

BAB 2

TINJAUAN PUSTAKA

2.1 Review Penelitian Terdahulu

Review penelitian terdahulu berguna untuk memberikan masukan dan ide yang terkait dengan penelitian yang akan dilakukan. Adapun beberapa penelitian terdahulu yang digunakan pada proyek akhir ini adalah:

Usman dkk., (2025) melakukan penelitian yang bertujuan untuk meningkatkan akurasi klasifikasi kualitas buah kelapa lokal menggunakan *Convolutional Neural Network* (CNN) dengan pendekatan *transfer learning*. Penelitian ini dikembangkan karena penilaian mutu kelapa oleh petani masih bersifat manual dan subjektif. Model yang digunakan adalah VGG16 yang diadaptasi dengan teknik augmentasi citra seperti rotasi dan perubahan pencahayaan untuk memperluas dataset. Hasil penelitian menunjukkan bahwa metode ini mampu menghasilkan akurasi tinggi, mencapai lebih dari 98%. Namun, penelitian ini memiliki keterbatasan karena hanya berfokus pada varietas kelapa lokal dengan jumlah data terbatas sehingga generalisasi model ke kondisi lapangan masih perlu diuji

Luthfie dkk., (2022) meneliti klasifikasi tingkat kematangan kelapa dengan menggabungkan CNN dan YOLOv8 untuk mendeteksi serta mengenali buah pada gambar. Tujuan penelitian ini adalah mempermudah proses identifikasi kelapa muda dan kelapa tua yang penting untuk kebutuhan konsumsi maupun produksi santan. CNN digunakan untuk klasifikasi kematangan, sedangkan YOLOv8 berperan dalam mendeteksi objek kelapa pada citra dengan latar belakang kompleks. Hasil penelitian menunjukkan kinerja yang sangat baik dengan precision 99,3%, *recall* 94,2%, dan mAP 90,5%. Kelemahan penelitian ini adalah fokus kelas yang masih terbatas pada dua kategori utama sehingga belum mencakup variasi kualitas lain seperti cacat fisik atau kerusakan kulit.

Aulia dkk., (2022) yang mengembangkan sistem klasifikasi kualitas apel berbasis web menggunakan CNN. Penelitian ini bertujuan untuk memberikan solusi digital terhadap penilaian kualitas buah apel secara otomatis. Sistem dibangun dalam bentuk website yang memungkinkan pengguna mengunggah

gambar apel dan mendapatkan hasil klasifikasi secara langsung. Meskipun sistem memiliki antarmuka pengguna yang baik dan mudah diakses, kekurangan dari penelitian ini adalah model CNN masih dibangun secara manual tanpa pemanfaatan *transfer learning*, sehingga kurang efisien dalam hal waktu pelatihan.

Parvathi dkk., (2020) mengembangkan penelitian untuk mendeteksi tahap kematangan kelapa dalam lingkungan kompleks menggunakan Faster R-CNN. Penelitian ini dilatarbelakangi oleh sulitnya mengidentifikasi kematangan kelapa langsung dari pohon secara visual, terutama karena gangguan latar belakang seperti daun dan cabang. Model Faster R-CNN digunakan dengan backbone *ResNet-50* untuk melakukan deteksi objek sekaligus klasifikasi tahap kematangan. Hasil penelitian menunjukkan tingkat akurasi yang baik, sekitar 89%, pada kondisi pencahayaan dan lingkungan nyata. Keterbatasan penelitian ini adalah kebutuhan perangkat keras yang relatif tinggi dan waktu komputasi yang lebih lama dibandingkan metode klasifikasi sederhana.

Kardena dkk., (2025) melakukan penelitian klasifikasi kualitas buah kelapa dengan memanfaatkan ekstraksi fitur warna, bentuk, dan tekstur, kemudian diklasifikasikan menggunakan algoritma K-Nearest Neighbor (KNN). Penelitian ini dikembangkan karena klasifikasi visual sederhana masih banyak digunakan di kalangan petani dan pedagang. Dataset yang digunakan terdiri dari gambar buah kelapa yang dikelompokkan menjadi kategori *standard* dan *non-standard*. Hasil penelitian menunjukkan tingkat akurasi mencapai sekitar 96% pada konfigurasi tertentu. Namun, kekurangan dari penelitian ini adalah pendekatan berbasis fitur manual kurang fleksibel saat menghadapi variasi kondisi pencahayaan dan latar belakang yang beragam.

Munandar dkk. (2022) meneliti penerapan pemrosesan citra untuk mengklasifikasikan kualitas buah kelapa berdasarkan warnanya. Penelitian ini dilatarbelakangi oleh kebutuhan industri untuk metode inspeksi sederhana dalam menentukan mutu kelapa sebelum dipasarkan. Metode yang digunakan berupa konversi citra RGB menjadi *grayscale*, kemudian dilakukan *thresholding* dan operasi morfologi untuk membedakan tingkat kualitas. Hasil penelitian menunjukkan metode ini mampu mengelompokkan buah ke dalam beberapa kategori mutu dengan cukup baik. Kelemahannya adalah pendekatan ini sangat

bergantung pada pencahayaan saat pengambilan gambar, sehingga akurasi-nya berkurang jika kondisi lingkungan berubah.

