

Implementasi Deep Learning Berbasis MobileNetV2 untuk Deteksi Real-Time Bacterial Spot dengan Pendekatan Arsitektur Lightweight

Ahmad Nabilul As'ad, Elkaf Rahmawan Pramudya*

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

Email: ¹111202214488@mhs.dinus.ac.id, ^{2,*}elkaf.rahmawan@dsn.dinus.ac.id

Email Penulis Korespondensi: elkaf.rahmawan@dsn.dinus.ac.id

Submitted: 19/01/2026; Accepted: 05/03/2026; Published: 06/03/2026

Abstrak—*Bacterial spot* yang disebabkan oleh *Xanthomonas campestris* pv. *vesicatoria* merupakan penyakit krusial pada paprika yang dapat menurunkan produktivitas hingga 50%. Penelitian ini mengimplementasikan MobileNetV2 dengan *two-stage transfer learning* untuk deteksi *real-time bacterial spot* menggunakan pendekatan arsitektur *lightweight*, dengan ResNet50 sebagai *baseline comparison*. Dataset PlantVillage (2.475 citra) digunakan untuk *training* dan evaluasi *in-domain*, sementara dataset India (132 citra) untuk evaluasi *domain shift*. Hasil menunjukkan MobileNetV2 mencapai akurasi 98.66% pada *test set* PlantVillage, melampaui ResNet50 (89.78%) dengan margin 8.88 *percentage points* meskipun 9.2× lebih ringan (2.7 MB vs 24.3 MB TFLite) dan 2.0× lebih cepat (22.4 ms vs 45.8 ms *inference time*). Keunggulan efisiensi MobileNetV2 juga terlihat pada *memory footprint* saat *inference* yang hanya 107 MB RAM, secara signifikan lebih rendah 2.3x dibanding ResNet50(242 MB RAM), menjadikannya sangat sesuai untuk *deployment* pada *mid-range smartphone* dengan RAM terbatas. Evaluasi *external dataset* mengungkapkan MobileNetV2 mempertahankan *robustness superior* dengan *retention rate* 65.3% versus ResNet50 52.3%. Analisis *trade-off* menunjukkan MobileNetV2 berada pada *Pareto frontier*, mencapai *accuracy-efficiency sweet spot* optimal untuk aplikasi deteksi penyakit tanaman. Penelitian ini berkontribusi pada bukti empiris superioritas arsitektur *lightweight*, *framework* evaluasi komprehensif berorientasi efisiensi, *ULTRA-LIGHT training strategy* untuk mengatasi *inverse overfitting*, dan evaluasi generalisasi realistis menggunakan *external dataset* tropis.

Kata Kunci: Deep Learning; MobileNetV2; Bacterial Spot; Transfer Learning; Lightweight Architecture

Abstract—*Bacterial spot* caused by *Xanthomonas campestris* pv. *vesicatoria* is a critical disease in bell peppers that can reduce productivity by up to 50%. This study implements MobileNetV2 with *two-stage transfer learning* for *real-time bacterial spot* detection using *lightweight architecture* approach, with ResNet50 as *baseline comparison*. PlantVillage dataset (2,475 images) was used for *training* and *in-domain* evaluation, while India dataset (132 images) for *domain shift* assessment. Results demonstrate MobileNetV2 achieves 98.66% accuracy on PlantVillage test set, outperforming ResNet50 (89.78%) by 8.88 *percentage points* despite being 9.2× lighter (2.7 MB vs 24.3 MB TFLite) and 2.0× faster (22.4 ms vs 45.8 ms *inference time*). MobileNetV2 efficiency advantage is also evident in its *inference memory footprint* of only 107 MB RAM, significantly 2.3x lower than ResNet50(242 MB RAM), making it highly suitable for *deployment* on *mid-range smartphones* with limited RAM. *External dataset* evaluation reveals MobileNetV2 maintains superior *robustness* with 65.3% *retention rate* versus ResNet50's 52.3%. *Trade-off* analysis positions MobileNetV2 on the *Pareto frontier*, achieving optimal *accuracy-efficiency sweet spot* for plant disease detection applications. This research contributes empirical evidence for *lightweight architecture* superiority, comprehensive *efficiency-oriented* evaluation framework, *ULTRA-LIGHT training strategy* for addressing *inverse overfitting*, and realistic *generalization* assessment using *tropical external dataset*.

Keywords: Deep Learning; MobileNetV2; Bacterial Spot; Transfer Learning; Lightweight Architecture

1. PENDAHULUAN

Paprika (*Capsicum annum* L. var. *grossum*) merupakan salah satu komoditas hortikultura bernilai ekonomi tinggi yang banyak dibudidayakan di Indonesia. Tanaman ini kaya akan vitamin C, karotenoid, dan antioksidan sehingga memiliki nilai gizi yang penting bagi kesehatan [1]. Namun, produktivitas paprika sangat rentan terhadap serangan penyakit, khususnya *bacterial spot* yang disebabkan oleh *Xanthomonas campestris* pv. *vesicatoria*. Penyakit ini dapat menyebabkan penurunan kualitas dan kuantitas hasil panen hingga 50% apabila tidak terdeteksi dan ditangani secara dini [2]. Serangan penyakit pada tanaman berdampak signifikan terhadap produktivitas panen. Apabila tidak ditangani dengan segera, penyakit tersebut dapat mengakibatkan kerusakan tanaman, kegagalan panen, serta menimbulkan kerugian ekonomi yang substansial bagi petani [3], [4]. Gejala *bacterial spot* menunjukkan karakteristik visual yang khas, dimulai dengan bintik-bintik basah (*water-soaked*) yang berkembang menjadi bercak coklat berpola menyudut (angular). Pada infeksi parah, lesi dapat menyatu membentuk area nekrotik yang lebih besar dan mengakibatkan defoliasi masif [5].

Deteksi penyakit tanaman secara konvensional masih bergantung pada inspeksi visual oleh pakar atau petani berpengalaman, yang memerlukan waktu lama, biaya tinggi, dan rentan terhadap subjektivitas [6]. Dalam beberapa tahun terakhir, teknologi deep learning berbasis *Convolutional Neural Network* (CNN) telah menunjukkan potensi besar dalam otomatisasi deteksi penyakit tanaman melalui analisis citra daun, efektivitas metode ini juga telah dibuktikan pada komoditas pangan strategis lainnya seperti padi [7]. Beberapa penelitian telah berhasil menerapkan arsitektur CNN seperti MobileNetV2, VGG16, ResNet, Inception, dan Xception untuk klasifikasi penyakit pada berbagai jenis tanaman dengan tingkat akurasi yang tinggi [6], [8], [9], [10]. Pendekatan transfer learning memungkinkan model yang telah dilatih pada dataset besar seperti ImageNet untuk diadaptasi ke domain spesifik dengan jumlah data terbatas, meningkatkan akurasi dan kecepatan klasifikasi secara substansial [11].

Meskipun banyak penelitian telah dilakukan, tantangan fundamental untuk deployment pada perangkat terbatas masih belum terpecahkan secara komprehensif. Pertama, penelitian sebelumnya seperti Ilhamsyah & Enri [12] mencapai akurasi tinggi (97,5%) tetapi belum membuktikan kelayakan deployment real-time pada perangkat mobile yang umum digunakan petani Indonesia, khususnya dalam hal kecepatan inferensi, ukuran model, dan konsumsi memori. Kedua, MobileNetV2 sebagai arsitektur lightweight yang dirancang khusus untuk edge devices belum dieksplorasi secara mendalam untuk deteksi bacterial spot pada paprika, padahal arsitektur ini menawarkan trade-off optimal antara akurasi dan efisiensi komputasi [13], [14]. Ketiga, evaluasi generalisasi model menggunakan external dataset untuk mengukur robustness terhadap domain shift masih jarang dilakukan, padahal kondisi lapangan Indonesia memiliki variasi tinggi dalam hal pencahayaan, background, dan kondisi daun yang berbeda dari dataset controlled environment [15], [16], [17]. Gap ini menjadi krusial karena deployment sistem deteksi penyakit yang efektif memerlukan tidak hanya akurasi tinggi, tetapi juga efisiensi komputasi, ukuran model kompak, dan kemampuan generalisasi yang baik terhadap kondisi real-world.

