

Deteksi Penyakit Daun Mangga menggunakan *Threshold* dengan Metode *CNN* Arsitektur *ResNet50*

Mango Leaf Disease Detection using Threshold with CNN ResNet50 Architecture

¹Aditya Dwi Baginda*, ²Alfiah Fajriani, ³Rifa'atus Shalihah

^{1,2}Program Studi Sistem Teknologi dan Informasi, Fakultas Teknik, Universitas Muhammadiyah Kendari

³Program Studi Ilmu Komputer, FMIPA, Universitas Halu Oleo

^{1,2}Jl. KH. Ahmad Dahlan No. 10, Kota Kendari Sulawesi Tenggara 93117

³Jl. H.E.A. Mokodompit, Kota Kendari, Sulawesi Tenggara

e-mail: Aditya.22221007@umkendari.ac.id, alfiah.fajriani@umkendari.ac.id, rifaatus.shalihah@uho.ac.id

(received: 27 January 2026, revised: 7 February 2026, accepted: 8 February 2026)

Abstrak

Penyakit daun mangga merupakan ancaman produktivitas bagi petani di Indonesia sebab diagnosis yang sulit serta tidak akurat diperiksa secara manual. Sistem deteksi penyakit daun mangga dikembangkan dengan optimalisasi *decision threshold* pada klasifikasi melalui *Convolutional Neural Network ResNet50*. Dataset *Kaggle* berupa 3.979 gambar daun mangga dalam 8 kelas, *healthy*, *anthracnose*, *bacterial canker*, *gall midge*, *cutting weevil*, *die back*, *sooty mould*, dan *powdery mildew*. Dataset mentah yang dimasukkan di *Roboflow* dengan rasio pembagian data 80:10:10 dan augmentasi data 3 pada data latih menghasilkan 9.600 gambar. Optimasi *decision threshold* lewat analisis kurva *precision-recall* menunjukkan nilai 0,85 sebagai titik optimal. Di sini, *precision* ada di angka 97,03%, sedangkan *recall* 94,36%. Angka ini bisa menjadi pertimbangan penting dalam aplikasi pertanian di Indonesia, apalagi dengan karakteristik lokal yang khas. Model ini mencatat *F1-score* 95,49% setelah divalidasi pada dataset yang sudah diaugmentasi secara spesifik untuk kondisi tropis.

Kata kunci: deteksi penyakit daun mangga, *decision threshold*, *CNN ResNet50*, augmentasi data, pertanian

Abstract

Mango leaf diseases pose a significant threat to farmers' productivity in Indonesia due to the difficulty and inaccuracy of manual diagnosis. A mango leaf disease detection system was developed by optimizing the decision threshold for classification using a ResNet50 Convolutional Neural Network (CNN). The Kaggle dataset consisted of 3,979 mango leaf images across eight classes: healthy, anthracnose, bacterial canker, gall midge, cutting weevil, dieback, sooty mold, and powdery mildew. The raw dataset was processed in Roboflow with an 80:10:10 train-validation-test split, and threefold data augmentation on the training set produced a total of 9,600 images. Decision threshold optimization using the precision-recall curve analysis identified 0.85 as the optimal threshold. At this threshold, precision reached 97.03%, while recall was 94.36%. These results provide a critical reference for agricultural applications in Indonesia, particularly considering local characteristics. The model achieved an F1-score of 95.49% after validation on the augmented dataset specifically tailored for tropical conditions.

Keywords: mango leaf disease detection, *decision threshold*, *CNN ResNet50*, data augmentation, agriculture

1 Pendahuluan

Mangga *Mangifera indica* L. merupakan komoditas hortikultura yang strategis, memiliki nilai ekonomi yang tinggi dan peran yang signifikan dalam ketahanan pangan nasional. Sebagai salah satu

<http://sistemasi.ftik.unisi.ac.id>

produsen mangga terbesar di dunia, Indonesia menghadapi ancaman yang signifikan dalam budidaya, khususnya terkait dengan penyakit daun yang cenderung memusnahkan hasil panen dalam jumlah yang signifikan. *Anthracoze, bacterial canker, gall midge, cutting weevil, die back, sooty mould*, dan *powdery mildew* adalah penyakit utama yang bukan hanya berada pada daun tetapi juga memperluas ke batang dan buah, menyebabkan kerugian ekonomi yang besar bagi petani [1]. Penyakit mangga hingga saat ini diidentifikasi dan didiagnosis secara manual, bergantung pada pengalaman spesialis patologi tumbuhan yang terkadang tidak bisa dijangkau oleh semua pihak yang berkepentingan di tingkat lapangan; ketidakterjangkauan ini menyebabkan diagnosis berlangsung lambat, kurang akurat, dan sering kali menyebarkan penyakit sebelum penanganan yang efisien. Oleh karena itu, pendekatan berbasis teknologi untuk mendukung deteksi dini dan deteksi dini penyakit yang cepat, akurat, dan praktis.

Pengolahan citra digital telah menjadi teknologi yang semakin populer dan efektif yang diterapkan di bidang pertanian, di mana kondisi tanaman dideteksi dan diklasifikasikan berdasarkan karakteristik visual seperti warna, tekstur, dan fitur [2]. Dalam konteks deteksi penyakit daun, teknologi pengolahan citra memungkinkan analisis dilakukan secara otomatis dalam langkah-langkah akuisisi citra, pra-pemrosesan, segmentasi, ekstraksi fitur, dan klasifikasi. Metode pengolahan citra memungkinkan algoritma komputer untuk mengidentifikasi pola visual gejala penyakit; sehingga mengurangi ketergantungan pada pengamatan manual, yang kemungkinan dapat menimbulkan kesalahan subjektif. Studi terbaru telah membuktikan bahwa teknik pengolahan citra, termasuk segmentasi berbasis *threshold* dan model pembelajaran mendalam seperti *Convolutional Neural Network*, dapat membedakan daun sehat dengan daun yang terinfeksi dengan *accuracy* tinggi pada berbagai tanaman, termasuk mangga, dan, lebih jauh lagi, merupakan dasar yang kuat untuk mengembangkan sistem deteksi yang lebih canggih [3].

Pada klasifikasi berbasis pembelajaran mendalam, *decision threshold* adalah parameter kritis yang menentukan keseimbangan antara *precision* dan *recall*. Dalam konteks deteksi penyakit daun mangga, menjadikan *decision threshold* menjamin optimasi model yang lebih memilih memprediksi kelas positif, yang mengurangi kemungkinan *false positive* pada status tanaman, yang dapat berdampak pada penggunaan pestisida tidak perlu [4]. Pendekatan ini memungkinkan keputusan lebih akurat dalam kondisi lapangan yang berbeda, dan dapat digunakan bersama dengan model CNN seperti *ResNet50* untuk beberapa klasifikasi terbaik.

