

Comparison of Transfer learning Models MobileNetV3-Large and EfficientNet-B0 for Rice Leaf Disease Classification

Ahmad Naufal Abiyyu ^{1*}, Majid Rahardi ^{2*}

* Informatika, Universitas Amikom Yogyakarta

naufalabiyyu@students.amikom.ac.id ¹, majid@amikom.ac.id ²

Article Info

Article history:

Received 2025-12-15

Revised 2026-01-15

Accepted 2026-01-20

Keyword:

*EfficientNet-B0,
Leaf disease,
MobileNetV3-Large,
Rice,
Transfer learning.*

ABSTRACT

Rice productivity strongly depends on early detection of leaf diseases, while manual identification is often delayed and subjective. This study investigates the use of lightweight CNN architectures MobileNetV3-Large and EfficientNet-B0 based on transfer learning to classify six rice leaf disease classes, namely bacterial leaf blight, brown spot, healthy, leaf blast, leaf scald, and narrow brown spot. The dataset is obtained from Kaggle and consists of 2,628 images with a balanced class distribution, stratified into training, validation, and test sets with a ratio of 80%:10%:10%. The images are resized to 224×224 pixels and data augmentation was applied to the training set. Pretrained ImageNet weights are first used as frozen feature extractors, followed by partial fine-tuning of the last 30% backbone layers, with custom classification layers trained using the Adam optimizer with an early stopping mechanism. Model performance is evaluated using accuracy, precision, recall, F1-score, and confusion matrices, while computational efficiency is assessed based on parameter count and inference speed measured in frames per second. The results show that under partial fine-tuning MobileNetV3-Large achieves 95.83% test accuracy and 95.80% macro F1-score with 3.12 million parameters, while EfficientNet-B0 obtains 93.18% accuracy and 93.02% macro F1-score with 4.21 million parameters. Both models achieve inference speeds above 50 frames per second, suggesting their potential suitability for deployment on resource-constrained devices. Bootstrap analysis suggests the performance gap is clear in the frozen stage but becomes less conclusive after partial fine-tuning. Overall, MobileNetV3-Large provides the best trade-off between accuracy and efficiency for rice leaf disease classification.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Padi merupakan komoditas pangan strategis di Indonesia karena beras menjadi makanan pokok masyarakat, sekaligus sumber kalori utama bagi sebagian besar penduduk di negara-negara Asia [1], [2]. Oleh karena itu, kestabilan produksi padi berperan penting dalam menjaga ketahanan pangan nasional. Berdasarkan data Badan Pusat Statistik (BPS), luas lahan panen padi di Indonesia pada tahun 2024 mencapai sekitar 10,05 juta hektare dengan produksi sekitar 53,14 juta ton gabah kering giling (GKG), yang masing-masing tercatat menurun sekitar 1,6 persen untuk luas lahan panen dan 1,5 persen untuk produksi dibandingkan tahun 2023. Jika dikonversi menjadi beras, produksi beras tahun 2024 sekitar

30,62 juta ton, turun sekitar 1,54 persen dari tahun sebelumnya [3]. Penurunan produksi ini terjadi di tengah pertumbuhan jumlah penduduk dan meningkatnya kebutuhan konsumsi, sehingga diperlukan upaya peningkatan produktivitas yang lebih efektif dan berkelanjutan [4]. Selain faktor makro seperti perubahan iklim dan berkurangnya lahan, gangguan organisme pengganggu tanaman, khususnya hama dan penyakit padi, juga berkontribusi terhadap penurunan hasil panen [5], [6]. Beberapa penyakit penting pada padi, seperti *bacterial leaf blight*, *brown spot*, *blast*, *scald*, dan *narrow brown spot*, dilaporkan sebagai ancaman serius di berbagai wilayah penghasil padi, termasuk Indonesia [7]. Namun, identifikasi dini penyakit di lapangan masih sering mengandalkan inspeksi visual yang subjektif dan

membutuhkan keahlian, sehingga berpotensi menghambat pengendalian yang cepat dan tepat [8].

Seiring dengan berkembangnya teknologi dalam mendeteksi penyakit tanaman secara dini, Convolutional Neural Network (CNN) menjadi pendekatan utama dalam pengolahan citra digital karena kemampuannya mengekstraksi fitur visual kompleks secara otomatis. CNN konvensional seperti VGG16 dan ResNet50 terbukti akurat, tetapi memerlukan daya komputasi besar sehingga tidak praktis untuk perangkat lapangan atau mobile [9]. Oleh karena itu, dikembangkan arsitektur ringan (*lightweight CNN*) yang lebih efisien secara komputasi namun tetap kompetitif dalam hal akurasi. Di antara model tersebut, MobileNetV3-Large dan EfficientNet-B0 termasuk yang paling populer karena memiliki parameter yang jauh lebih kecil, waktu inferensi cepat, dan ukuran model yang relatif ringan [10], [11]. Kedua model ini dirancang agar dapat dijalankan pada perangkat terbatas seperti ponsel atau edge device, yang menjadikannya relevan untuk implementasi di sektor pertanian dengan infrastruktur komputasi terbatas.

Sejumlah studi sebelumnya telah menerapkan arsitektur ini untuk klasifikasi penyakit tanaman. Studi oleh Pramudhita et al. [12] menggunakan MobileNetV3-Large dan EfficientNet-B0 untuk mengklasifikasikan penyakit daun stroberi, dan menemukan bahwa MobileNetV3-Large memberikan akurasi sebesar 92,14%, sedangkan EfficientNet-B0 mencapai 89,28%, menjadikan MobileNetV3-Large unggul dalam seluruh metrik evaluasi. Sementara itu, Dianto et al. [13] menerapkan kedua model tersebut pada klasifikasi penyakit daun jeruk dan juga melaporkan akurasi tinggi (>90%) untuk keduanya, terutama pada kondisi data terbatas dan distribusi kelas yang tidak seimbang. Dalam konteks tanaman padi, Putra et al. [6] menggunakan MobileNetV2 untuk mendeteksi penyakit dan hama daun padi, sedangkan Azis et al. [14] menerapkan EfficientNet-B3 dengan pendekatan *transfer learning*. Studi-studi tersebut menunjukkan bahwa arsitektur CNN ringan telah banyak dimanfaatkan dalam klasifikasi penyakit tanaman secara umum, dengan hasil yang kompetitif dalam berbagai kasus dan jenis komoditas.

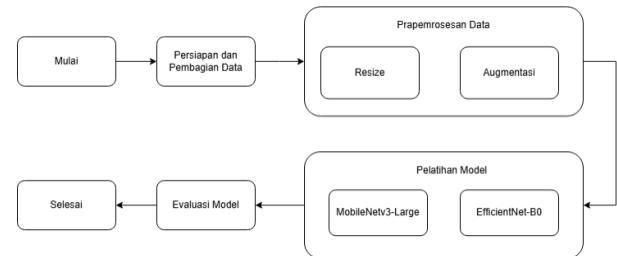
Berdasarkan temuan-temuan tersebut, penelitian ini difokuskan pada eksplorasi dan perbandingan dua arsitektur CNN ringan, yaitu MobileNetV3-Large dan EfficientNet-B0, dalam klasifikasi penyakit daun padi. Hingga saat ini, studi yang membahas perbandingan langsung antara kedua model ini dalam konteks tanaman padi masih sangat terbatas. Di sisi lain, sebagian besar penelitian sebelumnya belum memperhatikan distribusi data yang seimbang antar kelas, yang dapat memicu bias klasifikasi dan menurunkan keadilan evaluasi performa model [15]. Oleh karena itu, penelitian ini menekankan penggunaan dataset seimbang dengan enam kelas penyakit daun padi sebagai bagian dari kontribusi utama, guna memastikan bahwa model yang dibangun tidak hanya akurat, tetapi juga adil dan andal dalam mengenali seluruh kategori penyakit.

