

BAB III: METODOLOGI PENELITIAN

3.1. Objek Penelitian

Objek pada penelitian ini adalah berupa citra bunga yang terdiri dari 5 jenis bunga *dry flower*, yaitu *pampas*, *lagurus*, *baby breath*, *cotton flower*, dan *setaria*.

3.2. Alat dan Bahan Penelitian

3.2.1. 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	Jupyter Notebook		Aplikasi web gratis <i>open source</i> yang digunakan untuk menulis dan menjalankan kode program.
2	Python	3.7	Bahasa pemrograman yang digunakan untuk membuat kode program dalam membangun model klasifikasi bunga.
3	Tensorflow	2.8	<i>Library</i> untuk <i>machine learning</i> yang berfokus pada <i>deep learning</i> seperti klasifikasi citra.
3	Bing Image		<i>Search Engine</i> yang digunakan untuk pencarian citra bunga yang akan dijadikan sebagai dataset.
4	Fatkun Batch Download Image		Ekstensi yang digunakan untuk mendownload citra bunga yang akan dijadikan sebagai dataset.

3.2.2. Perangkat Keras (*Hardware*)

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

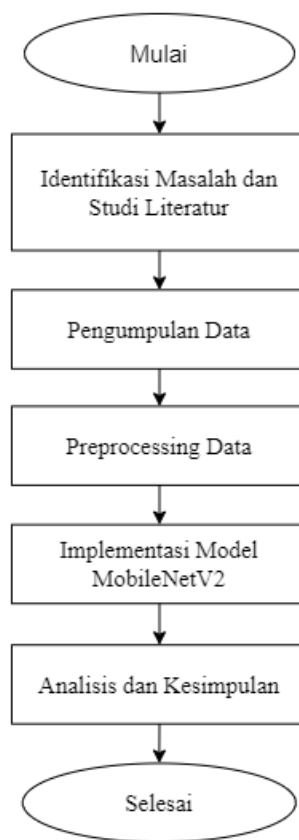
No	Perangkat	Jumlah	Kegunaan
1	Laptop HP Pavilion X360 (Processor: Intel ® Core™ i5-8250U CPU @1.60GHz 1.80 GHz)	1	Digunakan untuk proses pengerjaan tugas akhir.

3.2.3. Bahan Penelitian

Adapun bahan utama dari penelitian ini adalah data citra bunga *dry flower* yang terdiri dari 5 jenis bunga *dry flower*, yaitu *pampas*, *lagurus*, *baby breath*, *cotton flower*, dan *setaria* yang berjumlah 1.288 citra.

3.3. Diagram Alir Penelitian

Terdapat beberapa tahapan dalam melakukan penelitian untuk klasifikasi jenis bunga *dry flower* menggunakan MobileNetV2. Tahapan yang dilakukan diawali dengan identifikasi masalah dan studi literatur, kemudian dilanjutkan pengumpulan data, *preprocessing* data, implementasi model MobileNetV2, dan diakhiri dengan analisis dan kesimpulan. Pada tahap pertama, yaitu mengidentifikasi permasalahan dan studi literatur sebagai landasan dalam melakukan penelitian. Tahap kedua, dilakukan pengumpulan data yang akan digunakan pada penelitian. Tahap ketiga adalah *preprocessing* data sebelum diterapkan pada tahap berikutnya. Tahap keempat adalah menerapkan data yang telah dilakukan *preprocessing* pada proses pemodelan menggunakan MobileNetV2 untuk klasifikasi jenis bunga *dry flower*. Tahap terakhir adalah dilakukan analisis yang kemudian dari hasil analisis dapat diambil kesimpulan. Diagram alir tahapan penelitian yang dilakukan digambarkan pada Gambar 3.1.