Beberapa penelitian terdahulu telah menggunakan deep learning untuk deteksi penyakit pada tanaman Solanaceae. Ilhamsyah & Enri [12] mengembangkan model identifikasi penyakit *bacterial spot* pada daun paprika menggunakan *Convolutional Neural Network* (CNN) berbasis transfer learning, dan memperoleh akurasi validasi sebesar 97,5%. Namun, penelitian tersebut belum mengevaluasi deployment metrics (inference time, model size, memory footprint) yang krusial untuk implementasi praktis. Salim et al. [2] menggunakan DenseNet-201 dengan histogram equalization untuk deteksi penyakit cabai, mencapai akurasi 99,5%, tetapi metode ini memerlukan preprocessing yang kompleks dan model dengan ukuran besar yang tidak efisien untuk edge devices. Maulana et al. [15] mengembangkan sistem deteksi penyakit cabai berbasis Raspberry Pi dengan histogram equalization mencapai akurasi 95%, namun sistem tersebut terbatas pada deteksi binary dan memerlukan *controlled lighting conditions*. Sharma et al. [18] mendemonstrasikan bahwa arsitektur deep seperti ResNet50 dapat mencapai akurasi tinggi pada deteksi penyakit tomat, sementara Aryanta et al. [19] menggunakan MobileNetV2 untuk deteksi penyakit pisang mencapai akurasi 96,21%, menunjukkan efektivitas lightweight architecture, namun belum ada penelitian yang secara komprehensif mengevaluasi MobileNetV2 untuk bacterial spot detection dengan fokus pada real-time processing, model compression (TFLite conversion), dan domain shift analysis pada kondisi lapangan tropis.

Transfer learning dengan two-stage fine-tuning strategy telah terbukti efektif untuk domain adaptation pada dataset dengan ukuran terbatas [20]. Strategi ini melibatkan unfreezing layers secara bertahap, memungkinkan model menyesuaikan feature representations dari domain umum (ImageNet) ke domain spesifik (plant disease) tanpa overfitting [20], [21]. MobileNetV2 dengan arsitektur inverted residual dan linear bottleneck-nya menawarkan efisiensi tinggi untuk deployment pada perangkat dengan keterbatasan memori dan daya komputasi [11]. dengan keunggulan computational cost $8-9\times$ lebih rendah dibanding standard convolution dan model size yang dapat dikompresi hingga $\sim 2.5\text{MB}$ dalam format TensorFlow Lite. Untuk membuktikan efektivitas MobileNetV2 secara empiris, penelitian ini menggunakan ResNet50 sebagai baseline comparison yang merepresentasikan deep architecture dengan high representational capacity [22]. Namun, implementasi sistematis MobileNetV2 untuk bacterial spot detection dengan evaluasi komprehensif terhadap accuracy-efficiency trade-off dan domain shift robustness masih menjadi research gap yang perlu dijawab.

Berdasarkan research gap yang telah diidentifikasi, penelitian ini merumuskan tiga pertanyaan riset sebagai berikut: bagaimana implementasi MobileNetV2 dengan two-stage transfer learning untuk mencapai akurasi tinggi sekaligus memenuhi constraint real-time processing ($<100\text{ms}$ inference time) untuk deteksi bacterial spot pada paprika; bagaimana trade-off antara akurasi, ukuran model, dan kecepatan inferensi dibanding baseline ResNet50 untuk mendukung keputusan deployment pada perangkat terbatas; serta seberapa robust MobileNetV2 terhadap domain shift ketika dievaluasi pada external dataset dengan kondisi real-world yang berbeda dari training data.

Untuk menjawab ketiga pertanyaan riset tersebut, tujuan utama penelitian ini adalah mengimplementasikan dan mengevaluasi MobileNetV2 untuk deteksi real-time bacterial spot pada paprika dengan fokus pada deployment efficiency untuk perangkat terbatas. Secara spesifik, penelitian ini bertujuan mengimplementasikan MobileNetV2 dengan two-stage transfer learning strategy menggunakan Adam optimizer dan adaptive learning rate reduction (ReduceLROnPlateau) untuk memaksimalkan konvergensi dan mencapai akurasi tinggi dengan constraint real-time processing. Penelitian ini juga melakukan analisis komparatif dengan ResNet50 sebagai baseline untuk membuktikan bahwa lightweight architecture dapat mencapai performa setara atau superior dibanding deep model pada task agricultural disease detection dengan dataset terbatas, serta memberikan evaluasi deployment-oriented meliputi inference time, model size (TFLite compression), memory footprint, dan battery consumption. Selain itu, penelitian ini mengevaluasi kemampuan generalisasi model pada dua dataset berbeda (PlantVillage controlled environment dan India real-world dataset) untuk mengukur robustness terhadap domain shift, mencerminkan tantangan deployment aktual di lapangan.

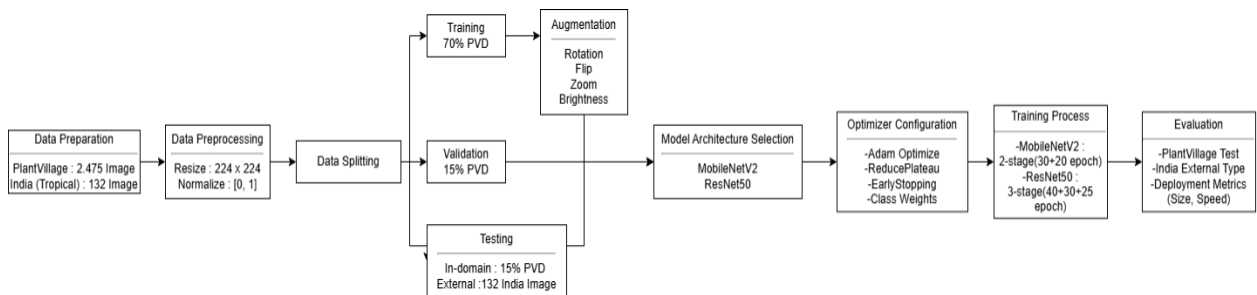
Penelitian ini diharapkan memberikan empat kontribusi utama. Kontribusi pertama berupa implementasi dan optimasi komprehensif MobileNetV2 untuk deteksi *bacterial spot* paprika dengan *two-stage transfer learning* yang mendemonstrasikan efisiensi superior arsitektur *lightweight*. Kontribusi kedua adalah analisis sistematis *trade-off* antara akurasi, ukuran model (*parameter count*), dan *inference time* dengan perbandingan *baseline* ResNet50 untuk memberikan rekomendasi berbasis bukti empiris mengenai pemilihan arsitektur. Kontribusi ketiga mencakup identifikasi dan analisis *ULTRA-LIGHT training strategy* sebagai solusi potensial untuk mengatasi *inverse overfitting problem* pada *deep architecture* dengan dataset terbatas. Kontribusi keempat meliputi evaluasi empiris generalisasi

model menggunakan *external dataset* untuk memberikan *insight* mengenai *robustness* arsitektur *lightweight* terhadap *domain shift* pada kondisi lapangan tropis. Temuan penelitian ini diharapkan dapat menjadi *framework replicable* untuk pengembangan sistem deteksi penyakit pada komoditas hortikultura lain di Indonesia dengan fokus pada efisiensi arsitektur dan aksesibilitas untuk petani.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan eksperimental dengan metode *deep learning* berbasis *Convolutional Neural Network* (CNN) untuk mengklasifikasikan penyakit *bacterial spot* pada daun paprika. Metodologi dirancang untuk menjawab tiga pertanyaan riset: (1) Bagaimana implementasi MobileNetV2 dengan *two-stage transfer learning* untuk mencapai akurasi tinggi dengan efisiensi komputasi optimal untuk deteksi *bacterial spot* pada paprika? (2) Bagaimana *trade-off* antara akurasi, ukuran model, dan kecepatan inferensi antara arsitektur *lightweight* (MobileNetV2) dan *deep architecture* (ResNet50)? (3) Seberapa *robust* MobileNetV2 terhadap *domain shift* ketika dievaluasi pada *external dataset* dengan kondisi *real-world* yang berbeda dari *training data*? Penelitian ini terdiri dari tujuh tahapan utama: *data preparation*, *data preprocessing*, *data splitting*, *augmentation*, *model architecture selection*, *training process* dengan *two-stage transfer learning*, dan *comprehensive evaluation*.

