

Analisis Komparatif CNN Ringan untuk Klasifikasi Penyakit Daun Tomat Menggunakan Visualisasi Grad-CAM

Sayuti Rahman^{1*}, Hartono², Arnes Sembiring³, Muhammad Khahfi Zuhanda⁴, Bayu Aditya Pratama⁵, Dewi Martini⁶

¹⁻⁶ Pascasarjana, Magister Informatika, Universitas Medan Area, Medan, Indonesia

^{1*}sayutirahman@staff.uma.ac.id, ²hartono@staff.uma.ac.id, ³arnessembiring@staff.uma.ac.id,

⁴khahfizuhanda@staff.uma.ac.id, ⁵Bayuaditya2257@gmail.com, ⁶heppytransportasindo@gmail.com

^{*}) sayutirahman@staff.uma.ac.id

Abstrak-Klasifikasi penyakit daun tomat berbasis citra digital menjadi salah satu pendekatan penting dalam mendukung pertanian cerdas, khususnya dalam deteksi dini serangan penyakit tanaman. Penelitian ini bertujuan membandingkan kinerja beberapa arsitektur lightweight Convolutional Neural Network (CNN), yaitu MobileNetV3-Small, MobileNetV2, dan EfficientNet-B0, dalam mengklasifikasikan penyakit daun tomat pada dataset PlantVillage. Dataset yang digunakan terdiri dari 3.628 citra yang terbagi ke dalam 10 kelas (9 penyakit dan 1 sehat) dengan skema pembagian data 80% untuk pelatihan dan 20% untuk validasi. Evaluasi kinerja dilakukan menggunakan metrik classification report, confusion matrix, serta analisis interpretabilitas menggunakan Grad-CAM dan feature map visualization. Hasil penelitian menunjukkan bahwa seluruh model mampu menghasilkan akurasi sangat tinggi di atas 99%. EfficientNet-B0 memperoleh performa terbaik dengan akurasi validasi 99,59%, diikuti MobileNetV2 sebesar 99,45% dan MobileNetV3-Small sebesar 99,04%. Meskipun demikian, kompleksitas model meningkat seiring akurasi, di mana EfficientNet-B0 memiliki parameter dan FLOPs terbesar. Analisis Grad-CAM menunjukkan bahwa model berakurasi tinggi memiliki fokus aktivasi yang lebih presisi pada area lesi daun. Penelitian ini menegaskan bahwa CNN ringan mampu memberikan kinerja klasifikasi yang sangat baik, sekaligus membuka peluang implementasi pada sistem deteksi penyakit tanaman berbasis perangkat terbatas.

Kata Kunci: PlantVillage; Klasifikasi Penyakit Daun; Lightweight CNN; MobileNet; EfficientNet; Grad-CAM

Abstract-Tomato leaf disease classification based on digital imagery has become an important approach in supporting smart agriculture, particularly for early detection of plant disease attacks. This study aims to compare the performance of several lightweight Convolutional Neural Network (CNN) architectures, namely MobileNetV3-Small, MobileNetV2, and EfficientNet-B0, in classifying tomato leaf diseases using the PlantVillage dataset. The dataset consists of 3,628 images distributed across 10 classes (9 disease classes and 1 healthy class), with a data split scheme of 80% for training and 20% for validation. Performance evaluation was conducted using classification reports, confusion matrices, and interpretability analysis through Grad-CAM and feature map visualization. The experimental results show that all models achieved very high accuracy, exceeding 99%. EfficientNet-B0 obtained the best performance with a validation accuracy of 99.59%, followed by MobileNetV2 at 99.45% and MobileNetV3-Small at 99.04%. However, model complexity increased along with accuracy, where EfficientNet-B0 had the largest number of parameters and FLOPs. Grad-CAM analysis revealed that higher-accuracy models demonstrated more precise activation focus on leaf lesion regions. This study confirms that lightweight CNN architectures are capable of delivering excellent classification performance while offering strong potential for deployment in plant disease detection systems on resource-limited devices..

Keywords: PlantVillage; Leaf Deases Classification; Lightweight CNN; MobileNet; EfficientNet; Grad-CAM

1. PENDAHULUAN

Penyakit pada tanaman tomat merupakan salah satu faktor utama yang menyebabkan penurunan produktivitas pertanian. Serangan penyakit seperti Early Blight, Late Blight, dan Septoria Leaf Spot dapat menyebar dengan cepat serta sulit diidentifikasi secara manual tanpa keahlian khusus. Metode identifikasi konvensional yang mengandalkan pengamatan visual oleh pakar seringkali memerlukan waktu lama, biaya tinggi, dan rentan terhadap subjektivitas. Oleh karena itu, diperlukan pendekatan otomatis berbasis teknologi yang mampu melakukan klasifikasi penyakit secara cepat, akurat, dan konsisten [1], [2], [3].

Perkembangan deep learning, khususnya Convolutional Neural Network (CNN), telah menunjukkan performa unggul dalam berbagai tugas klasifikasi citra, termasuk di bidang pertanian [4], [5], [6]. CNN mampu mengekstraksi fitur visual kompleks seperti tekstur, warna, dan pola bercak daun secara hierarkis. Namun demikian, arsitektur CNN konvensional umumnya memiliki jumlah parameter besar dan kebutuhan komputasi tinggi, sehingga kurang efisien untuk implementasi pada perangkat dengan sumber daya terbatas.



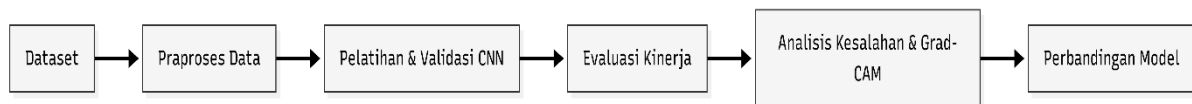
Untuk mengatasi keterbatasan tersebut, berbagai arsitektur lightweight CNN telah dikembangkan oleh peneliti sebelumnya, seperti keluarga MobileNet dan EfficientNet, yang dirancang untuk mempertahankan akurasi tinggi dengan kompleksitas komputasi lebih rendah. MobileNetV3-Small [7], [8] menekankan efisiensi melalui depthwise separable convolution dan modul squeeze-and-excitation, MobileNetV2 [9], [10], [11], [12] memanfaatkan inverted residual bottleneck untuk optimasi representasi fitur, sedangkan EfficientNet-B0 [13], [14], [15] menggunakan pendekatan compound scaling guna menyeimbangkan kedalaman, lebar, dan resolusi jaringan. Dalam penelitian ini, ketiga arsitektur tersebut tidak dimodifikasi, melainkan digunakan sebagai model pembandingan dalam skenario klasifikasi yang sama.

Meskipun berbagai penelitian telah memanfaatkan CNN untuk klasifikasi penyakit tanaman, kajian komparatif yang tidak hanya menilai akurasi tetapi juga kompleksitas model serta interpretabilitas visual masih relatif terbatas. Selain itu, pemahaman mengenai bagaimana model memfokuskan perhatian pada area lesi daun menjadi aspek penting dalam memvalidasi keandalan keputusan jaringan, terutama ketika akurasi yang dihasilkan sudah sangat tinggi [16], [17], [18].

Berdasarkan permasalahan tersebut, penelitian ini bertujuan membandingkan kinerja tiga arsitektur lightweight CNN seperti MobileNetV3-Small, MobileNetV2, dan EfficientNet-B0 dalam klasifikasi penyakit daun tomat menggunakan dataset PlantVillage. Evaluasi dilakukan secara kuantitatif melalui metrik klasifikasi, kompleksitas parameter, dan FLOPs, serta secara kualitatif menggunakan Grad-CAM dan visualisasi fitur berlapis. Novelty penelitian ini terletak pada analisis komprehensif yang mengintegrasikan perbandingan akurasi, efisiensi komputasi, dan interpretabilitas visual secara simultan untuk mengidentifikasi trade-off performa pada model CNN ringan, sehingga menghasilkan rekomendasi arsitektur yang paling optimal dan dapat dipertanggungjawabkan secara teknis maupun visual.