Threshold sering dikombinasikan dengan algoritma pembelajaran mendalam seperti CNN dengan arsitektur *ResNet50* untuk meningkatkan *accuracy* identifikasi dan klasifikasi penyakit [5]. Kombinasi ini memungkinkan untuk mengatasi tantangan seperti kemiripan visual antara daun sehat dan daun yang terinfeksi ringan serta variasi lingkungan yang umum ditemui di perkebunan Indonesia. Penelitian terbaru menunjukkan bahwa integrasi *threshold* sebagai pra-pemrosesan dengan CNN memungkinkan seseorang untuk mencapai *accuracy* yang sangat tinggi pada dataset multi-kelas, menjadikannya solusi yang kuat untuk aplikasi praktis di lapangan [6].

Dengan demikian, penelitian ini bertujuan untuk merancang sistem deteksi penyakit pada daun mangga dengan cara mengoptimalkan *threshold* dalam arsitektur *ResNet50*. Meskipun dataset yang digunakan berasal dari *Kaggle* yang mencakup variasi penyakit secara global, analisis mendalam mengenai keseimbangan antara *precision* dan *recall* pada berbagai *threshold* dilakukan untuk memastikan model dapat beradaptasi dengan ciri visual daun mangga di Indonesia. Studi ini juga melibatkan eksperimen inferensi waktu nyata menggunakan Google Colab untuk meniru situasi penggunaan di lapangan. Diharapkan hasil penelitian ini tidak hanya memberikan kontribusi algoritmik melalui optimasi *threshold*, tetapi juga menjadi dasar awal untuk mengembangkan aplikasi mobile yang praktis bagi petani dalam melakukan diagnosis awal, mengurangi ketergantungan pada ahli, serta mendukung efisiensi penggunaan pestisida.

2 Tinjauan Literatur

Setiawan [7] melakukan penelitian yang mengembangkan sistem klasifikasi penyakit daun mangga menggunakan Jaringan Saraf Konvolusional (CNN) dasar dan augmentasi data digunakan untuk memperkaya dataset pelatihan terbatas sebanyak 400 gambar menjadi 4.000 gambar, yang dibagi menjadi delapan kelas seperti *anthracnose* dan *bacterial canker*. Hasil penelitian tersebut menunjukkan *accuracy* pelatihan 93% dan *accuracy* validasi 94%, yang menunjukkan bahwa CNN

efektif untuk mengenali pola visual gejala penyakit pada daun mangga meskipun terdapat keterbatasan pada dataset pelatihan awal, namun masih bergantung pada augmentasi sederhana tanpa pra-pemrosesan lebih lanjut, misalnya segmentasi, yang dapat membatasi ketahanannya terhadap variasi pencahayaan di lapangan.

Nanda Cornelis Rasyid [8] membandingkan *ResNet* dan *MobileNet* untuk klasifikasi penyakit buah mangga dengan melakukan studi pada 1.600 citra yang terbagi dalam lima kelas, termasuk *healthy* dan *black mould rot*, dengan *transfer learning* menggunakan dari bobot *ImageNet*. *ResNet* 87,56% dan *MobileNet* 84,55%, mengatasi kelas *healthy*, tetapi mengalami kesulitan pada *black mould rot* karena kesamaan visual dan intuitif dalam top model *ImageNet*, mengindikasikan ketidakberhasilan algoritma mobile, berdasarkan pada *ResNet*, tapi belum berinvestasi dalam teknik pra-segmentasi untuk meningkatkan diskriminasi fitur halus tersebut.

Fiktor I. Tanesab [9] *InceptionV3* diterapkan menggunakan *transfer learning* ke jaringan pra-terlatih pada 4.000 gambar daun mangga dari delapan kelas, termasuk lalat empedu dan embun tepung, dengan *accuracy* 99,34% setelah augmentasi seperti rotasi dan geser. Pendekatan ini secara efektif mengatasi *overfitting* dengan menyempurnakan lapisan akhir, dengan *precision* dan *recall* rata-rata melebihi 99%, menunjukkan keunggulan *InceptionV3* dalam ekstraksi fitur kompleks meskipun masih bergantung pada dataset besar tanpa pendekatan berbasis *threshold* untuk ekstraksi gejala awal.

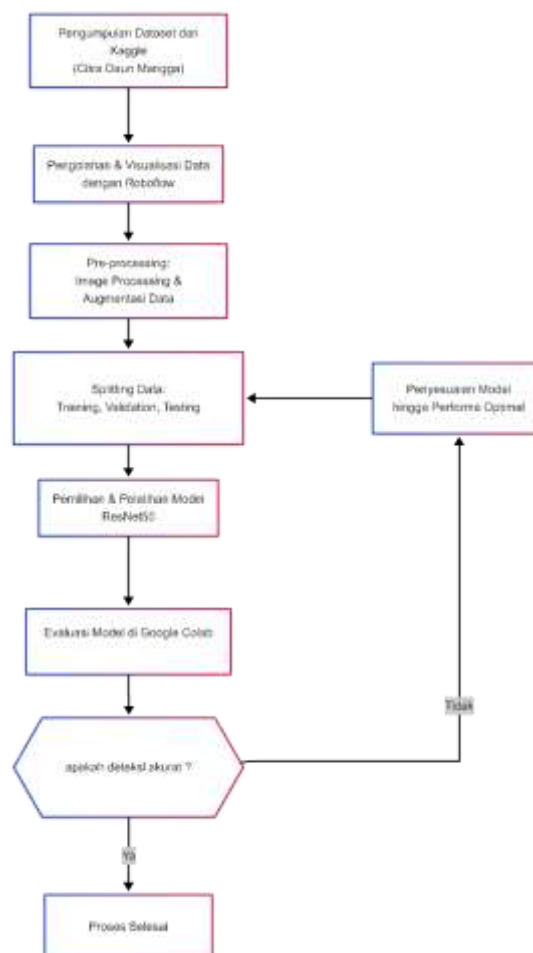
Dengan menggunakan CNN *AlexNet*, Theresia Romauli Siagian [10] mengklasifikasikan 325 gambar mangga yang telah di augmentasi menjadi 2.600, dengan *hyperparameter* dan rasio data, termasuk learning rate 10^{-5} dan rasio pelatihan-validasi 80:20 dengan *accuracy* 98,11%. Diimplementasikan melalui *Streamlit* pada situs web dengan deteksi waktu nyata, model ini berkinerja lebih rendah daripada model sebelumnya dalam menangani masalah ketidakseimbangan data tetapi kemudian menunjukkan variasi kinerja yang signifikan antar kelas karena variasi data, yang disoroti oleh, sambil terlalu menekankan potensi model berbasis web tanpa menekankan perlunya pra-pemrosesan yang sama untuk variasi lingkungan yang berbeda.