Penelitian ini bertujuan untuk membandingkan kinerja dua arsitektur *lightweight CNN*, yaitu MobileNetV3-Large dan

EfficientNet-B0, pada tugas klasifikasi enam kelas penyakit daun padi berbasis citra. Evaluasi kinerja dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score, serta aspek efisiensi komputasi yang diukur melalui jumlah parameter, ukuran berkas model, dan kecepatan inferensi (*frame per second/fps*). Selain itu, perbedaan kinerja antar model dianalisis menggunakan pendekatan statistik berbasis *confidence interval* (CI) untuk menilai konsistensi selisih performa pada data uji. Untuk menilai kemampuan generalisasi, model terbaik juga diuji pada dataset eksternal yang berbeda dari data pelatihan dengan subset tiga kelas yang tersedia. Dengan pendekatan tersebut, penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan sistem deteksi penyakit tanaman berbasis citra yang ringan, akurat, dan siap diterapkan di lingkungan pertanian dengan keterbatasan perangkat komputasi.

II. METODE

Metode penelitian ini terdiri dari beberapa tahap utama yang disusun secara berurutan, meliputi persiapan dan pembagian data, prapemrosesan citra, perancangan serta pelatihan model MobileNetV3-Large dan EfficientNet-B0 berbasis *transfer learning*, serta evaluasi performa model. Setiap tahap saling berkaitan untuk menghasilkan sistem klasifikasi penyakit daun padi yang teruji. Alur lengkap tahapan penelitian ditunjukkan pada gambar 1.



Gambar 1. Alur Tahapan Penelitian

A. Persiapan dan Pembagian Data

Dataset yang digunakan dalam penelitian ini adalah *Rice Leaf Disease Dataset* yang diperoleh dari platform publik Kaggle dan berisi citra daun padi sebanyak 2.628 dengan enam kelas didalamnya, yaitu *bacterial leaf blight*, *brown spot*, *healthy*, *leaf blast*, *leaf scald*, dan *narrow brown spot* [16]. Dataset ini disusun dalam struktur direktori yang telah memisahkan data pelatihan dan validasi, sehingga memudahkan proses pengaturan ulang data sesuai kebutuhan eksperimen. Contoh citra daun padi untuk masing-masing kelas ditunjukkan pada Gambar 2.



Gambar 2. Contoh Citra Per kelas

Dalam penelitian ini, dataset diorganisasi ulang menjadi tiga subset, yaitu data latih (*training set*), data validasi (*validation set*), dan data uji (*test set*) dengan rasio 80% : 10% : 10%. Seluruh citra pada direktori pelatihan awal dimanfaatkan sebagai data latih, sedangkan citra pada direktori validasi awal dibagi kembali secara *stratified* per kelas menjadi dua bagian yang sama besar untuk data validasi dan data uji. Pendekatan ini memastikan bahwa distribusi kelas tetap seimbang pada ketiga subset dan tidak terjadi kebocoran data antara tahap pelatihan dan pengujian.

Ringkasan distribusi jumlah citra pada setiap kelas setelah proses pembagian data disajikan pada Tabel I, yang menunjukkan bahwa keenam kelas memiliki jumlah sampel yang sama pada data latih, data validasi, maupun data uji.

TABEL I
DATA SPLITTING

Class	Training	Validation	Test
Bacterial leaf blight	350	44	44
Brown spot	350	44	44
Healthy	350	44	44
Leaf blast	350	44	44
Leaf scald	350	44	44
Narrow brown spot	350	44	44
Total	2100	264	264

B. Prapemrosesan Citra

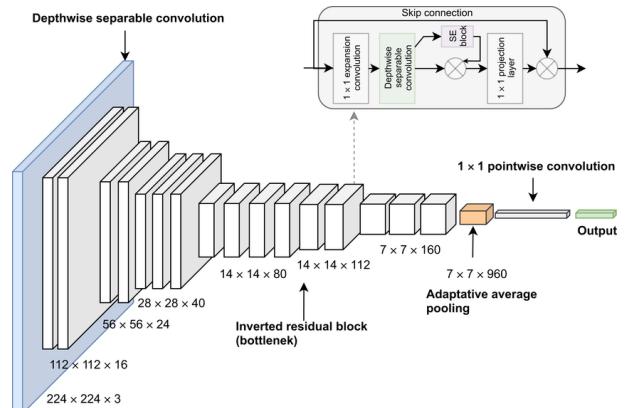
Pada tahap prapemrosesan, seluruh citra daun padi diubah ukurannya menjadi 224×224 piksel dengan tiga kanal warna (RGB). Untuk meningkatkan kemampuan generalisasi model, data latih kemudian diberikan augmentasi berupa rotasi acak dengan faktor 0,05 (sekitar $\pm 18^\circ$), *horizontal flip*, *random zoom* 0,1, penambahan *Gaussian noise* 0,1, serta penyesuaian kecerahan 0,1. Sementara itu, data validasi dan data uji hanya dikenai proses resize tanpa augmentasi tambahan.

C. Perancangan Model Berbasis Transfer learning

Penelitian ini memanfaatkan pendekatan *transfer learning* dengan dua arsitektur Convolutional Neural Network (CNN)

ringan, yaitu MobileNetV3-Large dan EfficientNet-B0, yang telah dilatih sebelumnya (pretrained) pada dataset ImageNet [17]. Proses pelatihan dilakukan dalam dua tahap, yaitu tahap *frozen* untuk melatih kepala klasifikasi pada fitur pralatih, serta tahap *partial fine-tuning* untuk meningkatkan adaptasi model terhadap karakteristik citra daun padi.

1) *MobileNetV3-Large*: MobileNetV3-Large merupakan pengembangan dari keluarga MobileNet yang dirancang untuk perangkat dengan sumber daya komputasi terbatas [10]. Seperti terlihat pada Gambar 3, arsitektur ini diawali dengan lapisan konvolusi awal yang diikuti depthwise separable convolution dan deretan blok inverted residual dengan linear bottleneck. Di dalam setiap blok juga disisipkan mekanisme *squeeze-and-excitation* (SE) untuk menonjolkan kanal fitur yang penting, serta skip connection yang menjaga aliran gradien selama pelatihan [18]. Pada bagian akhir jaringan terdapat global average pooling dan konvolusi 1×1 (pointwise convolution) sebelum menuju kepala klasifikasi.

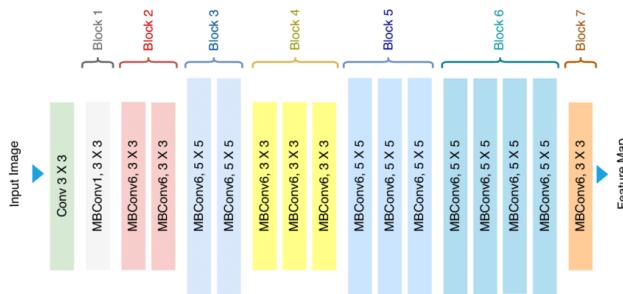


Gambar 3. Arsitektur MobileNetV3-Large

Dalam penelitian ini, bobot pralatih MobileNetV3-Large pada ImageNet dimanfaatkan sebagai dasar representasi fitur. Lapisan klasifikasi asli digantikan dengan rangkaian lapisan baru yang terdiri atas Global Average Pooling 2D, satu lapisan Dense dengan 128 neuron dan fungsi aktivasi ReLU, satu lapisan Dropout dengan rasio 0,5 sebagai regularisasi, dan lapisan Dense dengan 6 neuron dan fungsi aktivasi softmax untuk menghasilkan probabilitas tiap kelas penyakit daun padi. Seluruh lapisan konvolusional *backbone* MobileNetV3-Large dibekukan (*frozen*) pada tahap *frozen* sehingga hanya parameter pada kepala klasifikasi yang dioptimasi. Selanjutnya, dilakukan tahap *partial fine-tuning* dengan membuka 30% lapisan terakhir *backbone* untuk dilatih bersama kepala klasifikasi, sementara lapisan Batch Normalization tetap dibekukan untuk menjaga stabilitas statistik aktivasi. Konfigurasi ini menghasilkan model dengan jumlah parameter yang relatif kecil namun tetap mampu mengekstraksi fitur visual yang kaya dari citra daun padi.