2. METODE PENELITIAN

Metode penelitian yang digunakan dalam studi ini mengacu pada alur kerja klasifikasi citra berbasis *deep learning* yang terstruktur, sebagaimana ditunjukkan pada Gambar 1. Proses diawali dengan pengumpulan dataset citra daun tomat, kemudian dilanjutkan dengan tahap praproses data untuk menyiapkan kualitas dan konsistensi citra sebelum pelatihan model. Selanjutnya, beberapa arsitektur CNN ringan dilatih dan divalidasi untuk menghasilkan model terbaik. Kinerja model kemudian dievaluasi menggunakan metrik klasifikasi, diikuti analisis kesalahan prediksi serta visualisasi interpretabilitas menggunakan Grad-CAM. Tahap akhir adalah perbandingan performa antar model untuk menentukan arsitektur paling optimal.



Gambar 1. Metode Penelitian

Gambar 1 memperlihatkan tahapan penelitian yang terdiri dari enam proses utama, yaitu: (1) pengumpulan dataset, (2) praproses data, (3) pelatihan dan validasi CNN, (4) evaluasi kinerja model, (5) analisis kesalahan dan visualisasi Grad-CAM, serta (6) perbandingan model. Alur ini dirancang untuk memastikan bahwa evaluasi tidak hanya dilakukan secara kuantitatif melalui akurasi, tetapi juga secara kualitatif melalui interpretasi area perhatian model.

2.1 Dataset

Dataset yang digunakan dalam penelitian ini bersumber dari PlantVillage Dataset, salah satu dataset terbuka yang banyak dimanfaatkan dalam klasifikasi penyakit tanaman berbasis citra. Penelitian difokuskan pada daun tomat (tomato leaves) yang terdiri dari 10 kelas, yaitu 9 jenis penyakit dan 1 kelas sehat, dengan total 3.628 citra. Seluruh gambar melalui tahap praproses untuk menyeragamkan ukuran dan kualitas sebelum pelatihan model. Dataset kemudian dibagi menggunakan skema hold-out validation dengan komposisi 80% data latih (2.902 citra) dan 20% data validasi (726 citra) yang dipilih secara acak menggunakan fixed random seed guna menjaga konsistensi eksperimen antar model CNN. Distribusi jumlah data tiap kelas disajikan pada Tabel 1.

Tabel 1. Dataset

No	Kelas Penyakit Daun Tomat	Jumlah Citra
1	Tomato Bacterial_spot	425
2	Tomato Early_blight	200
3	Tomato Late_blight	381
4	Tomato Leaf_Mold	190

5	Tomato Septoria_leaf_spot	356
6	Tomato Spider_mites Two-spotted_spider_mite	335
7	Tomato Target_Spot	281
8	Tomato Tomato_Yellow_Leaf_Curl_Virus	1.071
9	Tomato Tomato_mosaic_virus	75
10	Tomato healthy	314
Total		3.628

2.2 Lightweight CNN

Convolutional Neural Network (CNN) ringan (lightweight CNN) merupakan arsitektur jaringan saraf konvolusional yang dirancang dengan jumlah parameter dan kompleksitas komputasi yang lebih rendah dibandingkan CNN konvensional, namun tetap mampu mempertahankan performa klasifikasi yang tinggi. Arsitektur ini umumnya memanfaatkan teknik efisiensi seperti depthwise separable convolution, bottleneck layer, dan compound scaling untuk mengurangi beban komputasi tanpa mengorbankan akurasi secara signifikan. Dalam penelitian ini, penggunaan CNN ringan menjadi penting karena memungkinkan proses pelatihan dan inferensi yang lebih cepat serta berpotensi diimplementasikan pada perangkat dengan sumber daya terbatas, khususnya dalam aplikasi deteksi penyakit tanaman berbasis citra.

2.2.1 MobileNetV3-Small

MobileNetV3-Small merupakan arsitektur CNN ringan yang dirancang untuk efisiensi tinggi melalui kombinasi depthwise separable convolution, inverted residual block, dan squeeze-and-excitation module. Model ini memiliki jumlah parameter yang relatif kecil sehingga cocok untuk klasifikasi citra dengan kebutuhan komputasi rendah namun tetap mampu menghasilkan akurasi tinggi [19], [20], [21].

2.2.2 MobileNetV2

MobileNetV2 mengusung konsep inverted residual dan linear bottleneck untuk meningkatkan efisiensi representasi fitur. Arsitektur ini lebih dalam dibanding MobileNetV3-Small dan mampu mengekstraksi fitur yang lebih kompleks, sehingga sering digunakan sebagai baseline kuat pada tugas klasifikasi citra ringan [22], [23], [24].

2.2.3. EfficientNetB0

EfficientNet-B0 dikembangkan menggunakan pendekatan compound scaling yang menyeimbangkan kedalaman, lebar, dan resolusi jaringan. Meskipun masih tergolong ringan, model ini memiliki kapasitas representasi fitur lebih tinggi sehingga berpotensi menghasilkan performa klasifikasi terbaik di antara model lightweight lainnya [25], [26], [27].

2.3 Grad-CAM

Gradient-weighted Class Activation Mapping (Grad-CAM) merupakan metode *visual explanation* yang digunakan untuk mengidentifikasi area citra yang paling berkontribusi terhadap keputusan klasifikasi CNN. Grad-CAM bekerja dengan memanfaatkan gradien dari kelas target terhadap *feature map* pada lapisan konvolusi terakhir, sehingga mampu menampilkan peta panas (*heatmap*) yang merepresentasikan tingkat kepentingan setiap wilayah citra [28], [29], [30]. Secara matematis, bobot kepentingan setiap kanal fitur dihitung menggunakan rata-rata global gradien seperti disajikan pada Persamaan 1.

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (1)$$

dimana:

α_k^c = bobot kepentingan kanal ke-k untuk kelas c

y^c = skor keluaran sebelum *softmax* untuk kelas c

A_{ij}^k = aktivasi *feature map* pada kanal k di posisi (i,j)

Z = jumlah piksel pada *feature map*

Selanjutnya, peta aktivasi kelas (*class activation map*) dihitung dengan mengombinasikan bobot dan *feature map*:

$$L_{Grad-CAM}^c = ReLU\left(\sum_k \alpha_k^c A^k\right) \quad (2)$$

Fungsi ReLU digunakan untuk mempertahankan kontribusi positif yang relevan terhadap kelas target dan menghilangkan pengaruh negatif.

3. HASIL DAN PEMBAHASAN

Berdasarkan metode penelitian yang dilakukan, studi ini membandingkan beberapa arsitektur lightweight Convolutional Neural Network (CNN) untuk klasifikasi citra daun tomat pada dataset PlantVillage. Evaluasi tidak hanya menilai akurasi, tetapi juga kemampuan model dalam mengenali pola visual penyakit. Pengujian dilakukan secara kuantitatif menggunakan classification report dan confusion matrix, serta secara kualitatif melalui visualisasi Grad-CAM dan layer-wise feature maps untuk memahami fokus perhatian model. Pendekatan ini memberikan gambaran komprehensif mengenai kinerja, interpretabilitas, serta potensi kesalahan klasifikasi pada tiap arsitektur CNN.