Ardiansyah [11] membandingkan kinerja *VGG16* dan *Xception* sebagai arsitektur konvolusi yang melatih model deteksi kadar klorofil dalam dataset daun mangga di dua skenario konsep klasifikasi. Dengan dua kelas lain, *VGG16* mencapai *accuracy* 100% yang diprediksi dengan model, sementara dengan tiga kelas lain, *accuracy*nya 98%, sedangkan *Xception*, dengan dua kategori lain, memiliki *accuracy* 99% dan *accuracy* 96% dengan lima kategori. Dari penelitian ini, dapat disimpulkan bahwa *VGG16* lebih konsisten jika diterapkan dalam penanganan klasifikasi gambar mangga atau klasifikasi deteksi variasi gejala persentase pada penyakit bercak daun hitam dan perubahan warna daun, yaitu jamur jelaga dan *klorosis*. Penelitian ini tidak memberikan analisis mendalam tentang penurunan parameter *decision threshold* dalam keseimbangan medium antara *precision* dan *recall* dalam konsep klasifikasi empat kelas.

Namun, analisis terhadap literatur terkait membuktikan bahwa berbagai penelitian yang menunjukkan peningkatan *accuracy* model CNN dalam mendeteksi infeksi penyakit tanaman dimungkinkan oleh penggunaan teknik *transfer learning* dan augmentasi data [12], belum ada yang melakukan analisis sistematis terhadap peran (*decision threshold*) dalam pipeline klasifikasi ini. Nyatanya, *threshold* klasifikasi tersebut merupakan parameter penting dengan kaitannya pada keseimbangan antara *precision* dan *recall*, terutama dalam konteks non-stationarity seperti ketidakseimbangan kelas di multi-class dataset, dan penelitian lain juga menunjukkan kemiripan visual antar kelas penyakit sebagai gejala, sesuai dengan multi-class dataset penyakit daun mangga yang digunakan dalam penelitian ini. Sehingga, dengan mengoptimalkan *threshold*, penelitian ini memiliki fokus pada integrasinya pada arsitektur *ResNet50* dalam multi-class leaf image dataset untuk pengembangan sistem diagnosis dini yang bisa digunakan

3 Metode Penelitian

3.1 Tahapan Penelitian



Gambar 1 Alur tahapan penelitian

Penelitian ini mengikuti beberapa tahap penelitian yang melewati alur atau sistematis (Gambar 1), yaitu pengumpulan dataset dari *Kaggle* sebagai sumber utama citra daun mangga, studi literatur dari jurnal dan sumber internet yang digunakan sebagai landasan teori, pengolahan data menggunakan platform *Roboflow* untuk visualisasi dan pengelolaan dataset, *pre-processing* dimulai dengan *image processing* dan augmentasi data, *splitting data set training, validation* dan *testing*, pemodelan dengan memilih, membangun dan melatih model *ResNet50*, dan evaluasi model dengan *confusion matrix* di *Google Colab*. Metrik pada evaluasi model antara lain *threshold*, *accuracy*, *precision*, *recall*, *F1-score* yang bertujuan untuk mendapatkan informasi yang akurat dengan proses deteksi, apabila hasil deteksi benar maka proses berhenti, jika salah maka dilakukan penyesuaian hingga mencapai performa yang optimal.

3.2 ResNet50

ResNet50 adalah varian CNN yang dikembangkan oleh *Microsoft Research* pada tahun 2015 dan menggantikan konsep residual blocks dan *skip connections* untuk menyelesaikan masalah *vanishing gradient* di jaringan yang sangat dalam [13]. Arsitektur terdiri dari lapisan konvolusi awal, dilanjutkan dengan serangkaian blok residu, lapisan *Pooling Global*, dan *fully connected* untuk output klasifikasi. *ResNet50* banyak digunakan dalam visi komputer karena stabil saat pelatihan dan mampu mengekstraksi fitur yang kaya.

3.3 Classification/Decision Threshold

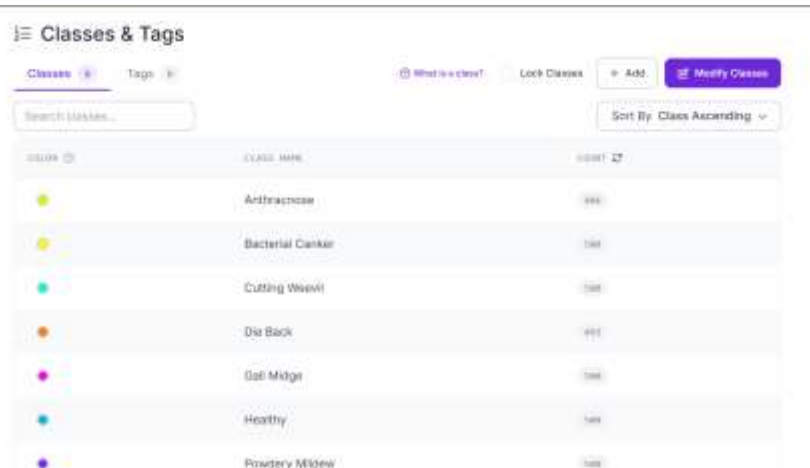
Decision threshold dalam penelitian ini merupakan nilai batas probabilitas output model CNN (dari fungsi *softmax*) yang digunakan untuk mengonversi skor probabilitas menjadi keputusan

klasifikasi. Pada skenario klasifikasi multi-kelas, *threshold* diterapkan dengan dua pendekatan bertahap:

1. **Tahap biner awal:** Model menentukan apakah daun termasuk kategori "sehat" atau "terinfeksi" berdasarkan probabilitas kelas tertinggi. Jika probabilitas tertinggi $\geq \text{threshold}$ (0,85), maka diputuskan sebagai "terinfeksi" dan dilanjutkan ke tahap kedua.
2. **Tahap diferensiasi multi-kelas:** Citra yang terdeteksi "terinfeksi" kemudian diklasifikasikan ke dalam satu dari tujuh kelas penyakit berdasarkan probabilitas terbesar di antara kelas penyakit.

Pendekatan yang digunakan memberikan kendali lebih baik terhadap kemunculan *false positive* sejak tahap awal, sehingga secara praktis dapat menekan potensi pemberian rekomendasi pestisida yang tidak perlu. Untuk menilai dampaknya secara sistematis, nilai *threshold* diuji dalam rentang 0,50 hingga 0,95 dengan kenaikan 0,05, kemudian dianalisis pengaruhnya terhadap metrik kinerja utama, meliputi *accuracy*, *precision*, *recall*, dan *F1-score*.

