

Klasifikasi Penyakit Daun Mangga Menggunakan CNN Berbasis Transfer Learning Dengan Model Arsitektur VGG16, DenseNet121, dan InceptionV3

Zaky Dwi Purnomo*, Wahyu Aji Eko Prabowo

Fakultas Ilmu Komputer, Teknik Informatika, Universitas Dian Nuswantoro, Semarang, Indonesia

Email: ¹111202214227@mhs.dinus.ac.id, ²prabowo@dsn.dinus.ac.id

Email Penulis Korespondensi: 111202214227@mhs.dinus.ac.id

Submitted 14-01-2026; Accepted 24-02-2026; Published 28-02-2026

Abstrak

Mangga adalah salah satu buah yang penting di Indonesia, tetapi produksinya sering terganggu karena adanya penyakit dan hama pada daun yang sulit dideteksi secara dini. Metode pengenalan penyakit secara manual biasanya tergantung pada pengamat dan tidak selalu akurat. Penelitian ini bertujuan membuat sistem otomatis untuk mengklasifikasikan penyakit daun mangga, dengan menggunakan teknik deep learning berbasis algoritma Convolutional Neural Network (CNN). Penelitian ini juga membandingkan tiga model, yaitu VGG16, DenseNet121, dan InceptionV3, dengan menerapkan metode transfer learning. Dataset yang digunakan terdiri dari 4.000 gambar yang dibagi merata ke dalam 8 kategori, terdiri dari 7 jenis penyakit (Anthracnose, Bacterial Canker, Cutting Weevil, Die Back, Gall Midge, Powdery Mildew, Sooty Mould) dan 1 kategori tanaman yang sehat. Evaluasi dilakukan dengan metode 5-Fold Cross-Validation untuk memastikan hasil yang valid. Hasil menunjukkan bahwa ketiga model mampu memberikan akurasi lebih dari 90%. Model VGG16 menunjukkan performa terbaik dan paling stabil, dengan akurasi 93,25%, Presisi 0,93, Recall 0,93, F1-Score 0,93, serta nilai AUC-ROC 0,98. Sementara InceptionV3 mencapai akurasi 92,38% dan DenseNet121 mencapai 91,25%. Oleh karena itu, VGG16 disarankan sebagai model utama karena kemampuannya yang lebih baik dalam mengekstrak fitur tekstur dan mengenali penyakit daun mangga secara tepat. Model arsitektur VGG16 yang sederhana mampu mengungguli model kompleks dalam ekstraksi fitur tekstur daun mangga secara efisien, sehingga sangat potensial dijadikan basis untuk aplikasi diagnosis penyakit tanaman real-time bagi petani.

Kata Kunci: Klasifikasi Penyakit Daun Mangga; Convolutional Neural Network; VGG16; DenseNet121; InceptionV3

Abstract

Mango is one of the important fruits in Indonesia, but its production is often disrupted by leaf diseases and pests that are difficult to detect early. Manual disease recognition methods usually depend on observers and are not always accurate. This study aims to create an automated system to classify mango leaf diseases, using deep learning techniques based on the Convolutional Neural Network (CNN) algorithm. This study also compares three models, namely VGG16, DenseNet121, and InceptionV3, by applying the transfer learning method. The dataset used consists of 4,000 images divided evenly into 8 categories, consisting of 7 types of diseases (Anthracnose, Bacterial Canker, Cutting Weevil, Die Back, Gall Midge, Powdery Mildew, Sooty Mold) and 1 category of healthy plants. Evaluation was carried out using the 5-Fold Cross-Validation method to ensure valid results. The results show that all three models are able to provide an accuracy of more than 90%. The VGG16 model showed the best and most stable performance, with an accuracy of 93.25%, a Precision of 0.93, a Recall of 0.93, an F1-Score of 0.93, and an AUC-ROC of 0.98. Meanwhile, InceptionV3 achieved an accuracy of 92.38% and DenseNet121 reached 91.25%. Therefore, VGG16 is recommended as the primary model due to its better ability to extract texture features and accurately recognize mango leaf diseases. VGG16 architecture is able to outperform complex models in efficiently extracting mango leaf texture features, making it very potential to be used as a basis for real-time plant disease diagnosis applications for farmers.

Keywords: Mango Leaf Disease Classification; Convolutional Neural Network; VGG16; DenseNet121; InceptionV3

1. PENDAHULUAN

Mangga adalah buah musiman yang sudah dikenal oleh banyak orang di Indonesia. Buah ini pertama kali berasal dari India dan memiliki berbagai jenis yang disukai oleh masyarakat Indonesia [1]. Sebagai negara tropis dengan keanekaragaman agroklimate yang tinggi, Indonesia memiliki potensi besar dalam produksi buah-buahan, termasuk mangga. Kondisi iklim yang mendukung memungkinkan tumbuhnya berbagai varietas mangga dengan kualitas yang baik. Mangga merupakan salah satu komoditas buah unggulan di Indonesia. Walaupun demikian, industri mangga Indonesia masih menghadapi berbagai tantangan, terutama terkait penyakit tanaman yang dapat mempengaruhi produktivitas dan kualitas buah. Proses budidaya tanaman mangga tidak selalu bebas dari serangan penyakit dan hama. Gejala-gejala yang muncul dapat menunjukkan jenis penyakit yang menyerang tanaman. Penyakit ini dapat disebabkan oleh faktor parasit (patogen, hama, gulma) atau non-parasit (air, suhu, cahaya, nutrisi). Selain itu, perubahan iklim juga memperparah situasi dengan menciptakan kondisi ideal bagi perkembangan patogen, yang mengakibatkan penurunan hasil panen [2].