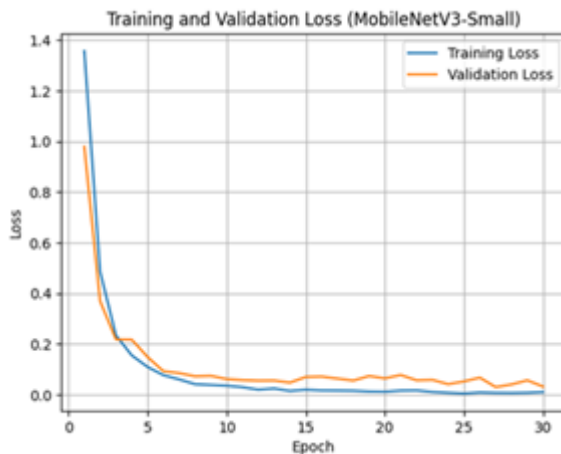
3.1 Hasil Klasifikasi Menggunakan MobileNet V3 Small

Subbagian ini menyajikan hasil klasifikasi citra daun tomat menggunakan arsitektur MobileNetV3-Small sebagai salah satu representasi lightweight CNN. Evaluasi kinerja model dilakukan secara kuantitatif melalui laporan klasifikasi yang mencakup metrik precision, recall, F1-score, dan accuracy pada seluruh kelas penyakit daun tomat. Pendekatan ini digunakan untuk menilai kemampuan MobileNetV3-Small dalam mengenali setiap kelas secara seimbang, sekaligus mengidentifikasi potensi perbedaan performa antar kelas yang disebabkan oleh variasi karakteristik visual maupun distribusi jumlah data. Hasil lengkap dari evaluasi klasifikasi MobileNetV3-Small disajikan pada Tabel 2 sebagai dasar analisis kinerja model pada subbagian ini.

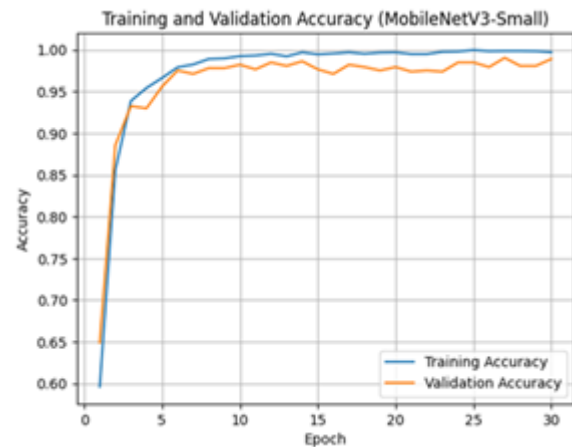
Tabel 2. Laporan Klasifikasi MobileNet V3 Small

Kelas	Precision	Recall	F1-Score	Support
Tomato Bacterial_spot	0.988	0.965	0.9762	85
Tomato Early_blight	0.9744	0.95	0.962	40
Tomato Late_blight	0.9737	0.974	0.9737	76
Tomato Leaf_Mold	1	1	1	38
Tomato Septoria_leaf_spot	1	1	1	71
Tomato Spider_mites Two-spotted_spider_mite	1	1	1	67
Tomato Target_Spot	0.9825	1	0.9912	56
Tomato Tomato_Yellow_Leaf_Curl_Virus	0.9907	1	0.9953	214
Tomato Tomato_mosaic_virus	1	1	1	15
Tomato healthy	1	1	1	64
Accuracy			0.9904	726
Macro Avg	0.9909	0.989	0.9898	726
Weighted Avg	0.9903	0.99	0.9903	726

Berdasarkan hasil pada tabel 2, MobileNetV3-Small menunjukkan performa klasifikasi yang sangat konsisten pada hampir seluruh kelas penyakit daun tomat. Nilai precision, recall, dan F1-score yang tinggi menandakan bahwa model mampu mengenali pola visual penyakit secara akurat dengan tingkat kesalahan yang sangat rendah. Beberapa kelas bahkan mencapai skor sempurna, yang mengindikasikan bahwa ciri visual pada kelas tersebut relatif mudah dibedakan oleh model. Namun, pada kelas seperti Bacterial spot, Early blight, dan Late blight, masih ditemukan kesalahan klasifikasi meskipun dalam jumlah kecil, yang diduga disebabkan oleh kemiripan tekstur, warna, dan pola bercak daun antar penyakit. Kondisi ini menunjukkan bahwa meskipun MobileNetV3-Small memiliki kemampuan ekstraksi fitur yang kuat, kompleksitas visual antar kelas tertentu tetap menjadi tantangan. Untuk memahami bagaimana proses pembelajaran model berlangsung selama pelatihan dan bagaimana stabilitas konvergensinya, analisis selanjutnya difokuskan pada kurva training loss dan validation loss yang disajikan pada Gambar 2.



Gambar 2. Training and Validation Loss MobileNetV3-Small



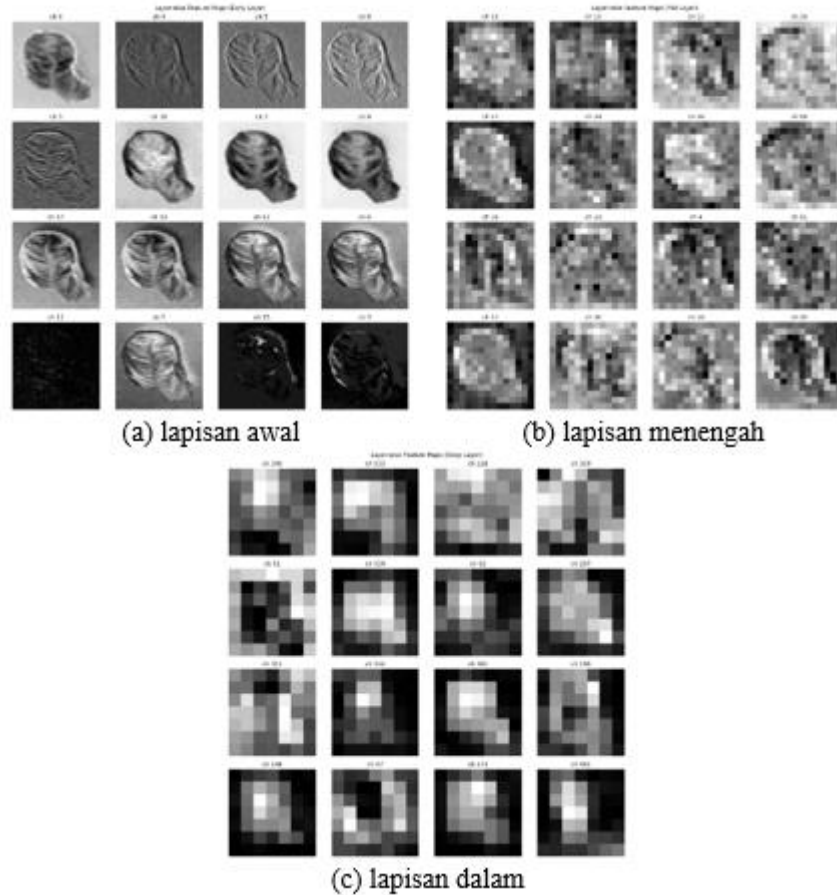
Gambar 3. Training and Validation Accuracy MobileNetV3-Small

Gambar 2 menunjukkan bahwa nilai training loss dan validation loss MobileNetV3-Small mengalami penurunan yang stabil seiring bertambahnya epoch, menandakan proses pembelajaran yang efektif. Kesenjangan loss yang kecil antara data latih dan validasi mengindikasikan bahwa model mampu melakukan generalisasi dengan baik tanpa mengalami *overfitting* yang signifikan. Berdasarkan Gambar 3, peningkatan *training accuracy* dan *validation accuracy* yang stabil menunjukkan bahwa MobileNetV3-Small mampu mempelajari pola visual penyakit daun tomat secara efektif tanpa mengalami penurunan kinerja pada data validasi. Konsistensi antara kedua kurva akurasi tersebut mengindikasikan bahwa model tidak mengalami *overfitting* yang signifikan dan memiliki kemampuan generalisasi yang baik. Untuk memahami bagaimana representasi fitur dibangun pada setiap tahap jaringan, analisis selanjutnya difokuskan pada visualisasi *layer-wise feature maps* yang ditampilkan pada Gambar 4.