3.4 Pengumpulan Data



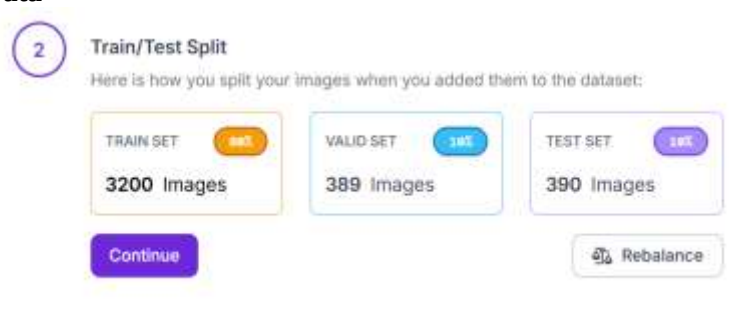
CLASS NAME	COUNT
Anthracnose	486
Bacterial Canker	500
Cutting Weevil	500
Die Back	493
Gall Midge	500
Powdery Mildew	500
Sooty Mould	500
Healthy	500

Gambar 2 Dataset Classes & Tags

Data diperoleh dengan mengakses dataset *Mango Leaf Disease* dari *Kaggle* (Gambar 2), terdiri dari 3.979 citra daun mangga berformat JPEG dengan resolusi tinggi. Dataset terdiri dari delapan kelas: *Anthracnose* dengan jumlah 486 citra, *Bacterial Canker* dengan jumlah 500 citra, *Cutting Weevil* dengan jumlah 500 citra, *Die Back* dengan jumlah 493 citra, *Gall Midge* dengan jumlah 500 citra, *Powdery Mildew* dengan jumlah 500 citra, *Sooty Mould* dengan jumlah 500 citra, dan *Healthy* dengan jumlah 500 citra. Data yang digunakan bersifat digital yang diambil secara langsung dari platform *Kaggle* yang dapat diakses melalui tautan berikut:

<https://www.kaggle.com/datasets/mahatimohammed/mango-leaf-disease-dataset>

3.5 Pengolahan Data



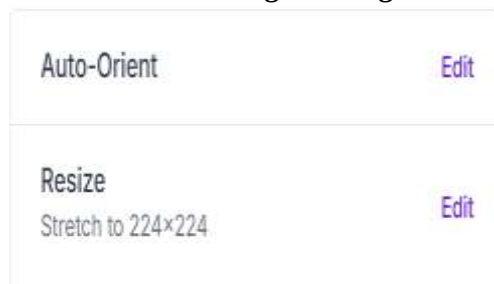
Gambar 3 Splitting data awal

Kemudian, dataset diimpor ke Roboflow untuk ulang labeling jika diperlukan, diikuti dengan pembagian rasio 80:10:10 untuk 3.200 citra untuk training dan 389 untuk citra validasi dan testing, 390 citra (Gambar 3). Alat utama yang digunakan adalah Google Colab dengan GPU untuk komputasi intensif. Sementara itu, variabel operasional yang digunakan adalah intensitas piksel dan definisi label

<http://sistemasi.ftik.unisi.ac.id>

kelas untuk klasifikasi multi-kelas. Lokasi terdiri dari lingkungan virtual berbasis cloud, sementara prosedur analisis awal berupa *explatory data analysis* untuk distribusi kelas dan identifikasi ketidakseimbangan [14].

3.6 Data Pre-Processing dan Augmentasi



Gambar 4 Proses *pre processing*



Gambar 5 Proses augmentasi

Pada data *pre-processing* dan augmentasi dilakukan sepenuhnya di Roboflow untuk meningkatkan kualitas dan variasi dataset. *Pre-processing* mencakup *image processing* seperti *auto-orient* (untuk koreksi orientasi gambar), *resize* dengan *stretch* ke ukuran 224x224 piksel sesuai input *ResNet50* [15], serta normalisasi untuk mengurangi distorsi pencahayaan. Augmentasi diterapkan pada data training dengan outputs per example sebanyak 3, termasuk *flip horizontal*, rotasi antara $\pm 15^\circ$, serta *brightness* adjustment antara $\pm 20\%$ (Gambar 4 dan Gambar 5).

3.7 Pemodelan Sistem

Pemodelan kemudian dijalankan dengan arsitektur *ResNet50* pada platform Roboflow diatur dengan konfigurasi default: learning rate 0,001, batch size 32 dan optimizer Adam dengan early-stopping mechanism. *ResNet50* dipilih berdasarkan residual learningnya yang hebat dalam mengambil fitur-fitur hierarkis pada citra multi-kelas, dan ini dilakukan *vanishing gradient* dengan memanfaatkan pretrained-weight dari *ImageNet* untuk mempercepat konvergen melalui *transfer learning* [16]. Proses pelatihan berjalan dan dilakukan otomatis pada dataset yang telah di pre-process pada tahap terdahulu, otomatis menghasilkan model yang optimal dan pencegah *overfitting* yang efisien secara komputasi.

Pada akhirnya, model diintegrasikan ke Google Colab dan uji real-time dilakukan melalui API endpoint model. Model dievaluasi dengan cara menghitung matriks performa pada berbagai *threshold* klasifikasi, yaitu sebesar 0,50, 0,55, 0,60, ..., 0,95. Dalam klasifikasi multi-kelas, *decision threshold* diterapkan pada probabilitas output fungsi *softmax* untuk setiap kelas. Mekanisme yang diimplementasikan adalah sebagai berikut: (1) Model menghasilkan probabilitas untuk masing-masing dari 8 kelas, (2) Probabilitas tertinggi dibandingkan dengan threshold yang telah ditentukan, (3) Jika probabilitas tertinggi $\geq$ *threshold*, maka kelas tersebut dipilih sebagai prediksi, (4) Jika probabilitas tertinggi $<$ *threshold*, citra dikategorikan sebagai 'tidak yakin' (*uncertain*) dan memerlukan verifikasi lebih lanjut. Pendekatan ini memungkinkan kontrol yang lebih baik terhadap *false positive rate* dalam aplikasi praktis.

3.8 Evaluasi

Pada tahap evaluasi, berdasar pemodelan dan pelatihan sistem telah terjadi pada tahap pemutakhiran model dan merupakan tahap terakhir untuk mendapatkan keluaran sejenis, maka "measure performance of the model" dalam penelitian ini menggunakan *confusion matrix* (Tabel 1), *accuracy*, *precision*, *recall*, dan *F1-score* (Persamaan 1-4) dengan *threshold* yang telah ditentukan [17]. Hasil penelitian yang dilakukan menggunakan teknik pembelajaran alih menunjukkan bahwa pelatihan model menggunakan *transfer learning* memberikan hasil yang lebih baik daripada pembelajaran biasa. Dengan demikian, *transfer learning* dapat membangkitkan kapabilitas model dalam mengenali suatu pola dan memutuskan hasil.

