

Analisis Komparasi Arsitektur Deep Learning untuk Klasifikasi Penyakit Daun Cabai

Bramudya Toguando Sitohang*, Wahyu Aji Eko Prabowo

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

Email: ¹*111202214782@mhs.dinus.ac.id, ²prabowo@dsn.dinus.ac.id

Email Penulis Korespondensi: 111202214782@mhs.dinus.ac.id

Submitted 07-01-2026; Accepted 24-02-2026; Published 28-02-2026

Abstrak

Produktivitas komoditas cabai (*Capsicum annum* L.) di Indonesia menghadapi tantangan signifikan akibat serangan penyakit daun, yang diperkirakan dapat memangkas hasil panen hingga 35%. Metode deteksi konvensional yang mengandalkan pengamatan visual seringkali kurang akurat akibat tingginya kemiripan gejala fisik antar penyakit. Studi ini difokuskan pada evaluasi komparatif terhadap performa tiga arsitektur Deep Learning terkemuka, yaitu VGG16, ResNet50, dan InceptionV3, dalam mengklasifikasikan 6 jenis penyakit daun cabai menggunakan dataset publik. Penelitian ini menerapkan strategi penggunaan citra resolusi tinggi (512 x 512 piksel) untuk memaksimalkan ekstraksi fitur tekstur penyakit. Metode penelitian menggunakan pendekatan Transfer Learning dengan skema hyperparameter tuning yang terstandarisasi. Hasil eksperimen menunjukkan bahwa penggunaan resolusi tinggi memberikan dampak signifikan terhadap akurasi model. Arsitektur VGG16 mencatat performa terbaik dengan akurasi pengujian mencapai 99,83% dan F1-Score 1,00, mengungguli ResNet50 (99,75%) dan InceptionV3 (84,00%). Analisis Confusion Matrix menunjukkan bahwa VGG16 memiliki stabilitas superior dalam membedakan kelas penyakit dengan kemiripan visual tinggi, seperti Bacterial Spot dan Cercospora. Penelitian ini menyimpulkan bahwa arsitektur yang mempertahankan informasi spasial secara mendalam seperti VGG16 lebih efektif untuk diagnosis penyakit tanaman berbasis citra resolusi tinggi dibandingkan arsitektur yang lebih kompleks namun melakukan kompresi fitur yang agresif.

Kata Kunci: Deep Learning; Klasifikasi Penyakit Cabai; VGG16; ResNet50; InceptionV3

Abstract

The productivity of chili (*Capsicum annum* L.) in Indonesia faces significant challenges due to leaf diseases, which are estimated to reduce harvest yields by up to 35%. Conventional detection methods relying on visual observation often lack accuracy due to the high visual similarity between disease symptoms. This study focuses on a comparative evaluation of three leading Deep Learning architectures VGG16, ResNet50, and InceptionV3 in classifying six types of chili leaf diseases using a public dataset. The research implements a high-resolution image strategy (512 x 512 pixels) to maximize the extraction of disease texture features. The methodology employs a Transfer Learning approach with a standardized hyperparameter tuning scheme. Experimental results indicate that the use of high-resolution images significantly impacts model accuracy. The VGG16 architecture achieved the best performance with a testing accuracy of 99.83% and an F1-Score of 1.00, outperforming ResNet50 (99.75%) and InceptionV3 (84.00%). Confusion Matrix analysis demonstrates that VGG16 possesses superior stability in distinguishing disease classes with high visual similarity, such as Bacterial Spot and Cercospora. The study concludes that architectures preserving deep spatial information, such as VGG16, are more effective for high-resolution image-based plant disease diagnosis compared to more complex architectures that perform aggressive feature compression.

Keywords: Deep Learning; Chili Disease Classification; VGG16; ResNet50; InceptionV3

1. PENDAHULUAN

Sektor agrikultur memegang posisi sentral sebagai pilar utama dalam menjamin ketersediaan pangan serta keseimbangan ekonomi di tingkat global. Urgensi peran ini semakin nyata mengingat estimasi populasi dunia yang diprediksi akan menembus angka 9,7 miliar jiwa pada tahun 2050. Sebagai konsekuensinya, sektor pertanian kini menghadapi tekanan besar untuk meningkatkan produktivitas tanaman setidaknya sebesar 70% guna memenuhi kebutuhan pangan yang terus melonjak [1]. Namun, upaya pemenuhan target produktivitas tersebut menghadapi kendala serius berupa kerentanan tanaman terhadap gangguan eksternal. Hambatan biotik utama yang secara konsisten mengganggu ketahanan pangan adalah invasi hama serta infeksi patogen pada tanaman. Wabah penyakit pada vegetasi seringkali menyebar layaknya epidemi biologis yang menyebar dengan cepat dan seringkali tidak terdeteksi hingga mencapai tahap lanjut. Laporan dari FAO menunjukkan bahwa hama dan penyakit tanaman bertanggung jawab atas kehilangan hasil produksi global antara 20% hingga 40% setiap tahunnya [2]. Kerugian ini tidak hanya berdampak pada kuantitas pangan yang tersedia, tetapi juga pada kualitas nutrisi dan keamanan konsumsi [3].

Di negara-negara beriklim tropis dan berproduksi menengah seperti Indonesia, risiko penyebaran penyakit tanaman semakin diperparah oleh faktor perubahan iklim. Transformasi pola cuaca yang tidak menentu, dikombinasikan dengan kelembaban udara yang tinggi sepanjang tahun, menciptakan lingkungan mikroklimat yang sangat ideal bagi perkembangbiakan patogen seperti jamur dan bakteri [4]. Kondisi ini meningkatkan kerentanan tanaman hortikultura terhadap wabah penyakit [5]. Eskalasi serangan penyakit ini tidak hanya berhenti pada kerusakan biologis tanaman, namun secara langsung memicu kerugian hasil panen yang membawa implikasi ekonomi sangat serius. Kerugian hasil panen akibat penyakit tanaman ini memiliki implikasi ekonomi yang sangat serius. Biaya ekonomi akibat kerugian pangan global diproyeksikan dapat mencapai triliunan dolar per tahun, yang berdampak langsung pada kesejahteraan petani kecil

[6]. Bagi negara agraris seperti Indonesia, penurunan produktivitas komoditas strategis dapat mengguncang stabilitas harga pasar domestik [7].