Gambar 3.1 Diagram Alir Penelitian

3.3.1. Identifikasi Masalah dan Studi Literatur

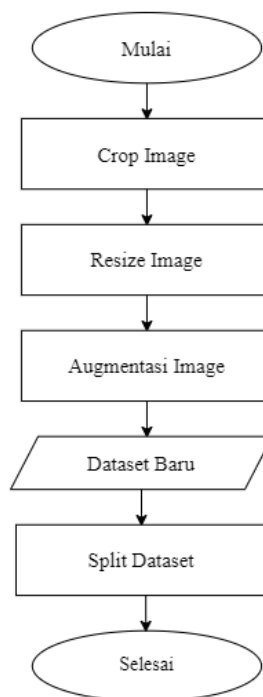
Langkah awal dalam pelaksanaan penelitian ini adalah melakukan identifikasi permasalahan dan studi literatur. Identifikasi masalah dilakukan dengan menentukan topik penelitian yang akan dilakukan. Selanjutnya dilakukan studi literatur dengan mengumpulkan beberapa referensi seperti jurnal nasional, jurnal internasional, website artikel, e-book dan kode program yang berhubungan dengan penelitian.

3.3.2. Pengumpulan Data

Data yang digunakan pada penelitian ini adalah berupa kumpulan data citra bunga *dry flower* yang terdiri dari 5 kelas jenis bunga, yaitu *pampas*, *lagurus*, *baby breath*, *cotton flower*, dan *setaria*. Pengumpulan data bunga diambil melalui situs *Bing Image* dan diunduh menggunakan bantuan ekstensi *Fatkun Batch Download Image*. Data yang terkumpul sebanyak 1.288 data citra dari 5 kelas jenis bunga,

yaitu *baby breath*, *cotton flower*, *lagurus*, *pampas*, dan *setaria* yang dikelompokkan pada folder yang bersesuaian.

3.3.3. Preprocessing Data



Gambar 3.2 Diagram Tahap *Preprocessing* Data

Sebelum dilakukan implementasi model, tahap yang dilakukan adalah *preprocessing* yang dimulai dengan *crop image*, *resize image*, *augmentasi image*, dan split dataset menjadi 3 bagian, yaitu data *train*, data *val*, dan data *test*. Berikut ini penjelasan dari tiap tahap *preprocessing* yang dilakukan seperti pada gambar 3.2.

3.3.3.1. Crop image

Tahap *preprocessing* yang pertama dilakukan adalah melakukan *cropping* ke dalam bentuk *square* dengan *ratio* 1:1 secara manual untuk memotong bagian inti dari citra bunga. Pemotongan dilakukan agar citra menjadi lebih fokus ke bagian bunga.

3.3.3.2. *Resize image*

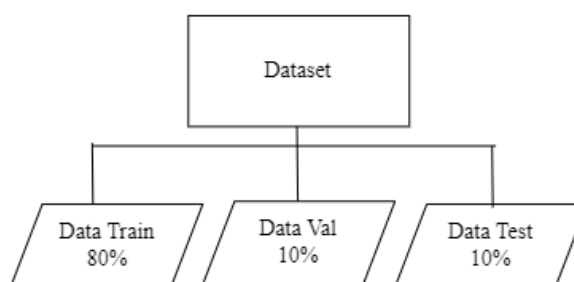
Tahap selanjutnya yaitu *resizing* dengan mengubah ukuran piksel citra asli menjadi 224 x 224 piksel. Pada penelitian ini, ukuran 224 x 224 dipilih karena resolusi tersebut merupakan resolusi standar dari MobileNetV2 yang telah dilatih pada dataset ImageNet. Perubahan ukuran sebelum dilakukan augmentasi bertujuan untuk mengurangi beban iterasi yang diperlukan saat mengubah ukuran.

3.3.3.3. Augmentasi Data

Tahap selanjutnya dilakukan data augmentasi untuk memperbanyak jumlah dan variasi data. Augmentasi data yang dilakukan diantaranya yaitu dengan *rotation_range*, *width_shift_range*, *height_shift_range*, *horizontal_flip*, *vertical_flip*, dan *fill_mode*.