Table 2.1 Tabel Perbandingan

No	Nama	Judul	Metode	kelebihan	kelemahan
1	Usman dkk. (2025)	Klasifikasi kualitas buah kelapa local menggunakan CNN dengan <i>transfer learning</i>	CNN (VGG16) + <i>Transfer Learning</i> + Augmentasi Citra	Akurasi sangat tinggi (98%), mampu mengatasi variasi cahaya dengan augmentasi	Hanya fokus pada kelapa lokal, dataset terbatas sehingga generalisasi rendah
2	Luthfie dkk. (2022)	Klasifikasi Tingkat Kematangan Kelapa Menggunakan CNN dan YOLOv8	CNN + YOLOv8	Precision 99,3%, <i>recall</i> 94,2%, mAP 90,5%; mampu deteksi objek di latar kompleks	Hanya dua kelas (kelapa muda & tua), belum mencakup cacat fisik atau kerusakan
3	Parvathi dkk. (2020)	Deteksi tahap kematangan kelapa menggunakan faster R-CNN	Faster R-CNN (<i>ResNet-50</i>)	Akurasi baik (89%) pada kondisi nyata, mampu deteksi + klasifikasi sekaligus	Membutuhkan perangkat keras tinggi, komputasi lebih lama

4	Karden a dkk. (2025)	Klasifikasi kualitas kelapa menggunakan ekstraksi fitur & KNN	Ekstraksi fitur (warna, bentuk, tekstur) + KNN	Akurasi 96%, sederhana untuk implementasi awal	Kurang fleksibel terhadap variasi pencahayaan & latar belakang
5	Munandar dkk. (2022)	Klasifikasi Kualitas Buah Kelapa Berdasarkan Warna dengan Pengolahan Citra	RGB : <i>Grayscale</i> + <i>Thresholding</i> + Morfolog	Metode sederhana, mudah diimplementasi kan	Akurasi bergantung pada pencahayaan, kurang stabil di kondisi berbeda

2.2 Landasan Teori

2.2.1 Buah kelapa

Menurut penelitian DebMandal & Mandal, (2011), kualitas buah kelapa sangat dipengaruhi oleh faktor fisiologis dan tingkat kematangan buah. Kelapa yang dipanen pada usia matang fisiologis menghasilkan daging buah yang lebih tebal serta memiliki kandungan minyak yang lebih tinggi sehingga ideal digunakan sebagai bahan baku pembuatan santan dan minyak kelapa. Penelitian tersebut juga menjelaskan bahwa karakteristik seperti warna sabut, kekerasan tempurung, dan kadar air merupakan indikator penting dalam menentukan mutu buah kelapa. Buah kelapa yang mengalami penurunan kualitas biasanya menunjukkan tanda seperti bercak hitam pada kulit, retakan tempurung, serta perubahan tekstur akibat kontaminasi jamur maupun serangan hama.

Menurut penelitian yang dilakukan oleh Wahyudi dkk., (2025) kualitas buah kelapa (*Cocos nucifera* L.) dipengaruhi oleh faktor fisiologis seperti umur tanaman dan karakteristik buah yang meliputi tebal daging buah (*endosperm*

thickness), rendemen minyak, serta kadar air. Hasil penelitian menunjukkan bahwa umur tanaman kelapa berpengaruh nyata terhadap ketebalan endosperm dan rendemen minyak pohon kelapa yang lebih tua menghasilkan buah dengan endosperm yang lebih tebal dan rendemen minyak yang lebih tinggi dibandingkan pohon yang lebih muda. Temuan ini menunjukkan bahwa umur tanaman merupakan salah satu parameter penting yang memengaruhi kualitas buah kelapa untuk diolah menjadi produk bernilai jual tinggi seperti santan atau minyak kelapa.

Selain itu, penelitian mutu buah kelapa untuk konsumsi segar yang dilakukan di Daerah Istimewa Yogyakarta menemukan bahwa sejumlah atribut morfologis seperti volume air kelapa, diameter buah, dan tebal daging buah berkorelasi dengan preferensi konsumen terhadap kualitas kelapa segar. Buah kelapa yang sesuai dengan preferensi konsumen umumnya memiliki diameter besar (>210 mm), warna sabut hijau, volume air banyak, serta tebal kernel yang mencukupi, sementara karakteristik visual seperti kulit buah yang retak atau bercak gelap cenderung menunjukkan penurunan mutu yang kurang disukai konsumen.