Di Indonesia, cabai (*Capsicum annuum L.*) merupakan salah satu komoditas hortikultura strategis yang memiliki nilai ekonomi tinggi dan permintaan pasar yang stabil [8]. Namun, budidaya cabai di Indonesia menghadapi tantangan berat berupa kerentanan yang tinggi terhadap berbagai jenis penyakit daun. Infeksi pada daun sangat krusial karena daun adalah organ utama fotosintesis; kerusakan pada daun secara langsung akan menghambat pertumbuhan tanaman dan menurunkan kualitas serta kuantitas buah yang dihasilkan [9]. Secara spesifik di wilayah Jawa Tengah, yang merupakan salah satu sentra produksi cabai nasional, petani rumah tangga seringkali menghadapi kerugian panen yang signifikan, berkisar antara 25% hingga 35%, akibat serangan penyakit daun [10]. Penyakit seperti *Bacterial Spot* (Bercak Bakteri), *Cercospora*, dan serangan virus seperti *Curl Virus* (Daun Keriting) adalah masalah endemik yang persisten [11]. Masalah utamanya adalah keterbatasan petani dalam melakukan deteksi dini. Metode inspeksi visual manual yang selama ini digunakan sangat subjektif, memakan waktu lama, dan seringkali tidak akurat, sehingga penanganan seringkali terlambat dilakukan [12].

Untuk mengatasi keterbatasan metode manual tersebut, teknologi *Computer Vision* yang didukung oleh kecerdasan buatan (*Artificial Intelligence*) telah muncul sebagai solusi revolusioner [13]. Secara khusus, algoritma *Deep Learning* berbasis *Convolutional Neural Network* (CNN) telah terbukti memiliki kemampuan luar biasa dalam mengenali pola visual yang kompleks pada citra digital [14]. Dalam dekade terakhir, teknologi ini telah banyak diadopsi dalam *precision agriculture* untuk mendeteksi penyakit tanaman dengan tingkat akurasi yang jauh melampaui kemampuan pengamatan mata manusia [15]. Dalam perkembangan riset *deep learning*, terdapat beberapa arsitektur model yang umum digunakan karena performanya yang tinggi dalam klasifikasi citra berskala besar, yaitu VGGNet, ResNet, dan InceptionV3. Arsitektur seperti VGGNet, yang diperkenalkan oleh Simonyan & Zisserman [16], dikenal dengan kedalamannya yang ekstensif menggunakan kernel 3x3. ResNet (Residual Network), yang dikembangkan oleh He dkk. [17], merevolusi pelatihan model yang sangat dalam dengan *residual connections* untuk mengatasi *vanishing gradient*. Sementara itu, InceptionV3, yang dirancang oleh Szegedy dkk. [18], dikenal dengan efisiensi ekstraksi fiturnya melalui modul *inception* paralel. Ketiga arsitektur ini sering digunakan sebagai basis dalam metode *Transfer Learning* untuk mencapai akurasi tinggi. Untuk melakukan perbandingan performa model-model tersebut, penelitian ini akan menggunakan dataset publik *Chili Plant Leaf Disease* dari Bangladesh (Mendeley Data) [19]. Penggunaan dataset regional ini memiliki justifikasi metodologis yang kuat untuk konteks Indonesia karena didukung oleh tiga faktor utama. Pertama, secara iklim, kedua negara berada di wilayah tropis monsoonal dengan suhu dan kelembaban tinggi yang menciptakan mikroklimat serupa bagi pertumbuhan patogen [20], [21], [22]. Kedua, secara genetik, varietas cabai rawit di kedua negara berasal dari garis keturunan Asia Tenggara [23] dengan koefisien similaritas fenotipik yang tinggi [24], [25]. Ketiga, dan yang paling krusial, adalah keseragaman patogen; penyakit utama dalam dataset ini (seperti *Bacterial Spot* dan *Curl Virus*) telah teridentifikasi sebagai ancaman endemik yang juga menyerang pertanaman cabai di Jawa Tengah dengan karakteristik gejala yang identik [11], [26].

Meskipun studi komparasi (*benchmarking*) terhadap ketiga arsitektur ini (VGG, ResNet, Inception) telah terbukti efektif untuk komoditas lain seperti kembang kol [27] dan padi [28], serta penggunaan VGG secara hibrida pada cabai telah dieksplorasi [29], masih terdapat *research gap* berupa kurangnya studi yang melakukan evaluasi kinerja komparatif dari ketiga model tersebut secara langsung (*head-to-head*) pada dataset penyakit cabai spesifik regional Asia Selatan [19]. Oleh karena itu, penelitian ini bertujuan untuk melakukan komparasi menyeluruh antara VGGNet, ResNet, dan InceptionV3 guna menentukan arsitektur mana yang memberikan akurasi tertinggi dalam mengklasifikasikan penyakit daun cabai. Hasil penelitian ini diharapkan menjadi rujukan ilmiah utama dalam pemilihan model terbaik untuk kasus penyakit tanaman hortikultura di Indonesia [30].

