

## BAB III

### METODE PENELITIAN

#### 3.1 Subjek Penelitian dan Objek Penelitian

Subjek dalam penelitian ini adalah buah pisang Cavendish dengan kualitas dan kematangan dibagi menjadi empat kategori, yaitu mentah kualitas bagus (MHBS), mentah kualitas buruk (MHBK), matang kualitas bagus (MGBS), dan matang kualitas buruk (MGBK) dan objek pada penelitian ini adalah hasil klasifikasi menggunakan CNN.

#### 3.2 Alat dan Bahan Penelitian

Alat yang digunakan dalam penelitian ini adalah laptop dengan spesifikasi:

1. Spesifikasi perangkat keras
  - Laptop Huawei MateBook D14 Intel(R) Core (TM) i3-10110U CPU @ 2.10GHz 2.59 GHz RAM 8 GB
  - Realme Narzo 50 5G 48mp Kamera Belakang
2. Spesifikasi perangkat lunak
  - Google Colaboratory
  - Python 3.10
  - Windows 11

Bahan yang digunakan dalam penelitian ini adalah kumpulan dataset citra buah pisang Cavendish berdasarkan kualitas dan kematangan, berikut tabel spesifikasi bahan data citra yang digunakan pada Tabel 3.1.

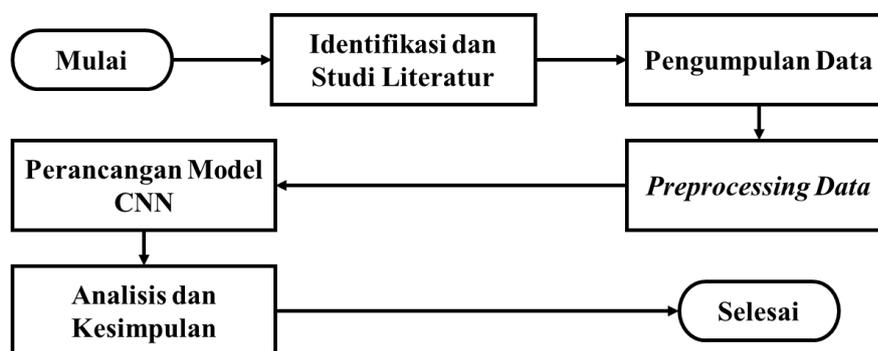
Tabel 3. 1 Spesifikasi data citra yang digunakan

No.	Spesifikasi Data	Keterangan
1	Jumlah data	1000
2	Jenis data dan <i>file type</i>	Image (.jpg)
3	Jumlah kategori/kelas	4 (mentah kualitas bagus (MHBS), mentah kualitas buruk (MHBK), matang kualitas bagus (MGBS), dan matang kualitas buruk (MGBK))
4	Ukuran citra	150 x 150 pixel

Bahan penelitian ini di dapat melalui dokumentasi pribadi yang telah dilakukan tahap *preprocessing* yang di mana setiap kategorinya memiliki 250 citra.

### 3.3 Diagram Alir Penelitian

Untuk menjalankan penelitian ini, penulis telah mengatur langkah-langkah yang harus diikuti untuk menggambarkan proses penelitian. Di bawah ini merupakan diagram alir yang dapat dilihat pada Gambar 3.1.



Gambar 3.1 Diagram Alir Penelitian

#### 3.3.1 Tahap Identifikasi

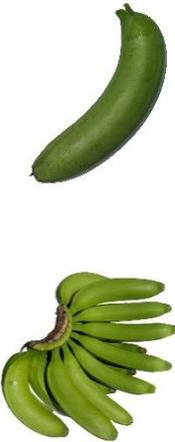
Penelitian dimulai dari identifikasi masalah yang ada di masyarakat sekitar di bidang pertanian khususnya untuk tanaman pisang Cavendish. Kemudian, dilanjutkan dengan melakukan studi literatur dengan membaca berbagai sumber seperti jurnal, buku, dan artikel online tentang topik klasifikasi citra dan pisang. Tujuannya adalah untuk memperoleh informasi dan pemahaman tentang objek atau metode yang dapat digunakan untuk mencari solusi yang berbeda dari penelitian sebelumnya, sehingga terdapat perkembangan dan peningkatan dalam segi performa dan hasil penelitian serta dapat menjadi acuan bagi penelitian selanjutnya.

