BAB II
TINJAUAN PUSTAKA

Bab ini menyajikan landasan teoretis yang komprehensif, tinjauan empiris yang mendalam, serta kerangka konseptual yang menjadi fondasi utama bagi penelitian ini. Pembahasan disusun secara sistematis, dimulai dari aspek biologis mangga, standar mutu SNI, landasan matematis *Deep Learning* dan CNN, spesifikasi teknis arsitektur ResNet50 dan MobileNetV2, hingga sintesis kerangka pemikiran.

2.1 Landasan Biologis: Botani dan Patologi Buah Mangga

Pengembangan sistem visi komputer yang *robust* untuk aplikasi pertanian menuntut pemahaman mendalam tentang karakteristik biologis objek. Algoritma *Deep Learning* tidak sekadar memproses matriks piksel secara abstrak, melainkan mengenali representasi digital dari fenomena biologis yang kompleks.

2.1.1 Karakteristik Morfologi Mangifera indica L.

Mangga (*Mangifera indica* L.) adalah tanaman dari famili *Anacardiaceae* yang berasal dari kawasan perbatasan India-Burma dan telah menyebar ke seluruh Asia Tenggara sejak ribuan tahun silam (Doan & Thi, 2023). Buah mangga diklasifikasikan sebagai buah batu (*drupe*) dengan struktur perikarp tiga lapisan:

1. Eksokarp (Kulit): Lapisan terluar yang tebal, glandular, dan berlilin. Mengandung pigmen klorofil, karotenoid, dan antosianin yang berubah dinamis selama proses pematangan (*ripening*) (Doan & Thi, 2023). Variabilitas tekstur permukaan dan densitas lentisel menjadi fitur pembeda utama antarvarietas.
2. Mesokarp (Daging Buah): Bagian berserat, kaya air, dan dapat dikonsumsi yang menjadi representasi nilai ekonomi utama buah mangga.
3. Endokarp: Lapisan pelindung biji yang keras dan berserabut.

Di Indonesia, varietas *Arumanis* memiliki bentuk lonjong-pipih dengan kulit hijau tua meski matang, dilapisi lilin (*pruinosa*) kelabu. Varietas *Gadung* serupa namun dengan lekukan paruh lebih jelas dan densitas lentisel lebih padat.

2.1.2 Patologi Penyakit Pascapanen Utama

1. Antraknosa (*Colletotrichum gloeosporioides*): Antraknosa merupakan penyakit pascapanen paling merusak. Patogen ini memiliki mekanisme infeksi laten (*quiescent infection*): spora menembus kutikula buah muda dan tetap dorman hingga buah matang (Dofuor, Asante, & Oppong, 2023). Saat pematangan, infeksi aktif dan bermanifestasi sebagai bercak cokelat tua/hitam (*necrotic lesions*) yang cekung (*sunken*). Pola visual lesi ini sangat halus pada tahap awal, menuntut kemampuan resolusi fitur tinggi dari model CNN (Bana, Kumar, & Bhatt, 2023).
2. Lalat Buah (*Bactrocera dorsalis*): *Bactrocera dorsalis* adalah hama karantina utama. Kerusakan ditimbulkan melalui proses peletakan telur (*oviposition*) yang meninggalkan titik hitam mikroskopis (*oviposition stings*) pada permukaan kulit. Aktivitas larva menyebabkan pembusukan jaringan mesokarp dan keluarnya cairan (*oozing*) (Cheseto, Rering, Broadhead, Torto, & Beck, 2023). Deteksi dini gejala ini sangat krusial untuk mencegah lolosnya buah terinfeksi ke pasar ekspor.

2.1.3 Standar Mutu Nasional (SNI 3164:2009)

Penelitian ini mengacu pada SNI 3164:2009 yang membagi mangga ke dalam tiga kelas mutu: Kelas Super (bebas cacat), Kelas A (cacat < 5% luas permukaan), dan Kelas B (cacat < 10%). Untuk keperluan klasifikasi otomatis, ketiga kelas disederhanakan menjadi: Layak (Fresh/Good, Kelas Super hingga B) dan Tidak Layak (Rotten/Defective, penyakit aktif atau cacat melebihi toleransi).