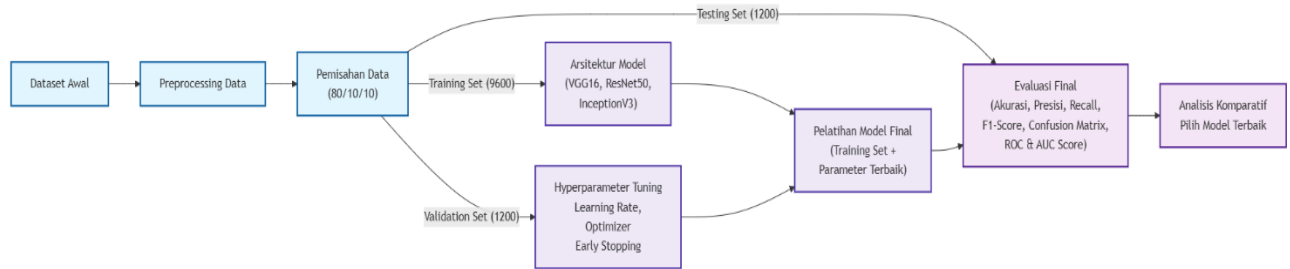
2. METODOLOGI PENELITIAN

Metodologi penelitian ini menerapkan pendekatan kuantitatif dengan desain eksperimental untuk melakukan analisis komparasi kinerja (*benchmarking*) terhadap tiga arsitektur *Deep Learning*. Fokus utama penelitian adalah mengevaluasi keandalan model dalam mengklasifikasikan citra penyakit daun cabai secara akurat pada resolusi citra tinggi.

2.1 Tahapan Penelitian

Penelitian ini dilaksanakan melalui serangkaian tahapan sistematis yang dirancang untuk memastikan validitas dan keterulangan hasil eksperimen. Sebagaimana diilustrasikan secara visual pada Gambar 1, alur penelitian diawali dengan proses akuisisi data mentah dari repositori publik, yang kemudian dilanjutkan dengan tahap pra-pemrosesan (*preprocessing*) untuk menyeragamkan format citra. Tahap selanjutnya melibatkan pembagian dataset ke dalam subset yang berbeda untuk keperluan pelatihan dan pengujian. Inti dari penelitian ini terletak pada tahap konfigurasi dan pelatihan model, di mana tiga arsitektur berbeda disiapkan menggunakan teknik transfer learning dan dilatih dengan skema hyperparameter yang terkontrol. Setelah model terbentuk, dilakukan evaluasi menyeluruh menggunakan berbagai

metrik kinerja statistik untuk mengukur akurasi prediksi. Tahapan ini diakhiri dengan analisis komparatif untuk memilih model terbaik berdasarkan keseimbangan antara akurasi, stabilitas, dan ukuran model.



Gambar 1. Tahapan Penelitian

2.2 Data Penelitian

Sumber data utama yang digunakan dalam penelitian ini adalah dataset publik "*Chili Plant Leaf Disease and Growth Stage Dataset*" yang diperoleh dari Mendeley Data [19]. Pemilihan dataset ini didasarkan pada relevansi agronomisnya dengan kondisi tanaman cabai di wilayah tropis, termasuk Indonesia, sebagaimana telah dipaparkan pada latar belakang. Dataset ini terdiri dari total 12.000 citra daun cabai yang telah melalui proses augmentasi awal untuk menyeimbangkan distribusi kelasnya. Tabel 1 menyajikan dataset yang mencakup enam kondisi daun yang merepresentasikan penyakit umum dan kondisi sehat, yaitu: *Bacterial Spot* (Bercak Bakteri), *Curl Virus* (Virus Daun Keriting), *Cercospora Leaf Spot* (Bercak Daun *Cercospora*), *Nutrition Deficiency* (Kekurangan Nutrisi), *White Spot* (Bercak Putih), dan *Healthy Leaf* (Daun Sehat). Keragaman kelas ini memungkinkan model untuk belajar membedakan fitur visual penyakit dengan tingkat kompleksitas yang bervariasi.

Tabel 1. Sampel *Dataset*

No	Nama Kelas (Label)	Deskripsi
1	Bacterial Spot	Daun dengan bercak bakteri.
2	Curl Virus	Daun yang terinfeksi virus keriting.
3	Cercospora Leaf Spot	Daun dengan bercak jamur <i>Cercospora</i> .
4	Nutrition Deficiency	Daun yang menunjukkan gejala kekurangan nutrisi.
5	White Spot	Daun dengan bercak putih.
6	Healthy Leaf	Daun cabai yang sehat (kelas kontrol).

2.3 Pembagian Data

Untuk menjamin objektivitas evaluasi dan mencegah terjadinya kebocoran data (*data leakage*), keseluruhan dataset yang berjumlah 12.000 citra didistribusikan ke dalam tiga subset utama dengan rasio 80:10:10. Bagian terbesar, yaitu 80% data atau setara dengan 9.600 citra, dialokasikan sebagai data latih (*training set*) yang berfungsi sebagai materi utama bagi model untuk mempelajari pola fitur penyakit. Selanjutnya, 10% data atau 1.200 citra digunakan sebagai data validasi (*validation set*). Detail distribusi data disajikan pada Tabel 2. Subset ini berperan krusial selama proses pelatihan untuk memantau kinerja model secara berkala dan menjadi acuan dalam penyetelan hyperparameter guna mencegah *overfitting*. Sisa 10% data terakhir, yang juga berjumlah 1.200 citra, disimpan sebagai data uji (*testing set*). Data uji ini merupakan data yang benar-benar baru dan tidak pernah dilihat oleh model selama proses pelatihan, sehingga hasil evaluasi pada subset ini merepresentasikan kemampuan generalisasi model yang sebenarnya di dunia nyata.

Tabel 2. Pembagian *Dataset*

Jenis Data	Persentase (%)	Jumlah Gambar
Data Latih	80	9600
Data Validasi	10	1200
Data Uji	10	1200
Total	100	12000