Berdasarkan data dari jurnal *National Library of Medicine*, produktivitas buah mangga sangat bergantung pada tingkat kesehatan pohon mangga, di mana salah satu indikator utamanya adalah kondisi daun. Kesehatan pohon ini sangat terancam oleh penyakit daun mangga, khususnya Antraknosa (MAD) yang disebabkan oleh kompleks jamur *Colletotrichum*. Penyakit ini berdampak langsung pada produktivitas, dengan potensi kehilangan hasil yang bisa mencapai 100% di perkebunan yang tidak dikelola. Tantangan utamanya adalah penyakit ini seringkali tidak terdeteksi (laten) hingga infeksi menyebar luas, sementara hampir semua kultivar mangga komersial rentan terhadapnya [3].

Menurunnya efektivitas pestisida sintesis akibat resistensi jamur semakin menegaskan bahwa pengelolaan kesehatan pohon secara terpadu dan ramah lingkungan sangat krusial untuk menjaga produktivitas buah mangga [3].

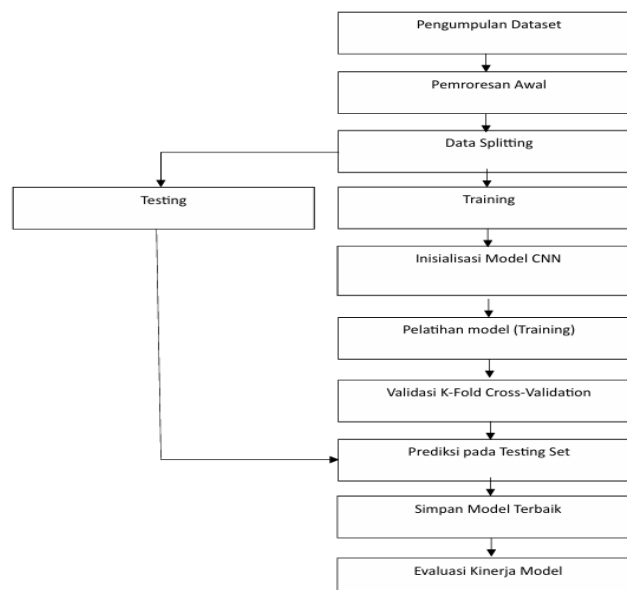
Untuk mengatasi kelemahan deteksi manual yang bersifat subjektif dan kurang akurat, penelitian ini mengusulkan pengembangan sistem klasifikasi otomatis penyakit daun mangga menggunakan pendekatan *deep learning* berbasis pengolahan citra digital. Perkembangan teknologi masa kini sudah mengalami perkembangan yang sangat pesat. Sehubungan dengan hal ini, kebutuhan akan mengelola teknologi modern menjadi salah satu kunci utama. Kemajuan dalam kecerdasan buatan (AI) dan visi komputer (*computer vision*) telah membuka jalan bagi deteksi penyakit secara dini dan akurat. kebutuhan akan pengelolaan yang lebih baik ini, teknologi modern kini memegang peranan kunci [4].

Secara khusus, kemajuan dalam kecerdasan buatan (AI) dan *computer vision* telah membuka jalan bagi deteksi penyakit secara dini dan akurat. Pendekatan berbasis teknologi ini memungkinkan intervensi yang cepat dan terarah prinsip dasar dari pertanian presisi sehingga mengurangi ketergantungan pada pestisida kimia dan pada akhirnya mendukung produktivitas pohon mangga secara lebih berkelanjutan. Metode yang akan diimplementasikan adalah Convolutional Neural Network (CNN). Secara spesifik, penelitian ini akan menggunakan arsitektur DenseNet121, InceptionV3, dan VGG16. Ketiganya adalah arsitektur canggih yang dikenal dengan efisiensi dan kinerjanya yang tinggi dalam tugas analisis citra [5]. DenseNet121 dikenal dengan strukturnya yang unik, di mana setiap lapisan terhubung ke semua lapisan berikutnya (konektivitas padat), yang mendorong penggunaan kembali fitur dan memperkuat propagasi gradien [6]. Arsitektur InceptionV3 juga terbukti sangat efektif untuk klasifikasi penyakit daun mangga, seringkali diimplementasikan menggunakan transfer learning yang dapat membantu mengatasi overfitting pada dataset yang terbatas [7]. Sementara itu, VGG16 dikenal dengan arsitektur berlapis sederhana yang menawarkan stabilitas tinggi dalam proses klasifikasi. Dalam studi komparasi deteksi penyakit daun mangga, VGG16 terbukti efektif dan konsisten memberikan performa terbaik dengan akurasi mencapai 100% pada skenario pengujian tertentu. Penggunaan kombinasi arsitektur ini sangat cocok untuk analisis citra pertanian yang kompleks [8].

Untuk mengatasi ancaman penyakit daun mangga, khususnya Antraknosa, yang dapat menyebabkan gagal panen dan sulit dideteksi secara akurat menggunakan metode manual, diperlukan sebuah sistem klasifikasi otomatis berbasis citra menggunakan pendekatan *deep learning* berbasis algoritma Convolutional Neural Network (CNN). Meskipun berbagai penelitian sebelumnya telah menerapkan metode CNN untuk klasifikasi penyakit daun mangga, kebanyakan dari studi tersebut memiliki kelemahan berupa evaluasi yang terbatas pada satu arsitektur tunggal dan kurangnya komparasi yang seragam pada dataset dengan jumlah kelas yang banyak. Berdasarkan celah riset tersebut, penelitian ini berkontribusi untuk mengembangkan sistem klasifikasi yang mengkomparasikan tiga arsitektur CNN secara adil (VGG16, DenseNet121, InceptionV3) menggunakan *transfer learning*. Diharapkan hasil komparasi ini dapat memberikan rekomendasi model paling objektif dan aplikatif sebagai alat bantu diagnosis langsung bagi para petani [9].