2) *EfficientNet-B0*: EfficientNet-B0 merupakan model dasar dari keluarga EfficientNet yang mengusulkan skema penskalaan terpadu (compound scaling) terhadap kedalaman,

lebar, dan resolusi jaringan [11]. Seperti ditunjukkan pada Gambar 4, arsitektur EfficientNet-B0 tersusun atas tujuh blok utama yang masing-masing terdiri atas beberapa *Mobile Inverted Bottleneck Convolution* (MBConv) dengan ukuran kernel 3×3 maupun 5×5 . Setiap blok MBConv dilengkapi dengan mekanisme *squeeze-and-excitation* (SE) untuk melakukan penimbangan ulang kanal fitur, sehingga model mampu mencapai keseimbangan antara akurasi dan efisiensi komputasi pada berbagai tingkat resolusi [18].



Gambar 4. Arsitektur EfficientNet-B0

Serupa dengan MobileNetV3-Large, lapisan klasifikasi bawaan EfficientNet-B0 diganti dengan kepala klasifikasi berupa Global Average Pooling diikuti Dense 128 (ReLU), Dropout 0,5, dan Dense 6 (softmax). Proses pelatihan EfficientNet-B0 mengikuti skema dua tahap (*frozen* dan *partial fine-tuning*) sebagaimana dijelaskan pada subbab Pelatihan Model. Dengan rancangan tersebut, penelitian ini dapat membandingkan kinerja dua arsitektur CNN ringan yang sama-sama efisien namun memiliki strategi desain blok yang berbeda pada tugas klasifikasi enam kelas penyakit daun padi.

D. Pelatihan Model

Kedua model dilatih menggunakan data latih dengan skema *transfer learning* dua tahap, yaitu tahap *frozen* dan tahap *partial fine-tuning*, di mana pada tahap *frozen backbone* pralatih dibekukan dan hanya kepala klasifikasi yang diperbarui. Proses pelatihan tahap *frozen* dilakukan selama maksimal 30 epoch dengan ukuran batch tetap 32, menggunakan optimizer Adam dengan learning rate sebesar 0,0001 dan *loss function sparse categorical crossentropy*. Selanjutnya dilakukan *partial fine-tuning* dengan membuka 30% lapisan terakhir pada *backbone* (setara 56 lapisan terakhir pada MobileNetV3-Large dan 71 lapisan terakhir pada EfficientNet-B0) menggunakan learning rate sebesar 0,00001 selama maksimal 15 epoch, dengan lapisan Batch Normalization tetap dibekukan. Selama pelatihan, performa model dipantau pada data validasi setiap epoch.

Untuk mencegah overfitting, diterapkan mekanisme early stopping berbasis nilai *validation loss* (loss validasi) dengan patience 5, sehingga pelatihan dihentikan lebih awal ketika kinerja pada data validasi tidak lagi membaik secara signifikan. Selain itu, digunakan pula model checkpoint yang menyimpan bobot model terbaik selama pelatihan, yaitu pada

saat nilai *validation accuracy* (akurasi validasi) mencapai maksimum. Penyesuaian *learning rate* juga dilakukan menggunakan ReduceLROnPlateau berdasarkan *validation loss* dengan faktor 0,5 dan patience 3. Seluruh proses pelatihan dan pengukuran waktu inferensi dilakukan pada lingkungan Google Colab dengan akselerator GPU NVIDIA T4.

E. Evaluasi dan Perbandingan Model

Setelah proses pelatihan selesai, model terbaik untuk setiap arsitektur dievaluasi menggunakan data uji yang sama sekali tidak digunakan selama pelatihan maupun validasi. Evaluasi diawali dengan pembuatan *confusion matrix*, yaitu tabel yang merangkum jumlah prediksi benar dan salah untuk setiap kelas [19]. Dari *confusion matrix* tersebut, diperoleh nilai dasar berupa True Positive (TP), False Positive (FP), False Negative (FN), dan True Negative (TN) yang dihitung untuk masing-masing kelas dengan definisi sebagai berikut:

- 1) True Positive (TP): sampel suatu kelas yang diprediksi benar sebagai kelas tersebut.
- 2) False Positive (FP): sampel dari kelas lain yang salah diprediksi sebagai kelas tersebut.
- 3) False Negative (FN): sampel suatu kelas yang gagal dikenali dan diprediksi sebagai kelas lain.
- 4) True Negative (TN): sampel yang bukan dari kelas tersebut dan diprediksi benar sebagai bukan kelas tersebut.

Rumus metrik evaluasi yang digunakan dapat dituliskan sebagai berikut:

- 1) Akurasi: Menunjukkan proporsi prediksi yang benar terhadap seluruh sampel.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- 2) Presisi: Mengukur ketepatan model ketika memprediksi suatu kelas.

$$\text{Precision} = \frac{TP}{TP + FP}$$

- 3) Recall: Mengukur kemampuan model menemukan kembali sampel yang benar-benar termasuk suatu kelas.

$$\text{Recall} = \frac{TP}{TP + FN}$$

- 4) F1-Score: Menyatakan rata-rata harmonis antara presisi dan recall.

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Untuk kasus klasifikasi multi-kelas, nilai presisi, recall, dan F1-score dirata-ratakan menggunakan pendekatan *macro average*, sehingga setiap kelas memiliki kontribusi yang seimbang terhadap metrik keseluruhan.

Selain metrik berbasis *confusion matrix*, penelitian ini juga mengukur efisiensi komputasi melalui waktu inferensi rata-rata per citra, yang kemudian dikonversi menjadi *frame per second* (fps). Waktu inferensi diukur dari rata-rata waktu prediksi per citra pada GPU NVIDIA T4 tanpa tahap prapemrosesan tambahan. Nilai-nilai evaluasi tersebut digunakan untuk membandingkan kinerja dan efisiensi relatif antara MobileNetV3-Large dan EfficientNet-B0 pada tugas klasifikasi enam kelas penyakit daun padi.

Untuk memperkuat perbandingan kinerja antar model, selisih performa juga dianalisis menggunakan metode *paired bootstrap* pada data uji dan dilaporkan dalam bentuk *confidence interval* (CI) 95% untuk selisih metrik (MobileNetV3-Large - EfficientNet-B0). Perbedaan dianggap bermakna apabila 95% CI tidak melintasi nol.

Sebagai validasi tambahan terhadap temuan pada data uji internal, model terbaik pada konfigurasi partial fine-tuning diuji pada dataset eksternal yang berbeda dari dataset pelatihan. Evaluasi eksternal dilakukan pada subset tiga kelas yang tersedia, yaitu brown spot, healthy, dan leaf blast, menggunakan akurasi dan F1-score (macro average), serta proporsi prediksi di luar subset tiga kelas.

III. HASIL DAN PEMBAHASAN

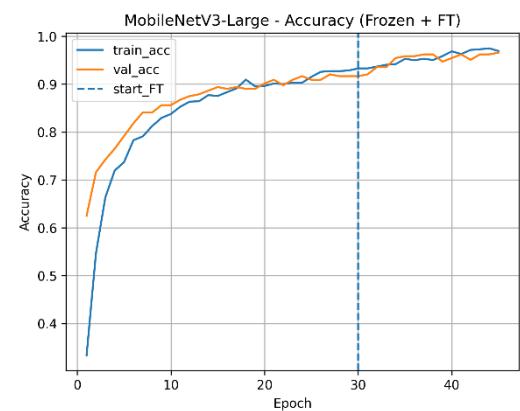
A. Hasil Pelatihan Model

Pelatihan model dilakukan menggunakan arsitektur MobileNetV3-Large dan EfficientNet-B0 yang termasuk keluarga CNN ringan. Kedua model diinisialisasi dengan bobot pralatin ImageNet dan digunakan dalam skema *transfer learning* dua tahap, yaitu tahap *frozen* dan tahap *partial fine-tuning*, dengan kepala klasifikasi baru yang dilatih pada dataset citra daun padi enam kelas. Dataset terdiri dari 2.628 citra yang dibagi menjadi 80:10:10 untuk data latih, validasi, dan uji. Pada tahap *frozen*, pelatihan dilakukan selama maksimal 30 epoch dengan optimisasi Adam, learning rate sebesar 0,0001, dan ukuran batch 32. Selanjutnya, dilakukan *partial fine-tuning* dengan membuka 30% lapisan terakhir pada *backbone* menggunakan learning rate 0,00001 selama maksimal 15 epoch.