2.4 Hyperparameter Tuning dan Konfigurasi

Proses optimalisasi kinerja model dilakukan melalui penyetelan hyperparameter (*hyperparameter tuning*) yang ketat dan terstandarisasi untuk ketiga model guna memastikan perbandingan yang adil (*apples-to-apples*). Pada tahap pra-pemrosesan, seluruh citra input diubah ukurannya (*resize*) menjadi dimensi resolusi tinggi 512 x 512 piksel. Pemilihan resolusi 512 x 512 ini bertujuan untuk mempertahankan detail visual halus pada bercak penyakit yang mungkin hilang pada resolusi rendah, sehingga diharapkan dapat meningkatkan akurasi ekstraksi fitur. Selain itu, nilai piksel dinormalisasi dengan membaginya dengan 255 (skala 0-1) untuk mempercepat konvergensi gradien. Selama proses pelatihan, algoritma optimasi yang digunakan adalah Adam (*Adaptive Moment Estimation*), yang dipilih karena efisiensinya dalam menangani data besar dan kemampuannya menyesuaikan laju pembelajaran secara adaptif. Laju

pembelajaran (*learning rate*) awal ditetapkan pada angka yang konservatif, yaitu 0,0001. Penggunaan *learning rate* yang kecil ini sangat krusial dalam metode *transfer learning* untuk mencegah kerusakan pada bobot pra-terlatih (*pretrained weights*) saat model menyesuaikan diri dengan dataset baru. Fungsi kerugian (*loss function*) yang diterapkan adalah *categorical crossentropy*, yang merupakan standar untuk klasifikasi multi-kelas. Pelatihan direncanakan berjalan hingga 50 epoch dengan ukuran batch 32. Untuk menjaga efisiensi dan mencegah model menghafal data latih (*overfitting*), mekanisme *early stopping* diterapkan. Mekanisme ini secara otomatis akan menghentikan pelatihan jika tidak terjadi penurunan nilai loss pada data validasi selama 10 epoch berturut-turut, dan akan mengembalikan bobot model ke kondisi terbaiknya.

2.5 Pelatihan Model

Strategi utama dalam pengembangan model pada penelitian ini adalah penerapan metode *transfer learning*. Pendekatan ini memanfaatkan pengetahuan yang telah dipelajari oleh model dari dataset berskala masif (ImageNet) untuk menyelesaikan tugas baru dengan data yang lebih terbatas. Tiga arsitektur Convolutional Neural Network (CNN) dipilih sebagai basis eksperimen, yaitu VGG16, ResNet50, dan InceptionV3. Arsitektur pertama yang dievaluasi adalah VGG16 [16]. Model ini dikenal dengan struktur arsitekturnya yang sederhana namun sangat dalam (*very deep*), yang tersusun dari tumpukan lapisan konvolusi dengan filter berukuran kecil 3 x 3. Karakteristik ini memungkinkan VGG16 untuk menangkap detail fitur tekstur penyakit secara mendalam. Dalam implementasinya, bagian pengekstraksi fitur (*base model*) VGG16 dibekukan bobotnya, dan lapisan klasifikasi (*head*) diganti dengan lapisan *Global Average Pooling* dan lapisan *Dense* baru yang sesuai dengan 6 kelas penyakit cabai. Model kedua adalah ResNet50 [17], yang memperkenalkan konsep inovatif berupa koneksi residual (*residual connections*) atau *skip connections*. Arsitektur ini dirancang untuk mengatasi masalah hilangnya gradien (*vanishing gradient*) yang sering terjadi pada jaringan saraf yang sangat dalam. Dengan kedalaman 50 *layer*, ResNet50 diharapkan mampu mempelajari fitur yang lebih kompleks dan abstrak dari citra daun cabai dibandingkan model yang lebih dangkal, namun tetap efisien dalam proses pelatihan. Model ketiga adalah InceptionV3 [18], yang dikembangkan dengan fokus pada efisiensi komputasi melalui penggunaan modul Inception. Modul ini memungkinkan jaringan untuk melakukan operasi konvolusi dengan berbagai ukuran filter secara paralel pada lapisan yang sama, sehingga model dapat menangkap fitur penyakit pada berbagai skala spasial sekaligus (baik bercak kecil maupun pola besar). Seperti kedua model lainnya, InceptionV3 juga diadaptasi menggunakan bobot ImageNet dan dimodifikasi lapisan akhirnya untuk tugas klasifikasi ini.

2.6 Evaluasi Model

Tahap evaluasi bertujuan untuk mengukur kinerja generalisasi model pada data uji yang belum pernah dilihat sebelumnya. Dasar utama evaluasi adalah Confusion Matrix (matriks kebingungan), sebuah tabel kontingensi yang memetakan prediksi model terhadap label sebenarnya untuk setiap kelas. Dari matriks ini, empat metrik statistik utama dihitung untuk memberikan gambaran performa yang komprehensif. Akurasi dihitung untuk melihat persentase total prediksi yang benar. Presisi (*Precision*) digunakan untuk mengukur ketepatan prediksi positif model, sementara Sensitivitas (*Recall*) mengukur kemampuan model dalam mendeteksi seluruh kejadian penyakit yang ada. F1-Score kemudian dihitung sebagai rata-rata harmonik dari Presisi dan Recall untuk memberikan penilaian tunggal yang menyeimbangkan kedua aspek tersebut. Selain metrik berbasis tabel kontingensi, penelitian ini juga menerapkan analisis kurva ROC (*Receiver Operating Characteristic*) dan nilai AUC (*Area Under Curve*) untuk mengukur kemampuan diskriminatif model secara lebih mendalam. Kurva ROC memvisualisasikan pertukaran (*trade-off*) antara tingkat positif benar dan tingkat positif palsu pada berbagai ambang batas keputusan. Mengingat kasus ini adalah klasifikasi multi-kelas, kurva ROC dan nilai AUC dihitung menggunakan pendekatan *One-vs-Rest* (OvR), di mana performa setiap kelas penyakit dianalisis secara terpisah dengan menganggap kelas lainnya sebagai negatif. Nilai AUC yang mendekati 1,00 mengindikasikan kemampuan model yang sempurna dalam membedakan antar kelas penyakit dengan tingkat kepercayaan yang tinggi.

Rumus matematis berikut akan digunakan untuk menghitung metrik performa dari model:

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

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

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

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

Keterangan:

True Positive (TP): Jumlah data kelas penyakit target yang berhasil diprediksi dengan benar.

True Negative (TN): Jumlah data kelas selain target yang berhasil diprediksi dengan benar sebagai bukan kelas target.

False Positive (FP): Jumlah data negatif yang salah diprediksi sebagai positif oleh model.

False Negative (FN): Jumlah data positif yang salah diprediksi sebagai negatif oleh model.

(Catatan: Untuk Presisi, Recall, dan F1-Score multikelas, akan digunakan "*Weighted Average*" untuk menghitung skor gabungan).