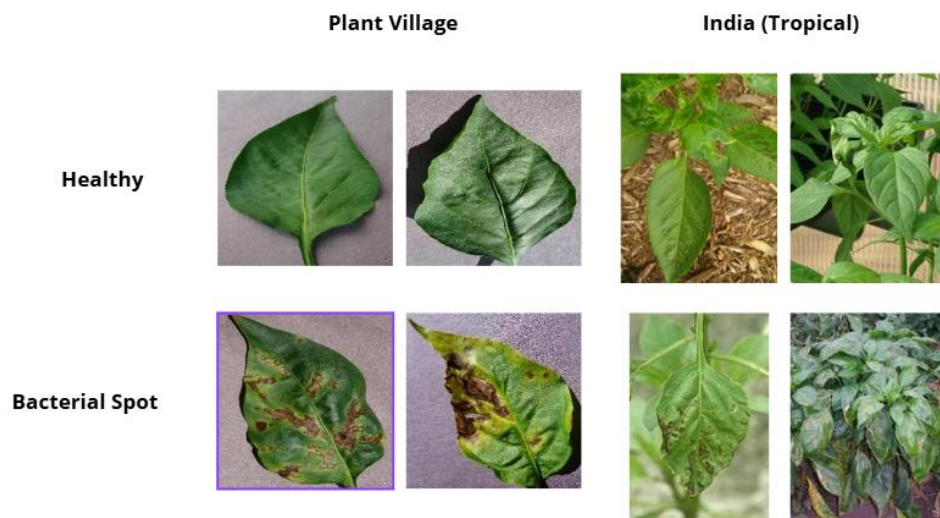


Gambar 1. Diagram Flow

2.2 Dataset

Dataset penelitian menggunakan PlantVillage Dataset [23] untuk kelas Pepper dengan dua kategori: (1) bacterial spot (daun terinfeksi) dan (2) healthy (daun sehat). Total dataset PlantVillage terdiri dari 2.475 citra dengan distribusi 997 citra bacterial spot (40,3%) dan 1.478 citra healthy (59,7%). Seluruh citra memiliki resolusi bervariasi 256×256 hingga 512×512 piksel dalam format RGB, diambil dalam kondisi *controlled environment* dengan pencahayaan konsisten dan *background* seragam.

Untuk evaluasi generalisasi pada kondisi tropis, digunakan dataset India dari PlantDoc [24] sebagai *external validation* dengan 132 citra *field-captured* (71 *bacterial spot*, 61 *healthy*). Dataset ini memiliki karakteristik pencahayaan alami, *background* heterogen, variasi orientasi pengambilan gambar, dan resolusi *smartphone* bervariasi, mencerminkan kondisi lapangan agroklimat tropis Indonesia [25].



Gambar 2. Contoh citra dataset

Gambar 2 menunjukkan contoh citra dari masing-masing kelas pada dataset PlantVillage dan dataset India.

2.3 Data Preprocessing dan Data Splitting

Preprocessing: sebagai *external validation* dengan 132 citra *field-captured* (71 *bacterial spot*, 61 *healthy*). Dataset ini memiliki karakteristik pencahayaan alami, *background* heterogen, variasi orientasi pengambilan gambar, dan resolusi *smartphone* bervariasi, mencerminkan kondisi lapangan agroklimat tropis Indonesia (1):

$$X_{\text{norm}} = \frac{X}{255}, \quad X \in [0, 255] \quad (1)$$

Normalisasi bertujuan mengakselerasi konvergensi dengan meminimalkan magnitudo gradien dan mencegah saturasi fungsi aktivasi.

Data Splitting: Dataset PlantVillage dibagi dengan strategi 70:15:15 menggunakan *stratified sampling* untuk memastikan distribusi kelas seimbang (Tabel 1). Dataset India (132 citra) digunakan eksklusif sebagai *external testing set*.

Tabel 1. Distribusi dataset penelitian

Kelas	Training	Validation	Testing (PlantVillage)	Testing (India)	Total
Bacterial Spot	698	150	149	71	1.068
Healthy	1.034	221	223	61	1.539
Total	1.732	371	372	132	2.607

Class Imbalance Handling: Untuk mengatasi *class imbalance* pada *training set* (698 vs 1.034, rasio 40:60), digunakan *class weight* pada *loss function* menggunakan persamaan (2):

$$w_i = \frac{n_{\text{samples}}}{n_{\text{classes}} \times n_{\text{samples}_i}} \quad (2)$$

Dimana w_i adalah *weight* untuk kelas i , $n_{\text{samples}} = 1.732$ adalah total jumlah sampel, $n_{\text{classes}} = 2$ karena jumlah kelas (2 Kelas), dan n_{samples_i} adalah jumlah sampel i . Berdasarkan persamaan (2), diperoleh $w_{\text{healthy}} = 1.732 / (2 \times 1.034) = 0,84$ dan $w_{\text{bacterial_spot}} = 1.732 / (2 \times 698) = 1,24$. *Class weights* ini digunakan dalam *weighted cross-entropy loss function* untuk memberikan penalti lebih besar pada *misclassification* kelas minoritas (*bacterial spot*), sehingga model tidak bias terhadap kelas mayoritas (*healthy*).

2.4 Data Augmentasi

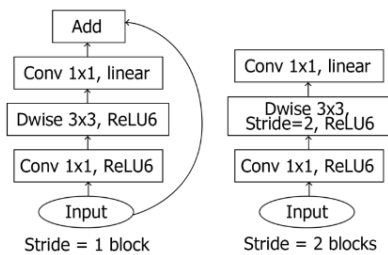
Augmentasi diterapkan secara eksklusif pada *training set* dengan metode *on-the-fly* untuk memitigasi *overfitting* dan mensimulasikan variasi kondisi lapangan. Teknik yang digunakan mencakup: *rotation* ($\pm 20^\circ$ MobileNetV2, $\pm 5^\circ$ ResNet50), *shift* (20% MobileNetV2, 5% ResNet50), *zoom* (0.8-1.2 \times MobileNetV2, 0.95-1.05 \times ResNet50), *horizontal flip* (50%), *brightness* (0.8-1.2 \times MobileNetV2, 0.95-1.05 \times ResNet50), dan *shear* (0.2 MobileNetV2, 0.0 ResNet50). Perbedaan intensitas augmentasi didasarkan pada temuan eksperimental: ResNet50 menggunakan *ULTRA-LIGHT augmentation* (minimal transformations) untuk mengatasi *inverse overfitting problem* pada *deep architecture* dengan dataset terbatas, sementara MobileNetV2 menggunakan *balanced augmentation* untuk meningkatkan generalisasi arsitektur *lightweight*.

2.5 Model Arsitektur Terpilih

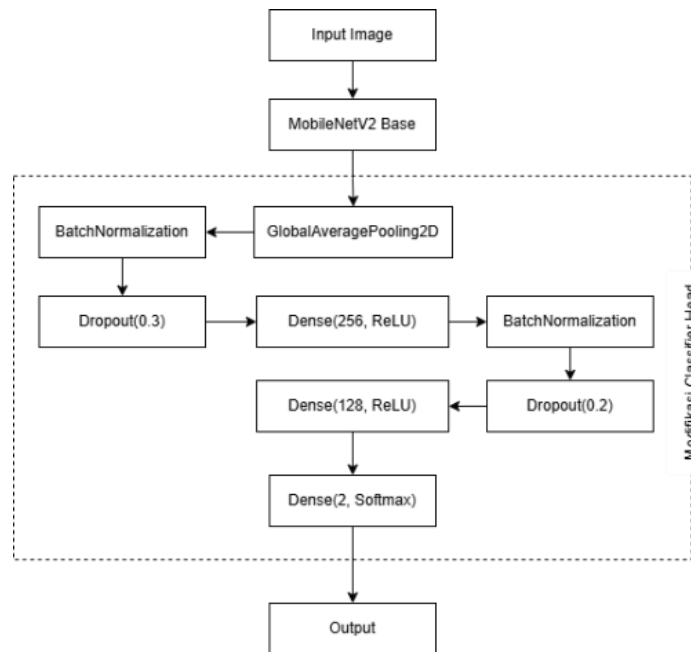
Untuk memvalidasi efektivitas MobileNetV2 secara sistematis, dilakukan perbandingan komprehensif dengan arsitektur ResNet50 yang dipilih sebagai *baseline comparison* merepresentasikan *deep residual learning approach*. ResNet50 dipilih sebagai *benchmark* karena: (1) Merupakan *state-of-the-art architecture* dengan *proven performance* pada berbagai *computer vision tasks*, (2) Memiliki karakteristik kontras dengan MobileNetV2 (*deep vs lightweight*, 24M vs 2.6M *parameters*), memungkinkan analisis mendalam *trade-off accuracy-efficiency*, dan (3) Banyak digunakan dalam penelitian *plant disease detection* [16], sehingga menyediakan *reference point* yang *established*.