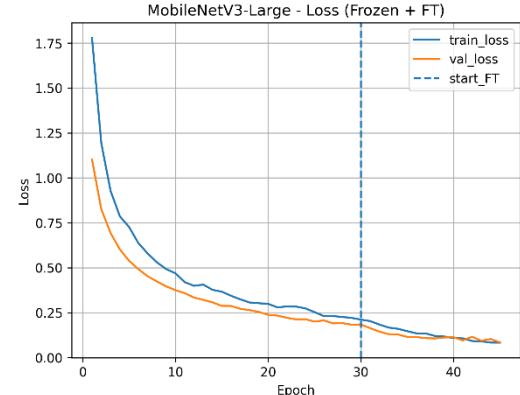
Selama pelatihan, kinerja model dipantau melalui metrik pada data validasi setiap epoch untuk memastikan proses pembelajaran tidak hanya baik pada data latih. Mekanisme *model checkpoint* digunakan untuk menyimpan bobot terbaik berdasarkan akurasi validasi tertinggi, sehingga evaluasi berikutnya menggunakan model yang benar-benar mewakili performa optimal dari masing-masing konfigurasi. Selain itu, *early stopping* berbasis *validation loss* dan penyesuaian *learning rate* (*ReduceLROnPlateau*) membantu menjaga stabilitas optimisasi ketika peningkatan kinerja pada validasi mulai melambat.

Perilaku pelatihan kedua model ditunjukkan pada kurva akurasi dan loss pada Gambar 5 hingga Gambar 8. Pada MobileNetV3-Large, akurasi pelatihan meningkat secara bertahap sejak epoch awal hingga mencapai nilai di atas 0,90 pada akhir tahap *frozen*. Kurva akurasi validasi mengikuti

pola yang serupa dengan selisih yang relatif kecil, sehingga tidak tampak adanya gap besar antara performa pada data latih dan data validasi. Hal ini mengindikasikan bahwa model memiliki kemampuan generalisasi yang baik dan tidak mengalami overfitting yang berarti. Kurva loss pelatihan dan validasi juga menurun secara konsisten hingga mendekati kisaran sekitar 0,2 pada epoch-epoch akhir tahap *frozen*, yang menunjukkan bahwa proses optimisasi berjalan stabil dan model telah mencapai kondisi mendekati konvergensi. Setelah dilanjutkan dengan tahap *partial fine-tuning*, tren akurasi dan loss tetap stabil dengan perbaikan yang lebih moderat. Grafik akurasi dan loss MobileNetV3-Large ditunjukkan pada Gambar 5 dan Gambar 6.



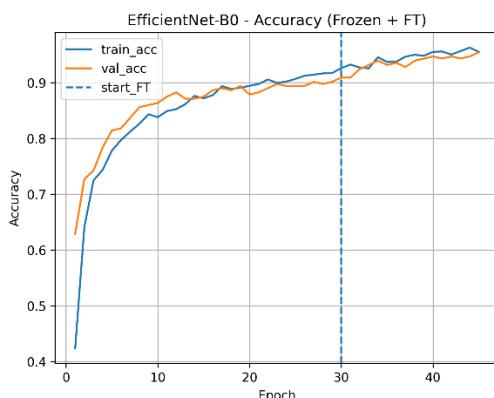
Gambar 5. Grafik akurasi model MobileNetV3-Large



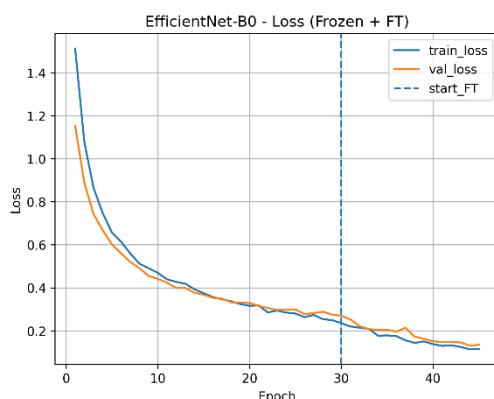
Gambar 6. Grafik loss model MobileNetV3-Large

EfficientNet-B0 menunjukkan pola pelatihan yang sejenis. Akurasi pelatihan dan validasi meningkat secara progresif, dengan akurasi validasi yang stabil di kisaran sekitar 0,88–0,90 pada akhir tahap *frozen*. Dibandingkan MobileNetV3-Large, kurva akurasi validasi EfficientNet-B0 tampak sedikit lebih fluktuatif pada beberapa epoch, namun tren globalnya tetap meningkat. Kurva loss untuk EfficientNet-B0 juga menurun secara bertahap dan mulai mendatar pada epoch-epoch akhir tahap *frozen*, menandakan bahwa model telah mencapai titik jenuh pelatihan pada kisaran 20–30 epoch. Setelah dilanjutkan dengan tahap *partial fine-tuning*, pola loss tetap stabil dan perubahan akurasi cenderung moderat. Grafik

akurasi dan loss EfficientNet-B0 ditunjukkan pada Gambar 7 dan Gambar 8.



Gambar 7. Grafik akurasi model EfficientNet-B0



Gambar 8. Grafik loss model EfficientNet-B0

Secara keseluruhan, hasil pelatihan menunjukkan bahwa kedua model mampu beradaptasi dengan baik terhadap karakteristik citra daun padi. MobileNetV3-Large cenderung menghasilkan kurva akurasi validasi yang sedikit lebih tinggi dan lebih stabil pada epoch-epoch akhir dibanding EfficientNet-B0, yang selaras dengan hasil evaluasi kuantitatif pada data uji yang akan dibahas pada subbab berikutnya.

B. Hasil Evaluasi Model

Evaluasi kinerja dilakukan menggunakan data uji yang sama sekali tidak digunakan selama tahap pelatihan maupun validasi. Tabel II menyajikan ringkasan hasil empat konfigurasi (tahap *frozen* dan *partial fine-tuning* untuk masing-masing model) berdasarkan akurasi, presisi, recall, dan F1-score yang dihitung berdasarkan nilai True Positive, False Positive, False Negative, dan True Negative pada *confusion matrix*. Untuk setiap model, nilai presisi, recall, dan F1-score dirata-ratakan dengan pendekatan *macro average*, sehingga setiap kelas memberikan kontribusi yang seimbang terhadap metrik keseluruhan.

TABEL II
HASIL EVALUASI MODEL PADA DATA UJI

Model	Akurasi	Presisi	Recall	F1-score
MobileNetV3-Large (<i>Frozen</i>)	94,70%	94,65%	94,70%	94,62%
MobileNetV3-Large (<i>Partial FT</i>)	95,83%	95,79%	95,83%	95,80%
EfficientNet-B0 (<i>Frozen</i>)	89,39%	89,42%	89,39%	88,83%
EfficientNet-B0 (<i>Partial FT</i>)	93,18%	93,06%	93,18%	93,02%

Ringkasan hasil evaluasi menunjukkan bahwa pada tahap *frozen*, MobileNetV3-L-L memperoleh akurasi uji sebesar 94,70% dengan presisi 94,65%, recall 94,70%, dan F1-score 94,62%. Sementara itu, EfficientNet-B0 pada tahap *frozen* menghasilkan akurasi 89,39% dengan presisi 89,42%, recall 89,39%, dan F1-score 88,83%. Setelah dilakukan *partial fine-tuning* (Partial FT), kedua model mengalami peningkatan kinerja. MobileNetV3-L mencapai akurasi 95,83% dengan F1-score 95,80%, sedangkan EfficientNet-B0 mencapai akurasi 93,18% dengan F1-score 93,02%.

Nilai presisi dan recall yang relatif seimbang pada keempat konfigurasi menunjukkan bahwa masing-masing arsitektur tidak hanya mampu mengenali sampel dengan baik, tetapi juga menjaga tingkat kesalahan prediksi pada level yang rendah. Secara umum, kedua model menunjukkan performa klasifikasi yang tinggi dengan F1-score di atas 88%. MobileNetV3-L cenderung memperoleh nilai yang lebih tinggi dibanding EfficientNet-B0 pada data uji internal, terutama pada tahap *frozen*.