2.7 Pemilihan Model Terbaik

Tahap akhir dari metodologi ini adalah pemilihan arsitektur model terbaik yang akan direkomendasikan. Kriteria pemilihan tidak hanya didasarkan pada skor Akurasi tertinggi, tetapi juga mempertimbangkan nilai F1-Score dan AUC untuk memastikan model bekerja baik pada setiap kelas. Selain itu, Ukuran Model (dalam Megabyte) juga menjadi pertimbangan sekunder sebagai indikator efisiensi penyimpanan. Model yang menunjukkan kombinasi terbaik antara performa klasifikasi yang tinggi, stabilitas prediksi (AUC tinggi), dan efisiensi penyimpanan akan ditetapkan sebagai model terbaik untuk klasifikasi penyakit daun cabai.

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan analisis komprehensif terhadap kinerja ketiga arsitektur *Deep Learning* (VGG16, ResNet50, dan InceptionV3) yang telah dilatih dan diuji menggunakan dataset penyakit daun cabai beresolusi tinggi (512 x 512 piksel). Evaluasi dilakukan secara berjenjang, mulai dari metrik statistik global, analisis stabilitas pelatihan, hingga diagnosis kesalahan spesifik melalui *Confusion Matrix*.

3.1 Hasil Evaluasi Model

Tabel 3 menjelaskan tentang rangkuman metrik evaluasi. Berdasarkan pengujian terhadap 1.200 citra pada data uji (*Testing Set*) yang tidak pernah dilihat model sebelumnya, hasil eksperimen menunjukkan variasi performa yang signifikan antar arsitektur.

Tabel 3. Hasil Evaluasi Performansi Model Pada Data Uji

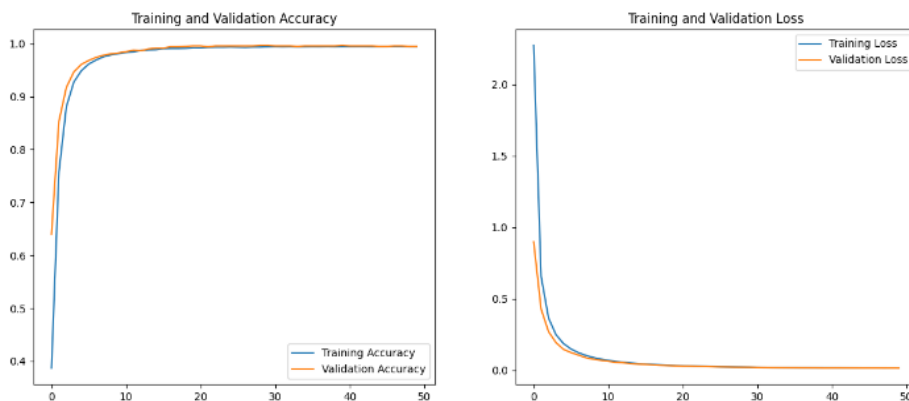
Arsitektur Model	Ukuran Input	Akurasi	Presisi (Avg)	Recall (Avg)	F1-Score (Avg)	AUC (Avg)
VGG16	512 x 512	99,83%	1,00	1,00	1,00	1,00
ResNet50	512 x 512	99,75%	1,00	1,00	1,00	1,00
InceptionV3	512 x 512	84,00%	0,84	0,84	0,84	0,97

Secara keseluruhan, VGG16 mendominasi sebagai arsitektur dengan performa terbaik, mencatat akurasi tertinggi sebesar 99,83%. Keunggulan ini diperkuat oleh pencapaian nilai Presisi dan *Recall* rata-rata yang sempurna yaitu 1,00, yang mengindikasikan bahwa model ini memiliki kemampuan absolut dalam meminimalkan kesalahan prediksi positif palsu (*False Positive*) maupun negatif palsu (*False Negative*). Di posisi kedua, ResNet50 menempel ketat dengan akurasi 99,75%. Meskipun memiliki selisih akurasi tipis, ResNet50 juga mencatatkan rata-rata Presisi dan *Recall* sebesar 1,00 (pada pembulatan dua desimal), menegaskan keandalannya yang setara dengan VGG16 dalam mendeteksi relevansi kelas. Sebaliknya, InceptionV3 tertinggal cukup jauh dengan akurasi 84,00%. Penurunan performa ini tercermin jelas pada metrik Presisi dan *Recall* yang hanya mencapai 0,84, menandakan bahwa model ini cukup sering mengalami kegagalan dalam menetapkan batas keputusan (*decision boundary*) yang tepat antar kelas penyakit, meskipun nilai AUC-nya (0,97) masih tergolong cukup baik.

3.2 Analisis Stabilitas Pelatihan

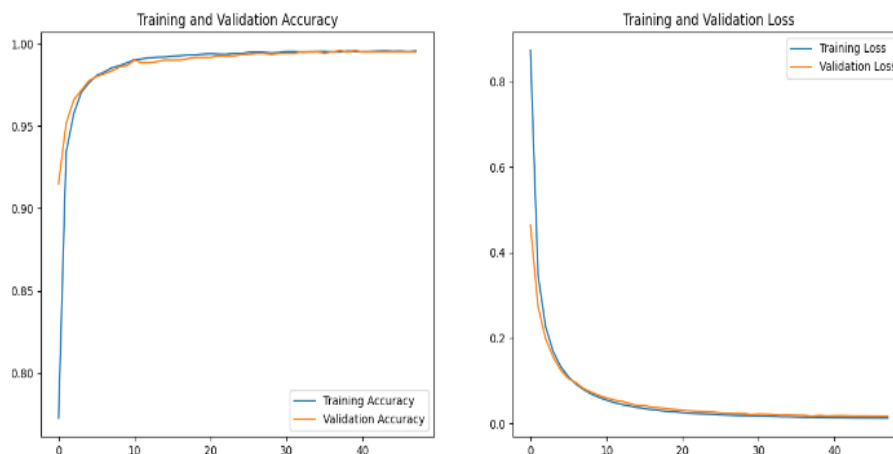
Analisis grafik riwayat pelatihan (*history plot*) dilakukan untuk mendiagnosis perilaku model selama proses pembelajaran (*learning curve*). Kurva akurasi dan loss untuk ketiga model dapat dilihat pada Gambar 2-4. Seperti ditunjukkan pada Gambar 2 dan 3, model VGG16 dan ResNet50 menunjukkan pola konvergensi yang sangat ideal atau *Good Fit*. Garis akurasi data latih dan data validasi bergerak saling berimpitan (menempel rapat) dari awal hingga akhir epoch. Tidak terlihat adanya kesenjangan (*gap*) yang melebar, yang membuktikan bahwa kedua arsitektur tersebut mampu menggeneralisasi fitur penyakit dengan sangat baik tanpa menghafal data latih (*overfitting*). Sebaliknya, grafik InceptionV3 pada Gambar 2 memperlihatkan gejala ketidakstabilan. Terdapat jarak (*gap*) yang terlihat jelas antara performa pada data latih yang terus membaik dan data validasi yang tertinggal. Kurva validasi juga menunjukkan fluktuasi, menandakan kesulitan model dalam menemukan pola yang konsisten untuk data baru.