3.3.3.4. *Split Dataset*

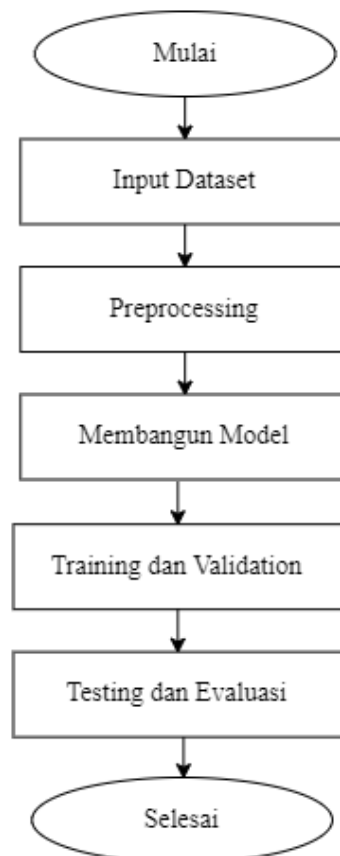
Tahap terakhir pada proses *preprocessing* adalah split dataset. Dataset yang telah diaugmentasi kemudian dibagi menjadi 3, yaitu 80% data *training* yaitu data yang digunakan untuk melatih model klasifikasi, 10% data *validation* untuk memvalidasi model klasifikasi, dan 10% data *testing* yaitu data yang digunakan sebagai data *testing* untuk menguji model klasifikasi. Pembagian dataset tersebut dilakukan dengan mempertimbangkan jumlah data yang sedikit. Diagram pembagian dataset pada penelitian ini digambarkan pada Gambar 3.3.



Gambar 3.3 Diagram Split Dataset

3.3.4. Implementasi Model MobileNetV2

Diagram alir penerapan model *transfer learning* MobileNetV2 ditunjukkan pada gambar 3.4.



Gambar 3.4 Diagram Alir Penerapan Model MobileNetV2

Berikut ini penjelasan dari diagram alir penerapan metode pada gambar 3.4.

3.3.4.1. *Input Dataset*

Tahap pertama yang dilakukan adalah menginputkan dataset berupa data citra bunga yang telah dilakukan *preprocessing* yang terdiri dari 5 kelas dengan ukuran $224 \times 224 \times 3$. Nilai 224×224 merupakan ukuran citra bunga yang digunakan untuk menyesuaikan input dari model MobileNetV2. Nilai 3 merupakan citra yang memiliki 3 buah channel, yaitu *Red*, *Green*, dan *Blue* (RGB).

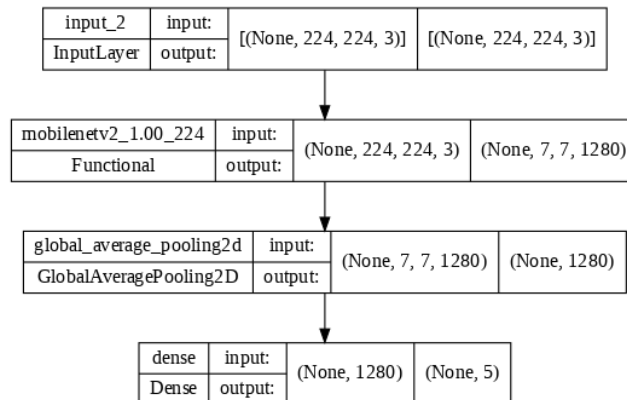
3.3.4.2. *Preprocessing*

Selanjutnya dilakukan *preprocessing* kedua pada data citra inputan dengan melakukan *rescaling* untuk mengubah ukuran data piksel RGB (0-255) menjadi rentang angka (0-1) untuk memudahkan proses *training* data.

3.3.4.3. Membangun Arsitektur Model

Tahap selanjutnya adalah membangun model dengan metode *transfer learning*. Model yang digunakan adalah MobileNetV2 yang telah dilatih sebelumnya menggunakan dataset ImageNet sebagai ekstraksi fitur. ImageNet adalah sekumpulan data yang besar yang terdiri dari 1,4 juta gambar dan 1000 kelas. Lapisan MobileNetV2 yang digunakan untuk ekstraksi fitur adalah lapisan terakhir sebelum *flatten operation (bottleneck layer)*. Inisialisasi model MobileNetV2 dilakukan menggunakan 3 saluran input dengan bobot *ImageNet*. *Include_top=False* digunakan karena lapisan yang digunakan untuk ekstraksi fitur adalah lapisan terakhir sebelum *flatten operation (bottleneck layer)*. Selanjutnya, semua lapisan pada MobileNetV2 dibekukan (*freeze*) untuk mencegah bobot lapisan tertentu diperbarui selama pelatihan [71]. *Pre-trained* model MobileNetV2 digunakan untuk ekstraksi fitur.

