu format yang umum digunakan adalah format berkas protobuf, yang bisa digunakan oleh berbagai jenis aplikasi. Dalam kesimpulannya, untuk melatih sebuah model CNN, perlu melakukan beberapa tahap, yaitu *preprocessing* data, pembuatan model, pelatihan model, validasi model, pengujian model, dan ekspor model. Tahap-

tahap ini penting untuk mendapatkan model yang akurat dan dapat digunakan pada aplikasi yang diinginkan.

a. Penerapan Model CNN dengan Arsitektur VGG-16



Gambar 3.5 Tahapan Klasifikasi Menggunakan VGG-16

Untuk mengklasifikasikan penyakit daun stroberi, Model *Convolutional Neural Network* (CNN) dengan arsitektur VGG-16 (VGGNet) dapat digunakan. Langkah-langkahnya dimulai dengan pengumpulan dataset berisi gambar daun stroberi yang mencakup kategori berbeda, termasuk daun yang sehat dan daun yang terkena penyakit.

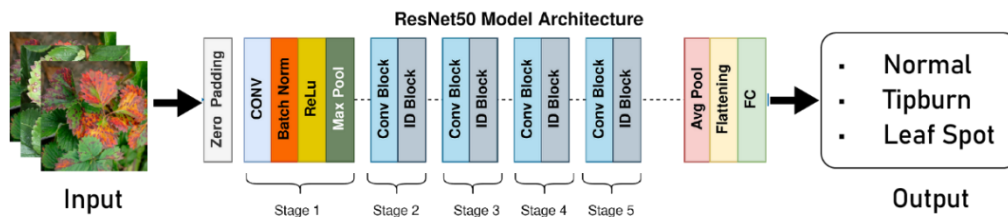
Gambar 3.5 merupakan arsitektur VGG-16 dibangun dengan lapisan konvolusi, lapisan pooling, dan lapisan terhubung penuh. Fungsi aktivasi ReLU digunakan pada lapisan konvolusi dan lapisan terhubung penuh, kecuali pada lapisan *output* yang menggunakan fungsi aktivasi *softmax* untuk menghasilkan probabilitas kelas.

Setelah model CNN dibangun, tahap pelatihan dimulai menggunakan subset pelatihan. Model akan memperbarui bobot dan biasanya melalui proses *backpropagation* dan optimisasi berbasis gradien, seperti SGD atau Adam, dengan tujuan meminimalkan fungsi kerugian. Selanjutnya, model dievaluasi menggunakan subset validasi untuk mengukur akurasi, presisi, *recall*, dan metrik evaluasi lainnya. Jika hasil validasi tidak memuaskan, dapat dilakukan penyesuaian *hyperparameter* atau penyempurnaan lainnya.

Setelah proses pelatihan dan validasi selesai, model diuji menggunakan subset pengujian yang sebelumnya tidak pernah dilihat. Akurasi dan metrik evaluasi lainnya dihitung untuk mengevaluasi performa model pada data yang belum pernah dikenal sebelumnya. Jika model tidak memberikan teknik hasil yang memuaskan, langkah penyempurnaan seperti penambahan regularisasi, perubahan arsitektur, atau penambahan data pelatihan tambahan dapat dilakukan.

Dengan mengikuti langkah-langkah ini, Model CNN dengan arsitektur VGG-16 dapat digunakan untuk mengklasifikasikan penyakit daun stroberi, membantu dalam deteksi dan pengendalian penyakit pada tanaman stroberi secara efektif.

b. Penerapan Model CNN dengan Arsitektur ResNet-50



Gambar 3. 6 Tahapan Klasifikasi Menggunakan ResNet-50

Pada Gambar 3.6 menunjukkan tahapan klasifikasi menggunakan model ResNet-50, yang dibangun dengan memanfaatkan framework deep learning seperti TensorFlow. Tahapan ini mencakup proses esensial dalam penerapan model klasifikasi, yang terutama didasarkan pada arsitektur ResNet-50. Model ResNet-50 dibangun dengan menggunakan framework *deep learning* seperti TensorFlow atau *PyTorch*. Arsitektur ini terdiri dari 50 lapisan, termasuk konvolusi, *max pooling*, dan blok identitas residual. Blok identitas residual merupakan komponen utama yang memungkinkan aliran informasi yang langsung melalui lapisan jaringan tanpa gangguan, mengatasi masalah degradasi performa pada jaringan yang dalam. Setelah model dibangun, langkah selanjutnya adalah melatihnya dengan dataset yang sesuai. Pelatihan melibatkan optimisasi fungsi kerugian dengan memperbarui bobot dan bias model berdasarkan gradien *loss* yang dihasilkan. Validasi dilakukan pada dataset validasi yang tidak digunakan selama pelatihan, untuk mengevaluasi generalisasi model. Jika diperlukan, penyetelan *hyperparameter* dapat dilakukan untuk meningkatkan performa model. Evaluasi akhir dilakukan pada dataset pengujian terpisah, dengan menghitung matrik seperti akurasi, presisi, *recall*, atau *F1-score*. Setelah dinilai berhasil, model dapat digunakan untuk inferensi pada data citra baru. Dengan demikian, penggunaan dan proses model CNN dengan arsitektur ResNet-50 melibatkan langkah-langkah ini untuk membangun, melatih, dan menggunakan model tersebut.

3.2.5 Confusion Matrix

Confusion matrix merupakan pengukuran yang dapat digunakan untuk menghitung kinerja atau tingkat kebenaran pada proses klasifikasi. Tabel *confusion matrix* ditunjukkan pada Tabel 3.3 berikut.

Tabel 3. 3 Confusion Matrix

| N = 3000 | | <i>Actual Values</i> | |
|-------------------------|------------------|----------------------|---------------------|
| | | <i>Positive (1)</i> | <i>Negative (0)</i> |
| <i>Predicted Values</i> | <i>True (1)</i> | TP | FP |
| | <i>False (0)</i> | FN | TN |

Keterangan:

- TP (*True Positive*) merupakan banyaknya data yang memiliki prediksi matang kemudian terbaca oleh sistem sebagai matang.
- FP (*False Positive*) merupakan banyaknya data yang memiliki prediksi matang kemudian terbaca oleh sistem sebagai belum matang.
- FN (*False negative*) merupakan banyaknya data yang memiliki prediksi belum matang kemudian terbaca oleh sistem sebagai belum matang.
- TN (*True Negative*) merupakan banyaknya data yang memiliki prediksi belum matang kemudian terbaca oleh sistem sebagai matang.