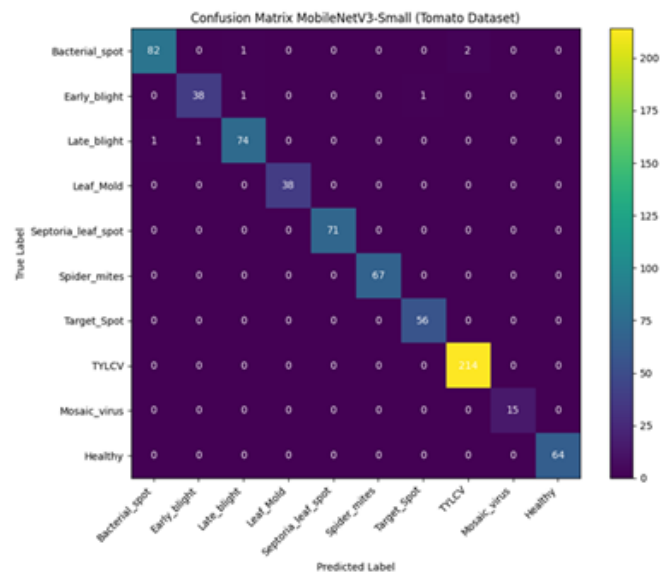
Visualisasi *layer-wise feature maps* pada Gambar 4 menunjukkan proses hierarkis MobileNetV3-Small dalam mengekstraksi informasi visual dari citra daun tomat. Pada lapisan awal (Gambar 4a), jaringan menangkap fitur-fitur dasar seperti tepi daun, tekstur permukaan, dan pola urat daun yang masih menyerupai bentuk asli citra. Selanjutnya, lapisan menengah (Gambar 4b) mulai merepresentasikan pola yang lebih kompleks dan bersifat lokal, seperti area bercak, perubahan tekstur, serta perbedaan intensitas yang berkaitan dengan karakteristik penyakit tertentu. Pada lapisan dalam (Gambar 4c), representasi fitur menjadi semakin abstrak dan terfokus pada area diskriminatif yang paling berkontribusi terhadap proses klasifikasi, meskipun detail spasial semakin berkurang. Pola ini menunjukkan bahwa MobileNetV3-Small mampu membangun representasi fitur yang relevan secara bertahap. Untuk mengevaluasi dampak representasi fitur tersebut terhadap kesalahan klasifikasi antar kelas, analisis selanjutnya difokuskan pada *confusion matrix* yang disajikan pada Gambar 5.

Gambar 5 menunjukkan *confusion matrix* MobileNetV3-Small yang memperlihatkan dominasi nilai pada diagonal utama, menandakan tingkat klasifikasi yang sangat baik pada sebagian besar kelas. Meski demikian, masih ditemukan kesalahan klasifikasi pada beberapa kelas dengan karakteristik visual yang mirip, seperti *Early Blight*, *Late Blight*, dan *Target Spot*. Hal ini mengindikasikan adanya tumpang tindih pola visual yang sulit dibedakan oleh model ringan. Untuk menganalisis area citra yang memengaruhi keputusan model, visualisasi Grad-CAM pada kasus salah klasifikasi disajikan pada Gambar 6.

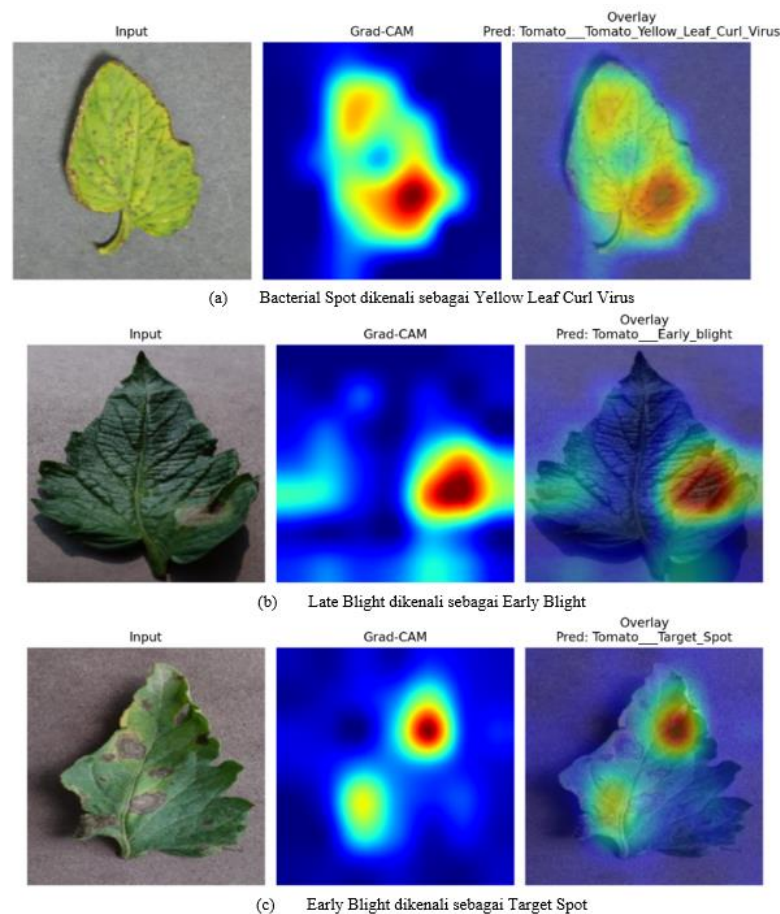
Gambar 6 memperlihatkan visualisasi Grad-CAM pada beberapa kasus salah klasifikasi yang dilakukan oleh MobileNetV3-Small. Pada kasus (a), area perhatian model terfokus pada bagian daun yang menguning dan mengalami perubahan tekstur sehingga menyerupai pola *Yellow Leaf Curl Virus*. Pada kasus (b), kemiripan distribusi bercak dan intensitas warna menyebabkan *Late Blight* terdeteksi sebagai *Early Blight*. Sementara itu, pada kasus (c), fokus aktivasi model berada pada area bercak melingkar yang menyerupai karakteristik *Target Spot*. Hasil ini menunjukkan bahwa kesalahan klasifikasi umumnya terjadi akibat kemiripan visual antar penyakit, sekaligus mengonfirmasi bahwa model telah mempelajari fitur yang relevan secara visual meskipun masih terbatas dalam membedakan pola penyakit yang sangat mirip.