Setelah *partial fine-tuning*, selisih performa kedua model mengecil. Untuk menguatkan temuan pada Tabel II, selisih performa tersebut dianalisis menggunakan uji statistik *paired bootstrap* dan dilaporkan dalam bentuk *confidence interval* (CI) 95%.

Pada konfigurasi *frozen*, selisih akurasi (MobileNetV3-L – EfficientNet-B0) sebesar 5,34% dengan 95% CI [1,89%; 8,71%] dan selisih F1-score makro sebesar 5,87% dengan 95% CI [2,37%; 9,54%], sehingga perbedaan kinerja pada tahap ini bermakna. Setelah *partial fine-tuning*, selisih akurasi menurun menjadi 2,64% dengan 95% CI [-0,38%; 5,68%] dan selisih F1-score makro menjadi 2,79% dengan 95% CI [-0,37%; 6,13%]. Karena 95% CI pada tahap *partial fine-tuning* masih melintasi nol, perbedaan kinerja pada tahap ini belum konklusif.

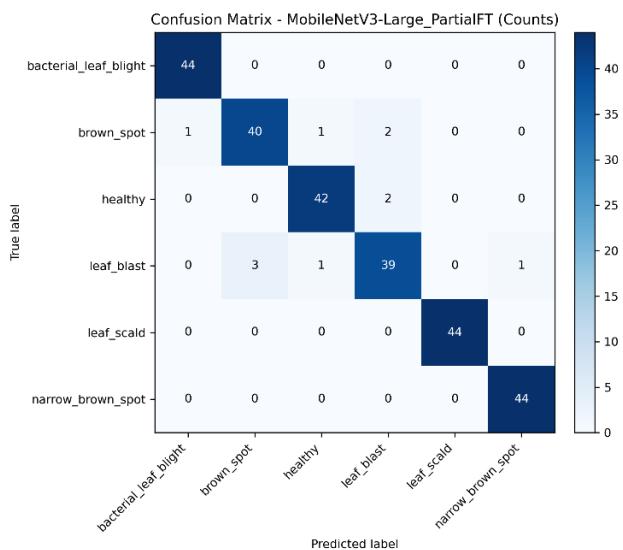
Selanjutnya, analisis kesalahan dilakukan menggunakan *confusion matrix* untuk mengidentifikasi pasangan kelas yang paling sering tertukar pada konfigurasi terbaik.

C. Analisis Confusion Matrix

Confusion matrix digunakan untuk menganalisis performa klasifikasi per kelas dan mengidentifikasi pola kesalahan yang terjadi pada kedua model. Matriks ini merepresentasikan jumlah prediksi benar dan salah untuk setiap kombinasi kelas aktual dan kelas prediksi pada data uji. Gambar 9 menampilkan *confusion matrix* untuk model MobileNetV3-Large (*Partial FT*), sedangkan Gambar 10 menunjukkan *confusion matrix* untuk EfficientNet-B0 (*Partial FT*). Masing-

masing baris pada matriks merepresentasikan kelas sebenarnya, sedangkan kolom merepresentasikan kelas hasil prediksi model.

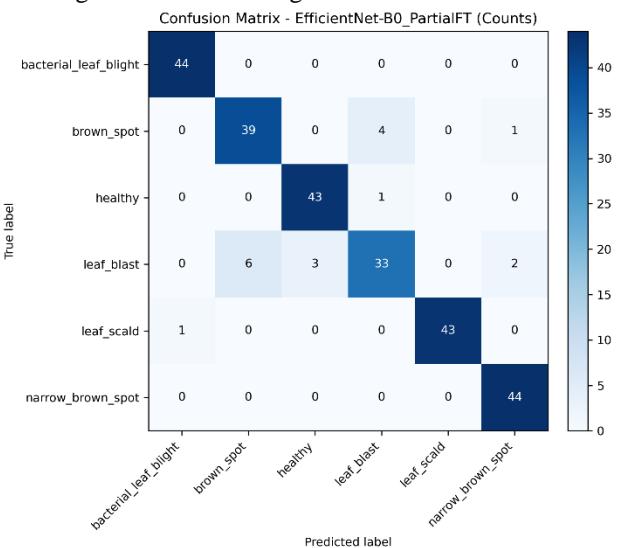
Pada MobileNetV3-Large, sebagian besar elemen diagonal memiliki nilai yang tinggi, yang menunjukkan bahwa model mampu mengklasifikasikan mayoritas sampel dengan benar pada hampir semua kelas. Kelas *bacterial leaf blight*, *leaf scald* dan *narrow brown spot* memiliki tingkat prediksi benar yang sangat tinggi tanpa adanya sampel yang salah klasifikasi (44/44). Kesalahan kecil terlihat pada kelas *healthy*, yaitu 2 sampel diprediksi sebagai *leaf blast*. Kesalahan klasifikasi terutama terjadi pada kelas-kelas dengan gejala bercak yang mirip, yakni *brown spot* dan *leaf blast*, di mana beberapa sampel *brown spot* teridentifikasi sebagai *leaf blast* dan sebaliknya. Pada kelas *brown spot*, terdapat 4 sampel yang salah prediksi (1 sebagai *bacterial leaf blight*, 1 sebagai *healthy*, dan 2 sebagai *leaf blast*). Pada kelas *leaf blast*, terdapat 5 sampel yang salah prediksi (3 sebagai *brown spot*, 1 sebagai *healthy*, dan 1 sebagai *narrow brown spot*). Meskipun demikian, jumlah kesalahan tersebut relatif kecil dibandingkan jumlah prediksi benar, sehingga kinerja per kelas masih tergolong tinggi secara keseluruhan.



Gambar 9. Confusion Matrix MobileNetV3-Large

Pada EfficientNet-B0, pola diagonal masih terlihat dominan, yang menunjukkan bahwa model ini juga mampu mengenali sebagian besar sampel dengan benar. Namun, dibandingkan MobileNetV3-Large, tingkat kesalahan klasifikasi pada kelas tertentu tampak lebih tinggi. Kelas *bacterial leaf blight* dan *narrow brown spot* menunjukkan prediksi benar yang sangat tinggi tanpa adanya salah klasifikasi (44/44). Untuk kelas *healthy*, sebagian besar sampel diprediksi benar (43/44) dengan 1 sampel salah prediksi sebagai *leaf blast*, sedangkan pada kelas *leaf scald* terdapat 1 sampel yang salah prediksi sebagai *bacterial leaf blight* (43/44). Kelas *leaf blast* menjadi kelas yang paling menantang, dengan cukup banyak sampel yang salah diprediksi sebagai *brown spot* atau bahkan *healthy*. Secara

rinci, pada kelas *leaf blast* terdapat 11 sampel yang salah prediksi (6 sebagai *brown spot*, 3 sebagai *healthy*, dan 2 sebagai *narrow brown spot*). Selain itu, beberapa kesalahan juga terjadi pada kelas *brown spot* yang sebagian sampelnya diprediksi sebagai *leaf blast* atau *narrow brown spot*. (4 sampel *brown spot* diprediksi sebagai *leaf blast* dan 1 sampel diprediksi sebagai *narrow brown spot*). Pola kebingungan ini mengindikasikan bahwa EfficientNet-B0 lebih sensitif terhadap kemiripan pola bercak antar penyakit, sehingga akurasi per kelas untuk kelompok penyakit dengan gejala visual yang saling mendekati menjadi sedikit lebih rendah dibanding MobileNetV3-Large.



Gambar 10. Confusion Matrix EfficientNet-B0

Secara umum, analisis *confusion matrix* menunjukkan bahwa kedua model memiliki kemampuan klasifikasi yang sangat baik untuk kelas-kelas dengan karakteristik visual yang jelas, seperti *bacterial leaf blight* dan *narrow brown spot*. Tantangan utama terdapat pada kelas-kelas dengan gejala bercak yang mirip, khususnya *brown spot*, *leaf blast*, yang menjadi sumber utama mis-klasifikasi pada kedua model. Namun, MobileNetV3-Large menghasilkan distribusi prediksi yang lebih seimbang dan tingkat kesalahan yang lebih rendah pada kelas-kelas tersebut, yang selaras dengan kecenderungan nilai akurasi dan F1-score makro yang lebih tinggi pada hasil evaluasi model.