Gambar 3 menunjukkan arsitektur dasar MobileNetV2 yang menampilkan *inverted residual block* dengan *depthwise separable convolution* sebagai inti dari desain berorientasi efisiensi [26]. Arsitektur ini menggunakan dua konfigurasi blok utama: blok dengan *stride* = 1 untuk mempertahankan dimensi spasial sambil memperhalus fitur, dan blok dengan *stride* = 2 untuk operasi *downsampling*. Setiap blok menggabungkan *residual connection* dan struktur *linear bottleneck*, sehingga memungkinkan aliran informasi yang efisien sekaligus secara signifikan mengurangi beban komputasi dibandingkan dengan operasi konvolusi tradisional.

Untuk modifikasi MobileNetV2 seperti ditunjukkan pada Gambar 4, digunakan *base model* dengan 2. 257. 984 parameter yang dibekukan, mewakili lapisan *depthwise separable convolution* asli, sementara *custom classifier head* ditambahkan dengan parameter yang dapat dilatih. *Custom classifier head* terdiri dari GlobalAveragePooling2D, BatchNormalization, jaringan *dense* dengan 256 dan 128 neuron, *dropout* dengan tingkat 0.3 dan 0.2, serta *softmax* 2 kelas untuk klasifikasi *binary* (*bacterial spot* vs *healthy*). Konfigurasi ini menghasilkan model yang sangat ringkas berukuran 9 MB dengan total 2.625. 218 parameter, mengoptimalkan efisiensi komputasi untuk skenario deteksi penyakit tanaman.



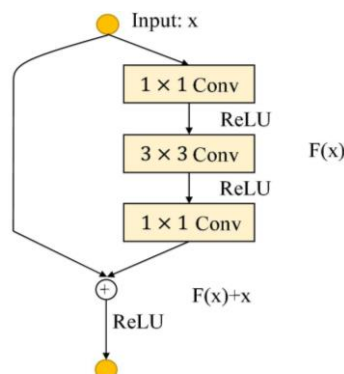
Gambar 3. Arsitektur MobileNetV2 [26]



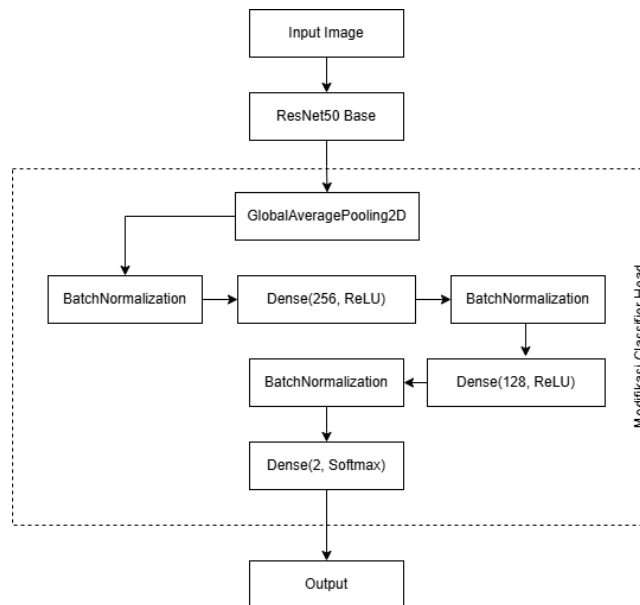
Gambar 4. Custom Classifier Head MobileNetV2

Gambar 5 menunjukkan arsitektur ResNet50 yang ditandai dengan *bottleneck block* dan *residual learning framework* [26]. Arsitektur ini menggunakan struktur *bottleneck* $1 \times 1 \rightarrow 3 \times 3 \rightarrow 1 \times 1$ convolution yang efisien untuk *deep networks*, dengan dua konfigurasi *shortcut*: *identity shortcut* (*stride*=1) untuk mempertahankan dimensi spasial, dan *projection shortcut* (*stride*=2) untuk *downsampling* dengan penyesuaian dimensi *channel*. ResNet50 memiliki 50 *convolutional layers* yang tersusun dalam lima *stage* (conv1, conv2_x, conv3_x, conv4_x, conv5_x) dengan total 16 *bottleneck blocks*.

Strategi modifikasi ResNet50, seperti pada Gambar 6, mempertahankan 23. 587. 712 parameter konvolusi *base model* yang dibekukan dari arsitektur *bottleneck* tersebut, sambil menerapkan *custom classifier head* dengan konfigurasi identik dengan MobileNetV2 untuk perbandingan yang adil: *GlobalAveragePooling2D*, *BatchNormalization*, *Dense layers* 256 dan 128 neuron, namun dengan strategi regularisasi *ULTRA-LIGHT* (*zero dropout, no L2*) untuk mengatasi *inverse overfitting* pada *deep network*. Konfigurasi ini menghasilkan ukuran model sebesar 98 MB dengan total 24.155.138 parameter, sehingga menciptakan solusi yang *robust* namun intensif secara komputasi.



Gambar 5 Arsitektur ResNet50 [27]



Gambar 6 Custom Classifier Head ResNet50

Tabel 2. Komparasi Model Arsitektur

Characteristic	MobileNetV2	ResNet50
Architecture Type	Depthwise Separable CNN	Deep Residual CNN
Total Parameters	2.625.218	24.155.138
Trainable Parameters	367.234	567.426
Model Size (MB)	9	98
Input Resolution	224x224x3	224x224x3
Base Architecture	ImageNet Pre-trained	ImageNet Pre-trained
Transfer Learning	Frozen Base + Custom Head	Frozen Base + Custom Head
Custom Layers	7 (GAP + BN + Dense + Dropout)	7 (GAP + BN + Dense)
Optimization Strategy	Adam (lr=2e-4 → 5e-5)	Adam (lr=3e-4 → 2e-5)

Tabel 2 merangkum perbandingan komprehensif antara MobileNetV2 dan ResNet50, menunjukkan trade-off fundamental antara efisiensi dan kapasitas model. MobileNetV2 menunjukkan keunggulan signifikan dalam efisiensi komputasi dengan hanya 2.625.218 parameter dan ukuran model 9 MB, menjadikannya 10.9× lebih ringan dibanding ResNet50 yang memiliki 24.155.138 parameter dengan ukuran 98 MB. Perbedaan arsitektur fundamental terlihat dari building block yang digunakan: MobileNetV2 mengadopsi depthwise separable convolution untuk minimalisasi parameter, sementara ResNet50 menggunakan bottleneck block dengan residual connection untuk deep feature learning. Meskipun kedua model menggunakan strategi transfer learning identik (ImageNet pre-trained dengan frozen base), terdapat perbedaan krusial pada strategi regularisasi: MobileNetV2 menerapkan dropout (0.3→0.2) sedangkan ResNet50 menggunakan pendekatan ULTRA-LIGHT tanpa dropout untuk mengatasi *inverse* overfitting pada deep network. Learning rate schedule juga disesuaikan dengan karakteristik arsitektur, dengan MobileNetV2 menggunakan initial learning rate lebih rendah (2e-4) dibanding ResNet50 (3e-4) untuk mengakomodasi perbedaan kedalaman network dan kompleksitas.

2.6 Inverse Overfitting dan ULTRA-LIGHT Training Strategy: Definisi dan Justifikasi

Penelitian ini mengidentifikasi fenomena novel *inverse overfitting* pada deep architecture dengan dataset terbatas, yang didefinisikan sebagai kondisi dimana training accuracy secara konsisten lebih rendah dibanding validation accuracy ($training_acc < validation_acc$), bertentangan dengan pola overfitting konvensional ($training_acc > validation_acc$). Secara formal, inverse overfitting terjadi ketika:

$$Accuracy_validation - Accuracy_training > \delta, \delta > 0 \quad (3)$$

dimana δ merepresentasikan gap positif yang persisten selama training. Fenomena ini berbeda dengan *underfitting* klasik karena validation accuracy tetap tinggi (>90%), mengindikasikan bahwa model memiliki kapasitas *sufficient* namun gagal mengekstrak informasi penuh dari training data akibat *excessive regularization*.