Tabel 1 Confusion matrix

Aktual	Prediksi	
	+	-
+	True Positive (TP)	False Negative (FN)
-	False Positive (FP)	True Negative (TN)

Berikut rumus untuk *accuracy*, *precision*, *recall* dan *F1-score* ;

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

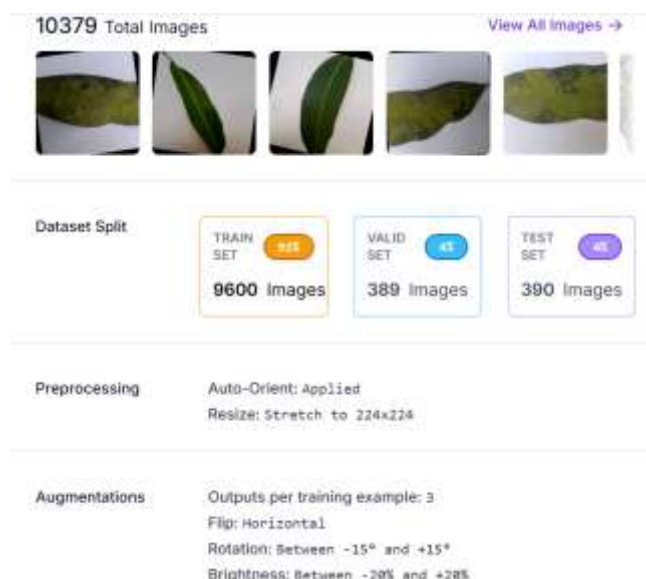
$$Precision = \frac{TP}{TP + FP} \quad (2)$$

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

$$F1 \text{ Score} = 2 \frac{\text{Presisi} \cdot \text{Recall}}{\text{Presisi} + \text{Recall}} \quad (4)$$

4 Hasil dan Pembahasan

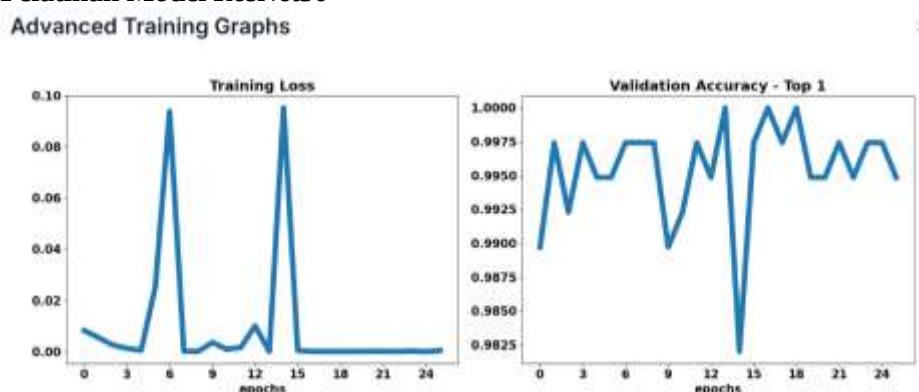
4.1 Hasil Pemrosesan Data dan Augmentasi



Gambar 6 Statistik dataset setelah pra-pemrosesan dan augmentasi

Berikutnya (Gambar 6), dataset awal yang terdiri dari 3.979 citra daun mangga dari Kaggle diproses melalui platform augmentation Roboflow dalam rasio training: validation: testing = 80:10:10. Melalui *auto-orient*, *resize* ke 224×224 piksel, *flip horizontal*, rotasi $\pm 15^\circ$, dan penyesuaian kecerahan $\pm 20\%$, data training telah di-augmentasi menjadi 9.600 citra. Hal ini meningkatkan variasi dataset dan mengurangi overfitting, terutama pada kelas seperti Anthracnose yang hanya memiliki 486 citra awal. Setelah *augmentasi*, semua kelas terdistribusi lebih merata, yang mana setiap kelas pada set training memiliki beberapa sampel yang cukup untuk mendukung pelatihan model yang stabil.

4.2 Hasil Pelatihan Model ResNet50



Gambar 7 Grafik training loss & validation accuracy

Model ResNet50 diperoleh dengan *transfer learning* menggunakan bobot awal dari ImageNet pada platform Roboflow, pada konfigurasi standar: learning rate 0,001, *batch size* 32, dan *early stopping*. Pelatihan menghasilkan konvergensi yang stabil, terlihat dari penurunan training loss hingga 0,02 dan *accuracy* validasi hingga 99,50% pada *epoch* ke-12. Kurva pelatihan (Gambar 7) tidak menunjukkan *overfitting* berkesan bahwa augmentasi data benar-benar berguna dan model dapat menggeneralisir baik untuk mengekstraksi fitur visual pada gambar penyakit daun mangga kelas multi.

4.3 Evaluasi Threshold terhadap Performa Klasifikasi

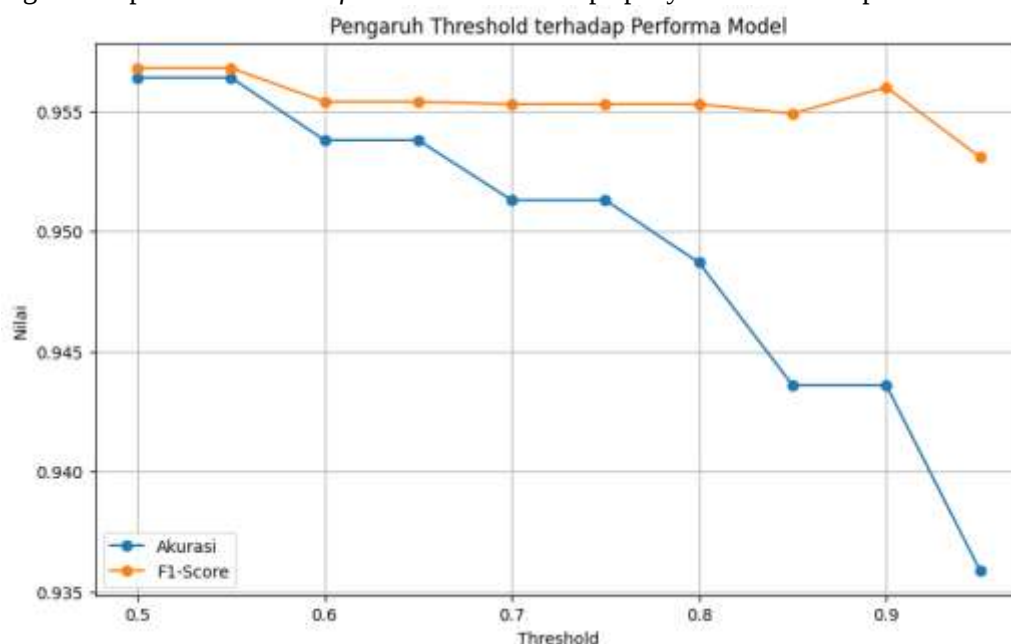
Pengujian terhadap berbagai nilai *threshold* dilakukan untuk menentukan titik optimal dalam mengubah probabilitas prediksi menjadi keputusan klasifikasi. Hasil pengujian menunjukkan variasi performa model pada rentang *threshold* 0,50 hingga 0,95 (Tabel 2). *Threshold* 0,85 dipilih sebagai nilai optimal berdasarkan pertimbangan berikut:

1. **Analisis trade-off precision–recall:** Pada *threshold* 0,85, *precision* mencapai 97,03% yang terbaik dalam rentang pengujian dengan *recall* 94,36% yang masih tinggi. Ini berarti model sangat akurat dalam mengidentifikasi daun yang benar-benar sakit (minim *false positive*) sambil tetap mampu mendeteksi sebagian besar kasus positif.
2. **Konteks aplikasi pertanian:** *False positive* (daun sehat didiagnosis sakit) berimplikasi pada penggunaan pestisida tidak perlu yang meningkatkan biaya produksi dan berdampak negatif terhadap lingkungan. Oleh karena itu, *precision* tinggi lebih diutamakan.
3. **Kurva F1-score:** *F1-score* pada *threshold* 0,85 adalah 95,49%, yang stabil pada rentang 0,80-0,90, menunjukkan titik kompromi terbaik antara *precision* dan *recall*.

Pemilihan *threshold* 0,85 dalam penelitian ini didasarkan pada pertimbangan *trade-off* antara *precision* dan *recall*, sekaligus didukung oleh temuan empiris dari studi lain. Misalnya, pada klasifikasi cacat kain, *threshold* 0,85 terbukti memberikan *accuracy* tertinggi dibandingkan nilai lain (0,25; 0,35; 0,50) di berbagai arsitektur *deep learning*, sekaligus menekan tingkat *false positive*[18]. Temuan ini memperkuat kesimpulan bahwa rentang *threshold* 0,80-0,90 umumnya sesuai untuk aplikasi klasifikasi visual yang menekankan *precesion* tinggi.

Walaupun penerapan *threshold* 0,85 menyebabkan *accuracy* sedikit lebih rendah (Gambar 8 & Tabel 2), yakni berada pada angka 94,36% jika dibandingkan dengan nilai *threshold* yang lebih kecil, peningkatan *precision* sebesar 0,85% dari *threshold* 0,80 ke 0,85 justru menjadi aspek yang lebih

signifikan ketika dilihat dari kebutuhan penggunaan di lapangan. Dalam konteks ini, pemilihan *threshold* tidak dilakukan secara generik, melainkan mempertimbangkan karakteristik domain aplikasi, sehingga analisis yang dilakukan mampu melengkapi keterbatasan studi sebelumnya yang cenderung menetapkan *threshold default* sebesar 0.50 tanpa penyesuaian terhadap kebutuhan spesifik.



Gambar 8 Analisis pengaruh threshold terhadap performa model klasifikasi merupakan grafik line chart yang menggambarkan hubungan antara nilai *threshold* (sumbu X) dengan metrik evaluasi performa model (sumbu Y), khususnya *Accuracy* dan *F1-score*.

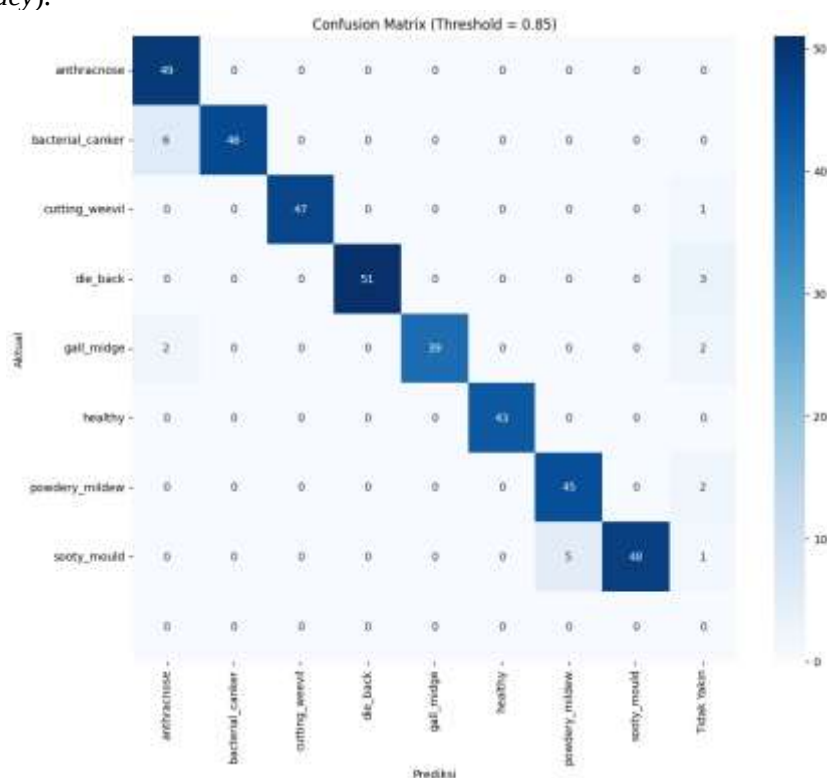
Tabel 2 Hasil evaluasi threshold terhadap performa model

Threshold	Accuracy	Precision	Recall	F1-score
0,50	0,9564	0,9617	0,9564	0,9568
0,55	0,9564	0,9617	0,9564	0,9568
0,60	0,9538	0,9617	0,9538	0,9554
0,65	0,9538	0,9617	0,9538	0,9554
0,70	0,9513	0,9641	0,9513	0,9553
0,75	0,9513	0,9641	0,9513	0,9553
0,80	0,9487	0,9666	0,9487	0,9553

Threshold	Accuracy	Precision	Recall	F1-score
0,85	0,9436	0,9703	0,9436	0,9549
0,90	0,9436	0,9722	0,9436	0,9560
0,95	0,9359	0,9737	0,9359	0,9531

4.4 Confusion Matrix dan Analisis Performa per Kelas

Confusion matrix pada threshold 0,85 (Gambar 9) mengungkapkan distribusi prediksi model pada 390 citra testing. Secara keseluruhan, model berhasil mengklasifikasikan 368 citra dengan benar (94,36% accuracy).



Gambar 9 Confusion matrix untuk evaluasi model klasifikasi pada threshold 0,85

merupakan representasi (*confusion matrix*) yang menunjukkan distribusi prediksi model klasifikasi penyakit daun mangga dibandingkan dengan label asli (*ground truth*).