#### 3.3.2 Pengumpulan Data

Data yang digunakan dalam penelitian ini terdiri dari citra pisang Cavendish dalam keadaan satuan dan sisir dengan 4 kategori kualitas dan kematangan yaitu: pisang mentah kualitas bagus (MHBS), pisang mentah kualitas buruk (MHBK), pisang matang kualitas bagus (MGBS), dan pisang matang kualitas buruk (MGBK). Pengumpulan data pisang dilakukan dengan

sudut pengambilan gambar sekitar 60° hingga 70° dan jarak sekitar 35 cm hingga 40 cm dari objek pisang. Pengambilan gambar dilakukan pada rentang jam 11 sampai 13 siang. Total ada 400 gambar yang terkumpul dalam format .jpg, dengan masing-masing kategori memiliki 100 gambar.

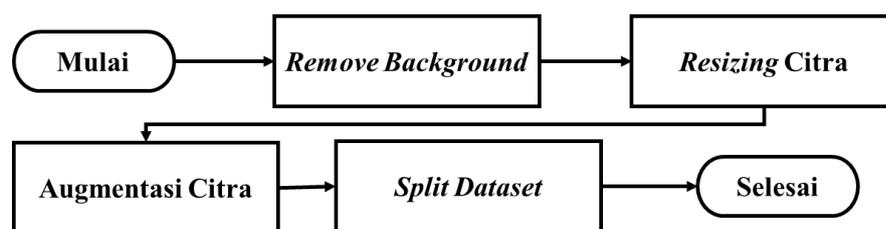
Tabel 3.2 Kategori pisang Cavendish

Kategori/Kelas	Gambar	Deskripsi
Mentah Kualitas Bagus (MHBS)		Memiliki warna hijau yang terang dan seragam. Kulit pisang tidak memiliki bercak hitam, coklat, atau noda lain, menandakan bahwa fisik buah tersebut dalam keadaan baik. Bagian ujung pisang tidak mengalami perubahan warna yang mencolok, seperti menjadi coklat atau hitam, menunjukkan bahwa buah masih dalam kondisi baik.
Mentah Kualitas Buruk (MHBK)		Pisang Cavendish mentah umumnya berwarna hijau cerah. Buah pisang yang masuk dalam kategori ini umumnya ditandai oleh keberadaan bercak coklat gelap, warna yang memudar, atau ketidakkonsistenan warna. Adanya noda hitam, coklat, atau bercak lain pada kulit dapat mengindikasikan kerusakan atau infeksi. Pada umumnya kategori ini digunakan untuk dibuat bubuk atau tepung pisang.
Matang Kualitas Bagus (MGBS)		Buah pisang yang termasuk dalam kategori ini umumnya memiliki warna kuning yang cerah dan seragam. Memiliki tekstur yang agak padat, tetapi

		tidak terlalu keras, bentuk yang normal, tidak memiliki tanda-tanda bercak-bercak coklat atau memar yang signifikan
Matang Kualitas Buruk (MGBK)		Pisang Cavendish matang umumnya berwarna kuning cerah. Buah dalam kategori ini memiliki ciri warnanya tidak merata, terlalu lunak atau terlalu lembut, memiliki bercak-bercak coklat.

### 3.3.3 Preprocessing Data

Untuk memulai tahap *preprocessing data*, peneliti telah merencanakan langkah-langkah yang harus dijalani untuk menggambarkan proses yang akan dilakukan. Di bawah ini merupakan diagram alir yang dapat dilihat pada Gambar 3.2.



Gambar 3. 2 Tahap *Preprocessing Data*

Sebelum melakukan implementasi model, langkah yang perlu dilakukan adalah *preprocessing data*. Tahap ini dimulai dengan *remove background*, *resizing*, *augmentasi*, dan pembagian dataset menjadi dua bagian, yakni data pelatihan (*train*) dan data pengujian (*test*). Berikut adalah penjelasan rinci tentang setiap tahap *preprocessing data* ini, sebagaimana yang digambarkan dalam Gambar 3.6.