Gambar 4. Visualisasi *layer-wise feature maps* pada MobileNetV3-Small



Gambar 5. Confusion Matrix MobileNetV3-Small



Gambar 6. Grad-CAM Misclassification MobileNetV3-Small

3.2 Hasil Klasifikasi Menggunakan MobileNet V2

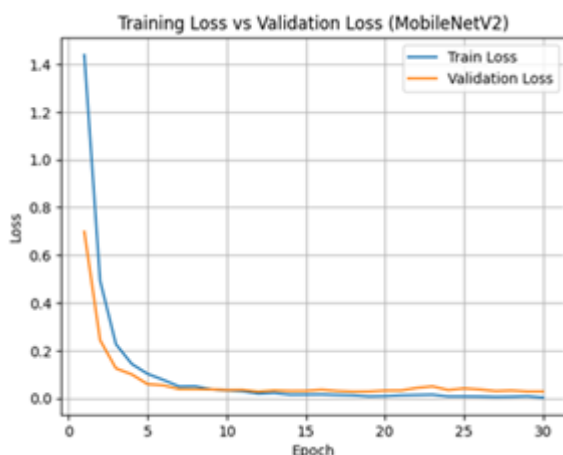
MobileNetV2 digunakan dalam penelitian ini sebagai representasi arsitektur lightweight CNN yang lebih matang dan stabil dibandingkan MobileNetV3-Small. Arsitektur ini mengandalkan konsep inverted residual block dan linear bottleneck yang dirancang untuk mempertahankan representasi fitur penting dengan jumlah parameter dan komputasi yang relatif rendah. Pemilihan MobileNetV2 menjadi penting untuk mengevaluasi sejauh mana arsitektur ringan generasi sebelumnya masih mampu memberikan kinerja klasifikasi yang sangat tinggi pada citra daun tomat, sekaligus menjadi pembanding langsung terhadap MobileNetV3-Small dan EfficientNet-B0. Hasil kinerja MobileNetV2 disajikan secara rinci melalui laporan klasifikasi pada Tabel 3.

Tabel 3. Laporan Klasifikasi MobileNet V2

Kelas Penyakit Tomat	Precision	Recall	F1-Score	Support
Tomato Bacterial_spot	1	1	1	85
Tomato Early_blight	1	1	1	40
Tomato Late_blight	1	0.974	0.9867	76
Tomato Leaf_Mold	1	1	1	38
Tomato Septoria_leaf_spot	1	1	1	71
Tomato Spider_mites Two-spotted_spider_mite	0.9706	0.985	0.9778	67
Tomato Target_Spot	0.9821	0.982	0.9821	56
Tomato Tomato_Yellow_Leaf_Curl_Virus	0.9953	1	0.9977	214
Tomato Tomato_mosaic_virus	1	1	1	15
Tomato healthy	1	1	1	64
Accuracy			0.9945	726

Macro Average	0.9948	0.994	0.9944	726
Weighted Average	0.9945	0.995	0.9945	726

Berdasarkan Tabel 3, MobileNetV2 menunjukkan kinerja klasifikasi yang sangat baik dengan akurasi keseluruhan sebesar 99,45%. Hampir seluruh kelas penyakit tomat memperoleh nilai precision, recall, dan F1-score mendekati atau sama dengan 1, yang menandakan kemampuan model dalam mengenali pola visual penyakit secara konsisten. Beberapa kelas seperti Late Blight dan Spider Mites masih menunjukkan sedikit penurunan pada nilai recall, namun tidak berdampak signifikan terhadap performa keseluruhan. Hasil ini mengindikasikan bahwa MobileNetV2 mampu menyeimbangkan efisiensi komputasi dan akurasi tinggi. Untuk memahami dinamika proses pembelajaran model selama pelatihan, Gambar 7 menyajikan kurva error pada data pelatihan dan validasi.



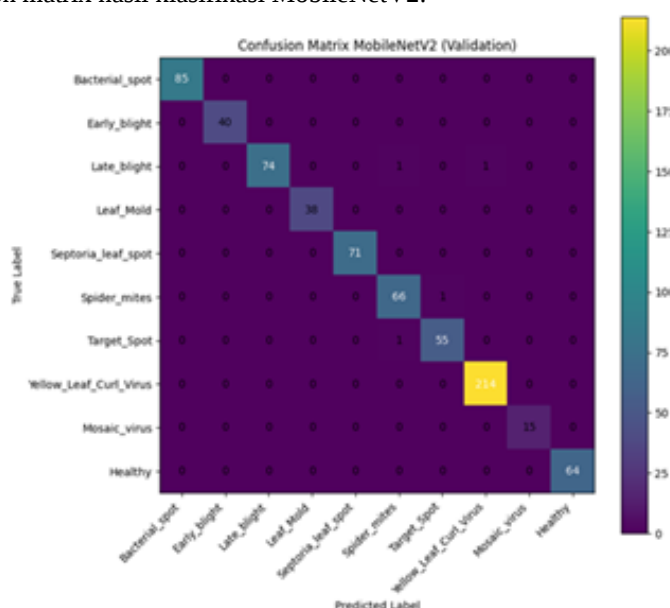
Gambar 7. Error pada Training dan Valisasi MobileNetV2



Gambar 8. Akurasi Pelatihan dan Validasi MobileNetV2

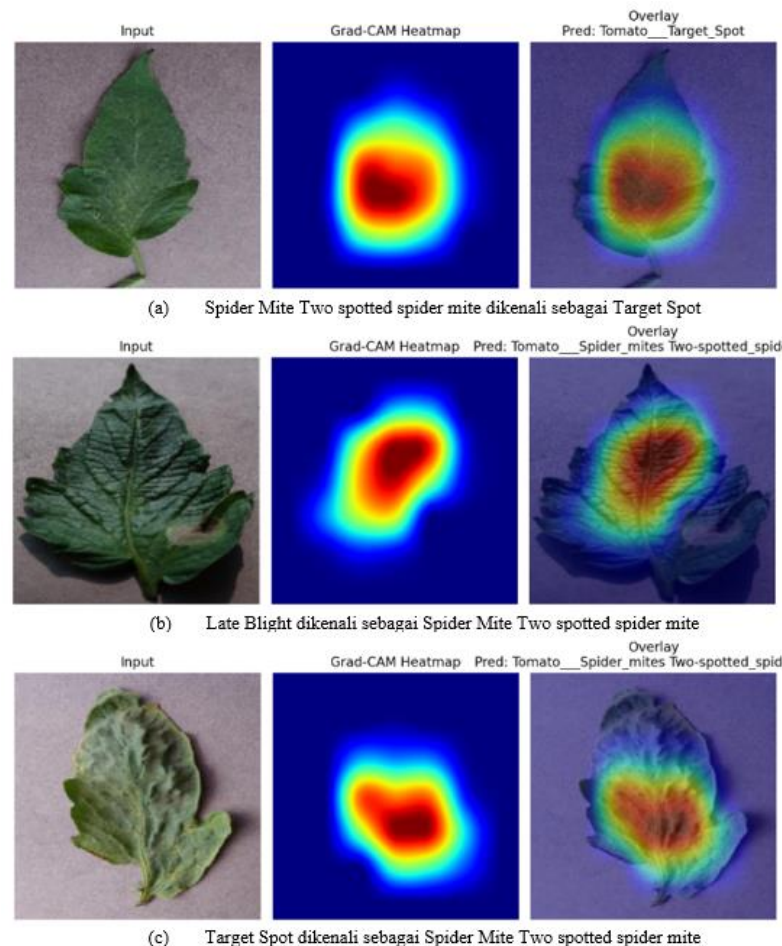
Gambar 7 memperlihatkan bahwa error pelatihan dan validasi MobileNetV2 menurun secara stabil seiring bertambahnya epoch, tanpa fluktuasi yang signifikan. Pola ini menunjukkan proses pembelajaran yang konvergen dan minim indikasi *overfitting*. Untuk melengkapi analisis tersebut, Gambar 8 menyajikan kurva akurasi pelatihan dan validasi guna menunjukkan konsistensi performa model selama proses training.

Pada Gambar 8 terlihat bahwa akurasi pelatihan dan validasi MobileNetV2 meningkat secara cepat dan cenderung stabil pada nilai yang sangat tinggi. Kesenjangan yang kecil antara kedua kurva menunjukkan kemampuan generalisasi model yang baik. Untuk melihat distribusi prediksi model secara lebih detail pada setiap kelas, Gambar 9 menampilkan confusion matrix hasil klasifikasi MobileNetV2.