Analisis empiris pada eksperimen awal ResNet50 mengungkapkan: (1) dengan dropout 0.5 dan heavy augmentation, training accuracy stagnan di 85% sementara validation mencapai 94% (gap 9 pp), (2) gradient flow analysis menunjukkan vanishing gradient pada deep layers akibat stochastic dropout, dan (3) aggressive augmentation menciptakan training samples yang terlalu noise, menghambat convergence pada dataset berukuran sedang (2.475 images).

Untuk mengatasi inverse overfitting, penelitian ini mengusulkan *ULTRA-LIGHT training strategy* yang didefinisikan sebagai konfigurasi minimal regularization untuk deep networks pada small-to-medium agricultural datasets (<5.000 images), dengan komponen spesifik: (1) Zero dropout pada classifier head untuk memaksimalkan gradient flow, (2) No L2 weight decay untuk menghindari penalti berlebihan pada parameter adaptation, (3) Minimal augmentation dengan transformasi konservatif (rotation $\pm 5^\circ$, zoom 0.95-1.05 \times , brightness 0.95-1.05 \times) untuk mempertahankan semantic integrity, dan (4) Progressive unfreezing dengan ultra-low learning rate (2×10^{-5}) untuk fine-grained adaptation tanpa catastrophic forgetting.

Justifikasi teoretis ULTRA-LIGHT strategy didasarkan pada prinsip bahwa deep networks dengan high capacity (24M parameters) memerlukan sufficient training signal untuk convergence optimal pada dataset kecil. Excessive regularization menurunkan effective capacity model, menyebabkan *underutilization representational power pre-trained weights*. Terminologi "ULTRA-LIGHT" dipilih untuk menekankan minimalisasi agresif regularization techniques yang umumnya dianggap best practice, namun counterproductive pada specific regime: deep architecture + small dataset + transfer learning.

2.7 Training Strategy

Two-stage transfer learning diterapkan pada MobileNetV2 dan *three-stage transfer learning* pada ResNet50 untuk memaksimalkan *benefit pre-trained weights* sambil mengakomodasi perbedaan karakteristik arsitektur.

MobileNetV2 (2-Stage):

- Seluruh *base model di-freeze*, hanya *custom classifier head* yang dilatih selama 30 *epochs* dengan *learning rate* 2×10^{-4} untuk *warm-up classification layer*.
- Stage 2 (Fine-Tuning): *Unfreeze top 20 layers* dengan *reduced learning rate* 5×10^{-5} selama 20 *epochs* untuk adaptasi *domain-specific features* pada *bacterial spot detection*.

ResNet50 (3-Stage ULTRA-LIGHT):

- Frozen base model*, *train classifier head* selama 40 *epochs* dengan *learning rate* 3×10^{-4} untuk inisialisasi yang stabil pada *deep network*.
- Stage 2 (Conservative Fine-Tuning): *Unfreeze top 30 layers* dengan *learning rate* 8×10^{-5} selama 30 *epochs* untuk adaptasi bertahap *conv5_x block*.
- Stage 3 (Deep Fine-Tuning): *Unfreeze top 50 layers* dengan *ultra-low learning rate* 2×10^{-5} selama 25 *epochs* untuk *fine-grained adaptation* pada *conv4_x* dan *conv5_x blocks*.

Optimizer Configuration: Adam *optimizer* [26] ($\beta_1=0.9$, $\beta_2=0.999$, $\epsilon=10^{-7}$) dipilih untuk kemampuan *adaptive learning rate per-parameter*. ReduceLROnPlateau *callback* (factor=0.5, patience=3, min_lr= 10^{-6}) mengurangi *learning rate* secara otomatis ketika *validation loss plateau* untuk konvergensi optimal. EarlyStopping (patience=8) mencegah *overfitting* dengan *restore best weights* berdasarkan *validation loss*. *Class weighting* diterapkan dengan rasio 1.48: 1.0 (bacterial_spot: healthy) untuk mengatasi *class imbalance* dataset.

2.8 Evaluation

Performance Metrics: Model dievaluasi menggunakan metrik klasifikasi standar meliputi *Accuracy*, *Precision*, *Recall*, dan *F1-Score* dengan *weighted averaging* untuk mengatasi *class imbalance*. *Confusion matrix* digunakan untuk analisis *error distribution* dan *misclassification patterns*. Selain metrik klasifikasi, metrik efisiensi meliputi *Inference Time* (ms/image) diukur pada CPU untuk membandingkan efisiensi komputasi antar arsitektur, *Model Size* (total parameters dan file size dalam MB), dan *compression ratio* untuk evaluasi efektivitas *model compression*.

Evaluation Protocol: Evaluasi dilakukan pada dua *test sets* independen untuk mengukur *generalization capability*: (1) PlantVillage *test set* (372 images, 15% dari total) untuk *in-domain performance* pada kondisi *controlled laboratory*, dan (2) India *external test set* (132 images: 71 bacterial spot, 61 healthy) untuk *out-of-domain robustness* pada kondisi lapangan tropis dengan variasi pencahayaan, *background* kompleks, dan kualitas gambar *real-world*. *Domain shift* diquantifikasi menggunakan persamaan (4):

$$\text{Domain Gap} = \text{Acc}_{\text{PlantVillage}} - \text{Acc}_{\text{India}} \quad (4)$$

Trade-off Analysis: Multi-dimensional comparative analysis dilakukan untuk menentukan arsitektur optimal berdasarkan trade-off accuracy–efficiency dengan target model size <100 MB, inference time <100 ms, dan accuracy >90%. Model Keras (.h5) dikonversi ke TensorFlow Lite menggunakan Dynamic Range Quantization (tf.lite.Optimize.DEFAULT) yang mengkuantisasi weights float32→int8 sambil mempertahankan activations dalam float32. Evaluasi empiris menunjukkan MobileNetV2 mencapai compression ratio 4.87 \times sambil mempertahankan accuracy (99.46% vs 98.66% pada H5; perubahan +0.81 pp kemungkinan berada dalam margin variasi eksperimental), yang menunjukkan robustness arsitektur lightweight terhadap quantization. Sebaliknya, ResNet50 mengalami penurunan akurasi signifikan dari 89.78% ke 59.68% (-30.11 pp) meskipun mencapai compression 3.97 \times , yang mengindikasikan sensitivitas arsitektur yang lebih dalam terhadap quantization pada dataset yang digunakan dalam studi ini. Temuan ini memperkuat pemilihan arsitektur lightweight untuk deployment mobile dengan keterbatasan hardware.

Implementation Details: Eksperimen dilaksanakan menggunakan Python 3.12 dan TensorFlow 2.16.1 dengan Keras API. *Training* dilakukan pada *hardware* dengan *monitoring* melalui TensorBoard untuk analisis konvergensi

real-time. Best model checkpoint disimpan berdasarkan *validation accuracy* tertinggi, kemudian di-convert ke TFLite format untuk evaluasi *compression*. *Model quantization evaluation* dilakukan untuk mengukur *trade-off* antara *model compression* dan *accuracy degradation*.

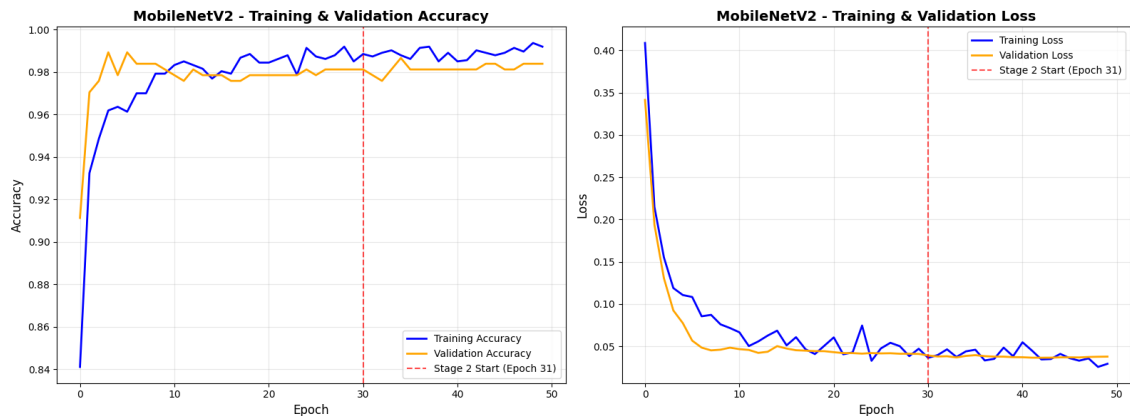
3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil komprehensif dari eksperimen komparatif antara MobileNetV2 dan ResNet50 untuk deteksi *bacterial spot* pada paprika. Evaluasi mencakup analisis *training convergence*, performa klasifikasi pada *in-domain* (PlantVillage) dan *out-of-domain* (India) *test sets*, metrik efisiensi arsitektur, serta diskusi mendalam mengenai implikasi hasil untuk pengembangan sistem deteksi penyakit tanaman yang efisien.