#### 1. *Remove Background*

Setelah dataset terkumpul dilakukan *remove background* objek gambar dengan menggunakan *adobe express remove background* guna meningkatkan kualitas data dan mengurangi noise agar mudah diolah oleh model. Berikut hasil dari *remove background* yang ditunjukkan pada gambar 3.3.



Gambar 3. 3 Setelah dilakukan *remove background*.

## 2. *Resizing* citra

Setelah penghapusan *background* dilakukan perubahan ukuran gambar menjadi ukuran yang lebih kecil dengan ukuran gambar 150x150 pixel. Berikut adalah hasil *resizing* gambar yang dilakukan dalam penelitian ini yang ditunjukkan pada gambar 3.4.

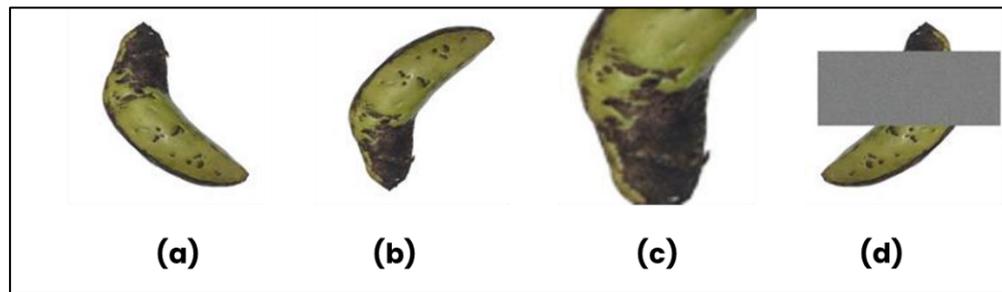


Gambar 3. 4 Setelah dilakukan *Resizing* citra.

**Algoritma 3.1:** *Resizing* Citra

1. Tentukan nilai lebar gambar (*width*), tinggi gambar (*height*).
  2. Inisialisasi objek *ImageDataGenerator* dengan *rescale=1./255* untuk melakukan normalisasi intensitas piksel.
  3. Buat *data generator* untuk data pelatihan dengan memanggil metode *flow\_from\_directory* pada objek *ImageDataGenerator*:
    - a. Tentukan direktori pelatihan (*train\_dir*).
    - b. Atur ukuran target gambar menjadi (*width, height*) dengan parameter *target\_size*.
    - c. Tentukan ukuran batch menggunakan parameter *batch\_size*.
    - d. Atur mode kelas menjadi '*categorical*' jika memiliki kelas-kelas yang merupakan kategori.
  4. *Data generator* siap digunakan untuk memuat gambar-gambar dari direktori pelatihan dengan *batch* yang telah ditentukan.
3. Augmentasi Citra

Tahap augmentasi adalah tahap yang penting untuk menciptakan keragaman data pelatihan yang lebih besar dari gambar yang ada. Dalam penelitian ini, augmentasi yang dilakukan adalah *random zoom 50%*, *flip top-bottom*, *flip left-right*, dan *random erasing*. Berikut hasil augmentasi yang ditunjukkan pada gambar 3.5.



Gambar 3. 5 (a) *flip left-right*, (b) *flip top-bottom*.