D. Perbandingan Kinerja Model

Hasil evaluasi pada data uji menunjukkan bahwa kedua model CNN yang digunakan mampu memberikan performa klasifikasi yang tinggi. Secara umum, MobileNetV3-Large menghasilkan nilai akurasi, presisi, recall, dan F1-score yang sedikit lebih tinggi dibandingkan EfficientNet-B0 pada data uji internal. Perbedaan ini terlihat lebih jelas pada tahap *frozen*, sedangkan setelah *partial fine-tuning* selisih performa keduanya mengecil.

Perbandingan pola kesalahan pada *confusion matrix* memperkuat temuan tersebut. MobileNetV3-Large cenderung lebih stabil dalam membedakan kelas-kelas

dengan gejala bercak yang mirip, seperti *brown spot*, *leaf blast*, dan *narrow brown spot*, dengan jumlah mis-klasifikasi yang lebih sedikit. EfficientNet-B0 masih menunjukkan kebingungan yang lebih besar pada kelompok kelas tersebut, misalnya ketika beberapa sampel *leaf blast* diprediksi sebagai *brown spot* atau *healthy*. Kondisi ini menjelaskan mengapa nilai F1-score makro MobileNetV3-Large pada konfigurasi *partial fine-tuning* tetap lebih tinggi, meskipun selisihnya tidak sebesar pada tahap *frozen*.

TABEL III
PERBANDINGAN KINERJA KEDUA MODEL

Model	F1-score (%)	Parameter (Juta)	Ukuran (MB)	FPS
MobileNetV3-Large	95,80	3,12	13,99	81,88
EfficientNet-B0	93,02	4,21	18,81	51,60

Dari sisi efisiensi komputasi, perbedaan kedua arsitektur terlihat pada jumlah parameter dan ukuran berkas model. Ringkasan F1-score dan aspek efisiensi ditunjukkan pada Tabel III. MobileNetV3-Large memiliki sekitar 3,12 juta parameter dengan ukuran model sekitar 13,99 MB, sedangkan EfficientNet-B0 memiliki sekitar 4,21 juta parameter dengan ukuran model sekitar 18,81 MB. Kedua model sama-sama mampu melakukan inferensi lebih dari 50 frame per detik, namun MobileNetV3-Large menunjukkan kecepatan inferensi yang lebih tinggi (81,88 fps) dibanding EfficientNet-B0 (51,60 fps), sehingga memberikan keuntungan tambahan untuk penerapan real-time. Namun, kombinasi F1-score yang lebih tinggi, jumlah parameter yang lebih sedikit, dan ukuran model yang lebih kecil menjadikan MobileNetV3-Large sebagai pilihan yang lebih sesuai untuk implementasi pada perangkat dengan sumber daya terbatas, sementara EfficientNet-B0 dapat diposisikan sebagai alternatif yang tetap kompetitif pada lingkungan dengan kapasitas komputasi yang lebih longgar.

E. Uji Generalisasi pada Dataset Eksternal

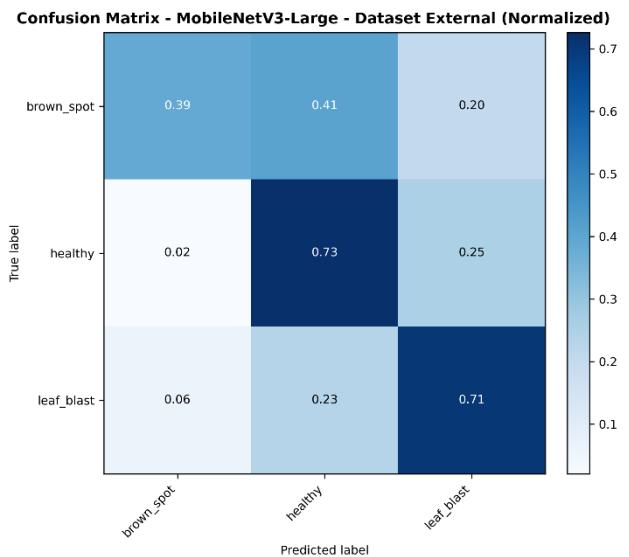
Pengujian eksternal dilakukan sebagai validasi tambahan untuk menilai konsistensi kinerja model pada domain data yang berbeda dari data uji internal. Dataset eksternal diperoleh dari Roboflow Universe [20] menggunakan folder test dan dievaluasi pada tiga kelas yang konsisten dengan label pada dataset internal, yaitu *brown spot*, *healthy*, dan *leaf blast*, dengan total 276 citra, menggunakan model terbaik pada konfigurasi *partial fine-tuning*.

TABEL IV
RINGKASAN HASIL UJI EKSTERNAL

Model	Akurasi	F1-score	Pred. luar 3 kelas
MobileNetV3-Large	65,22%	61,10%	1,09%
EfficientNet-B0	61,59%	61,72%	4,71%

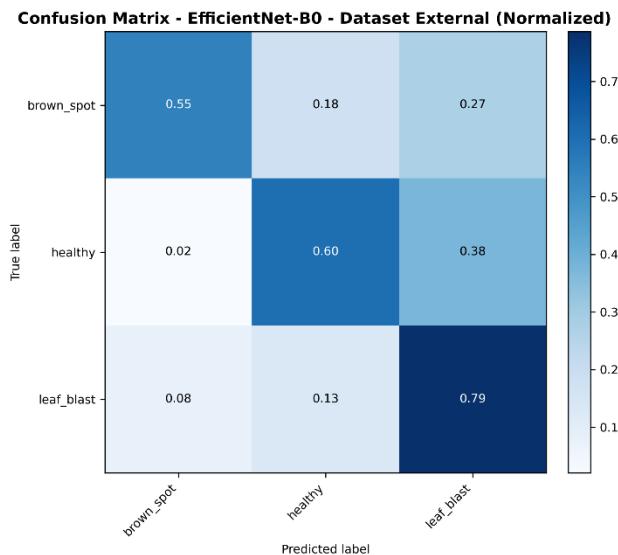
Hasil pada Tabel IV menunjukkan bahwa performa kedua model menurun dibandingkan evaluasi data uji internal, yang mengindikasikan adanya perbedaan karakteristik domain antara dataset internal dan eksternal. Pada metrik akurasi, MobileNetV3-Large memperoleh nilai lebih tinggi, sedangkan EfficientNet-B0 menghasilkan F1-score makro yang sedikit lebih tinggi. Selain itu, proporsi prediksi di luar subset tiga kelas pada EfficientNet-B0 lebih besar dibanding MobileNetV3-Large, yang menunjukkan bahwa pada sebagian citra eksternal model lebih sering memetakan prediksi ke kelas lain di luar subset yang diuji.

Pola kesalahan prediksi pada MobileNetV3-Large ditunjukkan pada confusion matrix ternormalisasi pada Gambar 11. Kelas *healthy* merupakan kelas yang paling stabil dikenali (nilai diagonal 0,73). Namun, kelas *brown spot* cukup sering tertukar menjadi *healthy* (0,41) dan sebagian menjadi *leaf blast* (0,20). Untuk kelas *leaf blast*, sebagian sampel masih keliru diprediksi sebagai *healthy* (0,23). Pola ini menunjukkan bahwa kemiripan gejala bercak antar kelas masih menjadi sumber utama mis-klasifikasi pada pengujian lintas-dataset.



Gambar 11. Confusion matrix MobileNetV3-Large pada dataset eksternal.