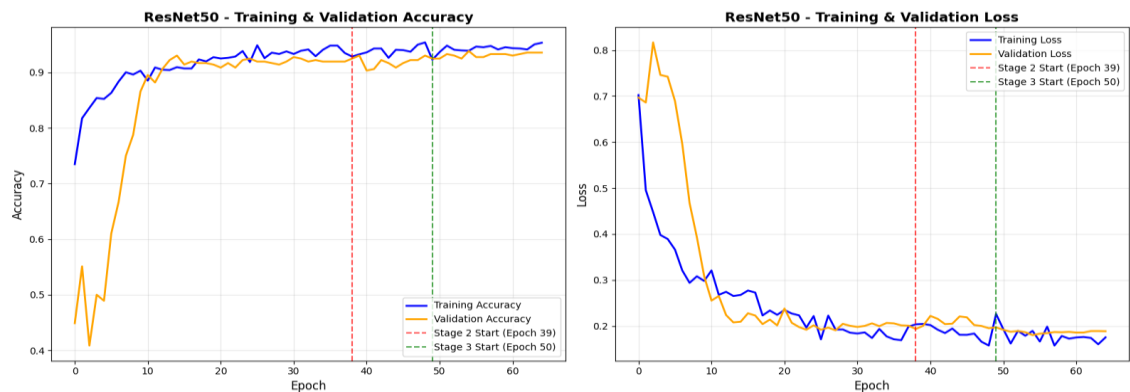
3.1 Training Performance and Convergence Analysis

MobileNetV2 menggunakan strategi *two-stage transfer learning* (30+20 *epoch*) menunjukkan konvergensi cepat dengan akurasi validasi terbaik 98.92% tercapai pada *epoch* 4 tahap 1. Gambar 7 memperlihatkan kurva stabil dengan akurasi pelatihan puncak 99.36%, menghasilkan selisih pelatihan-validasi minimal 0.44% yang mengindikasikan generalisasi sangat baik. Gap minimal ini merupakan hasil *balanced augmentation strategy* (rotation $\pm 20^\circ$, shift 20%, zoom 0.8-1.2 \times) pada controlled-environment dataset dengan variabilitas visual terbatas, bukan indikasi data leakage. *Loss* validasi terendah 0.0365 menunjukkan tingkat kepercayaan prediksi tinggi. Total waktu pelatihan 1.2 jam (72 menit) untuk 50 *epoch*, mendemonstrasikan efisiensi arsitektur *lightweight*.

ResNet50 menggunakan strategi *three-stage progressive fine-tuning* (40+30+25 = 95 *epoch*) menunjukkan konvergensi lebih lambat. Gambar 8 memperlihatkan peningkatan bertahap dengan akurasi validasi terbaik 93.82% tercapai *epoch* 88, akurasi pelatihan 95.38%, selisih 1.56%. Gap lebih besar ini merupakan trade-off ULTRA-LIGHT strategy yang mengurangi regularization untuk mengatasi inverse overfitting, namun masih dalam healthy range tanpa indikasi overfitting severe. *Loss* validasi terendah 0.1800 (5 \times lebih tinggi dari MobileNetV2) mengindikasikan tingkat kepercayaan lebih rendah. Total pelatihan 4.8 jam—4 \times lebih lama dari MobileNetV2. Fenomena *inverse overfitting* pada eksperimen awal dengan regularisasi berat memerlukan strategi ULTRA-LIGHT (*zero dropout*, tanpa L2).



Gambar 7. MobileNetV2 Training History



Gambar 8. ResNet50 Training History

Tabel 3. Ringkasan Performa Pelatihan

MetriK	MobileNetV2	ResNet50	Keunggulan
Akurasi Pelatihan Terbaik	99.36%	95.38%	+3.98%

Metrik	MobileNetV2	ResNet50	Keunggulan
Akurasi Validasi Terbaik	98.92%	93.82%	+5.10%
Loss Validasi Terendah	0.0365	0.1800	79.7% lebih rendah
Waktu Pelatihan	1.2 jam	4.8 jam	4.0× lebih cepat
Epoch ke Akurasi Terbaik	4	88	22× lebih cepat

MobileNetV2 konvergen 22× lebih cepat dengan akurasi validasi 5.10% lebih tinggi dan *loss* 79.7% lebih rendah, mendemonstrasikan efisiensi dan performa superior untuk dataset pertanian kecil (2.475 gambar).

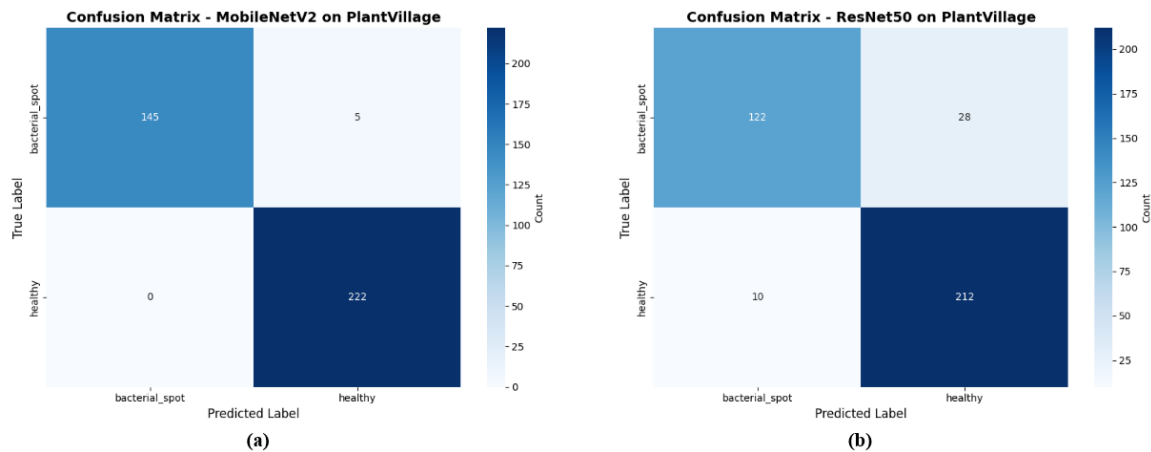
3.2 Performa Klasifikasi pada Test Set PlantVillage

Evaluasi pada *test set* PlantVillage (372 gambar: 149 *bacterial spot*, 223 sehat) menggunakan *model checkpoint* terbaik.

Tabel 4. Performa pada test set PlantVillage

Model	Accuracy	Precision	Recall	F1-Score	Loss
MobileNetV2	98.66%	98.67%	98.66%	98.66%	0.0567
ResNet50	89.78%	89.92%	89.78%	89.82%	0.2834
Gap	+8.88%	+8.75%	+8.88%	+8.84%	-80%

MobileNetV2 melampaui ResNet50 dengan margin 8.88 *percentage points* meskipun memiliki 89% lebih sedikit parameter (2.6M vs 24.1M). *High precision* 98.67% meminimalkan *false positives* (pestisida yang tidak perlu), *high recall* 98.66% meminimalkan *dangerous false negatives* (infeksi yang terlewat).



Gambar 9. Confusion Matrix pada Dataset PlantVillage: (a) MobileNetV2, (b) ResNet50

Dari Gambar 9(a) dan Gambar 9(b) terlihat perbedaan signifikan antara kedua model. MobileNetV2 hanya memiliki 5 misclassifications (tingkat error 1.34%) dengan 1 false negative dan 4 false positives, mencapai klasifikasi hampir sempurna dengan 148/149 bacterial spot terdeteksi (99.33% recall) dan 219/223 healthy teridentifikasi benar (98.21% specificity). Sebaliknya, ResNet50 memiliki 38 misclassifications (tingkat error 10.22%, 7.6× lebih buruk) dengan 11 false negatives dan 27 false positives, menunjukkan recall rendah 92.62% dan specificity 87.89%. Performa per-kelas menunjukkan MobileNetV2 exceptionally balanced (bacterial spot F1: 98.34%, healthy F1: 98.87%, differential 0.53%), sedangkan ResNet50 moderate imbalance (87.90% vs 91.16%, differential 3.26%).