2. METODOLOGI PENELITIAN

Penelitian ini menerapkan pendekatan eksperimental menggunakan metode *deep learning* untuk klasifikasi penyakit daun mangga. Tahapan penelitian disusun secara sistematis untuk membandingkan kinerja tiga arsitektur Convolutional Neural Network (CNN) yang berbeda [10], [11]. Secara umum, alur penelitian (*research flowchart*) yang dilakukan dapat dilihat pada Gambar 1. Berdasarkan Gambar 1, tahapan penelitian secara berurutan diawali dengan pengumpulan dataset, pra-pemrosesan citra, pembagian data latih dan uji, hingga proses pelatihan dan evaluasi model CNN. Masing-masing tahapan ini akan diuraikan lebih rinci pada sub-bab berikut.



Gambar 1. Alur Penelitian Klasifikasi Penyakit Daun Mangga

2.1 Dataset

Dataset yang digunakan dalam penelitian ini terdiri dari 4.000 citra daun mangga yang dibagi secara merata ke dalam 8 kategori, sehingga setiap kategori memiliki 500 gambar dan distribusi data bersifat seimbang. Delapan kategori tersebut mencakup 7 jenis penyakit dan 1 kategori daun sehat. Jenis penyakit yang digunakan meliputi *Anthraxnose*, *Bacterial Canker*, *Cutting Weevil*, *Die Back*, *Gall Midge*, *Powdery Mildew*, dan *Sooty Mould*, sedangkan kategori sehat merepresentasikan daun mangga yang tidak terinfeksi penyakit. Contoh citra dari masing-masing kelas penyakit dan daun sehat pada dataset ini ditunjukkan pada Gambar 2. Dataset daun mangga tersebut diperoleh dari platform Kaggle dan telah banyak digunakan dalam penelitian klasifikasi penyakit daun berbasis pengolahan citra dan *deep learning* [11]. Seperti yang terlihat pada Gambar 2, setiap kelas penyakit memiliki karakteristik visual berupa pola bercak, tekstur, dan warna yang berbeda pada permukaan daun. Perbedaan ciri visual inilah yang akan diekstraksi dan dipelajari oleh model jaringan saraf tiruan.



Gambar 2. Contoh Jenis-Jenis Penyakit yang Ada pada Dataset

2.2 Pra-pemrosesan Data

Citra asli dari dataset memiliki resolusi 240 x 320 piksel. Sebelum dimasukkan ke dalam model, citra melewati tahapan pra-pemrosesan (*preprocessing*) untuk menyeragamkan format input. Tahapan ini meliputi: tahapan pra-pemrosesan (*pre-processing*) data citra yang dilakukan sebelum gambar daun mangga diproses oleh model *Convolutional Neural Network* (CNN), yang di tunjukkan oleh Tabel 1. Tahapan pertama dalam *pre-processing* adalah *resizing*, yaitu mengubah dimensi spasial seluruh citra masukan menjadi ukuran seragam sebesar 128×128 piksel agar sesuai dengan standar input arsitektur CNN seperti VGG16, DenseNet121, dan InceptionV3, sekaligus mengurangi beban komputasi selama proses pelatihan. Tahapan berikutnya adalah *rescaling*, yaitu normalisasi nilai piksel dari rentang 0-255 menjadi 0-1, yang bertujuan untuk menyeragamkan skala fitur data dan mempercepat konvergensi gradien selama proses pelatihan model. Dengan menerapkan kedua tahapan ini, data citra menjadi lebih terstruktur, efisien untuk diproses, dan mendukung peningkatan stabilitas serta kinerja model dalam mengklasifikasikan penyakit daun mangga.

Tabel 1. Tahapan Preprocessing

No	Tahapan	Deskripsi Proses	Tujuan/Manfaat
1	Resizing	Mengubah dimensi spasial seluruh citra masukan menjadi ukuran seragam, yaitu 128×128 piksel.	1. Menyesuaikan standar input arsitektur CNN (VGG16, DenseNet121, InceptionV3). 2. Mengurangi beban komputasi (<i>computational cost</i>).
2	Rescaling	Menormalisasi nilai piksel citra dari rentang integer 0–255 menjadi rentang pecahan 0–1.	1. Mempercepat konvergensi gradien (<i>gradient convergence</i>) selama proses pelatihan model. 2. Menyeragamkan skala fitur data.

2.3 Data Splitting

Dataset dibagi menjadi dua bagian besar, yaitu data latihan dan data uji. Pembagian dilakukan dengan perbandingan 80:20, di mana 80% dari data atau sebanyak 3.200 gambar digunakan untuk mengajarkan model mengenali ciri-ciri penyakit, sedangkan 20% yang tersisa atau sebanyak 800 gambar digunakan untuk menguji apakah model bekerja dengan baik pada data yang belum pernah dilihat sebelumnya [12].

2.4 Arsitektur *Convolutional Neural Network*

Penelitian ini membandingkan tiga arsitektur CNN populer menggunakan teknik *transfer learning* [13]. Bobot awal model diambil dari pelatihan pada dataset ImageNet, kemudian lapisan *fully connected* dimodifikasi untuk klasifikasi 8 kelas penyakit daun mangga [14], [15].

2.4.1 VGG16

VGG16 adalah arsitektur CNN yang dikembangkan oleh *Visual Geometry Group* (VGG) dari Universitas Oxford [16]. Arsitektur ini dikenal dengan kesederhanaannya yang menggunakan filter konvolusi berukuran kecil (3 x 3) dengan *stride* 1 dan *padding* 1, serta lapisan *max-pooling* 2 x 2 dengan *stride* 2. Struktur ini memungkinkan jaringan untuk memiliki kedalaman hingga 16 lapisan (*weight layers*) sambil menjaga jumlah parameter tetap terkendali. Penggunaan filter kecil yang bertumpuk ini terbukti efektif dalam menangkap fitur detail pada citra penyakit tanaman, seperti tekstur dan tepi bercak penyakit, yang krusial untuk klasifikasi akurat [17].