(c) *random zoom 50%* (d) *random erasing*

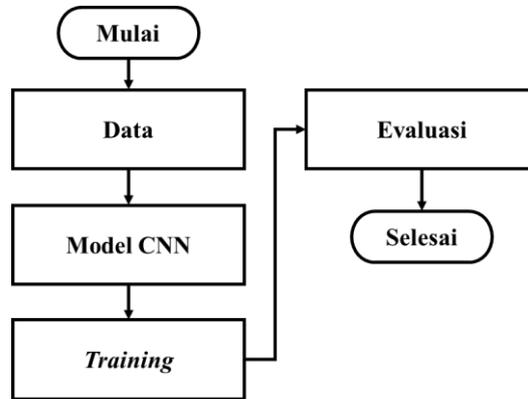
### Algoritma 3.2: Augmentasi Citra

1. Masukkan direktori sumber gambar (*dir*) dan jumlah sampel yang ingin diproses (*sebanyak\_ini*).
  2. Tentukan direktori *output* untuk menyimpan gambar hasil *preprocessing* (*output\_dir*).
  3. Inisialisasi *Augmentor Pipeline* dengan direktori sumber (*source\_dir*) dan direktori *output* (*output\_dir*).
  4. Terapkan operasi *flip* kiri-kanan dengan probabilitas 0.8.
  5. Terapkan operasi *flip* atas-bawah dengan probabilitas 0.8.
  6. Terapkan operasi *zoom* acak dengan probabilitas 0.8 dan persentase area 0.5.
  7. Tambahkan operasi *RandomErasing* dengan probabilitas 1, *short length* = 0.02, *short height* = 0.4 dan *aspect ratio* = 0.3.
  8. Sampelkan sebanyak '*sebanyak\_ini*' gambar dari *pipeline*.
4. *Split Dataset*

Dataset yang telah di augmentasi kemudian dibagi menjadi 90:10. Pemilihan skenario ini didasarkan pada penelitian sebelumnya [9] yang menunjukkan bahwa semakin besar proporsi data latih dalam pembagian data, semakin tinggi akurasi model yang dihasilkan. Meskipun penelitian sebelumnya menggunakan dataset citra daun kentang, dalam penelitian ini, skenario pembagian data yang sama diterapkan pada dataset citra pisang Cavendish. Dalam skenario ini, proporsi 90:10 berarti 90% data digunakan untuk melatih model klasifikasi, sementara 10% sisanya digunakan sebagai data uji untuk mengevaluasi kinerja.

### 3.3.4 Perancangan Model CNN

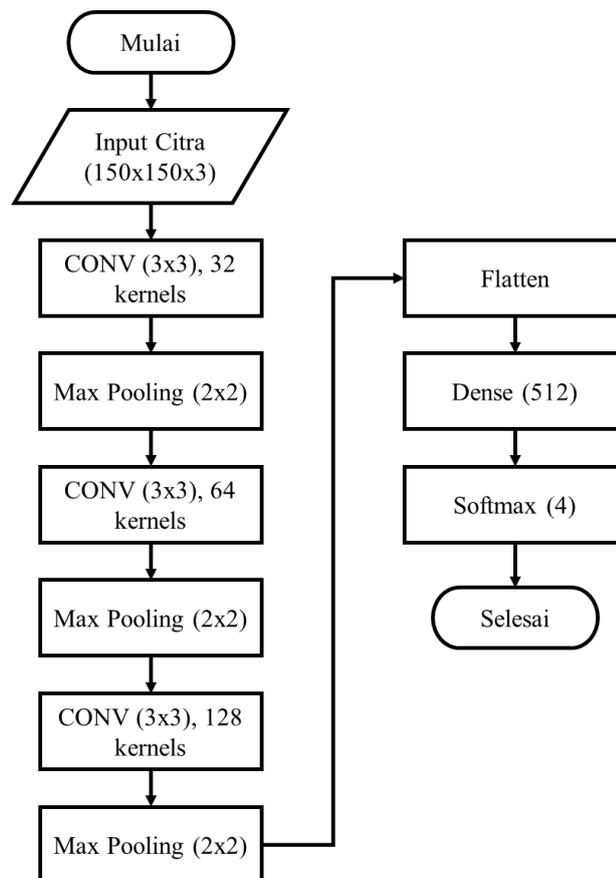
Langkah-langkah selanjutnya setelah *preprocessing data* dapat dilihat pada gambar 3.5 di bawah ini.



Gambar 3.5 Perancangan Model

#### 1. Model CNN

Algoritma *Convolutional Neural Network* (CNN) dalam penelitian ini bertujuan untuk mengklasifikasikan kualitas dan kematangan pisang Cavendish guna memperoleh hasil akurasi yang baik. Beberapa tahapan dalam proses model klasifikasi CNN ditunjukkan pada Gambar 3.6 di bawah ini.



Gambar 3.6 Arsitektur CNN

Proses pembuatan model ini menggunakan *input shape* dengan ukuran (150, 150, 3), yang berarti menggunakan citra dengan resolusi 150x150 piksel dan tiga channel warna (R, G, B) karena citra berwarna. Data yang digunakan untuk membuat model terdiri dari data pelatihan (*training*) dan data pengujian (*testing*).