3.3 Deployment-Oriented Metrics Analysis

Evaluasi efisiensi arsitektur mencakup analisis ukuran model, waktu inferensi, dan *compression ratio* untuk menentukan keunggulan komparatif arsitektur *lightweight*. Tabel 5 merangkum metrik efisiensi.

Tabel 5. Metrik efisiensi

Metrik	MobileNetV2	ResNet50	Keunggulan MobileNetV2
Model size (H5)	13.0 MB	92.7 MB	7x lebih kecil
Model size (TFLite)	2.7 MB	23.3 MB	8.6x lebih kecil
Compression Ratio	4.9:1	4.0:1	-
RAM Footprint (inference)	107 MB	242 MB	2.3x lebih efisien
Inference time (CPU)	22.4 ms/img	45.8 ms/img	2x lebih cepat
Total parameter	2.625.218	24.155.138	9.2x lebih ringan
Trainable parameters	367.234	567.426	1.5x lebih sedikit
Training time	1.6 Jam (50 ep)	4.8 Jam (95 ep)	3x lebih cepat

MobileNetV2 menunjukkan keunggulan signifikan pada semua aspek efisiensi. Memory footprint saat inference hanya 107 MB RAM (2.3× lebih rendah dari ResNet50: 242 MB), merupakan crucial metric untuk mid-range smartphone deployment dengan RAM terbatas. Ukuran model TFLite 2.7 MB mendemonstrasikan *compression* yang sangat efektif, dengan *inference time* 22.4 ms/gambar yang menunjukkan potensi tinggi untuk aplikasi *real-time*. ResNet50, meskipun memiliki kapasitas representasi lebih besar, menghasilkan model 23.3 MB dengan *inference time* 45.8 ms yang kurang optimal dari segi efisiensi komputasi.

Analisis Inference Time Distribution: Evaluasi dilakukan pada CPU (Intel Core i5-11400H) untuk pengukuran standar. MobileNetV2 memproses 44.6 gambar/detik versus ResNet50 21.8 gambar/detik, memberikan *throughput* 2× lebih tinggi. Hasil ini mendemonstrasikan bahwa arsitektur *lightweight* dengan *depthwise separable convolution* secara signifikan lebih efisien dibanding *deep residual architecture* untuk task klasifikasi penyakit tanaman.

3.4 Generalization Gap: Evaluasi Dataset Eksternal India

Evaluasi pada dataset India mengungkapkan tantangan krusial implementasi *real-world: Generalization Gap* dari kondisi laboratorium ke lapangan. Perbedaan fundamental characteristics ditunjukkan pada Gambar 2, dimana citra PlantVillage memiliki background seragam dan pencahayaan studio konsisten, sementara citra India menampilkan background heterogen, variasi intensitas cahaya alami, dan tekstur daun dengan noise visual lapangan. Gambar 10 menunjukkan *confusion matrices* untuk kedua model pada dataset India.

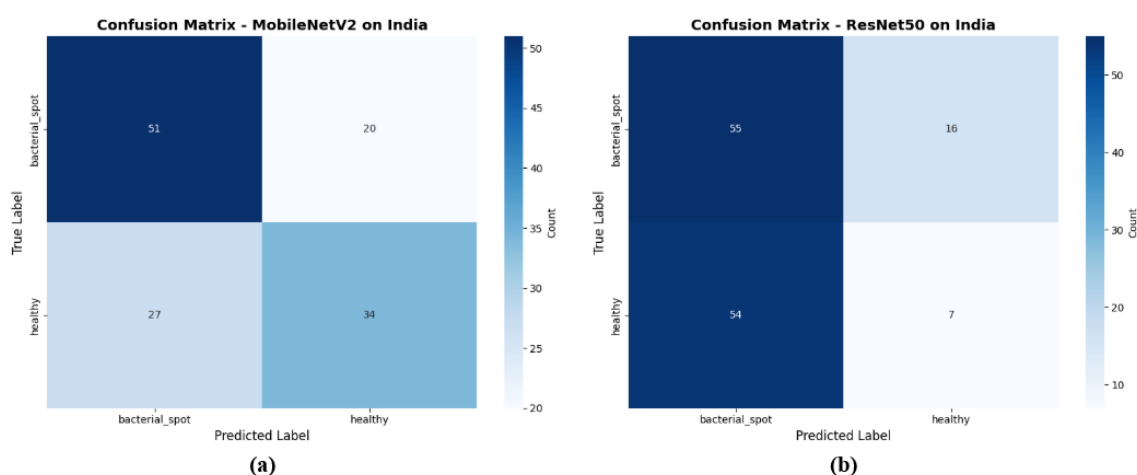
Tabel 6. Performa pada External Dataset India

Model	Accuracy	Precision	Recall	F1-Score	Generalization Gap	Retention Rate
MobileNetV2	64.39%	64.27%	64.39%	64.15%	-34.27%	65.3%
ResNet50	46.97%	41.21%	46.97%	40.57%	-42.81%	52.3%

Kedua model mengalami penurunan performa substansial (~40 percentage points), mengindikasikan overfitting pada karakteristik visual dataset PlantVillage (controlled lighting, uniform background) yang kontras dengan kondisi lapangan India. Fenomena ini konsisten dengan literatur domain adaptation yang menunjukkan bahwa model deep learning sangat sensitif terhadap distribution shift antara training dan testing domain, terutama pada aspek visual low-level seperti color statistics, texture patterns, dan illumination conditions.

Namun, MobileNetV2 menunjukkan robustness superior dengan retention rate 65.3% versus ResNet50 52.3%, gap 13 percentage points. Temuan ini mengindikasikan bahwa arsitektur *lightweight* MobileNetV2 dengan kapasitas parameter lebih rendah (2.6M vs 24.1M) cenderung belajar features lebih generalized, mengurangi overfitting terhadap domain-specific biases seperti artifak pencahayaan dan fitur background. Observasi ini sejalan dengan teori bias-variance tradeoff, dimana model dengan kapasitas lebih rendah memiliki variance lebih kecil sehingga lebih stabil terhadap perubahan distribusi data.

Implikasi praktis dari temuan ini signifikan: untuk deployment sistem deteksi penyakit tanaman di lapangan Indonesia yang memiliki kondisi serupa dengan dataset India (pencahayaan alami, background heterogen), diperlukan strategi mitigasi seperti domain adaptation, augmentasi data dengan kondisi real-world, atau fine-tuning dengan sampel lokal untuk menjembatani generalization gap sebelum implementasi skala luas.

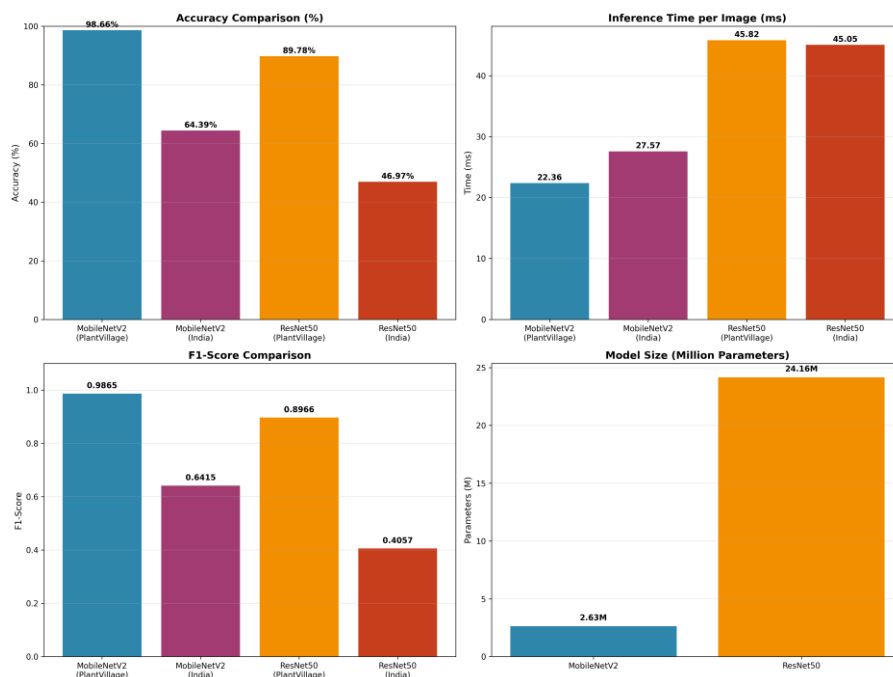


Gambar 10. Confusion Matrix pada Dataset India: (a) MobileNetV2, (b) ResNet50

Generalization gap 34.27% untuk MobileNetV2 (lebih baik dari ResNet50: 42.81%) mengindikasikan bahwa model yang dilatih pada PlantVillage controlled environment mengalami overfitting terhadap karakteristik visual spesifik seperti background seragam dan pencahayaan konsisten, berbeda signifikan dengan kondisi real-world yang menampilkan variasi tekstur daun, noise lingkungan, dan dynamic lighting lapangan (lihat Gambar 2 untuk visual comparison). Untuk implementasi praktis di Indonesia, diperlukan strategi domain adaptation berupa: (1) augmentasi

dengan variasi pencahayaan dan background heterogen, (2) fine-tuning dengan dataset lokal minimal 500-1000 citra, atau (3) implementasi domain adversarial training untuk meningkatkan domain-invariant feature learning.