Visualisasi kurva pembelajaran (*learning curve*) untuk arsitektur VGG16 selama 30 *epoch* ditampilkan secara rinci pada Gambar 2. Grafik ini memetakan pergerakan metrik akurasi dan *loss* pada data latih maupun validasi. Terlihat konvergensi yang sangat stabil dan cepat, di mana garis akurasi validasi berimpit rapat dengan akurasi pelatihan, mengindikasikan bahwa model VGG16 mampu mempelajari fitur penyakit tanpa mengalami gejala *overfitting* yang signifikan.



Gambar 2. Grafik Akurasi/Loss VGG16

Selanjutnya, dinamika kinerja pelatihan untuk arsitektur ResNet50 dapat diamati pada Gambar 3. Grafik tersebut memperlihatkan tren peningkatan akurasi yang konsisten mendekati angka 1.0. Meskipun terdapat sedikit fluktuasi minor pada kurva *validation loss* dibandingkan VGG16, secara umum model ini tetap menunjukkan stabilitas yang baik dalam meminimalkan *error* seiring bertambahnya iterasi pelatihan.



Gambar 3. Grafik Akurasi/Loss ResNet50

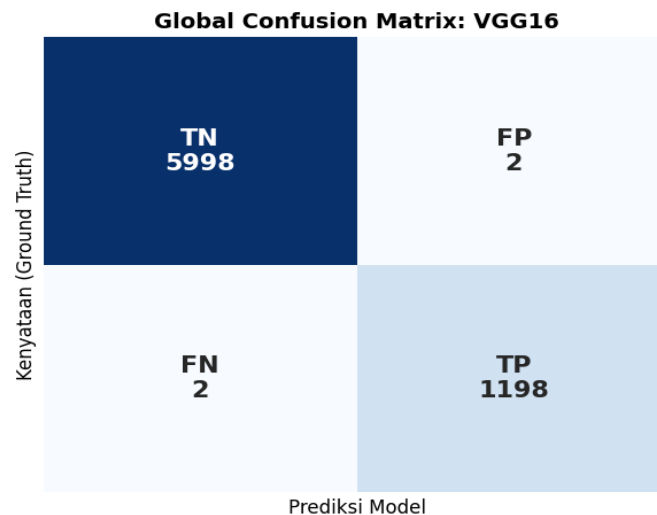
Berbeda dengan dua arsitektur sebelumnya, karakteristik pelatihan model InceptionV3 yang disajikan pada Gambar 4 menunjukkan pola yang kurang optimal. Terlihat adanya kesenjangan (*gap*) yang cukup lebar antara akurasi pelatihan dan validasi, serta fluktuasi tajam pada *validation loss*. Hal ini mengindikasikan bahwa arsitektur InceptionV3 mengalami kesulitan dalam menggeneralisasi fitur penyakit pada data validasi, yang berdampak pada performa keseluruhan yang lebih rendah.



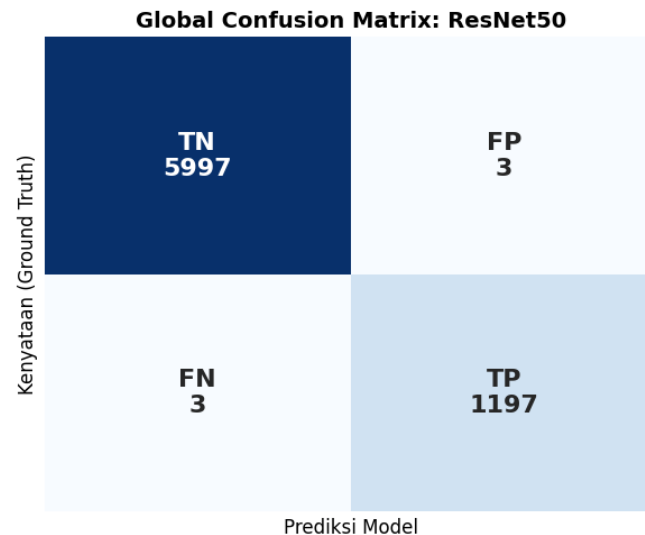
Gambar 4. Grafik Akurasi/Loss Inception

3.3 Visualisasi *Confusion Matrix*

Untuk membedah performa generalisasi model secara global, dilakukan analisis menggunakan *Global Confusion Matrix* yang memetakan total prediksi benar (*True Positive/Negative*) dan salah (*False Positive/Negative*) dari total 1.200 data uji, sebagaimana ditampilkan pada Gambar 5-7. Berdasarkan visualisasi pada Gambar 5, arsitektur VGG16 menunjukkan performa yang nyaris sempurna dengan berhasil memprediksi 1.198 citra secara tepat sebagai *True Positive* (TP). Tingkat kesalahannya sangat minim, di mana hanya terdapat 2 citra yang salah diklasifikasikan, menjadikannya model dengan konsistensi global tertinggi. Pola stabilitas yang sangat mirip juga ditunjukkan oleh ResNet50 pada Gambar 6, yang menempel ketat sebagai runner-up dengan mencatatkan 1.197 prediksi benar dan hanya 3 kesalahan prediksi, mengonfirmasi kemampuan kedua model ini dalam menangkap fitur penyakit dengan presisi tinggi. Sebaliknya, perbedaan performa yang signifikan terlihat pada arsitektur InceptionV3 sebagaimana ditunjukkan pada Gambar 7.