2.2.2 *Deep Learning*

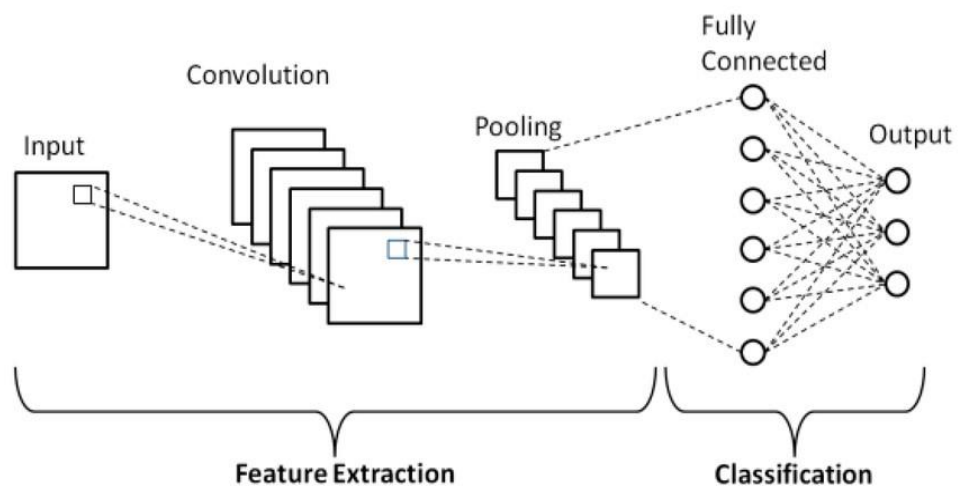
Deep Learning (DL) adalah cabang dari *machine learning* yang menggunakan jaringan saraf tiruan dengan banyak lapisan (*deep neural networks*). Menurut Lecun dkk., (2015), DL unggul dalam mengekstrak fitur secara otomatis dari data mentah.

Perbedaan utama antara ML tradisional dan DL adalah pada cara pengolahan fitur. ML masih memerlukan rekayasa fitur manual (misalnya menggunakan GLCM untuk tekstur), sedangkan DL mampu mempelajari fitur langsung dari data citra melalui proses pembelajaran hierarkis. Hal ini membuat DL sangat efektif dalam klasifikasi gambar, termasuk dalam bidang pertanian.

Beberapa aplikasi DL di bidang pertanian meliputi deteksi penyakit tanaman, klasifikasi kualitas buah, serta prediksi kematangan. Dalam penelitian ini, DL digunakan melalui arsitektur *Convolutional Neural Network* (CNN) yang telah terbukti efektif dalam pengolahan citra.

2.2.3 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) adalah salah satu arsitektur deep learning yang dirancang khusus untuk pengolahan data berbentuk grid, terutama citra digital. CNN pertama kali diperkenalkan oleh LeCun pada tahun 1990-an melalui arsitektur *LeNet-5*, dan semakin populer setelah Krizhevsky dkk., (2012) memperkenalkan *AlexNet* yang berhasil mendominasi kompetisi *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)* dengan akurasi yang jauh lebih tinggi dibanding metode konvensional. Arsitektur dasar *Convolutional Neural Network (CNN)* dapat dilihat pada Gambar 2.1.

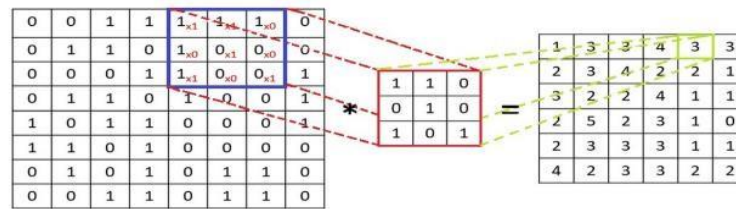


Gambar 2.1 Arsitektur dasar *Convolutional Neural Network (CNN)*

Sumber: Phung & Rhee, (2019)

Arsitektur CNN secara umum terdiri dari tiga jenis lapisan utama, yaitu convolutional layer, pooling layer, dan fully connected layer.

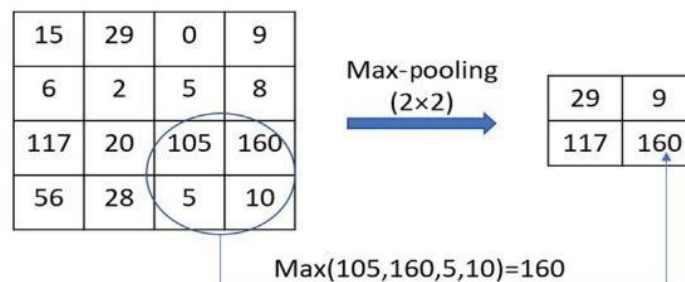
1. *Convolutional Layer* : Lapisan ini berfungsi untuk mengekstraksi fitur dari citra dengan melakukan operasi konvolusi menggunakan kernel atau filter. Setiap filter akan bergerak melintasi citra untuk mendeteksi pola-pola visual seperti tepi, garis, atau tekstur tertentu. Hasil dari proses ini berupa *feature map* yang merepresentasikan ciri khas dari citra input. Proses operasi konvolusi pada citra dilakukan dengan menggeser kernel untuk mengekstraksi fitur visual, seperti ditunjukkan pada Gambar 2.2.