Pada EfficientNet-B0, confusion matrix ternormalisasi pada Gambar 12 menunjukkan bahwa kelas *leaf blast* memiliki tingkat pengenalan yang paling tinggi (nilai diagonal 0,79), diikuti *healthy* (0,60) dan *brown spot* (0,55). Kesalahan terbesar terjadi ketika *healthy* diprediksi sebagai *leaf blast* (0,38). Pada kelas *brown spot*, sebagian sampel juga tertukar menjadi *leaf blast* (0,27) dan *healthy* (0,18), sedangkan kesalahan pada *leaf blast* relatif lebih kecil, terutama ketika diprediksi sebagai *brown spot* (0,08) dan *healthy* (0,13). Dibanding MobileNetV3-Large, EfficientNet-B0 lebih kuat dalam mengenali *leaf blast*, namun menunjukkan kebingungan yang lebih besar pada kelas *healthy*, yang selaras dengan perbedaan kinerja ringkas pada evaluasi eksternal.



Gambar 12. Confusion matrix EfficientNet-B0 pada dataset eksternal.

Secara umum, uji eksternal ini memberikan validasi tambahan bahwa kinerja model dapat berubah ketika diuji pada domain data yang berbeda dari data pelatihan dan uji internal. Perubahan performa ini mengindikasikan adanya perbedaan karakteristik citra antara dataset internal dan eksternal, terutama pada kelas-kelas dengan gejala bercak yang mirip. Ke depan, *robustness* model berpotensi diperkuat melalui penggunaan data pelatihan yang lebih besar dan lebih beragam, misalnya dengan menggabungkan beberapa sumber dataset atau menambah variasi kondisi pengambilan citra, sehingga model lebih stabil ketika diterapkan pada skenario lapangan.

F. Diskusi

Hasil eksperimen menunjukkan bahwa kedua arsitektur CNN ringan yang digunakan, yaitu MobileNetV3-Large dan EfficientNet-B0, mampu memberikan performa klasifikasi yang tinggi pada enam kelas penyakit daun padi. Nilai akurasi dan F1-score makro yang tinggi pada evaluasi data uji internal mengindikasikan bahwa pendekatan *transfer learning* dengan *backbone* pralatih ImageNet efektif untuk klasifikasi citra daun. Selain itu, performa yang relatif merata pada seluruh kelas menunjukkan bahwa dataset dengan distribusi seimbang membantu mengurangi kecenderungan bias prediksi terhadap kelas tertentu pada skenario klasifikasi multi-class.

Selisih performa antar model juga dianalisis menggunakan *paired bootstrap* berbasis *confidence interval* (CI) 95%. Hasilnya menunjukkan bahwa pada tahap *frozen*, selisih performa MobileNetV3-Large terhadap EfficientNet-B0 tampak konsisten. Namun, setelah *partial fine-tuning* selisih performa mengecil dan untuk beberapa metrik CI melintasi nol, sehingga perbedaan kinerja pada tahap ini belum dapat dinyatakan konklusif pada tingkat keyakinan 95%. Oleh karena itu, keunggulan MobileNetV3-Large pada konfigurasi akhir diinterpretasikan secara hati-hati tanpa menyatakan keunggulan signifikan untuk seluruh metrik.

Uji eksternal yang ditambahkan menunjukkan adanya perbedaan performa dibanding pengujian internal, yang mengindikasikan perbedaan karakteristik citra antar dataset. Temuan ini menguatkan pentingnya data pelatihan yang lebih beragam agar model lebih stabil pada skenario lapangan.

Hasil penelitian ini juga konsisten dengan temuan pada studi klasifikasi penyakit daun stroberi dan daun jeruk yang menggunakan MobileNetV3-Large dan EfficientNet-B0, di mana kedua arsitektur tersebut mampu mencapai akurasi di atas 90% pada berbagai jenis tanaman [12], [13]. Pola bahwa MobileNetV3-Large sedikit lebih unggul dibanding EfficientNet-B0 juga terlihat pada hasil penelitian ini, meskipun selisih kinerjanya relatif kecil. Di sisi lain, sejumlah penelitian penyakit daun padi juga telah mengeksplorasi arsitektur lain, termasuk MobileNetV2 dan varian EfficientNet yang lebih besar [6], [14]. Namun, perbedaan karakteristik dataset, jumlah kelas, serta protokol pembagian data dan evaluasi membuat perbandingan nilai performa antar studi tersebut tidak sepenuhnya setara. Oleh karena itu, kontribusi utama penelitian ini terletak pada perbandingan langsung dua arsitektur *lightweight CNN* dalam satu skenario eksperimen yang seragam, serta evaluasi efisiensi komputasi melalui jumlah parameter, ukuran model, dan kecepatan inferensi.

Analisis confusion matrix pada konfigurasi terbaik (*partial fine-tuning*) menunjukkan bahwa kedua model sangat baik dalam mengenali kelas dengan karakteristik visual yang jelas, seperti bacterial leaf blight, healthy, dan leaf scald, dengan tingkat kesalahan yang rendah. Kesalahan klasifikasi terutama muncul pada kelas-kelas dengan gejala bercak yang memiliki kemiripan visual, yaitu *brown spot*, *leaf blast*, dan *narrow brown spot*. Perbedaan kinerja pada kelompok kelas ini kemungkinan berkaitan dengan karakteristik arsitektur masing-masing model, di mana MobileNetV3-Large dirancang untuk ekstraksi fitur yang efisien pada resolusi relatif rendah sehingga berpotensi lebih stabil dalam membedakan tekstur bercak yang saling berdekatan pada pengaturan pelatihan yang sama.

G. Implikasi Praktis

Implikasi praktis dari penelitian ini dapat diperluas pada beberapa skenario penerapan di lingkungan pertanian. Kecepatan inferensi yang diperoleh pada pengujian (di atas 50 fps pada GPU NVIDIA T4) menunjukkan potensi penggunaan real-time, misalnya untuk klasifikasi cepat melalui aplikasi smartphone saat pemantauan tanaman, integrasi pada drone untuk inspeksi area sawah, maupun pada perangkat IoT pertanian sebagai sistem pemantauan berbasis kamera. Pada skenario tersebut, citra daun dapat diambil menggunakan kamera perangkat, diproses melalui prapemrosesan sederhana, lalu model menghasilkan prediksi kelas penyakit sebagai keluaran untuk membantu keputusan awal di lapangan. Dari sisi efisiensi, MobileNetV3-Large memiliki ukuran model yang lebih kecil dan kecepatan inferensi yang lebih tinggi dibanding EfficientNet-B0 pada lingkungan pengujian yang sama, sehingga lebih

menguntungkan untuk perangkat dengan sumber daya terbatas. Namun, pengukuran fps pada penelitian ini dilakukan pada lingkungan GPU sehingga performanya dapat berbeda pada perangkat nyata. Oleh karena itu, pengujian langsung pada perangkat target (smartphone, drone, atau edge device) diperlukan sebagai tahap lanjutan untuk memvalidasi performa implementasi.

H. Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan. Dataset internal bersumber dari dataset publik sehingga variasi kondisi lapangan, seperti pencahayaan, latar belakang, sudut pengambilan, dan tingkat keparahan penyakit, belum sepenuhnya terwakili. Sebagai validasi tambahan, penelitian ini telah menyertakan uji eksternal pada dataset berbeda, dan perbedaan domain data terlihat dari perubahan kinerja pada subset tiga kelas yang tersedia. Strategi pelatihan menggunakan partial fine-tuning dengan membuka 30% lapisan terakhir backbone dan membukukan lapisan Batch Normalization dipilih sebagai kompromi untuk meningkatkan adaptasi model sekaligus menekan risiko overfitting mengingat ukuran dataset pelatihan relatif terbatas, sedangkan full fine-tuning berpotensi meningkatkan kebutuhan komputasi dan risiko overfitting. Selain itu, kecepatan inferensi diukur pada GPU NVIDIA T4 di lingkungan Google Colab sehingga nilai fps belum merepresentasikan performa pada perangkat smartphone atau edge device, dan pengujian langsung pada perangkat target diperlukan untuk memvalidasi efisiensi implementasi.