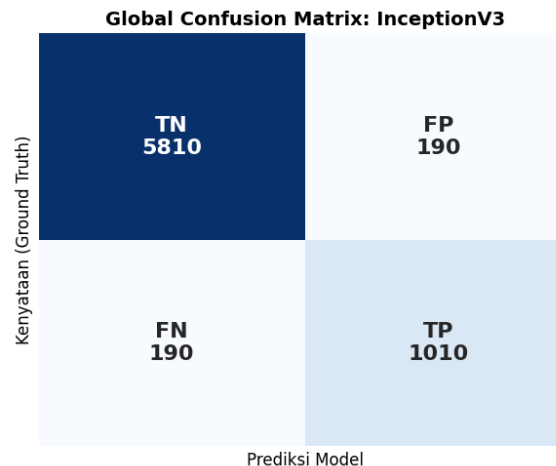


Gambar 5. Confusion Matrix VGG16



Gambar 6. Confusion Matrix ResNet50

Berbeda dengan dua model sebelumnya, InceptionV3 memperlihatkan jumlah kesalahan yang substansial dengan lonjakan angka kesalahan (*False Positive/Negative*) mencapai 190 citra, sementara jumlah prediksi benar (*True Positive*) hanya mencapai 1.010 citra. Tingginya angka kesalahan ini mengindikasikan bahwa arsitektur InceptionV3 memiliki kesulitan yang lebih besar dalam melakukan generalisasi fitur pada dataset penyakit daun cabai resolusi tinggi ini dibandingkan arsitektur VGG16 dan ResNet50 yang jauh lebih stabil. Temuan penelitian ini menegaskan bahwa untuk kasus klasifikasi penyakit daun cabai dengan resolusi citra tinggi 512 x 512, arsitektur yang lebih "sederhana" namun dalam seperti VGG16 justru mengungguli arsitektur yang lebih kompleks seperti InceptionV3. Keunggulan VGG16 dan ResNet50 diduga kuat karena penggunaan filter konvolusi berukuran kecil 3 x 3 yang konsisten, yang sangat efektif dalam mengekstraksi detail tekstur halus pada bercak penyakit. Sebaliknya, InceptionV3 tampak kesulitan mempertahankan koherensi fitur visual pada dataset ini, terutama dalam membedakan gradasi warna halus pada kasus kekurangan nutrisi. Berdasarkan akurasi tertinggi (99,83%), stabilitas pelatihan, dan ukuran model yang efisien, VGG16 direkomendasikan sebagai model terbaik.



Gambar 7. Confusion Matrix Inception

4. KESIMPULAN

Penelitian ini telah berhasil melakukan komparasi komprehensif terhadap tiga arsitektur *Deep Learning*, yaitu VGG16, ResNet50, dan InceptionV3, untuk klasifikasi penyakit daun cabai menggunakan citra resolusi tinggi (512 x 512 piksel). Berdasarkan hasil eksperimen dan evaluasi, dapat ditarik beberapa kesimpulan utama. Pertama, arsitektur VGG16 terbukti menjadi model paling unggul dengan akurasi pengujian tertinggi mencapai 99,83%, Recall dan F1-Score dengan nilai sempurna 1,00. Hal ini menunjukkan bahwa arsitektur yang mempertahankan informasi spasial secara mendalam melalui filter konvolusi kecil (3 x 3) sangat efektif dalam menangkap detail tekstur penyakit pada citra resolusi tinggi. Kedua, ResNet50 menunjukkan performa yang sangat kompetitif sebagai runner-up dengan akurasi 99,75%, menegaskan bahwa penggunaan residual connections mampu menjaga stabilitas pelatihan pada jaringan yang dalam. Ketiga, arsitektur InceptionV3 menunjukkan performa yang paling rendah dengan akurasi 84,00% dan tingkat kesalahan klasifikasi yang signifikan pada kelas dengan kemiripan visual tinggi. Hal ini mengindikasikan bahwa arsitektur yang terlalu kompleks dengan faktorisasi konvolusi agresif mungkin kurang optimal untuk tugas klasifikasi penyakit tanaman spesifik ini dibandingkan pendekatan VGGNet yang lebih *straightforward*. Sebagai saran untuk pengembangan penelitian selanjutnya, disarankan untuk: (1) Menguji ketahanan model VGG16 ini pada kondisi lingkungan nyata dengan pencahayaan yang bervariasi melalui pengembangan aplikasi mobile; (2) Memperluas dataset dengan varietas cabai lokal Indonesia untuk meningkatkan relevansi geografis; dan (3) Mengeksplorasi arsitektur yang lebih ringan namun efisien seperti EfficientNet atau MobileNetV3 jika tujuan implementasi bergeser ke perangkat dengan sumber daya terbatas (*edge devices*).