Gambar 9. Confusion Matrik MobileNetV2

Confusion matrix pada Gambar 9 menunjukkan bahwa sebagian besar sampel diklasifikasikan dengan benar ke kelas yang sesuai. Kesalahan klasifikasi hanya terjadi pada sejumlah kecil sampel dan melibatkan kelas-kelas penyakit yang memiliki kemiripan visual tinggi. Untuk menganalisis penyebab kesalahan tersebut secara lebih mendalam dan interpretatif, Gambar 10 menyajikan visualisasi Grad-CAM pada beberapa kasus salah klasifikasi MobileNetV2.



Gambar 10. Grad-CAM Misclassification MobileNetV2

Gambar 10 memperlihatkan visualisasi Grad-CAM pada beberapa kasus misklasifikasi MobileNetV2. Pada kasus (a), citra Spider Mites diprediksi sebagai Target Spot, dan Grad-CAM menunjukkan perhatian model terkonsentrasi pada area daun yang relatif homogen tanpa penekanan yang jelas pada tekstur khas tungau, sehingga pola aktivasi menyerupai area bercak. Pada kasus (b), Late Blight terdeteksi sebagai Spider Mites, di mana aktivasi model kuat pada satu area dominan yang tampak seperti perubahan tekstur/warna lokal, sehingga model keliru menafsirkan pola kerusakan jaringan sebagai ciri infestasi. Pada kasus (c), Target Spot diprediksi sebagai Spider Mites, dan aktivasi Grad-CAM kembali terfokus pada area bercak utama, menandakan bahwa model cenderung mengandalkan intensitas dan kontras bercak tanpa membedakan bentuk bercak melingkar (Target Spot) versus pola kerusakan halus yang sering dikaitkan dengan Spider Mites.

3.3 Hasil Klasifikasi Menggunakan EfficientNet B0

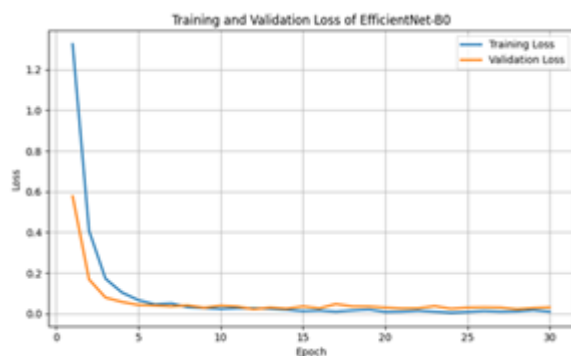
Arsitektur EfficientNet-B0 digunakan dalam penelitian ini karena dikenal memiliki keseimbangan yang optimal antara akurasi dan kompleksitas model melalui pendekatan compound scaling yang mengatur kedalaman, lebar, dan resolusi jaringan secara proporsional. Dibandingkan lightweight CNN sebelumnya, EfficientNet-B0 mampu mengekstraksi fitur visual yang lebih representatif tanpa meningkatkan jumlah parameter secara signifikan. Oleh karena itu, model ini dipilih untuk mengevaluasi sejauh mana peningkatan kapasitas representasi fitur berdampak

terhadap performa klasifikasi penyakit daun tomat. Hasil evaluasi kinerja klasifikasi EfficientNet-B0 disajikan pada Tabel 3.

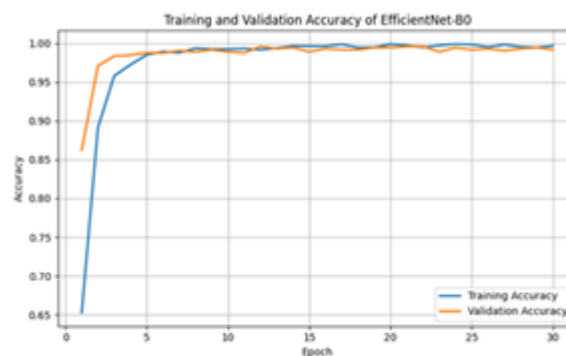
Tabel 4. Laporan Klasifikasi EfficientNet B0

Kelas Penyakit Daun Tomat	Precision	Recall	F1-Score	Support
Tomato Bacterial_spot	1	1	1	85
Tomato Early_blight	0.975	0.975	0.975	40
Tomato Late_blight	1	0.987	0.9934	76
Tomato Leaf_Mold	0.9744	1	0.987	38
Tomato Septoria_leaf_spot	0.9859	0.986	0.9859	71
Tomato Spider_mites Two-spotted_spider_mite	1	1	1	67
Tomato Target_Spot	1	1	1	56
Tomato Tomato_Yellow_Leaf_Curl_Virus	1	1	1	214
Tomato Tomato_mosaic_virus	1	1	1	15
Tomato healthy	1	1	1	64
Accuracy			0.9959	726
Macro Average	0.9935	0.995	0.9941	726
Weighted Average	0.9959	0.996	0.9959	726

Berdasarkan Tabel 4, EfficientNet-B0 menunjukkan performa klasifikasi yang sangat tinggi dengan akurasi keseluruhan mencapai 99,59%. Sebagian besar kelas memperoleh nilai precision, recall, dan F1-score sempurna, menandakan kemampuan model dalam membedakan pola penyakit secara konsisten. Kesalahan klasifikasi hanya muncul pada beberapa kelas dengan kemiripan visual tinggi seperti Early Blight, Septoria Leaf Spot, dan Leaf Mold. Tingginya performa ini disebabkan oleh kemampuan EfficientNet dalam mengekstraksi fitur multi-skala secara lebih efisien dibandingkan arsitektur lightweight sebelumnya. Dampaknya, model menjadi lebih stabil dalam mengenali tekstur bercak, pola nekrosis, serta distribusi kerusakan daun. Untuk melihat stabilitas proses pelatihan, kurva error training dan validasi ditampilkan pada Gambar 10.



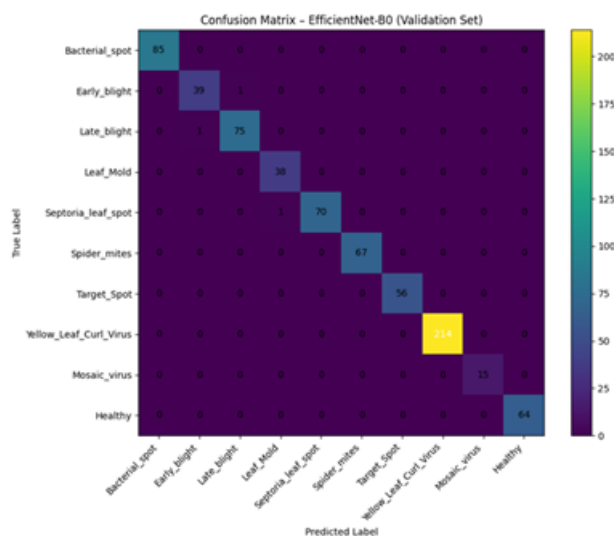
Gambar 11. Error pelatihan dan validasi EfficienNet B0