2.4.2 DenseNet121

Jaringan saraf berlapis padat (DenseNet) menggunakan konsep koneksi yang rapat, di mana setiap lapisan terhubung langsung ke semua lapisan lainnya dalam satu blok (*dense block*) [18]. Arsitektur DenseNet121 terdiri dari 121 lapisan yang dirancang untuk mengatasi masalah gradient yang menghilang, yang sering terjadi pada jaringan saraf yang sangat dalam. Setiap lapisan menerima peta fitur dari semua lapisan sebelumnya sebagai masukan, serta meneruskan peta fitur yang dimilikinya ke semua lapisan berikutnya. Cara kerja ini memungkinkan penggunaan kembali fitur secara efisien, sehingga mampu menghasilkan prediksi yang akurat dengan jumlah parameter yang lebih sedikit dibandingkan dengan arsitektur jaringan saraf tradisional [4].

2.4.3 InceptionV3

InceptionV3 merupakan pengembangan dari arsitektur GoogLeNet yang memperkenalkan modul Inception. Modul ini memungkinkan jaringan untuk melakukan proses konvolusi dengan berbagai ukuran filter (1 x 1, 3 x 3, dan 5 x 5) serta pooling secara paralel pada satu lapisan input yang sama [19]. Keunggulan utama InceptionV3 terletak pada kemampuannya menangkap fitur citra pada berbagai skala (*multi-scale feature extraction*), baik fitur yang bersifat lokal (detail kecil) maupun global (bentuk keseluruhan daun). Pendekatan ini sangat relevan untuk dataset penyakit daun mangga yang memiliki variasi ukuran dan bentuk gejala penyakit yang beragam [20].

2.5 Pelatihan (*Training*)

Seluruh model arsitektur dilatih menggunakan skenario dan hyperparameter yang seragam. Proses pelatihan diimplementasikan dalam bahasa *Python* dengan menggunakan library TensorFlow/Keras [21] yang dieksekusi pada Google Colab. Model dikompilasi menggunakan algoritma optimasi Adam (*Adaptive Moment Estimation*) [22]. Adam dipilih karena kemampuannya yang efisien dalam menangani data citra berskala besar, serta mampu menyesuaikan learning rate secara adaptif sehingga mempercepat konvergensi parameter jaringan. Berdasarkan pengujian awal, pelatihan dilakukan secara optimal selama 20 epoch. Jumlah iterasi ini terbukti cukup untuk mencapai keseimbangan antara konvergensi akurasi yang maksimal dan pencegahan overfitting tanpa membebani komputasi secara berlebihan..

2.6 Validasi K-Fold *Cross-Validation*

Sesuai dengan alur penelitian pada Gambar 1, penelitian ini menggunakan metode K-Fold *Cross-Validation* [23] untuk memeriksa kemampuan model dan memastikan hasil akurasi yang didapat tidak tergantung pada cara pembagian data tertentu. Metode yang digunakan adalah 5-Fold Cross-Validation dengan nilai k = 5. Dengan demikian, setiap gambar dalam dataset memiliki kesempatan untuk digunakan sebagai data uji sebanyak satu kali. Akhirnya, hasil kemampuan model dihitung berdasarkan rata-rata akurasi dari lima kali pengujian tersebut. Pendekatan ini memberikan penilaian performa yang lebih kuat dan mengurangi kemungkinan model terlalu cocok dengan data latihan dibandingkan dengan metode pemisahan data secara statis biasa [24], [25].

2.7 Evaluasi Kerja

Evaluasi model dilakukan dengan menggunakan Confusion Matrix, yaitu metode yang menampilkan hasil prediksi model terhadap label sebenarnya. Dari hasil pemetaan tersebut, diperoleh empat jenis utama, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Berdasarkan keempat jenis ini, kinerja model diukur dengan beberapa metrik penting. Metrik pertama adalah Akurasi, yang menunjukkan perbandingan jumlah prediksi yang benar, baik positif maupun negatif, terhadap total data. Rumus untuk menghitung akurasi adalah:

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

Selain akurasi, ada aspek lain yang dinilai yaitu Presisi dan Recall. Presisi menunjukkan seberapa tepat model dalam mengklasifikasikan data positif yang diprediksi, sehingga bisa mengurangi kesalahan prediksi yang mengatakan sesuatu positif padahal sebenarnya bukan. Rumus untuk menghitung Presisi adalah sebagai berikut:

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

Sementara itu, Recall mengukur kemampuan model dalam menemukan kembali seluruh data positif yang sebenarnya ada di dalam dataset, dengan persamaan:

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

Untuk menyeimbangkan evaluasi terutama jika distribusi data antar kelas tidak merata digunakan F1-Score. Metrik ini merupakan rata-rata harmonik antara presisi dan recall, yang memberikan gambaran performa model secara lebih komprehensif melalui rumus:

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

Selain metrik berbasis ambang batas tunggal seperti akurasi dan presisi, evaluasi model klasifikasi juga sering menggunakan analisis *Receiver Operating Characteristic* (ROC) dan *Area Under the Curve* (AUC). Kurva ROC menunjukkan bagaimana kinerja model berubah saat nilai ambang batas klasifikasi berbeda. Kurva ini menampilkan dua parameter, yaitu *True Positive Rate* (TPR) di sumbu vertikal dan *False Positive Rate* (FPR) di sumbu horizontal. TPR, yang sama dengan Recall, dihitung menggunakan rumus:

$$TPR = \frac{TP}{TP + FP} \quad (5)$$

FPR mengukur persentase data negatif yang salah diprediksi sebagai positif, dengan rumus:

$$FPR = \frac{FP}{TN + FP} \quad (6)$$

Nilai AUC menggambarkan luas area di bawah kurva ROC. Nilai AUC berada dalam rentang 0 hingga 1, dimana nilai yang mendekati 1 menunjukkan bahwa model sangat mampu membedakan antara kelas positif dan kelas negatif, terlepas dari nilai threshold yang digunakan.