IV. KESIMPULAN

Penelitian ini membandingkan kinerja dua arsitektur CNN ringan, yaitu MobileNetV3-Large dan EfficientNet-B0, pada tugas klasifikasi enam kelas penyakit daun padi berbasis citra menggunakan pendekatan *transfer learning* dengan *backbone* pralatih ImageNet. Proses pelatihan dilakukan dalam dua tahap, yaitu tahap *frozen* dan tahap *partial fine-tuning*. Kedua model dilatih pada 2.628 citra yang dibagi menjadi data latih, validasi, dan uji dengan rasio 80%:10%:10% dan dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score (*macro average*). Pada konfigurasi terbaik (*partial fine-tuning*), MobileNetV3-Large mencapai akurasi uji 95,83% dan F1-score makro 95,80%, sedangkan EfficientNet-B0 memperoleh akurasi 93,18% dan F1-score makro 93,02%. Analisis *confusion matrix* mengindikasikan bahwa keduanya sangat baik dalam mengenali kelas dengan karakteristik visual yang jelas, sementara kesalahan utama terjadi pada kelas-kelas bercak dengan gejala mirip, di mana MobileNetV3-Large menghasilkan jumlah mis-klasifikasi yang sedikit lebih rendah.

Perbandingan kinerja antar model juga diperkuat melalui analisis *paired bootstrap* berbasis *confidence interval* (CI) 95%, yang menunjukkan selisih performa tampak konsisten pada tahap *frozen*, namun setelah *partial fine-tuning* selisih mengecil dan untuk beberapa metrik CI melintasi nol

sehingga interpretasi perbedaan kinerja dilakukan secara hati-hati.

Dari sisi efisiensi komputasi, MobileNetV3-Large memiliki jumlah parameter 3,12 juta, ukuran model 13,99 MB, dan kecepatan inferensi 81,88 fps, sedangkan EfficientNet-B0 memiliki 4,21 juta parameter, ukuran 18,81 MB, dan kecepatan inferensi 51,60 fps pada pengujian GPU NVIDIA T4. Pada uji eksternal berbasis tiga kelas, performa kedua model menurun dibanding pengujian internal yang mengindikasikan adanya perbedaan domain data. Kombinasi kinerja internal yang lebih tinggi, kompleksitas model yang lebih rendah, dan kecepatan inferensi yang memadai menjadikan MobileNetV3-Large lebih sesuai sebagai kandidat utama untuk implementasi sistem deteksi penyakit daun padi pada perangkat dengan sumber daya terbatas, sementara EfficientNet-B0 dapat diposisikan sebagai alternatif yang kompetitif. Ke depan, penelitian ini dapat diperluas dengan pengujian model pada citra daun padi yang dikumpulkan langsung dari lapangan dengan variasi kondisi pencahayaan, latar belakang, dan tingkat keparahan penyakit yang lebih beragam, serta dengan memanfaatkan hasil yang diperoleh sebagai dasar pengembangan prototipe aplikasi pendukung keputusan berbasis perangkat mobile atau edge device agar manfaat model dapat dirasakan secara langsung oleh petani dan penyuluh pertanian.

DAFTAR PUSTAKA

- [1] N. K. Fukagawa and L. H. ZISKA, "Rice: Importance for Global Nutrition," *J Nutr Sci Vitaminol (Tokyo)*, vol. 65, no. Supplement, pp. S2–S3, Oct. 2019, doi: 10.3177/jnsv.65.S2.
- [2] "Feeding the world: revolutionary rice," University of Oxford, Queen's College. Accessed: Dec. 05, 2025. [Online]. Available: <https://www.queens.ox.ac.uk/research-at-queens/how-do-we-feed-the-world-2/>
- [3] Badan Pusat Statistik, "Luas Panen dan Produksi Padi di Indonesia 2024." Accessed: Nov. 05, 2025. [Online]. Available: <https://www.bps.go.id/id/pressrelease/2025/02/03/2414/pada-2024-luas-panen-padi-mencapai-sekitar-10-05-juta-hektare-dengan-produksi-padi-sebanyak-53-14-juta-ton-gabah-kering-giling--gkg-.html>
- [4] Badan Pusat Statistik, "Jumlah Penduduk Pertengahan Tahun 2025." Accessed: Nov. 05, 2025. [Online]. Available: <https://www.bps.go.id/id/statistics-table/2/MTk3NSMy/jumlah-penduduk-pertengahan-tahun--ribu-jawa-.html>
- [5] R. Ramli, E. Evanita, and A. Akbar Riadi, "Classification of Rice Leaf Diseases Using Support Vector Machine with HSV and GLCM-Based Feature Extraction," *Journal of Applied Informatics and Computing*, vol. 9, no. 5, pp. 2329–2337, Oct. 2025, doi: 10.30871/jaic.v9i5.10403.
- [6] O. Virgantara Putra, M. Zaim Mustaqim, and D. Muriatmoko, "Transfer Learning untuk Klasifikasi Penyakit dan Hama Padi Menggunakan MobileNetV2," *Techno.com*, vol. 22, no. 3, pp. 562–575, 2023.
- [7] B. Naresh kumar and S. Sakthivel, "Rice leaf disease classification using a fusion vision approach," *Sci Rep*, vol. 15, no. 8692, Mar. 2025, doi: 10.1038/s41598-025-87800-3.
- [8] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Front Plant Sci*, vol. 7, no. September, Sep. 2016, doi: 10.3389/fpls.2016.01419.
- [9] X. Zhang, N. Han, and J. Zhang, "Comparative analysis of VGG, ResNet, and GoogLeNet architectures evaluating performance, computational efficiency, and convergence rates," *Applied and*

- Computational Engineering*, vol. 44, no. 1, pp. 172–181, Mar. 2024, doi: 10.54254/2755-2721/44/20230676.
- [10] A. Howard *et al.*, “Searching for MobileNetV3,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Nov. 2019, pp. 1314–1324. [Online]. Available: <http://arxiv.org/abs/1905.02244>
- [11] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, Sep. 2019, pp. 6105–6114. [Online]. Available: <http://arxiv.org/abs/1905.11946>
- [12] D. A. Pramudhita, F. Azzahra, I. K. Arfat, R. Magdalena, and S. Saidah, “Strawberry Plant Diseases Classification Using CNN Based on MobileNetV3-Large and EfficientNet-B0 Architecture,” *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 9, no. 3, pp. 522–534, Jul. 2023, doi: 10.26555/jiteki.v9i3.26341.
- [13] A. R. Dianto, F. T. Anggraeny, and H. Maulana, “Analisis efektifitas algoritma MobileNetV3-Large dan EfficientNet-B0 untuk klasifikasi citra penyakit daun jeruk,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3, Jul. 2025, doi: 10.23960/jitet.v13i3.6956.
- [14] A. Azis, A. Fadlil, and T. Sutikno, “Optimization of Convolutional Neural Network (CNN) Using Transfer Learning for Disease Identification in Rice Leaf Images,” *Jurnal E-Komtek (Elektro-Komputer-Teknik)*, vol. 8, no. 2, pp. 504–515, Dec. 2024, doi: 10.37339/e-komtek.v8i2.2132.
- [15] M. Buda, A. Maki, and M. A. Mazurowski, “A systematic study of the class imbalance problem in convolutional neural networks.” *Neural Networks*, vol. 106, pp. 249–259, Oct. 2018, doi: 10.1016/j.neunet.2018.07.011.
- [16] D. Saputra, “Rice Leaf Disease Dataset,” Kaggle. Accessed: Nov. 03, 2025. [Online]. Available: <https://www.kaggle.com/datasets/dedeikhsandwisaputra/rice-leaves-disease-dataset>
- [17] O. Russakovsky *et al.*, “ImageNet Large Scale Visual Recognition Challenge,” *International Journal of Computer Vision (IJCV)*, vol. 115, no. 3, pp. 211–252, Jan. 2015, [Online]. Available: <http://arxiv.org/abs/1409.0575>
- [18] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, “Squeeze-and-Excitation Networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, May 2018. [Online]. Available: <http://arxiv.org/abs/1709.01507>
- [19] S. A. Hicks *et al.*, “On evaluation metrics for medical applications of artificial intelligence,” *Sci Rep*, vol. 12, no. 1, Apr. 2022, doi: 10.1038/s41598-022-09954-8.
- [20] crop disease, “Rice,” Roboflow Universe. Accessed: Jan. 05, 2026. [Online]. Available: <https://universe.roboflow.com/crop-disease-wntdr/rice-khbng/dataset/1>