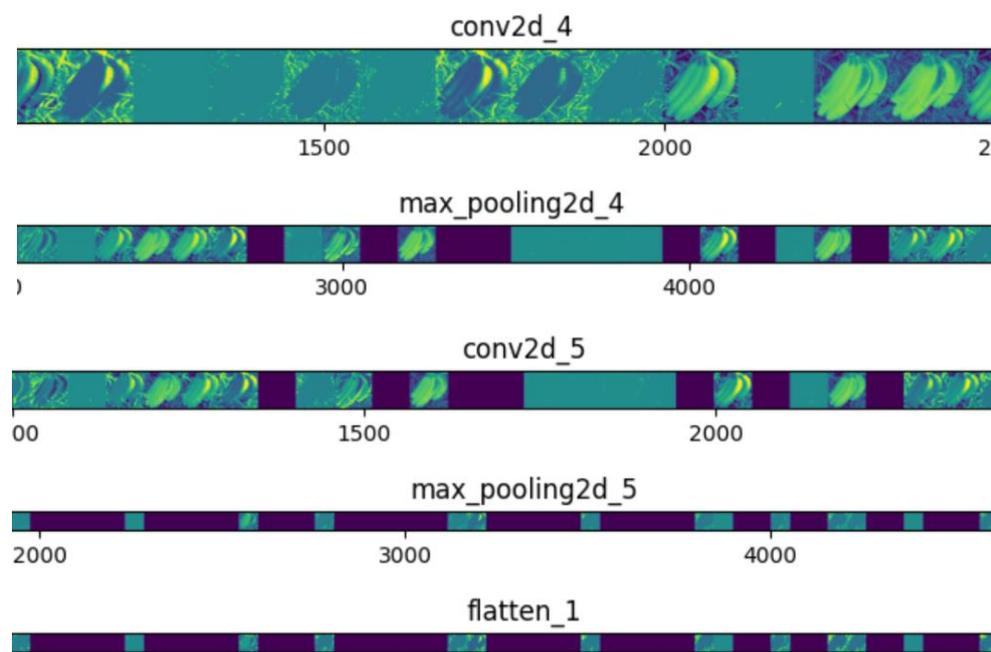
Pada lapisan konvolusi (*convolutional layer*), nilai filter yang digunakan adalah 32, 64, dan 128. Filter ini berdimensi 3x3 dan digunakan untuk deteksi fitur pada citra. Fungsi aktivasi yang digunakan pada lapisan konvolusi adalah ReLU (*Rectified Linear Unit*), yang umum digunakan dalam *Convolutional Neural Networks* (CNN). Total terdapat tiga lapisan konvolusi yang secara otomatis mengekstraksi fitur/ciri pada citra pisang. Setiap lapisan konvolusi dalam CNN dapat mempelajari fitur-fitur yang semakin kompleks dan abstrak seiring dengan kedalaman model. Fitur/ciri

yang digunakan melibatkan karakteristik bentuk, warna dan tekstur buah pisang Cavendish.

Pada konvolusi pertama pada umumnya digunakan untuk mendeteksi tepi pada gambar, dan tekstur sederhana, konvolusi selanjutnya mempelajari fitur tekstur dan pola yang lebih kompleks, konvolusi yang terakhir mempelajari fitur seperti objek atau bagian objek[37].

Pada lapisan *Pooling*, digunakan *MaxPooling2D* untuk mereduksi ukuran citra dengan mengambil nilai maksimum dari matrix 2x2. *MaxPooling* membantu mempertahankan informasi penting dalam citra dengan mengambil nilai maksimum[23].

Proses *Flattening* digunakan untuk mengubah array dua dimensi yang dihasilkan dari lapisan *pooling* menjadi satu dimensi (*single vector*), yang nantinya akan menjadi input untuk lapisan *neural network* selanjutnya[23].



Gambar 3.7 Proses di dalam CNN

Lapisan *Dense* digunakan untuk menjalankan *fully connected layer* dari *neural network*. Lapisan *Dense* pertama memiliki 512 unit, yang merupakan jumlah *node* di antara jumlah *node input* dan *output*. Fungsi aktivasi yang digunakan pada lapisan *Dense* ini adalah ReLU.