Gambar 12. Akurasi Training dan Validasi EfficientNet B0

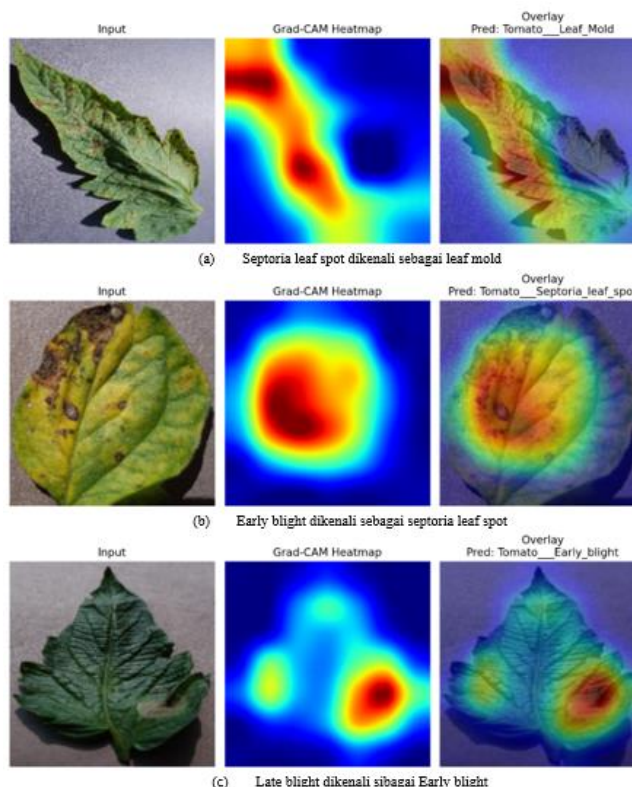
Gambar 11 menunjukkan bahwa nilai error training dan validasi menurun secara konsisten seiring bertambahnya epoch, menandakan proses pembelajaran berlangsung optimal tanpa gejala overfitting yang signifikan. Stabilitas konvergensi ini mengindikasikan bahwa arsitektur EfficientNet-B0 mampu melakukan generalisasi dengan baik. Selanjutnya, performa pembelajaran ditinjau melalui kurva akurasi pada Gambar 12.

Pada Gambar 12 terlihat bahwa akurasi training dan validasi meningkat cepat sejak epoch awal dan mencapai kondisi stabil mendekati 100%. Pola ini menunjukkan bahwa model mampu mempelajari representasi fitur penyakit secara efektif. Untuk melihat distribusi prediksi benar dan salah secara detail, disajikan confusion matrix pada Gambar 13.



Gambar 13. Confusion matrix EfficientNet B0

Confusion matrix pada Gambar 13 memperlihatkan dominasi nilai diagonal yang tinggi, menandakan sebagian besar citra berhasil diklasifikasikan dengan benar. Kesalahan prediksi hanya terjadi pada beberapa pasangan kelas yang memiliki kemiripan gejala visual. Untuk memahami penyebab kesalahan tersebut, dilakukan analisis visual menggunakan Grad-CAM yang ditampilkan pada Gambar 14.



Gambar 14. Grad-CAM Misclassification EfficientNet B0

Gambar 14 memperlihatkan beberapa kasus misklasifikasi pada EfficientNet-B0 berdasarkan visualisasi Grad-CAM. Pada kasus (a), Septoria Leaf Spot dikenali sebagai Leaf Mold karena model berfokus pada area klorosis memanjang di tepi daun yang secara visual mirip infeksi jamur. Pada kasus (b), Early Blight diklasifikasikan sebagai Septoria Leaf Spot akibat kemiripan bercak nekrotik melingkar. Sementara pada kasus (c), Late Blight

dikenali sebagai Early Blight karena aktivasi model terpusat pada area nekrosis yang memiliki pola serupa. Hasil ini menunjukkan bahwa misklasifikasi terutama disebabkan oleh kemiripan morfologi bercak antar penyakit, meskipun model tetap mampu mengidentifikasi area lesi yang relevan.

3.4 Pembahasan

Pada bagian ini dilakukan analisis komparatif terhadap tiga arsitektur lightweight CNN yang diuji, yaitu MobileNetV3-Small, MobileNetV2, dan EfficientNet-B0. Perbandingan tidak hanya difokuskan pada tingkat akurasi, tetapi juga mempertimbangkan kompleksitas model yang direpresentasikan melalui jumlah parameter dan kebutuhan komputasi (FLOPs). Evaluasi multi-metrik seperti macro precision, recall, dan F1-score digunakan untuk memberikan gambaran performa klasifikasi yang lebih menyeluruh, khususnya dalam membedakan kemampuan model mengenali setiap kelas penyakit daun tomat. Ringkasan hasil perbandingan kinerja ketiga model disajikan pada Tabel 5.

Tabel 5. Perbandingan Metode Lightweight CNN

Model CNN	Jumlah Parameter	FLOPs ($\approx$)	Akurasi Validasi	Macro Precision	Macro Recall	Macro F1-score
MobileNetV3-Small	2,54 M	66 M	99,04%	99,09%	98,88%	98,98%
MobileNetV2	3,50 M	300 M	99,45%	99,48%	99,41%	99,44%
EfficientNet-B0	4,02 M	827 M	99,59%	99,35%	99,48%	99,41%

Berdasarkan Tabel 5, seluruh model lightweight CNN menunjukkan performa klasifikasi yang sangat tinggi dengan akurasi di atas 99%. EfficientNet-B0 memperoleh akurasi tertinggi sebesar 99,59% dengan nilai macro recall 99,48%, menunjukkan kemampuan generalisasi paling baik dalam mengenali seluruh kelas penyakit. Hal ini dipengaruhi oleh mekanisme compound scaling yang meningkatkan kualitas representasi fitur. Namun demikian, model ini memiliki kompleksitas terbesar dengan 4,02 juta parameter dan kebutuhan komputasi ± 827 M FLOPs sehingga kurang efisien untuk perangkat terbatas.

MobileNetV2 menempati posisi kompromi terbaik dengan akurasi 99,45% serta kompleksitas menengah (3,50 juta parameter; 300M FLOPs). Arsitektur inverted residual dan linear bottleneck memungkinkan ekstraksi fitur yang efektif dengan efisiensi komputasi yang masih terjaga. Sementara itu, MobileNetV3-Small merupakan model paling ringan (2,54 juta parameter; 66M FLOPs) dengan akurasi 99,04%, menjadikannya sangat potensial untuk implementasi real-time meskipun sedikit di bawah dua model lainnya.

Analisis Grad-CAM memperkuat hasil kuantitatif tersebut. EfficientNet-B0 menghasilkan peta aktivasi paling terfokus pada area lesi, diikuti MobileNetV2 yang cukup presisi, sedangkan MobileNetV3-Small cenderung memiliki aktivasi lebih menyebar pada kasus tertentu. Misklasifikasi yang terjadi umumnya disebabkan kemiripan morfologi bercak penyakit, bukan karena kegagalan model dalam melokalisasi area daun.

3. KESIMPULAN

Penelitian ini bertujuan membandingkan kinerja tiga arsitektur lightweight Convolutional Neural Network (CNN), yaitu MobileNetV3-Small, MobileNetV2, dan EfficientNet-B0, dalam mengklasifikasikan penyakit daun tomat pada dataset PlantVillage. Evaluasi dilakukan tidak hanya berdasarkan akurasi klasifikasi, tetapi juga mempertimbangkan kompleksitas model serta interpretabilitas visual menggunakan Grad-CAM dan layer-wise feature visualization. Hasil eksperimen menunjukkan bahwa seluruh model mampu menghasilkan performa klasifikasi yang sangat tinggi. EfficientNet-B0 memperoleh akurasi validasi terbaik sebesar 99,59%, diikuti MobileNetV2 sebesar 99,45%, dan MobileNetV3-Small sebesar 99,04%. Meskipun demikian, perbedaan akurasi antar model relatif kecil, sehingga aspek efisiensi komputasi menjadi faktor pembeda penting. MobileNetV3-Small memiliki jumlah parameter dan FLOPs paling rendah, menjadikannya model paling ringan, sedangkan EfficientNet-B0 memiliki kompleksitas tertinggi namun memberikan performa paling optimal. Analisis confusion matrix menunjukkan bahwa kesalahan klasifikasi yang terjadi umumnya berada pada kelas penyakit dengan karakteristik visual serupa, seperti Early Blight, Late Blight, dan Septoria Leaf Spot. Temuan ini diperkuat melalui visualisasi Grad-CAM yang memperlihatkan bahwa model memfokuskan perhatian pada area lesi daun, meskipun pada beberapa kasus terjadi tumpang tindih fokus fitur antar penyakit yang mirip secara morfologis.