3. HASIL DAN PEMBAHASAN

3.1 Tahapan Penerapan Algoritma

Tahapan algoritma klasifikasi ini dimulai dengan mengubah citra input berukuran 128x128 piksel menjadi format matriks angka (*tensor*) agar bisa dibaca oleh mesin. *Tensor* ini kemudian dimasukkan ke dalam lapisan konvolusi dari arsitektur model (VGG16, DenseNet121, atau InceptionV3) untuk proses ekstraksi fitur secara bertahap, mulai dari mengenali bentuk visual dasar seperti garis tepi daun, hingga mendeteksi pola yang lebih spesifik seperti tekstur bercak penyakitnya. Setelah peta fitur (*feature map*) ini didapatkan, ukurannya direduksi melalui tahap *pooling* dan diratakan (*flatten*) menjadi vektor satu dimensi agar bisa diproses oleh *Fully Connected Layer* [26]. Pada tahap akhir klasifikasi ini, fungsi *Softmax* akan menghitung persentase probabilitas untuk menentukan citra tersebut paling cocok masuk ke label mana dari total 8 kelas yang ada. Selama proses training yang berlangsung sebanyak 20 epoch, model secara terus-menerus belajar mengevaluasi tebakannya melalui mekanisme *backpropagation* [27]. Setiap kali model menghitung seberapa besar nilai kesalahannya (*loss*) dibandingkan dengan label asli, *optimizer* Adam [22] akan merespons dengan mengupdate dan menyesuaikan bobot (*weights*) parameter jaringan secara otomatis, sehingga dari satu epoch ke epoch berikutnya, model menjadi semakin presisi dalam mengenali penyakit dan mencapai akurasi yang maksimal.

3.2 Hasil Evaluasi Model

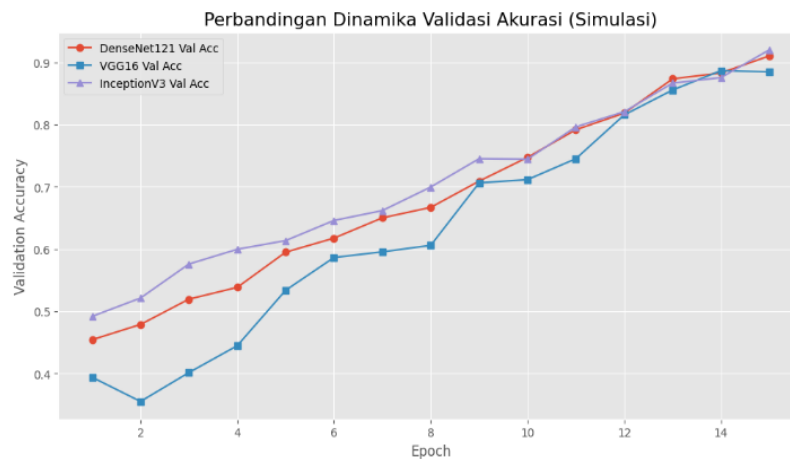
Evaluasi pertama dilakukan pada arsitektur InceptionV3. Hasil uji menunjukkan model ini memiliki performa yang cukup baik dengan tingkat akurasi mencapai 92,38%. Meskipun arsitektur ini terkenal sangat canggih dengan fitur modul inception yang mampu mengekstrak ciri dari berbagai skala, dalam kasus ini, InceptionV3 berada di peringkat kedua. Meskipun bukan yang terbaik, model ini tetap menunjukkan kemampuan generalisasi yang cukup baik dan kompetitif dalam mengenali pola penyakit pada daun mangga. Selanjutnya, evaluasi dilakukan pada arsitektur VGG16. Model ini menunjukkan stabilitas yang sangat baik dan justru mencatatkan kinerja terbaik dan paling superior dalam penelitian ini dengan akurasi sebesar 93,25%. Walaupun VGG16 memiliki struktur yang lebih sederhana dengan tumpukan lapisan konvolusi standar dibandingkan DenseNet atau Inception, arsitektur ini terbukti paling efektif dalam menangkap fitur visual tekstur pada dataset daun mangga ini. Tingginya akurasi ini menjadikan VGG16 sebagai model rekomendasi utama dalam penelitian ini. Terakhir, arsitektur DenseNet121 mencatatkan akurasi sebesar 91,25%. Meskipun memiliki mekanisme dense connectivity yang memungkinkan aliran informasi maksimal, dalam eksperimen ini kinerjanya berada di posisi terendah dibandingkan InceptionV3 dan VGG16. Hal ini kemungkinan disebabkan oleh kompleksitas model yang mungkin memerlukan data yang lebih besar atau penyetelan hyperparameter yang lebih spesifik agar dapat melampaui kinerja arsitektur lainnya. Meskipun demikian, angka di atas 90% tetap menunjukkan bahwa model ini bekerja dengan baik, walau tidak setinggi VGG16.

Rangkuman perbandingan kinerja ketiga model CNN pada data uji disajikan pada Tabel 2. Tabel tersebut menampilkan hasil evaluasi setiap arsitektur berdasarkan metrik akurasi, presisi, recall, F1-score, dan AUC-ROC. Berdasarkan Tabel 2, model VGG16 menunjukkan kinerja paling unggul dengan nilai akurasi dan F1-score tertinggi dibandingkan DenseNet121 dan InceptionV3, yang menandakan kemampuan klasifikasi yang lebih stabil dan konsisten dalam mengenali penyakit daun mangga.

Tabel 2. Matrik Evaluasi pada Data Testing

No	Model	Akurasi	Presisi	Recall	F1-Score	AUC-ROC
1	DenseNet121	0,9125	0,9128	0,9125	0,9124	0,9781
2	VGG16	0,9325	0,9333	0,9325	0,9326	0,9831
3	InceptionV3	0,9238	0,9260	0,9238	0,9242	0,9809

Berdasarkan Tabel 2, dapat dilihat bahwa VGG16 mencatatkan kinerja terbaik dibandingkan arsitektur lainnya, dengan akurasi mencapai 93,25%. Tingginya nilai F1-Score pada model ini menunjukkan keseimbangan yang baik antara nilai Presisi dan Recall, yang berarti model mampu meminimalkan kesalahan deteksi (positif palsu) sekaligus mampu mengenali sebagian besar sampel penyakit yang ada. Adapun dinamika pergerakan nilai akurasi dari ketiga model selama proses pelatihan dan validasi dapat dilihat secara visual pada Gambar 3 berikut. Gambar 3 menunjukkan grafik akurasi pada masing-masing model dalam mengenali pola penyakit daun mangga selama proses pelatihan.