REFERENCES

- [1] FAO, "Global agriculture towards 2050," *FAO*, 2012, [Online]. Available: https://www.fao.org/fileadmin/templates/wsfs/docs/Issues_papers/HLEF2050_Global_Agriculture.pdf
- [2] FAO, "Cooperation at all levels and funding 'critical' for plant health and food security," Rome, Italy, 2025. [Online]. Available: <https://www.fao.org/newsroom/detail/cooperation-at-all-levels-and-funding--critical--for-plant-health-and-food-security--fao-says/en>
- [3] J. Li, "Compound impact of COVID-19, economy and climate on the spatial distribution of global agriculture and food security," *Sci. Total Environ.*, vol. 876, p. 162796, 2023, [Online]. Available: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10039698/>
- [4] R. and S. Asia, "EL NIÑO IMPACT ON AGRICULTURE IN SOUTHEAST ASIA," 2023. [Online]. Available: <https://rassea.org/2023/10/03/el-nino-impact-on-agriculture-in-southeast-asia/>
- [5] "Begomovirus: Potensi ancaman pertanian hortikultura Lampung," *J. Pertan. Agros*, vol. 45, no. 1, pp. 1–10, 2023, [Online]. Available: <https://jpa.fp.unila.ac.id/index.php/jpa/article/view/31>
- [6] FAO, "The state of food and agriculture," Rome, Italy, 2023. [Online]. Available: <https://www.fao.org/documents/card/en/c/cc7724en>
- [7] "Changing food systems and infectious disease risks in low-income and middle-income countries," *Lancet Planet. Heal.*, vol. 5, no. 9, pp. e588–e597, 2021, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2542519622001164>
- [8] B. P. S. S. I. K. Regency, "Cabai Besar (Large Chili)," 2023. [Online]. Available: <https://kebumenkab.bps.go.id/en/statistics-table/2/NzMjMg==/cabai-besar.html>
- [9] M. A. Khan, "Chili cultivars vulnerability: a multi-factorial examination of disease impact and yield reduction," *Front. Plant Sci.*, vol. 15, p. 1522985, 2024, [Online]. Available: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11488149/>
- [10] "Mitigation and Adaptation of Chili Farmers to Climate Change," *Int. J. Des. Nat. Ecodynamics*, vol. 19, no. 4, pp. 567–574, 2024, [Online]. Available: <https://www.iieta.org/journals/ijdne/paper/10.18280/ijdne.190432>
- [11] "Molecular Diversity of Pepper yellow leaf curl Indonesia virus," *J. Penelit. Tanam. Ind. dan Penyegar*, vol. 10, no. 2, pp. 89–

- 98, 2023, [Online]. Available: <https://jurnal.ugm.ac.id/v3/jtbb/article/download/15670/5546/>
- [12] S. Savary, "Certification, good agricultural practice and smallholder farmers," *Glob. Food Sec.*, vol. 21, pp. 50–58, 2019, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0959378018306897>
- [13] "A Mobile Application for Plant Disease Detection," in *Global Media Publication International Conference on Science*, 2023, pp. 120–130. [Online]. Available: <https://journal.gmpionline.com/index.php/gmpics/article/view/173>
- [14] "Deep Learning Approaches for Plant Disease Diagnosis Systems: A Review," *J. Agric. Appl. Stat.*, vol. 1, no. 1, pp. 1–15, 2025, [Online]. Available: <https://jaast.org/index.php/jaast/article/download/308/208>
- [15] "Machine learning techniques for plant disease detection," *J. Big Data*, vol. 10, no. 1, pp. 1–25, 2023, [Online]. Available: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-023-00863-9>
- [16] K. . Z. Simonyan A., "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv*, 2014, [Online]. Available: <https://arxiv.org/pdf/1409.1556.pdf>
- [17] K. . Z. He X.; Ren, S.; Sun, J., "Deep Residual Learning for Image Recognition," *CVPR*, 2016, [Online]. Available: <https://doi.org/10.1109/CVPR.2016.90>
- [18] C. . V. Szegedy V.; Ioffe, S.; Shlens, J.; Wojna, Z., "Rethinking the Inception Architecture for Computer Vision," *CVPR*, 2016, [Online]. Available: <https://doi.org/10.1109/CVPR.2016.308>
- [19] M. Data, "Chili Plant Leaf Disease and Growth Stage Dataset from Bangladesh," 2025. [Online]. Available: <https://data.mendeley.com/datasets/w9mr3vf56s/1>
- [20] S. H. Kim, "Microclimates growing and watering volumes influences the physiological traits of chili pepper cultivars in combating abiotic stress," *PMC*, 2025, [Online]. Available: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11794617/>
- [21] A. K. Das, "Assessing the climate adaptive potential of imported Chili in comparison with local cultivars through germination performance analysis," *PMC*, 2024, [Online]. Available: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11177443/>
- [22] R. K. Sharma, "Strategies to develop climate-resilient chili peppers," *PMC*, 2025, [Online]. Available: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12174194/>
- [23] A. Mumtaz, "Global range expansion history of pepper (*Capsicum* spp.) revealed by over 10,000 genebank accessions," *Proc. Natl. Acad. Sci.*, vol. 118, no. 34, p. e2104315118, 2021.
- [24] M. S. Islam, "Agro-morphological characterization and genotypic diversity of chilli (*Capsicum frutescens*) landraces of Bangladesh," *Int. J. Sci. Res. Anal.*, vol. 12, no. 3, pp. 265–275, 2024.
- [25] E. L. Arumingtyas, "Similarity Among 10 Cayenne Pepper (*Capsicum frutescens* L.) Genotypes Based on Quantitative Characters," *J. Trop. Plant Physiol.*, vol. 15, no. 2, pp. 89–98, 2023.
- [26] A. Supyani, "Determinants of symptom variation of Pepper yellow leaf curl Indonesia virus on chili pepper," *Biodiversitas*, vol. 24, no. 3, pp. 1234–1245, 2023.
- [27] P. T. et al., "Comparative result analysis of cauliflower disease classification," *Decis. Anal. J.*, 2025, [Online]. Available: <https://doi.org/10.1016/j.dajour.2025.100552>
- [28] N. E. K. et al., "Deep Learning Models Comparisons on Rice Image Disease Classification," *IEEE Access*, 2025, [Online]. Available: <https://doi.org/10.1109/ACCESS.2025.10842812>
- [29] V. T. et al., "VGG-EffAttnNet: Hybrid Deep Learning Model for Chili Plant Disease Dataset," *Front. Plant Sci.*, 2025, [Online]. Available: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12288623/>
- [30] J. Haines, "Call for emergency action to limit global temperature increases, restore biodiversity, and protect health," *J. Heal. Popul. Nutr.*, vol. 40, no. 1, p. 25, 2021, [Online]. Available: <https://jhpnm.biomedcentral.com/articles/10.1186/s41043-021-00262-x>