Gambar 2.2 Operasi Konvolusi

(Sumber: Purwono dkk., (2022))

2. *Pooling Layer* : Lapisan ini berfungsi untuk melakukan *downsampling* pada *feature map* hasil konvolusi dengan tujuan mengurangi dimensi data dan jumlah parameter yang harus dihitung. Teknik pooling yang umum digunakan adalah *max pooling* dan *average pooling*, di mana nilai representatif seperti maksimum diambil untuk menjaga fitur penting sekaligus mempercepat komputasi dan mengurangi risiko *overfitting*. Selain *max pooling* dan *average pooling*, terdapat pula teknik Global Average Pooling yang meringkas setiap *feature map* menjadi satu nilai rata-rata, sehingga menghasilkan vektor fitur berdimensi kecil sekaligus mengurangi jumlah parameter secara signifikan dibandingkan lapisan Flatten. Proses subsampling atau *pooling layer* yang berfungsi mengurangi dimensi fitur ditunjukkan pada Gambar 2.3.

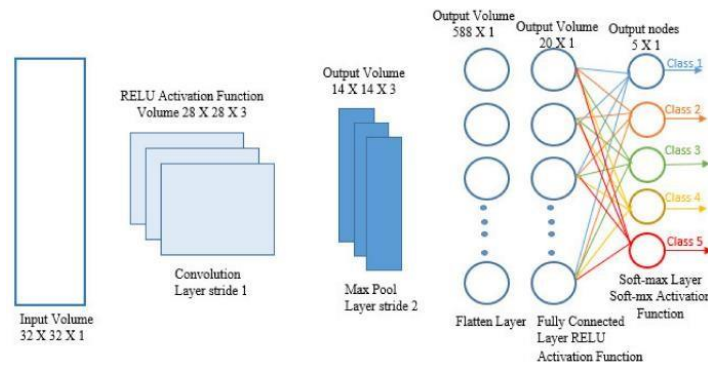


Gambar 2.3 *Pooling Layer*

(Sumber: Purwono dkk., (2022))

3. *Fully Connected Layer (Dense Layer)* : Lapisan ini merupakan tahap akhir dari CNN yang menghubungkan seluruh *neuron* pada lapisan sebelumnya secara penuh (*fully connected*). Lapisan ini berfungsi untuk melakukan proses klasifikasi berdasarkan fitur-fitur yang telah diekstraksi pada lapisan sebelumnya, dengan menghasilkan output berupa probabilitas dari

setiap kelas. Lapisan fully connected yang berfungsi sebagai proses klasifikasi akhir dapat dilihat pada Gambar 2.4.



Gambar 2.4 *Fully Connected Layer (Dense Layer)*

(Sumber Purwono et al., (2022))

Ketiga lapisan utama tersebut bekerja secara berurutan untuk mengekstraksi, mereduksi, dan mengklasifikasikan fitur dari citra. Dengan kombinasi ini, CNN mampu mengenali pola kompleks pada citra kelapa.

2.2.3.1 Feature Learning

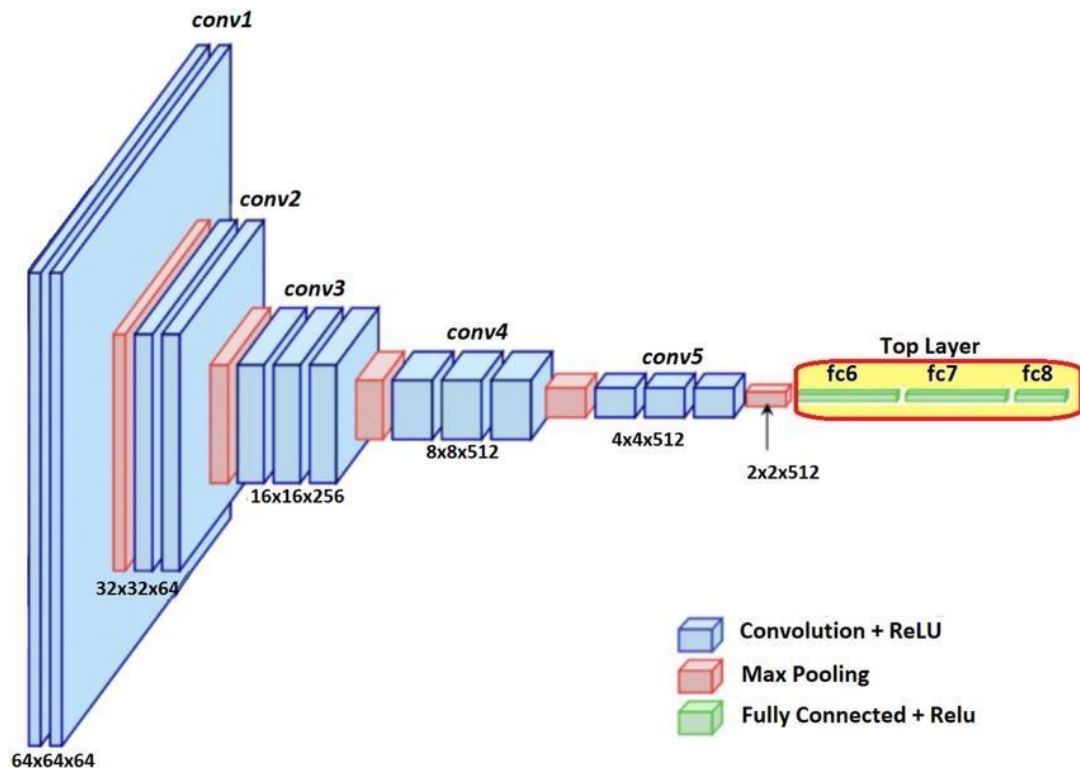
Feature learning merupakan kemampuan *Convolutional Neural Network* (CNN) dalam mempelajari representasi fitur dari data secara otomatis tanpa memerlukan rekayasa fitur secara manual. Proses ini terjadi pada *convolutional layer* yang menggunakan filter atau kernel untuk mendeteksi pola tertentu pada citra (Lecun dkk., 1998); Goodfellow dkk., (2016)