Secara keseluruhan, penelitian ini menegaskan bahwa pemilihan arsitektur CNN ringan tidak hanya bergantung pada akurasi, tetapi juga harus mempertimbangkan efisiensi parameter, kebutuhan komputasi, dan transparansi keputusan model. Kontribusi utama penelitian terletak pada penyajian analisis komparatif yang terintegrasi antara kinerja klasifikasi, kompleksitas model, dan interpretabilitas visual, sehingga dapat menjadi referensi dalam pengembangan sistem deteksi penyakit tanaman berbasis kecerdasan buatan yang efisien dan andal.

REFERENSI

- [1] S. Ahmed, M. B. Hasan, T. Ahmed, M. R. K. Sony, and M. H. Kabir, "Less is more: Lighter and faster deep neural architecture for tomato leaf disease classification," *IEEE Access*, vol. 10, pp. 68868–68884, 2022.
- [2] A. Islam, T. Tahsin, Z. Anjum, M. B. Hasan, and M. H. Kabir, "A Domain-Adapted Lightweight Ensemble for Resource-Efficient Few-Shot Plant Disease Classification," *arXiv Prepr. arXiv2512.13428*, 2025.
- [3] S. Thuseethan, P. Vigneshwaran, J. Charles, and C. Wimalasooriya, "Siamese network-based lightweight framework for tomato leaf disease recognition," *Computers*, vol. 13, no. 12, p. 323, 2024.
- [4] S. Rahman, M. Elvany, and M. Ramli, "MobileChiliNet: convolutional neural network for chili leaves classification," *Iaes Int. J. Artif. Intell.*, vol. 14, no. 5, pp. 3757–3770, 2025.
- [5] S. Rahman, A. Indrawati, and M. Zen, "Enhanced RegNetY-400MF for Fruit Fly Species Classification: Fine-Tuning Strategies and Data Balancing for Improved Accuracy," vol. 9, no. 6, pp. 1463–1473, 2025, doi: <https://doi.org/10.29207/resti.v9i6.6973>.
- [6] S. Rahman, R. A. Setyadi, A. Indrawati, A. Sembiring, and M. Zen, "Improving the Accuracy of Chili Leaf Disease Classification with ResNet and Fine-Tuning Strategy," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 10, 2024.
- [7] A. Gao, A. Geng, Y. Song, L. Ren, Y. Zhang, and X. Han, "Detection of maize leaf diseases using improved MobileNet V3-small," *Int. J. Agric. Biol. Eng.*, vol. 16, no. 3, pp. 225–232, 2023.
- [8] R. Zehra, O. Sharma, and R. B. Singh, "Mobilenet-v3: A comprehensive survey of object detection algorithms using CNN," *Emerg. Trends Comput. Sci. Its Appl.*, pp. 362–366, 2025.
- [9] Y. Gulzar, "Fruit image classification model based on MobileNetV2 with deep transfer learning technique," *Sustainability*, vol. 15, no. 3, p. 1906, 2023.
- [10] E. Elfatimi, R. Eryigit, and L. Elfatimi, "Beans leaf diseases classification using mobilenet models," *IEEE Access*, vol. 10, pp. 9471–9482, 2022.
- [11] X. Shi, P. Li, H. Wu, Q. Chen, and H. Zhu, "A lightweight image splicing tampering localization method based on MobileNetV2 and SRM," *IET Image Process.*, vol. 17, no. 6, pp. 1883–1892, 2023.
- [12] J. Chen, D. Zhang, M. Suzaiddola, and A. Zeb, "Identifying crop diseases using attention embedded MobileNet-V2 model," *Appl. Soft Comput.*, vol. 113, p. 107901, 2021.
- [13] M. Hossain and M. A. Rahman, "Brain tumor detection using deep network EfficientNet-B0," in *International Conference on Machine Intelligence and Emerging Technologies*, Springer, 2022, pp. 213–225.
- [14] N. V. Hiremath, "Breast cancer detection and classification using EfficientNet B0 and EfficientNet B0-HSV," 2023, *Dublin, National College of Ireland*.
- [15] H. Ali, N. Shifa, R. Benlamri, A. A. Farooque, and R. Yaqub, "A fine tuned EfficientNet-B0 convolutional neural network for accurate and efficient classification of apple leaf diseases," *Sci. Rep.*, vol. 15, no. 1, p. 25732, 2025.
- [16] A. M. Abdu, M. M. Mokji, and U. U. Sheikh, "Automatic vegetable disease identification approach using individual lesion features," *Comput. Electron. Agric.*, vol. 176, p. 105660, 2020, doi: <https://doi.org/10.1016/j.compag.2020.105660>.
- [17] J. G. A. Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification," *Comput. Electron. Agric.*, vol. 153, pp. 46–53, 2018, doi: <https://doi.org/10.1016/j.compag.2018.08.013>.
- [18] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336–359, 2020, doi: [10.1007/s11263-019-01228-7](https://doi.org/10.1007/s11263-019-01228-7).
- [19] F. Zhu, Y. Sun, Y. Zhang, W. Zhang, and J. Qi, "An improved MobileNetV3 mushroom quality classification model using images with complex backgrounds," *Agronomy*, vol. 13, no. 12, p. 2924, 2023.
- [20] L. Chen *et al.*, "Evaluation of MobileNetV3-Large for crack classification across low-and high-resolution images," *Front. Built Environ.*, vol. 11, p. 1724879, 2025.
- [21] L. Zhao and L. Wang, "A new lightweight network based on MobileNetV3," *KSII Trans. Internet Inf.*



- Syst., vol. 16, no. 1, pp. 1–15, 2022.
- [22] R. Indraswari, R. Rokhana, and W. Herulambang, “Melanoma image classification based on MobileNetV2 network,” *Procedia Comput. Sci.*, vol. 197, pp. 198–207, 2022.
- [23] Q. Zhu, H. Zhuang, M. Zhao, S. Xu, and R. Meng, “A study on expression recognition based on improved mobilenetV2 network,” *Sci. Rep.*, vol. 14, no. 1, p. 8121, 2024.
- [24] Y. Gulzar, “Fruit image classification model based on MobileNetV2 with deep transfer learning technique,” *Sustainability*, vol. 15, no. 3, p. 1906, 2023.
- [25] H. Guo, J. Zhang, Y. Li, X. Pan, and C. Sun, “Advanced pathological subtype classification of thyroid cancer using efficientNetB0,” *Diagn. Pathol.*, vol. 20, no. 1, p. 28, 2025.
- [26] A. Kumar, L. Nelson, and D. Arumugam, “Deep Learning-Based Classification of Brain Tumours on MRI Images Using EfficientNetB0,” in *2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS)*, IEEE, 2024, pp. 219–225.
- [27] K. Gencer, “A Comparative Analysis of EfficientNetB0 and EfficientNetV2 Variants for Brain Tumor Classification Using MRI Images,” *Int. J. Innov. Eng. Appl.*, vol. 9, no. 1, pp. 1–7, 2025.
- [28] S. H. Kim, J. S. Park, H. S. Lee, S. H. Yoo, and K. J. Oh, “Combining CNN and Grad-CAM for profitability and explainability of investment strategy: Application to the KOSPI 200 futures,” *Expert Syst. Appl.*, vol. 225, p. 120086, 2023.
- [29] H. Moujahid *et al.*, “Combining CNN and Grad-Cam for COVID-19 Disease Prediction and Visual Explanation,” *Intell. Autom. Soft Comput.*, vol. 32, no. 2, 2022.
- [30] W. Siripattanadilok and T. Siriborvornratanakul, “Recognition of partially occluded soft-shell mud crabs using Faster R-CNN and Grad-CAM,” *Aquac. Int.*, vol. 32, no. 3, pp. 2977–2997, 2024.