Kemudian menambahkan beberapa parameter lainnya seperti *GlobalAveragePooling2D* yang digunakan untuk menghitung nilai rata-rata dari setiap baris matriks pada hasil konvolusi sebelumnya dan *Dense Layer* menggunakan fungsi aktivasi *softmax*. Fungsi aktivasi *softmax* digunakan untuk klasifikasi citra ke dalam beberapa kelas sebagai output layer. Pada gambar 3.5 merupakan gambar perancangan arsitektur menggunakan model MobileNetV2.



Gambar 3.5 Arsitektur Model.

3.3.4.4. *Training dan Validation*

Proses *training* dilakukan untuk melatih arsitektur MobileNetV2 yang telah dibangun untuk mendapatkan model. Data *train* digunakan pada proses ini yang kemudian akan diperoleh hasil akurasi. Pada saat *training* model, beberapa *hyperparameter* yang digunakan diantaranya adalah *Adam Optimizer* dengan *learning rate (lr)* yang bernilai 0,0001 dan 0,0003. Parameter *epoch* yang digunakan bernilai 10, 15, dan 30 pada setiap model untuk melihat akurasi dan *loss*. Proses *validation* bertujuan untuk mengukur performa yang diperoleh dengan melihat *training loss* dan *validation loss*.

3.3.4.5. *Testing dan Evaluasi*

Model yang telah dilakukan *training* dan *validation* selanjutnya dilakukan *testing*. *Testing* dilakukan untuk memastikan model dapat memprediksi dengan baik. Data *test* digunakan untuk melakukan pengujian atau *testing*. Evaluasi model menggunakan *confusion matrix* dan *classification report*. Baris pada matriks mewakili kelas yang sebenarnya, sedangkan kolom mewakili kelas yang diprediksi. Performa metrics yang dievaluasi yaitu *Accuracy*, *Precision*, *Recall*, dan *F1-score*. Tujuan *confusion matrix* adalah untuk membandingkan nilai prediksi dengan nilai aktual dimana nilai ini terbagi ke dalam 4 kategori, yaitu:

1. *True Positif (TP)*, yaitu prediksi positif dan nilai sebenarnya positif.
2. *True Negatif (TN)*, yaitu hasil prediksi negatif dan nilai sebenarnya negatif.

3. *False Positif* (FP), yaitu prediksi positif dan nilai sebenarnya negatif.
4. *False Negatif* (FN), yaitu prediksi negatif dan nilai sebenarnya positif.

Berdasarkan *confusion matrix* tersebut, maka *performance metrics* yang direpresentasikan sebagai berikut:

1. *Accuracy*, yaitu untuk menghitung seberapa akurat model dalam mengklasifikasikan dengan benar dan direpresentasikan dengan Persamaan (3,1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3,1)$$

2. *Precision*, yaitu mengukur akurasi model berdasarkan prediksi positif dengan berapa banyak diantaranya yang sebenarnya positif. *Precision* direpresentasikan dengan Persamaan (3,2).

$$Precision = \frac{TP}{TP+FP} \quad (3,2)$$

3. *Recall*, yaitu mengukur kesesuaian dengan keakuratan contoh positif dan mengacu pada berapa banyak contoh kelas positif diberi label dengan benar. *Recall* direpresentasikan dengan Persamaan (3,3).

$$Recall = \frac{TP}{TP+FN} \quad (3,3)$$

4. *F1-score*, yaitu perbandingan rata-rata *precision* dan *recall* yang dibobotkan. *F1-score* direpresentasikan dengan Persamaan (3,4).

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (3,4)$$

3.3.5. Analisis dan Kesimpulan

Selanjutnya akan dilakukan analisis terhadap implementasi metode yang telah dilakukan pada penelitian sehingga dapat diambil kesimpulan. Kemudian diberikan saran yang dapat dijadikan sebagai referensi penelitian selanjutnya. Tahap ini kemudian akan ditulis pada sub bab kesimpulan dan saran.