Pada lapisan awal, CNN umumnya mempelajari *low-level features* seperti tepi, garis, dan sudut. Selanjutnya, pada lapisan menengah, CNN mulai mengenali *mid-level features* seperti tekstur dan pola bentuk. Pada lapisan yang lebih dalam, CNN mampu mengekstraksi *high-level features* yang lebih kompleks, seperti bentuk objek secara utuh atau ciri khas tertentu (Zeiler & Fergus, (2014))

Pada konteks penelitian ini, *feature learning* memiliki peran penting dalam membedakan kualitas buah kelapa. Buah kelapa dengan kualitas baik umumnya ditandai oleh permukaan sabut yang mulus, warna yang seragam, serta bebas dari bercak hitam. Sebaliknya, kelapa dengan kualitas rendah dapat ditandai oleh adanya retakan atau noda gelap pada permukaan kulit. CNN mampu

mempelajari karakteristik tersebut secara otomatis berdasarkan citra input tanpa memerlukan intervensi manusia.

Keunggulan *feature learning* dibandingkan metode tradisional terletak pada kemampuannya menghasilkan fitur yang lebih *robust* terhadap variasi pencahayaan, rotasi, dan sudut pengambilan gambar. Hal ini sangat relevan dalam kondisi nyata di lapangan, di mana citra buah kelapa dapat diambil dalam berbagai situasi dan kondisi lingkungan yang berbeda (Goodfellow dkk., 2016)



Gambar 2.5 VGG16 Model

(Sumber: Simonyan & Zisserman, (2015))

2.2.3.2 Classification

Tahap klasifikasi adalah proses pengambilan keputusan akhir dari fitur yang telah diekstraksi. Pada CNN, klasifikasi dilakukan oleh *Fully Connected Layer* dengan fungsi aktivasi *Softmax* yang mengubah output menjadi probabilitas dari setiap kelas.

Sebagai contoh, jika sistem klasifikasi kelapa memiliki tiga kelas (baik, sedang, kurang baik), maka fungsi *Softmax* akan menghasilkan tiga nilai probabilitas yang

jumlahnya 1. Kelas dengan probabilitas tertinggi akan dipilih sebagai hasil prediksi.

Fungsi kerugian (*loss function*) yang umum digunakan dalam klasifikasi adalah *categorical cross-entropy*. Fungsi ini menghitung seberapa jauh prediksi model dari label sebenarnya, kemudian digunakan untuk memperbarui bobot jaringan melalui proses *backpropagation*.

Dalam penelitian ini, tahap klasifikasi berperan untuk menentukan kualitas buah kelapa berdasarkan citra yang telah melalui preprocessing dan *feature learning*. CNN dengan kombinasi *transfer learning* mampu menghasilkan klasifikasi yang akurat meskipun dataset terbatas.

2.2.4 Confusion Matrix

Confusion Matrix merupakan salah satu metode evaluasi yang paling umum digunakan dalam sistem klasifikasi. *Confusion matrix* menyajikan perbandingan antara hasil prediksi model dengan label sebenarnya dalam bentuk tabel dua dimensi, sehingga memudahkan analisis kinerja model klasifikasi (Suginta, 2025)

Untuk klasifikasi dua kelas, *confusion matrix* terdiri dari empat komponen utama, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN), seperti ditunjukkan pada Tabel 2.2.

Tabel 2.2 Contoh *Confusion Matrix* untuk 2 kelas

	Prediksi <i>Standard</i>	Prediksi <i>non-standard</i>
<i>Standard</i>	True Positive (TP)	False Negative (FN)
<i>Non-standard</i>	False Positive (FP)	True Negative (TN)
Total Prediksi	TP + FP	FN + TN

Keterangan :

1. TP (*True Positive*) Model benar mendeteksi kelapa standar sebagai standar.
2. TN (*True Negative*) Model benar mendeteksi kelapa non-standar sebagai non-standar.
3. FP (*False Positive*) Model salah mengira kelapa non-standar sebagai standar.