2.2 Dataset Penelitian

2.2.1 Spesifikasi Dataset Citra

Dataset terdiri dari citra digital buah mangga varietas Arumanis dan Gadung serta varietas internasional bermorfologi serupa. Citra diambil dengan variabilitas latar belakang tinggi (*cluttered background*) untuk mensimulasikan lingkungan penyortiran nyata yang tidak selalu ideal.

2.2.2 Prapemrosesan Data

Strategi augmentasi data diterapkan untuk menangani potensi *class imbalance*. Seluruh citra masukan diubah ukurannya (*resize*) menjadi dimensi standar 224×224 piksel. Penetapan dimensi ini memiliki landasan teknis yang kuat: ukuran 224×224 piksel merupakan dimensi *input* spasial yang ditetapkan sebagai standar baku untuk arsitektur ResNet50 maupun MobileNetV2 ketika diinisialisasi dengan bobot *pretrained* ImageNet. Penggunaan dimensi yang berbeda akan menciptakan ketidaksesuaian (*mismatch*) dengan representasi fitur internal bobot tersebut. Selain itu, dimensi 224×224 piksel merepresentasikan titik keseimbangan optimal: cukup tinggi untuk melestarikan fitur patologis halus seperti bercak antraknosa awal dan jejak *oviposition* lalat buah, sekaligus efisien secara komputasional untuk inferensi *real-time* pada perangkat (He, Zhang, Ren, & Sun, 2016) (Sandler, Howard, Zhu, Zhmoginov, & Chen, 2018). Nilai piksel kemudian dinormalisasi ke rentang [0,1] melalui operasi *rescaling*.

2.3 Landasan Teori: Deep Learning dan CNN

Deep Learning adalah subbidang *Machine Learning* yang berbasis pada *Artificial Neural Networks* (ANN) berlapis banyak, yang mampu mempelajari representasi fitur secara hierarkis langsung dari data mentah. Perbedaan fundamental dengan *Machine Learning* konvensional terletak pada mekanisme ekstraksi fitur: konvensional membutuhkan rekayasa fitur manual (*handcrafted features*) oleh pakar, sementara *Deep Learning* melakukan ekstraksi fitur otomatis (*automatic feature extraction*) secara terintegrasi dalam proses pembelajaran.

Arsitektur dasar ANN terdiri dari tiga komponen: (1) *Input Layer* yang menerima data mentah, (2) *Hidden Layers* yang melakukan serangkaian transformasi matematis untuk mengenali pola, dan (3) *Output Layer* yang menghasilkan keputusan klasifikasi akhir. Istilah "Deep" merujuk pada jumlah *hidden layers* yang ekstensif di dalam topologi jaringan.

Relevansi *Deep Learning* dengan *Computer Vision* sangat kuat. Melalui lapisan-lapisan pembelajaran bertingkat, model memetakan fitur dari tingkat rendah (tepi, gradien) menuju fitur tingkat tinggi (representasi abstrak objek). Paradigma

ini ideal untuk klasifikasi objek visual dengan variabilitas morfologi tinggi seperti pada deteksi kelayakan mangga (Purwono, et al., 2022).

2.3.1 Convolutional Neural Networks (CNN)

MultiLayer Perceptron (MLP) konvensional mewajibkan transformasi matriks citra menjadi vektor satu dimensi (*flattening*), yang menghilangkan korelasi spasial dan topologis antarpiksel. Kehilangan informasi spasial ini fatal untuk klasifikasi mangga karena fitur penentu kualitas pola lesi antraknosa, kurvatura paruh varietas Gadung, distribusi lentisel sangat bergantung pada struktur relasional dua dimensi. Oleh karena itu, diimplementasikan *Convolutional Neural Networks* (CNN) yang mempertahankan informasi spasial melalui operasi konvolusi diskrit:

$$S(i,j) = (I * K)(i,j) = \sum\sum I(i-m, j-n) \cdot K(m,n) + b \quad (2.1)$$

di mana S(i,j) adalah nilai pada *feature map* pada koordinat (i,j); I adalah matriks citra masukan; K adalah matriks kernel/filter; m,n adalah indeks iterasi kernel; dan b adalah nilai bias. Melalui mekanisme *sliding kernel* dan *weight sharing*, CNN mengekstraksi fitur secara hierarkis: dari tepi dan gradien pada lapisan awal hingga representasi objek kompleks pada lapisan akhir (Purwono et al., 2022).