Gambar 3. Grafik Jumlah Epoch terhadap Validasi Akurasi

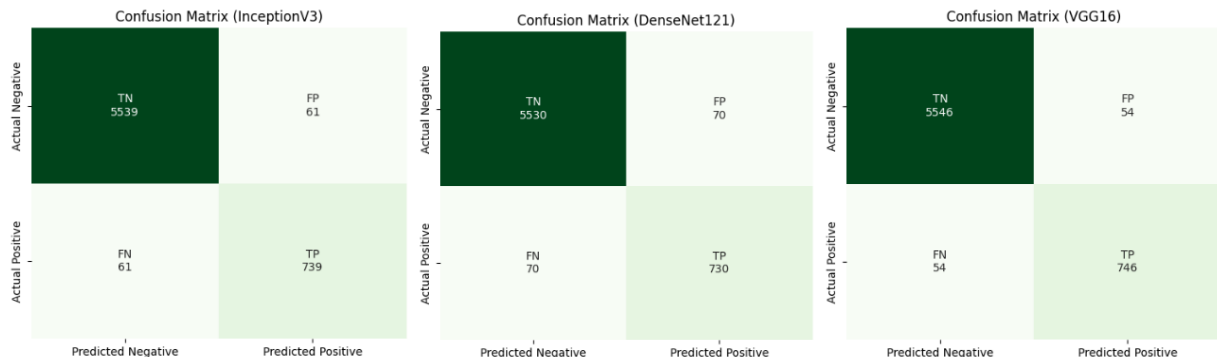
Terlihat bahwa nilai akurasi pada data pelatihan dan validasi mengalami peningkatan yang konsisten seiring bertambahnya jumlah *epoch*. Hal ini menandakan bahwa model secara bertahap berhasil mempelajari fitur-fitur penting dari citra daun mangga. Tidak terlihat adanya fluktuasi tajam atau penurunan akurasi yang signifikan, sehingga dapat disimpulkan bahwa proses pembelajaran berlangsung stabil. Selain itu, jarak antara kurva akurasi pelatihan dan validasi relatif kecil, yang menunjukkan bahwa model tidak mengalami *overfitting* dan memiliki kemampuan generalisasi yang baik terhadap data yang belum pernah dilihat sebelumnya. Selain metrik akurasi, performa stabilitas model selama pelatihan juga dievaluasi berdasarkan tingkat penurunan nilai kesalahan (*loss*), yang divisualisasikan pada Gambar 4. Gambar 4 menampilkan grafik nilai *loss* selama proses pelatihan dan validasi model. Dari grafik tersebut terlihat bahwa nilai *loss* menurun secara bertahap seiring bertambahnya *epoch*, baik pada data pelatihan maupun data validasi. Penurunan *loss* ini menunjukkan bahwa kesalahan prediksi model semakin kecil dan proses optimasi berjalan dengan efektif. Kurva *loss* pelatihan dan validasi juga cenderung mengikuti pola yang serupa, tanpa adanya kenaikan *loss* validasi yang signifikan di akhir pelatihan. Kondisi ini mengindikasikan bahwa model mampu menyesuaikan bobot jaringan dengan baik tanpa mengalami *overfitting*.



Gambar 4. Grafik Loss

3.3 Confusion Matrix

Confusion Matrix digunakan untuk memahami jenis kesalahan yang terjadi saat model melakukan klasifikasi. Gambar 5 menunjukkan confusion matrix dari InceptionV3, DenseNet121, dan VGG16. Confusion matrix ini menunjukkan bagaimana prediksi model dibandingkan dengan label sebenarnya untuk setiap jenis penyakit.



Gambar 5. Confusion Matrix InceptionV3, DenseNet121, VGG16

Confusion Matrix pada model InceptionV3 menunjukkan bahwa sebagian besar prediksi berada pada diagonal utama, yang berarti model mampu mengklasifikasikan mayoritas kelas penyakit daun mangga dengan benar. Keunggulan InceptionV3 dalam mengekstraksi fitur multi-skala terlihat dari kemampuannya mengenali variasi ukuran bercak penyakit dan pola global daun. Namun demikian, masih ditemukan kesalahan klasifikasi pada beberapa kelas penyakit yang memiliki karakteristik visual yang hampir serupa, seperti perbedaan tekstur halus atau intensitas warna yang tipis. Hal ini mengindikasikan bahwa meskipun InceptionV3 cukup baik dalam menangkap fitur global, model ini terkadang kurang optimal dalam membedakan detail lokal yang sangat spesifik antar penyakit.

Confusion Matrix pada DenseNet121 memperlihatkan distribusi prediksi yang cukup baik di sepanjang diagonal utama, tetapi dengan tingkat kesalahan klasifikasi yang relatif lebih tinggi dibandingkan VGG16 dan InceptionV3. Kesalahan ini umumnya terjadi pada kelas penyakit yang memiliki pola bercak dan warna daun yang mirip. Mekanisme dense connectivity pada DenseNet121 memungkinkan penggunaan ulang fitur dari lapisan sebelumnya, namun pada dataset berukuran sedang seperti penelitian ini, kompleksitas arsitektur tersebut berpotensi menyebabkan kebingungan antar kelas yang memiliki ciri visual hampir identik. Hal ini menunjukkan bahwa DenseNet121 memerlukan data yang lebih besar atau penyesuaian parameter yang lebih spesifik agar mampu memaksimalkan keunggulan arsitekturnya.