4. FN (*False Negative*) Model salah mengira kelapa standar sebagai nonstandar.

Dalam penelitian klasifikasi kelapa, *confusion matrix* digunakan untuk melihat seberapa akurat model dalam membedakan kelapa dengan kualitas standar dan nonstandar. Melalui analisis ini, dapat diketahui kinerja dan kelemahan model, misalnya jika model masih sering salah mengklasifikasikan kelapa non-standar sebagai standar, atau sebaliknya. (Suginta, (2025).

a. Accuracy

Accuracy menunjukkan proporsi prediksi yang benar (baik kelas standar maupun non-standar) terhadap keseluruhan data yang diprediksi. Accuracy dihitung menggunakan persamaan berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad 2-1$$

b. precision

Precision menunjukkan seberapa tepat model dalam memprediksi kelas standar, yaitu proporsi prediksi standar yang benar-benar merupakan kelapa standar. Precision dihitung menggunakan persamaan berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad 2-2$$

c. recall

Recall (sensitivity) menunjukkan kemampuan model dalam mendeteksi seluruh data kelapa standar yang sebenarnya, yaitu proporsi kelapa standar yang berhasil diprediksi dengan benar oleh model. Recall dihitung menggunakan persamaan berikut:

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad 2-3$$

d. F1-score

F1-Score merupakan nilai rata-rata harmonik (harmonic mean) antara precision dan recall, yang digunakan untuk memberikan gambaran keseimbangan antara kedua metrik tersebut, terutama pada kondisi distribusi data antar kelas yang tidak seimbang. F1-Score dihitung menggunakan persamaan berikut:

$$F1 - Score = 2 \times \frac{\text{Precision} \times \text{recall}}{\text{Precision} + \text{recall}} \times 100\% \quad (2.4)$$

Keempat metrik evaluasi ini digunakan secara bersamaan dalam penelitian ini agar kinerja model CNN dalam mengklasifikasikan kualitas buah kelapa dapat dinilai secara lebih komprehensif, tidak hanya berdasarkan tingkat akurasi semata.

2.2.5 Streamlit

Streamlit adalah kerangka kerja (framework) sumber terbuka berbasis Python yang dirancang untuk mempermudah pengembangan aplikasi web interaktif, khususnya pada bidang data science dan machine learning, tanpa memerlukan penguasaan mendalam terhadap HTML, CSS, maupun JavaScript. Melalui sekumpulan fungsi bawaan seperti file uploader, penampil gambar, serta komponen sidebar untuk navigasi, Streamlit memungkinkan sebuah model yang telah dilatih untuk langsung dijalankan dan ditampilkan pada antarmuka berbasis web hanya dengan beberapa baris kode Python. Karakteristik inilah yang menjadikan Streamlit dipilih pada penelitian ini untuk mengintegrasikan model klasifikasi Transfer Learning VGG16 ke dalam sebuah sistem yang dapat langsung digunakan oleh pengguna.

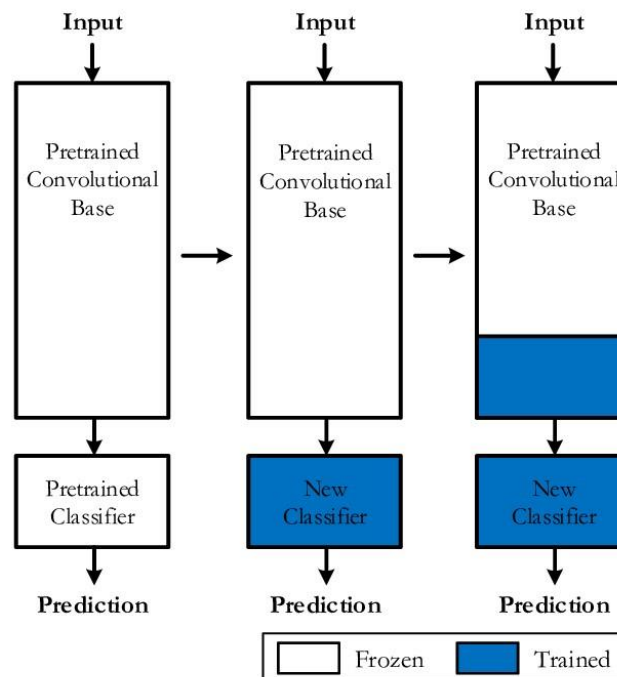
2.2.6 Transfer Learning

Transfer learning merupakan pendekatan dalam *deep learning* yang memanfaatkan model pralatih (*pre-trained model*) yang sebelumnya telah dilatih menggunakan dataset berskala besar seperti ImageNet. Model tersebut kemudian disesuaikan kembali (*fine-tuned*) dengan dataset baru yang lebih kecil dan spesifik terhadap permasalahan tertentu.

Proses *transfer learning* umumnya dilakukan dengan dua pendekatan utama, yaitu:

1. *Feature Extraction* - bobot pada model pralatih digunakan sebagai ekstraktor fitur tetap (*feature extractor*), kemudian ditambahkan lapisan klasifikasi baru yang dilatih pada dataset target.
2. *Fine - Tuning* – sebagian atau seluruh bobot model pralatih disesuaikan kembali agar lebih relevan dengan karakteristik dataset baru.”

Proses umum *transfer learning* dapat dilihat pada Gambar 2.6, yang memperlihatkan alur pemanfaatan model pralatih untuk menghasilkan model baru dengan domain berbeda. Proses umum *transfer learning* dapat dilihat pada Gambar 2.6.