2.3.2 Fungsi Aktivasi ReLU (NonLinear)

Fungsi aktivasi *Rectified Linear Unit* (ReLU) didefinisikan sebagai

$$f(x) = \max(0, x) \quad (2.2)$$

ReLU memperkenalkan nonlinearitas secara komputasional efisien: nilai negatif di-zero-out, nilai positif dilewatkan. Keunggulan krusialnya adalah mitigasi masalah *vanishing gradient*: gradiennya konstan bernilai 1 untuk semua input positif, memungkinkan pelatihan jaringan sangat dalam seperti ResNet50 menjadi stabil dan konvergen.

2.3.3 Pooling, Stride, Epoch, Batch Size, dan Hyperparameter Tuning

Pooling Layer mereduksi dimensi spasial *feature map* untuk mengurangi beban komputasi dan menginduksi invariansi translasi lokal. *Max Pooling*

mengambil nilai maksimum dalam setiap jendela, sedangkan *Global Average Pooling* meratakan dimensi spasial keseluruhan (Purwono et al., 2022).

Stride adalah langkah pergeseran filter konvolusi. *Stride* > 1 mereduksi dimensi spasial secara efisien tanpa lapisan *pooling* tambahan. ResNet50 memanfaatkan *stride* pada blok konvolusi tertentu untuk mengontrol reduksi dimensi secara terstruktur (He et al., 2016). Epoch adalah satu iterasi lengkap seluruh data latih melalui jaringan. Jumlah *epoch* optimal ditentukan adaptif menggunakan *Early Stopping* dengan *patience* 10 epoch.

Batch Size adalah jumlah sampel yang diproses sebelum pembaruan bobot. *Batch size* 32 digunakan sebagai kompromi optimal antara efisiensi komputasi dan kualitas estimasi gradien (Doan & Thi, 2023). Hyperparameter Tuning adalah optimasi parameter yang ditetapkan sebelum pelatihan. Parameter paling kritis adalah *Learning Rate* (η) yang mengontrol besaran pembaruan bobot:

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla L(\theta_t) \tag{2.3}$$

Nilai *learning rate* awal 10^{-4} ditetapkan untuk fase *Transfer Learning* menggunakan algoritma Adam (Kingma & Ba, 2015). Strategi *ReduceLROnPlateau* diterapkan untuk menurunkan *learning rate* secara otomatis saat validasi stagnan.

2.4 Arsitektur ResNet50

ResNet50 diperkenalkan oleh He et al. (2016) dan memenangkan kompetisi ILSVRC 2015. Inovasi fundamentalnya adalah *Residual Block* dengan *skip connection*: alih-alih mempelajari pemetaan H(x) secara langsung, setiap blok mempelajari fungsi residu $F(x) = H(x) - x$, sehingga output menjadi F(x) + x. Formulasi ini memungkinkan gradien mengalir langsung melalui *shortcut connection* selama *backpropagation*, mengatasi masalah *vanishing gradient* dan memungkinkan pelatihan jaringan 50 lapisan secara stabil (He, Zhang, Ren, & Sun, 2016).

ResNet50 mengimplementasikan *Bottleneck Block* dengan tiga lapisan berurutan: konvolusi 1×1 (reduksi dimensi kanal), konvolusi 3×3 (ekstraktor fitur utama), dan konvolusi 1×1 (ekspansi kembali). Arsitektur ini memiliki sekitar 25,6

juta parameter, menjadikannya ideal secara teoritis untuk mendeteksi fitur patologis halus seperti bintik antraknosa awal (Jayanthi & Kumar, 2024).