Lapisan *Dense* terakhir, yang berbeda dengan lapisan *Dense* sebelumnya, digunakan sebagai lapisan *output* dan memiliki 4 *node*, sesuai dengan jumlah kelas yang digunakan untuk klasifikasi (mode *categorical*). Fungsi aktivasi pada lapisan *output* adalah *softmax*, yang cocok untuk tugas klasifikasi *multiclass*.

**Algoritma 3.3:** Arsitektur CNN

1. Buat objek model *Sequential* sebagai wadah untuk lapisan-lapisan model:
  2. Tambahkan lapisan *Convolutional* 2D pertama dengan 32 filter, ukuran kernel (3,3), fungsi aktivasi ReLU, dan bentuk input (150,150,3).
  3. Tambahkan lapisan *MaxPooling* 2D dengan ukuran *pool* (2,2).
  4. Tambahkan lapisan *Convolutional* 2D kedua dengan 64 filter, ukuran kernel (3,3), dan fungsi aktivasi ReLU.
  5. Tambahkan lapisan *MaxPooling* 2D dengan ukuran *pool* (2,2).
  6. Tambahkan lapisan *Convolutional* 2D ketiga dengan 128 filter, ukuran kernel (3,3), dan fungsi aktivasi ReLU.
  7. Tambahkan lapisan *MaxPooling* 2D dengan ukuran *pool* (2,2).
  8. Tambahkan lapisan *Flatten* untuk meratakan output dari lapisan-lapisan sebelumnya menjadi vektor satu dimensi.
  9. Tambahkan lapisan *Dense* dengan 512-unit dan fungsi aktivasi ReLU.
  10. Tambahkan lapisan *Dense* terakhir dengan 4-unit dan fungsi aktivasi *softmax* untuk klasifikasi multi kelas.
  11. Model selesai dibuat dan siap untuk dilatih dengan data.
2. *Training*

*Training* model pada penelitian ini bertujuan untuk melatih model dimana model dapat mempelajari kualitas dan kematangan pisang Cavendish. Berdasarkan pertanyaan penelitian maka akan dilakukan beberapa parameter berbeda agar mendapatkan akurasi yang terbaik menggunakan skema yang disajikan pada tabel berikut ini.

Tabel 2. 3 Tabel skema percobaan *training model*

Percobaan	Skema
<i>Batch size</i>	8, 16, dan 32
<i>Epoch</i>	10, 20, dan 30

Pemilihan nilai *epoch* dan *batch size* didasarkan pada penelitian sebelumnya [11] yang menunjukkan bahwa peningkatan jumlah *epoch* dapat meningkatkan akurasi model tetapi memperpanjang waktu pelatihan. Oleh karena itu, hanya nilai 10, 20, dan 30 *epoch* yang digunakan dalam penelitian ini untuk mempercepat proses pelatihan. *Epoch* bertujuan untuk menentukan berapa kali model CNN telah melihat dataset secara menyeluruh. Dengan kata lain *epoch* menandakan sebuah model CNN telah mempelajari data *training* secara keseluruhan. Jumlah *epoch* dapat ditentukan sendiri dan tidak memiliki aturan khusus yang mengikat [38].

Selain variasi jumlah *epoch*, *batch size* juga bertujuan mengatur seberapa banyak sampel data yang akan diproses oleh model dalam satu iterasi selama proses pelatihan. *Batch size* yang lebih besar cenderung mempercepat proses pelatihan dengan memproses lebih banyak sampel data dalam satu iterasi, tetapi dapat menyebabkan peningkatan penggunaan memori dan kemungkinan pengurangan kestabilan sedangkan, *batch size* yang lebih kecil dapat menghasilkan pembelajaran yang lebih stabil karena pembaharuan bobot yang lebih sering, tetapi proses pelatihan mungkin memakan waktu lebih lama [39].

Setelah proses pelatihan data selesai, langkah berikutnya adalah menentukan nilai akurasi tertinggi.

### 3. Evaluasi

Pada tahap evaluasi penelitian ini menggunakan *Confusion Matrix* untuk memprediksi kondisi data yang dianalisis oleh model yang telah dibangun. Parameter pengujian yang digunakan untuk mengevaluasi adalah nilai *Accuracy*, *Precision*, *Recall* dan *F1-Score*.