Gambar 2.6 Proses *Transfer Learning*

Sumber: Keras Team (2023)

Transfer learning memiliki beberapa kelebihan, antara lain menghemat waktu pelatihan karena model tidak dibangun dari awal, melainkan memanfaatkan bobot dari model pralatih yang telah dilatih pada dataset berskala besar (Pan & Yang, (2010). Selain itu, *transfer learning* mampu memberikan

performa yang baik meskipun dataset yang digunakan terbatas, karena fitur-fitur umum telah dipelajari pada tahap pelatihan sebelumnya (Yosinski dkk., (2014). Kelebihan lainnya adalah meningkatkan akurasi pada domain tertentu dengan memanfaatkan pengetahuan yang diperoleh dari dataset besar, sehingga model dapat beradaptasi lebih cepat dan efektif pada permasalahan baru (Goodfellow et al., 2016).

Beberapa model pralatih populer dalam *transfer learning* antara lain VGG16, yang memiliki arsitektur sederhana dan efektif untuk klasifikasi citra (Simonyan & Zisserman, 2015); *ResNet*, yang menggunakan residual block untuk memungkinkan jaringan yang lebih dalam (He dkk., 2016); MobileNet, yang dirancang ringan dan efisien untuk perangkat dengan keterbatasan sumber daya (Howard dkk., 2017); serta EfficientNet, yang mengoptimalkan keseimbangan antara akurasi dan jumlah parameter melalui compound scaling (Tan & Le, 2019).

Dalam bidang pertanian, *transfer learning* telah berhasil digunakan untuk mengklasifikasikan kematangan mangga, kualitas cabai, serta kematangan kelapa sawit. (Usman dkk., 2025) berhasil mencapai akurasi 98% pada klasifikasi kualitas kelapa lokal menggunakan VGG16 dengan augmentasi citra. Hal ini menunjukkan bahwa *transfer learning* merupakan pendekatan yang tepat untuk penelitian klasifikasi kualitas kelapa berbasis CNN.

2.2.7 VGG16

VGG16 merupakan salah satu arsitektur Convolutional Neural Network (CNN) yang dikembangkan oleh Karen Simonyan dan Andrew Zisserman dari Visual Geometry Group (VGG), University of Oxford pada tahun 2014. Arsitektur ini diperkenalkan dalam penelitian berjudul *Very Deep Convolutional Networks for Large-Scale Image Recognition* dan berhasil memperoleh performa yang sangat baik pada kompetisi ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014. Keunggulan utama VGG16 terletak pada penggunaan susunan lapisan konvolusi yang sederhana namun mampu menghasilkan akurasi tinggi dalam berbagai tugas klasifikasi citra (Simonyan & Zisserman, 2015). VGG16 terdiri atas 16 lapisan yang memiliki parameter pembelajaran (learnable parameters), yaitu 13 convolutional layer dan 3 fully connected layer. Seluruh convolutional layer menggunakan kernel berukuran 3×3 piksel dengan nilai

stride sebesar 1 dan padding yang menjaga ukuran feature map tetap konsisten. Setelah beberapa convolutional layer, digunakan Max Pooling Layer berukuran 2×2 untuk mengurangi dimensi feature map sekaligus mempertahankan fitur-fitur penting dari citra. Struktur tersebut memungkinkan VGG16 mengekstraksi fitur secara bertahap, mulai dari fitur sederhana seperti tepi dan tekstur hingga fitur yang lebih kompleks seperti bentuk objek.

Pada awal proses, citra masukan harus memiliki ukuran 224×224 piksel dengan tiga kanal warna (RGB). Selanjutnya, citra diproses melalui beberapa convolutional layer yang menggunakan fungsi aktivasi Rectified Linear Unit (ReLU) untuk meningkatkan kemampuan model dalam mempelajari pola nonlinier. Setelah proses ekstraksi fitur selesai, hasil feature map diteruskan ke Fully Connected Layer yang berfungsi sebagai proses klasifikasi. Pada lapisan keluaran, digunakan fungsi aktivasi Softmax untuk menghasilkan probabilitas setiap kelas sehingga model dapat menentukan kelas dengan nilai probabilitas tertinggi sebagai hasil prediksi.

Meskipun memiliki jumlah parameter yang relatif besar, yaitu sekitar 138 juta parameter, VGG16 tetap menjadi salah satu arsitektur yang paling banyak digunakan dalam penelitian klasifikasi citra karena memiliki struktur yang mudah dipahami, stabil selama proses pelatihan, dan mampu memberikan performa yang tinggi ketika digunakan pada pendekatan Transfer Learning. Dengan memanfaatkan bobot hasil pelatihan pada dataset ImageNet, model tidak perlu dilatih dari awal sehingga waktu pelatihan menjadi lebih singkat dan kebutuhan dataset yang besar dapat dikurangi.

Pada penelitian ini, VGG16 digunakan sebagai model pralatih (*pre-trained model*) dalam pendekatan Transfer Learning untuk mengklasifikasikan kualitas buah kelapa menjadi dua kelas, yaitu kelapa standar dan kelapa non-standar. Lapisan konvolusi pada VGG16 dimanfaatkan sebagai feature extractor untuk mengenali karakteristik visual buah kelapa, seperti warna, tekstur permukaan, serta kondisi fisik kulit buah. Selanjutnya, lapisan klasifikasi bawaan VGG16 diganti dengan beberapa lapisan baru yang disesuaikan dengan kebutuhan penelitian sehingga model mampu melakukan klasifikasi terhadap dua kategori kualitas buah kelapa secara lebih optimal.