2.5 Arsitektur MobileNetV2

MobileNetV2 dirancang oleh Sandler et al. (2018) untuk memaksimalkan kinerja pada perangkat *mobile* dan *embedded systems*. Dua inovasi arsitektural utamanya adalah:

1. Depthwise Separable Convolution

   Dekomposisi konvolusi standar menjadi: (1) *Depthwise Convolution* satu filter per kanal input secara independen, dan (2) *Pointwise Convolution* (1×1) menggabungkan output lintas kanal. Dekomposisi ini mereduksi operasi hingga 8-9 kali lipat dibanding konvolusi standar (Sandler, Howard, Zhu, Zhmoginov, & Chen, 2018).

2. Inverted Residual Block with Linear Bottleneck

   Setiap blok: (1) memperluas dimensi kanal 6x, (2) menerapkan *depthwise convolution*, (3) mengkompres kembali ke dimensi rendah. Lapisan akhir menggunakan aktivasi linear (bukan ReLU) untuk mencegah hilangnya informasi pada dimensi rendah. MobileNetV2 hanya memiliki sekitar 3,4 juta parameter kurang dari 14% ResNet50 menjadikannya kandidat utama untuk implementasi pada *smartphone* bertenaga rendah (Hermanto, Aziz, & Sudianto, 2024).

2.6 Metrik Evaluasi Kinerja

Evaluasi kinerja model dilakukan menggunakan metrik berbasis Confusion Matrix. Untuk klasifikasi biner (Layak vs. Tidak Layak), empat komponen dasar didefinisikan sesuai pemetaan label model: TP (True Positive) yaitu buah Layak yang diprediksi benar sebagai Layak; TN (True Negative) yaitu buah Tidak Layak yang diprediksi benar sebagai Tidak Layak; FP (False Positive) yaitu buah Tidak Layak yang lolos dan salah diprediksi sebagai Layak (Kesalahan Tipe I); serta FN (False Negative) yaitu buah Layak yang salah diprediksi sebagai Tidak Layak (Kesalahan Tipe II).

1. Akurasi: $\frac{(TP+TN)}{(TP+TN+FP+FN)}$ proporsi prediksi benar secara keseluruhan pada kedua kelas.
2. Presisi: $\frac{(TP)}{(TP+FP)}$ mengukur ketepatan prediksi kelas positif serta meminimalkan lolosnya buah cacat (False Positive) ke dalam kategori layak.
3. Recall: $\frac{(TP)}{(TP+FN)}$ mengukur sensitivitas model dalam mengenali sampel pada masing-masing kelas. Dalam evaluasi per kelas (per-class evaluation), Recall Kelas Tidak Layak dan Presisi Kelas Layak merupakan indikator paling krusial karena kegagalan mendeteksi buah cacat (tingginya False Positive pada kelas Layak) menyebabkan buah terinfeksi lolos ke rantai pasok ekspor dan berpotensi mengakibatkan penolakan kontainer di pasar ekspor premium.
4. F1-Score: $2 \times \frac{(Presisi \times Recall)}{(Presisi + Recall)}$ rata-rata harmonik antara Presisi dan Recall yang memberikan penalti tegas apabila salah satu komponen bernilai rendah.
5. Efisiensi Komputasi: *Training time (detik) dan inference time rata-rata per citra (milidetik).*

2.7 Kerangka Pemikiran

1. Masalah: Klasifikasi manual mangga subjektif dan rentan kesalahan. Antraknosa laten dan bekas lalat buah sulit dideteksi dini, mengakibatkan kerugian ekspor signifikan.
2. Kesenjangan: Belum terdapat studi komparatif ResNet50 vs. MobileNetV2 pada mangga Indonesia dengan standar SNI 3164:2009, dataset *cluttered background*, dan orientasi implementasi *mobile*
3. Solusi: Analisis komparatif sistematis ResNet50 (~25,6 juta parameter) vs. MobileNetV2 (~3,4 juta parameter) menggunakan teknik *Transfer Learning*.
4. Validasi: Pengujian dengan metrik statistik (Akurasi, Presisi, Recall, F1-Score) dan efisiensi komputasi (waktu pelatihan dan inferensi) untuk rekomendasi teknis berbasis data.