Tabel 3 Performa klasifikasi per kelas pada threshold 0,85

Kelas	TP	FP	FN	Precision	Recall	F1-score	Jumlah Data
(Antraknosa) Anthracnose	49	2	1	0,961	0,980	0,970	50
(Kanker Bakteri) BacterialCanker	47	3	0	0,940	1.000	0,969	47

Kelas	TP	FP	FN	Precision	Recall	F1-score	Jumlah Data
(Kumbang Pemotong) Cutting Weevil	48	1	2	0,980	0,960	0,970	50
(Mati Pucuk) Die Back	46	0	4	1.000	0,920	0,958	50
(Lalat Puru) Gall Midge	44	2	4	0,957	0,917	0,936	48
(Sehat) Healthy	51	0	0	1.000	1.000	1.000	51
(Embun Tepung) Powdery Mildew	43	1	6	0,977	0,878	0,925	49
(Embun Jelaga) Sooty Mould	45	0	5	1.000	0,900	0,947	50

Berdasarkan Performa Klasifikasi per Kelas pada *Threshold* 0,85 (Tabel 3) dapat disimpulkan sebagai berikut:

- Class *Healthy*: Menampilkan performa sempurna dengan *precision*, *recall*, dan *F1-Score* 100%. Ini berarti bahwa model benar-benar mampu secara sempurna membedakan daun yang sehat dan daun yang sakit, yang merupakan kasus ideal dalam sistem deteksi dini.
- Class *Die back* dan *Sooty Mould*: Menampilkan *precision* 100% tetapi *recall* yang lebih rendah (92% dan 90%). Ini merupakan indikasi bahwa ketika model memprediksi kelas-kelas ini, sepenuhnya benar, tetapi ada beberapa kasus yang diabaikan model yaitu sekitar delapan data. Ciri visual dari kedua penyakit ini adalah kejelasan mati dari jaringan dan beberapa daun yang terkontaminasi pada *Sooty Mould* dicemari lapisan jelaga hitam memudahkan identifikasi ini, tetapi tidak dalam setiap kondisi penerangan atau dalam jarak yang sangat dekat.
- Class *Powdery Mildew*: Mempunyai *recall* terendah (87,8%) dengan 6 *false negative*. Penyakit ini sering kali hanya berada sebagai lapisan zat putih atau debu pada permukaan daun dan kadang-kadang berubah-ubah tergantung penerangan atau jika terdapat variasi pencahayaan.
- Class *Anthrachnose*, *Bacterial Canker* dan *Cutting Weevil*: Menampilkan performa sangat baik dengan *F1-score* sekitar 0,96% dan sama perbandingan *precision* dan *recall* di atas 0,99%. Hanya sejumlah kecil data termasuk di dalam kelas ini tetapi mereka memiliki gejala visual yang cukup khas.

4.5 Keunggulan dan Kontribusi Penelitian

Penelitian ini memberikan kontribusi inovatif melalui integrasi sistematis analisis *threshold* dengan arsitektur *ResNet50*, mengisi celah dalam penelitian sebelumnya. Berbeda dengan Setiawan [7] yang menggunakan CNN dasar tanpa optimasi *threshold*, analisis komprehensif pada rentang *threshold* 0,50–0,95 mengungkap nilai optimal 0,85 yang meningkatkan *precision* hingga 97,03% tanpa mengorbankan *recall* secara signifikan aspek kritis dalam pertanian untuk menghindari *false positive* dan penggunaan pestisida tidak perlu. Strategi augmentasi yang diterapkan lebih komprehensif daripada Theresia Romauli Siagian [10], dengan rotasi unik, *flip*, dan penyesuaian kecerahan pada kombinasi datasetnya untuk generalisasi lebih besar terhadap variasi kondisi

lapangan. Sementara itu, Nanda Cornelis Rasyid [8] fokus pada porositas arsitektur yang kompleks, alternatif ini menunjukkan bahwa *ResNet50* dengan kombinasi *threshold* optimasi dapat mencapai *F1-score* 0,9549 dengan biaya komputasi yang memungkinkan infrastruktur rendah. Berbeda dengan Fiktor I. Tanesab [9] yang hanya mempublikasikan hasil statis, pendekatan ini mengimplementasikan lintasan lengkap dari pelatihan hingga validasi real-time, mensimulasikan kondisi penggunaan aktual di lapangan. Oleh karena itu, metode ini mengungguli implementasi sistem berbasis Streamlit pada Ryo Ardiansyah [11] yang belum mengadopsi validasi real-time melalui API cloud. Secara keseluruhan, penelitian ini mengkonsolidasikan analisis *threshold* sistematis, augmentasi baru, efisiensi komputasi, dan validasi *real-time* ke dalam kerangka kerja yang dapat diterapkan untuk pertanian.

4.6 Generalisasi Model Dan Validasi Eksternal

Meskipun dataset utama bersumber dari platform global (*Kaggle*), penelitian ini memanfaatkan augmentasi data untuk menyiapkan model menghadapi beragam variasi visual khas perkebunan Indonesia. Rotasi $\pm 15^\circ$, *flip horizontal*, dan penyesuaian kecerahan $\pm 20\%$ digunakan untuk meningkatkan *robustness* terhadap kondisi lapangan yang tidak ideal. Perubahan kecerahan meniru fluktuasi pencahayaan akibat cuaca tropis, sedangkan transformasi geometris menambah variasi sudut pandang dan orientasi daun. Dengan mensimulasikan kondisi ini selama pelatihan, model tidak hanya mempelajari fitur penyakit utama, tetapi juga menguatkan ketahanannya terhadap nuansa lingkungan yang biasanya menurunkan performa model berbasis data terkurasi. Pendekatan ini secara efektif memperkecil *domain gap* antara dataset internasional dan kondisi nyata di Indonesia, sekaligus membangun dasar yang lebih kokoh untuk validasi dan penerapan praktis.