Confusion Matrix pada model VGG16 menunjukkan pola diagonal yang paling dominan dibandingkan dua arsitektur lainnya, menandakan tingkat klasifikasi yang paling akurat dan konsisten. Jumlah kesalahan klasifikasi relatif kecil dan tersebar merata, menunjukkan bahwa VGG16 mampu membedakan fitur tekstur, tepi, dan pola bercak penyakit dengan baik. Arsitektur VGG16 yang sederhana namun dalam membuat model ini fokus pada ekstraksi fitur lokal yang detail, sehingga sangat efektif untuk membedakan penyakit daun mangga yang memiliki kemiripan visual. Hasil ini memperkuat alasan mengapa VGG16 menjadi model dengan performa terbaik dalam penelitian ini berdasarkan evaluasi confusion matrix.

3.4 Pembahasan

Evaluasi menunjukkan bahwa VGG16 mampu mengenali pola visual pada daun mangga dengan cara yang paling stabil dan konsisten. VGG16 merupakan model terbaik dari ketiga model yang dipakai dengan akurasi mencapai 93.25% dan nilai AUC-ROC sebesar 0.9831. Selain itu, angka presisi dan recall yang tinggi menunjukkan bahwa model tidak hanya benar dalam memprediksi kelas yang tepat, tetapi juga mampu mendeteksi sebagian besar contoh penyakit yang terdapat dalam dataset. Nilai F1-score sebesar 0.9326 menegaskan bahwa VGG16 memiliki keseimbangan yang sangat baik antara presisi dan recall. Metrik ini penting pada kasus klasifikasi multi-kelas seperti penyakit daun, karena setiap kelas memiliki gejala visual yang berbeda namun kadang saling menyerupai. Kemampuan VGG16 mempertahankan keseimbangan ini menunjukkan bahwa model dapat menangani variasi gejala penyakit dengan konsisten. Selain itu, VGG16 mempunyai kemampuan pemisahan (*separation ability*) yang sangat baik antar kelas. Artinya, model mampu membedakan secara jelas antara fitur penyakit satu dengan lainnya di ruang fitur (*feature space*), yang menjadi indikator kuat kualitas generalisasi. Hasil penelitian ini menunjukkan bahwa model CNN dengan arsitektur VGG16 memiliki nilai akurasi yang lebih baik dibandingkan dengan model naive bayes sebesar 81% pada penelitian terdahulu [28]. Secara praktis, keunggulan performa VGG16 ini membawa implikasi positif bagi sektor pertanian. Model ini cukup efisien untuk diintegrasikan ke dalam sistem cerdas atau aplikasi seluler (*mobile*) pada penelitian berikutnya. Dengan demikian, petani di lapangan dapat mendeteksi penyakit sejak dini secara mandiri hanya dengan mengambil foto daun mangga melalui gawai mereka, sehingga intervensi pengendalian hama dapat dilakukan lebih cepat dan akurat.

4. KESIMPULAN

Penelitian ini membandingkan tiga jenis jaringan saraf tiruan berbasis konvolusi CNN, yaitu VGG16, DenseNet121, dan InceptionV3, untuk mengklasifikasikan delapan jenis penyakit daun mangga. Penelitian ini menggunakan dataset Mango Leaf Disease dari Kaggle. Semua model dilatih dengan pendekatan transfer learning, proses pengolahan data yang sudah diatur secara standar, serta metode validasi K-Fold Cross-Validation agar hasil evaluasinya lebih akurat dan dapat diandalkan. Berdasarkan hasil pengujian, ketiga model menunjukkan kinerja yang baik dengan akurasi di atas 90%. Namun, VGG16 tampil sebagai model dengan performa terbaik, mencapai akurasi 93,25%, presisi 0,93, recall 0,93, F1-score 0,93, serta AUC-ROC 0,98. Nilai-nilai ini menunjukkan bahwa VGG16 memiliki kemampuan paling stabil dalam mengenali pola visual penyakit, menjaga keseimbangan presisi dan recall, serta mampu membedakan antar kelas penyakit secara lebih akurat dibandingkan dua model lainnya. Keunggulan VGG16 dipengaruhi oleh arsitekturnya yang sederhana namun efektif dalam mengekstraksi fitur tekstur dan detail kecil pada daun, struktur jaringan yang tidak terlalu kompleks sehingga cocok untuk dataset berukuran sedang, serta konvergensi pelatihan yang stabil. Dengan performa ini, VGG16 direkomendasikan sebagai model utama untuk implementasi awal sistem deteksi otomatis penyakit daun mangga. Sebagai implikasi praktis, keandalan model VGG16 yang diusulkan sangat berpotensi untuk diintegrasikan ke dalam perangkat lunak berbasis *computer vision*, sehingga dapat dimanfaatkan langsung oleh tenaga penyuluh dan petani sebagai alat diagnosis dini untuk mencegah kerugian akibat gagal panen pada penelitian berikutnya.