Penggunaan VGG16 dalam penelitian ini dipilih karena beberapa alasan. Pertama, arsitektur VGG16 telah terbukti mampu menghasilkan akurasi yang tinggi pada berbagai penelitian klasifikasi citra. Kedua, struktur model yang sederhana memudahkan proses implementasi dan analisis. Ketiga, penggunaan Transfer Learning memungkinkan model memperoleh performa yang baik meskipun jumlah dataset relatif terbatas. Oleh karena itu, VGG16 dinilai sesuai untuk mendukung pengembangan sistem klasifikasi kualitas buah kelapa berbasis website yang dikembangkan pada penelitian ini.

Kelebihan VGG16 :

1. Arsitektur yang relatif sederhana dan terstruktur

VGG16 memiliki struktur jaringan yang sistematis dengan 16 *weight layers*, yang terdiri atas 13 *convolutional layers* dan 3 *fully connected layers*. Model ini menggunakan filter konvolusi berukuran kecil, yaitu 3×3 , secara berulang pada setiap *convolutional layer*. Struktur tersebut membuat arsitektur VGG16 lebih mudah dipahami dibandingkan arsitektur yang memiliki struktur koneksi yang lebih kompleks (Simonyan & Zisserman, 2015).

2. Mampu menghasilkan representasi fitur yang baik untuk klasifikasi citra.

VGG16 dikembangkan untuk meningkatkan kemampuan *Convolutional Neural Network* dalam pengenalan citra dengan meningkatkan kedalaman jaringan. Penelitian Simonyan dan Zisserman menunjukkan bahwa peningkatan kedalaman hingga 16–19 *weight layers* dapat memberikan peningkatan kinerja dibandingkan konfigurasi sebelumnya. Model VGG juga berhasil memperoleh posisi pertama dan kedua pada beberapa kategori dalam kompetisi ImageNet 2014, sehingga menunjukkan kemampuannya dalam tugas pengenalan dan klasifikasi citra berskala besar (Simonyan & Zisserman, 2015a).

3. Mendukung penggunaan Transfer Learning

VGG16 tersedia dalam bentuk model yang telah dilatih menggunakan dataset ImageNet. Dokumentasi Keras menyediakan VGG16 dengan bobot *pre-trained* ImageNet dan secara khusus menyebutkan penggunaannya untuk *transfer learning* dan *fine-tuning*. Dengan demikian, model tidak harus dilatih

sepenuhnya dari awal ketika digunakan untuk tugas klasifikasi pada dataset tertentu (Keras, 2026).

4. Mampu mempelajari fitur citra secara bertahap

Struktur VGG16 yang terdiri dari beberapa *convolutional layers* memungkinkan jaringan mempelajari representasi citra secara bertingkat. Lapisan awal dapat menangkap karakteristik visual yang lebih sederhana, sedangkan lapisan yang lebih dalam membentuk representasi yang lebih kompleks. Penggunaan beberapa lapisan konvolusi berukuran 3×3 secara berurutan menjadi salah satu karakteristik utama VGG yang memungkinkan pembelajaran representasi visual tersebut (Simonyan & Zisserman, 2015).

Kekurangan VGG16

1. Memiliki jumlah parameter yang sangat besar

Salah satu kelemahan utama VGG16 adalah jumlah parameternya yang besar. Berdasarkan dokumentasi Keras, VGG16 memiliki sekitar 138,4 juta parameter dengan ukuran model sekitar 528 MB. Jumlah parameter tersebut membuat VGG16 membutuhkan memori yang relatif besar dibandingkan model CNN yang lebih ringan (Keras, 2026).

2. Membutuhkan komputasi yang lebih besar

Banyaknya parameter dan operasi komputasi pada VGG16 menyebabkan model membutuhkan sumber daya komputasi yang lebih besar. Dalam perbandingan dengan MobileNetV1, VGG16 memiliki sekitar 138,62 juta parameter dan 27,17 GFLOPs, sedangkan MobileNetV1 hanya memiliki sekitar 4,25 juta parameter dan 1,08 GFLOPs. Hal tersebut menunjukkan bahwa VGG16 memiliki kompleksitas komputasi yang jauh lebih tinggi dibandingkan MobileNetV1 (Lyu dkk., 2022).

3. Kurang efisien untuk perangkat dengan sumber daya terbatas.

Ukuran model dan kebutuhan komputasi VGG16 membuatnya kurang efisien untuk penerapan pada perangkat yang memiliki keterbatasan sumber daya, seperti perangkat seluler atau sistem *edge computing*. Sebagai perbandingan, penelitian MobileNet menunjukkan bahwa MobileNet dirancang dengan jumlah parameter dan kebutuhan komputasi yang jauh lebih rendah dibandingkan VGG16. MobileNet dilaporkan hampir setara dengan