5 Kesimpulan

Secara keseluruhan, performa sistem yang dikembangkan dalam penelitian ini menunjukkan hasil yang cukup solid ketika diterapkan untuk mendeteksi penyakit pada daun mangga. Model berbasis CNN *ResNet50* tersebut tidak hanya mengandalkan arsitektur jaringan, tetapi juga diperkuat lewat pengaturan *decision threshold* yang tepat, sehingga tingkat keandalannya tetap terjaga. Dalam praktik pengujian, berbagai nilai *threshold* dicoba, mulai dari 0,50 hingga 0,95. Di antara rentang tersebut, angka 0,85 justru memberikan titik temu yang paling masuk akal. Pada nilai ini, sistem mampu mencatat *accuracy* sebesar 94,36%, dengan *precision* mencapai 97,03%, *recall* 94,36%, dan *F1-score* berada di angka 95,49%. Pemilihan *threshold* tersebut tentu bukan hasil coba-coba semata. Ada pertimbangan nyata di baliknya, terutama soal bagaimana menjaga keseimbangan antara kemampuan sistem dalam mengenali penyakit dan meminimalkan kesalahan deteksi positif. Hal ini menjadi penting karena, di lapangan, praktik pertanian cenderung lebih sensitif terhadap kesalahan alarm palsu dibandingkan dengan risiko terlewatnya kasus tertentu. Keterbatasan dan Arah Penelitian Selanjut (1) beberapa hal masih menjadi keterbatasan dalam penelitian ini. Pertama, data yang digunakan berasal dari sumber internasional (*Kaggle*), sehingga diperlukan pengujian lanjutan menggunakan data lokal Indonesia agar hasilnya lebih relevan dengan kondisi nyata di lapangan, (2) ke depan, sistem juga dapat dikembangkan dengan *threshold* yang bersifat adaptif, sehingga mampu menyesuaikan secara otomatis terhadap perbedaan pencahayaan maupun kualitas citra saat pengambilan gambar, (3) selain itu, cakupan dataset sebaiknya diperluas dengan menambahkan lebih banyak varietas mangga khas Indonesia serta jenis penyakit yang spesifik pada wilayah tertentu, sehingga model yang dihasilkan menjadi lebih representatif.

Referensi

- [1] S. I. Ahmed *et al.*, "MangoLeafBD: A Comprehensive Image Dataset to Classify Diseased and Healthy Mango Leaves," *Data Br.*, Vol. 47, p. 108941, 2023, DOI: 10.1016/j.dib.2023.108941.
- [2] J. R. Hidayat and Jemakmum, "Implementasi Klasifikasi Citra berbasis *Tensorflow* untuk mendeteksi Penyakit Tanaman pada Aplikasi Agroscan," *J. Fasilkom*, Vol. 15, No. 1, pp. 124–130, 2025, DOI: 10.37859/jf.v15i1.8536.
- [3] S. Khandelwal, A. Raut, H. Vyawahare, D. Theng, and S. Dhande, "Optimizing Performance in Mango Plant Leaf Disease Classification through Advanced Machine Learning Techniques," *Eng. Technol. Appl. Sci. Res.*, Vol. 14, No. 6, pp. 18476–18480, 2024, DOI: 10.48084/etasr.8220.

- [4] O. Rainio et al., “Comparison of Thresholds for a Convolutional Neural Network Classifying Medical Images,” *Int. J. Data SCI. Anal.*, Vol. 20, No. 3, pp. 2093–2099, 2025, DOI: 10.1007/s41060-024-00584-z.
- [5] S. Jayashree and V. Sumalatha, “Plant Leaf Disease Detection using ResNet-50 based on Deep Learning BT - Proceedings of the International Conference on Digital Transformation in Business: Navigating the New Frontiers Beyond Boundaries (DTBNNF 2024),” Atlantis Press, 2024, pp. 150–166. DOI: 10.2991/978-94-6463-433-4_12.
- [6] H. Rogers, B. De La Iglesia, T. Zebin, G. Cielniak, and B. Magri, “Advancing Precision Agriculture: Domain-Specific Augmentations and Robustness Testing for Convolutional Neural Networks in Precision Spraying Evaluation,” *Neural Comput. Appl.*, Vol. 36, No. 32, pp. 20211–20229, 2024, DOI: 10.1007/s00521-024-10142-0.
- [7] A. Setiawan, A. Nursafitri, E. V. Afnarista, H. Mulyasari, V. H. Zahwa, and Z. Suryadi, “Klasifikasi Penyakit Tanaman Mangga menggunakan *Algorithma CNN* pada Citra Daun Daun,” *Sentimeter. (Seminar Nas. Teknol. Informasi, Mekatronika dan Ilmu Komputer)*, Vol. 4, pp. 30–41, 2025.
- [8] N. C. Rasyid, J. Karman, A. T. Hidayat, H. Oktavia, and L. Wijaya, “Klasifikasi Jenis Penyakit Buah Mangga berbasis *Deep Learning* menggunakan Arsitektur *ResNet* dan *MobileNet*,” No. 31626, pp. 166–176, 2026.
- [9] F. I. Tanesab, C. Y. Laatrehe, and M. A. Wuarlela, “Penerapan Metode *CNN* dalam mengklasifikasi Jenis Penyakit pada Daun Mangga menggunakan Arsitektur *InceptionV3*,” Vol. 1, pp. 265–275, 2025.
- [10] S. T. R. Siagian, I. Taufik, Z. Indra, “Klasifikasi Penyakit dan Hama pada Daun Mangga menggunakan Metode *Convolutional Neural Network*,” *JATI (Jurnal Mhs. Tek. Inform.)*, Vol. 9, No. 3, pp. 4135–4143, 2025, DOI: 10.36040/jati.v9i3.13571.
- [11] R. Ardiansyah, M. A. 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.
- [12] E. A. Rosyidah, A. F. Hadi, and Y. S. Dewi, *The Classification of Tea Leaf Disease using CNN Image Classifier*. Atlantis Press International BV, 2023. DOI: 10.2991/978-94-6463-174-6_10.
- [13] E. Suherman, B. Rahman, D. Hindarto, and H. Santoso, “Implementation of *ResNet-50* on End-to-End Object Detection (*DETR*) on Objects,” *SinkrOn*, Vol. 8, No. 2, pp. 1085–1096, 2023, DOI: 10.33395/sinkron.v8i2.12378.
- [14] M. A. Akif, I. Butun, A. Williams, and I. Mahgoub, “Hybrid Machine Learning Models for Intrusion Detection in IoT: Leveraging a Real-World IoT Dataset,” pp. 1–9, 2025, [Online]. Available: <http://arxiv.org/abs/2502.12382>
- [15] N. IBRAHIM et al., “Klasifikasi Tingkat Kematangan Pucuk Daun Teh menggunakan Metode *Convolutional Neural Network*,” *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, Vol. 10, No. 1, p. 162, 2022, DOI: 10.26760/elkomika.v10i1.162.
- [16] S. A. Adzkia and T. Arifin, “Perbandingan Arsitektur *CNN* berbasis *Transfer Learning* untuk Klasifikasi pada *BreastMNIST*,” *J. Nas. Teknol. dan Sist. Inf.*, Vol. 11, No. 2, pp. 192–200, 2025, DOI: 10.25077/teknosi.v11i2.2025.192-200.
- [17] K. F. 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.
- [18] J. E. Adio et al., “Development of an Optimized Deep Learning Technique for Fabric Defect Classification using *Osprey Optimization Algorithm*,” *FUOYE J. Eng. Technol.*, Vol. 9, No. 4, pp. 643–648, 2025, DOI: 10.4314/fuoyejet.v9i4.12.