REFERENCES

- [1] E. J. Warschefsky, "Population genomic analysis of mango (*Mangifera indica*) suggests a complex history of domestication," no. 2019, pp. 2023–2037, 2023, doi: 10.1111/nph.15731.
- [2] K. Putri Ramadhani and M. Tarigan, "Implementasi Metode Convolutional Neural Network Untuk Klasifikasi Penyakit Daun Mangga Menggunakan Arsitektur Efficientnetv2-S Dan Resnet50," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 3, pp. 4135–4143, 2025, doi: 10.36040/jati.v9i3.13571.
- [3] A. K. Dofuor *et al.*, "Mango anthracnose disease: the current situation and direction for future research," *Front. Microbiol.*, vol. 14, no. August, pp. 1–18, 2023, doi: 10.3389/fmicb.2023.1168203.
- [4] rahayu deny danar dan alvi furwanti Alwie, A. B. Prasetyo, R. Andespa, P. N. Lhokseumawe, and K. Pengantar, "Tugas Akhir Tugas Akhir," *J. Ekon. Vol. 18, Nomor 1 Maret 201*, vol. 2, no. 1, pp. 41–49, 2024.
- [5] B. Paneru, B. Paneru, and K. B. Shah, "Analysis of Convolutional Neural Network-based Image Classifications: A Multi Featured Application for Rice Leaf Disease Prediction and Recommendations for Farmers," vol. 6, no. 3, pp. 174–187, 2024.
- [6] S. Dwi Anggraeni, E. Yulia Puspaningrum, and A. Lina Nurlaili, "Klasifikasi Penyakit Daun Mangga Menggunakan Anfis," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 3, pp. 5467–5473, 2025, doi: 10.36040/jati.v9i3.14311.
- [7] A. Rajbongshi, T. Khan, M. Rahman, A. Pramanik, and S. Tanvir, "Recognition of mango leaf disease using convolutional neural network models : a transfer learning approach," vol. 23, no. 3, pp. 1681–1688, 2021, doi: 10.11591/ijeecs.v23.i3.pp1681-1688.
- [8] Y. Agegnehu, A. Melese, B. Mulugeta, T. Nigussie, E. Ayenew, and T. Endeshaw, "Smart Agricultural Technology Classification of mango disease using ensemble convolutional neural network," *Smart Agric. Technol.*, vol. 8, no. May, p. 100476, 2024, doi: 10.1016/j.atech.2024.100476.
- [9] T. Ayu, V. Dwi, and A. E. Minarno, "Pendiagnosa Daun Mangga Dengan Model Convolutional Neural Network," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 6, no. 2, p. 230, 2021, doi: 10.24114/cess.v6i2.22857.
- [10] P. D. Rinanda, D. N. Aini, and T. A. Pertiwi, "Implementation of Convolutional Neural Network (CNN) for Image Classification of Leaf Disease In Mango Plants Using Deep Learning Approach," vol. 1, no. January, pp. 55–61, 2024.
- [11] A. R. Auni and E. Sugiharti, "Optimization of Mango Plant Leaf Disease Classification Using Concatenation Method of MobileNetV2 and DenseNet201 CNN Architectures," vol. 11, no. 4, pp. 1023–1034, 2024, doi: 10.15294/sji.v11i4.15169.
- [12] M. J. Setiawan, B. Nugroho, and A. P. Sari, "Klasifikasi Penyakit Daun Tanaman Menggunakan Algoritma CNN dan Random Forest," vol. 12, no. 1, pp. 1–7, 2023.
- [13] 6 · Xia Zhao1 · Limin Wang1 · Yufei Zhang2 · Xuming Han3 · Muhammet Deveci4, 5 and M. Parmar7, "A review of convolutional neural networks in computer vision," 2024.
- [14] B. U. Mahmud, A. Al Mamun, J. Hossen, G. Y. Hong, and B. Jahan, "Light-Weight Deep Learning Model for Accelerating the Classification of Mango-Leaf Disease," vol. 8, no. 1, pp. 28–42, 2024.
- [15] E. Rayed, J. R. Jim, J. Islam, and M. F. M. Senior, "MangoLeafXNet : An Explainable Deep Learning Model for Accurate Mango Leaf Disease Classification," *IEEE Access*, vol. PP, p. 1, 2025, doi: 10.1109/ACCESS.2025.3571450.
- [16] Z. Jiang, Y. Liu, Z. Shao, and K. Huang, "applied sciences An Improved VGG16 Model for Pneumonia Image Classification," 2021.
- [17] R. Ardiansyah, M. Ayu, D. Widyadara, and U. Mahdyah, "Deteksi Penyakit Daun Mangga Menggunakan Convolutional Neural Network Untuk Analisis Komperasi Arsitektur VGG16, Xception," *Inotek*, vol. 9, pp. 2549–7952, 2025.
- [18] Y. Zhang, C. Ning, and W. Yang, "An automatic cervical cell classification model based on improved DenseNet121," pp. 1–18, 2025.
- [19] A. Sharma and R. Parvathi, "Enhancing Cervical Cancer Classification : Through a Hybrid Deep Learning Approach Integrating," *IEEE Access*, vol. 13, no. January, pp. 9868–9878, 2025, doi: 10.1109/ACCESS.2025.3527677.
- [20] A. Nurdin, D. Satria, Y. Kartika, A. Rezha, and E. Najaf, "Klasifikasi Penyakit Daun Tomat Dengan Metode Convolutional Neural Network Menggunakan Arsitektur Inception-V3," no. 1, pp. 1–6, 2024.
- [21] S. N. & A. M. Ferdin Joe John Joseph, "Keras and TensorFlow: A Hands-On Experience," 2021.
- [22] M. Reyad and A. M. Sarhan, "A modified Adam algorithm for deep neural network optimization," *Neural Comput. Appl.*, vol. 35, no. 23, pp. 17095–17112, 2023, doi: 10.1007/s00521-023-08568-z.
- [23] J. Li, "Asymptotics of K-Fold Cross Validation," vol. 78, pp. 491–526, 2023.
- [24] Z. A. Sejuti and S. Islam, "A hybrid CNN – KNN approach for identi fi cation of COVID-19 with 5-fold cross validation,"

- Sensors Int.*, vol. 4, no. November 2022, p. 100229, 2023, doi: 10.1016/j.sintl.2023.100229.
- [25] M. S. Uzer, “Deep Learning-Based Classification Consisting of Pre-Trained Models and Proposed Model Using K-Fold Cross-Validation for Pistachio Species,” 2025.
 - [26] J. K. & S. Mukherjee, “Impact of Autotuned Fully Connected Layers on Performance of Self-supervised Models for Image Classification,” 2024.
 - [27] L. Anghel, Manon Dampfhoffer, Thomas Mesquida, Alexandre Valentian, “Backpropagation-Based Learning Techniques for Deep Spiking Neural Networks: A Survey,” 2023.
 - [28] L. Risqia and Z. Fatah, “Jamastika, volume 4 nomor 1 april 2025,” vol. 4, no. April, pp. 102–107, 2025.