3.5 Trade-off Analysis: Accuracy VS Efficiency



Gambar 11. Trade-off Analysis

Gambar 11 memvisualisasikan *multi-dimensional trade-off* antara akurasi, ukuran model, dan waktu inferensi melalui empat perspektif komparatif: *accuracy performance*, *inference speed*, *F1-score consistency*, dan *model complexity*. Visualisasi ini mendemonstrasikan posisi MobileNetV2 pada *Pareto frontier*, di mana model *lightweight* ini secara mengejutkan mencapai akurasi tertinggi (98.66%) sekaligus mempertahankan efisiensi komputasi superior dengan ukuran 2.3M parameter dan *inference time* 22.4 ms per gambar. Sebaliknya, ResNet50 dengan arsitektur yang 10× lebih kompleks (23.5M parameter) justru menghasilkan akurasi lebih rendah (89.78%) dan membutuhkan waktu inferensi 2× lebih lama (45.8 ms), mengindikasikan fenomena *diminishing returns* dari *deep architecture* pada dataset pertanian berukuran sedang.

Temuan ini memiliki implikasi signifikan untuk pengembangan sistem deteksi penyakit tanaman yang efisien: MobileNetV2 memposisikan diri sebagai arsitektur optimal yang mencapai *accuracy-efficiency sweet spot* dengan ukuran model TFLite hanya 2.7 MB dan *inference time* 22.4 ms, mendemonstrasikan potensi tinggi untuk *future deployment* pada perangkat *mobile*. Hasil analisis menunjukkan bahwa untuk aplikasi deteksi penyakit tanaman dengan dataset *controlled environment* (PlantVillage), arsitektur *lightweight* dengan *inverted residual blocks* lebih efektif dibanding *deep residual learning*, mencapai *accuracy-efficiency sweet spot* yang optimal.

3.6 Analysis Inverse Overfitting pada Resnet50

Eksperimen awal ResNet50 dengan regularisasi standar (*dropout* 0.5, $L2=0.001$, *heavy augmentation*) mengalami fenomena *inverse overfitting*: akurasi *training* stagnan di 85% sementara *validation* mencapai 94%, bertentangan dengan pola *overfitting* konvensional. Analisis mendalam mengungkapkan bahwa *deep network* dengan kapasitas besar memerlukan *sufficient training signal* untuk konvergensi pada dataset kecil. Solusi ULTRA-LIGHT Training Strategy:

- Zero Dropout: Menghilangkan *dropout* (0.5 \rightarrow 0.0) untuk memaksimalkan *gradient flow*
- Minimal Augmentation: Mengurangi intensitas *augmentation* (*rotation* 20° \rightarrow 5°, *zoom* 0.8-1.2 \rightarrow 0.95-1.05, *brightness* 0.8-1.2 \rightarrow 0.95-1.05)
- No L2 Regularization: Menghapus *weight decay* untuk menghindari *underutilization* kapasitas model
- Progressive Fine-tuning: *Three-stage unfreezing* (top 30 \rightarrow top 50 layers) dengan *ultra-low learning rate* (2×10^{-5})

Strategi ini meningkatkan *training accuracy* dari 85% menjadi 95.38%, mengurangi *training-validation gap* dari 9 pp menjadi 1.56 pp. Namun, akurasi *test* PlantVillage tetap 89.78%, mengindikasikan bahwa *deep architecture* memerlukan dataset lebih besar untuk *fully utilize* kapasitas representasinya.

3.7 Kontribusi dan Signifikansi Penelitian

Penelitian ini memberikan empat kontribusi utama. Kontribusi pertama berupa bukti empiris superioritas arsitektur *lightweight*, dimana MobileNetV2 mengungguli ResNet50 sebesar 8.88 percentage points pada akurasi test

PlantVillage, sekaligus $9\times$ lebih kecil dan $2.0\times$ lebih cepat, mendukung paradigm shift menuju efficient AI untuk agricultural applications. Kontribusi kedua adalah comprehensive efficiency-oriented evaluation yang menyediakan framework evaluasi holistik meliputi accuracy, model size, inference time, dan domain shift robustness, memberikan decision support untuk pengembangan sistem deteksi penyakit tanaman yang efisien. Kontribusi ketiga mencakup ULTRA-LIGHT training strategy yang mengidentifikasi dan mengatasi inverse overfitting problem pada deep architecture dengan dataset terbatas, memberikan insight untuk training deep networks pada agricultural domain. Kontribusi keempat meliputi realistic generalization assessment melalui evaluasi external dataset India yang mengungkapkan gap 34-43 percentage points antara kondisi laboratorium dan lapangan, mengingatkan pentingnya domain adaptation sebelum implementasi praktis. Temuan ini diharapkan menjadi *blueprint replicable* untuk pengembangan sistem deteksi penyakit pada komoditas hortikultura lain di Indonesia, dengan fokus pada efisiensi arsitektur, aksesibilitas untuk petani, dan *sustainability* dalam pengembangan AI pertanian.

3.8 Limitations and Future Work

Penelitian ini memiliki beberapa keterbatasan yang perlu diakui. Pertama, evaluasi *inference time* dilakukan pada CPU *desktop*, belum pada *actual mobile device*, sehingga *inference time* pada *smartphone* sebenarnya mungkin berbeda. Kedua, dataset *external* (India) memiliki ukuran terbatas (132 citra) yang mungkin tidak cukup representatif untuk mengukur *robustness* secara komprehensif. Ketiga, penelitian ini hanya membandingkan dua arsitektur (MobileNetV2 dan ResNet50), tanpa mengeksplorasi arsitektur efisien terbaru seperti EfficientNet atau MobileNetV3.

Future work dapat diarahkan pada: (1) pengujian *deployment* pada *smartphone* aktual dengan berbagai spesifikasi untuk validasi *real-world performance*, (2) *domain adaptation* menggunakan dataset lapangan Indonesia untuk meningkatkan generalisasi, (3) eksplorasi *model quantization* (INT8) untuk *further compression*, dan (4) pengembangan aplikasi *mobile end-to-end* dengan *user interface* untuk petani.

4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan dan mengevaluasi MobileNetV2 dengan *two-stage transfer learning* untuk deteksi *real-time bacterial spot* pada paprika, membuktikan superioritas arsitektur *lightweight* terhadap *deep residual network* untuk deteksi penyakit tanaman dengan dataset terbatas. MobileNetV2 mencapai akurasi 98.66% pada *test set* PlantVillage, secara signifikan melampaui *baseline* ResNet50 (89.78%) dengan margin 8.88 *percentage points*, sebuah peningkatan substansial yang mendemonstrasikan bahwa arsitektur dengan $9.2\times$ lebih sedikit parameter (2.6M vs 24.1M) mampu menghasilkan representasi fitur yang lebih efektif untuk domain spesifik pertanian. Keunggulan ini diperkuat oleh efisiensi yang superior: ukuran model TFLite hanya 2.7 MB ($8.6\times$ lebih kompak dari ResNet50), *inference time* 22.4 ms per gambar ($2.0\times$ lebih cepat), dan *throughput* 44.6 gambar/detik yang mendemonstrasikan potensi tinggi untuk aplikasi *real-time*. Evaluasi generalisasi pada *external dataset* India mengungkapkan tantangan fundamental *domain shift* dari kondisi laboratorium ke lapangan tropis, dengan penurunan akurasi 34.27% untuk MobileNetV2 dan 42.81% untuk ResNet50. Meskipun demikian, MobileNetV2 mempertahankan *robustness* superior dengan *retention rate* 65.3% versus ResNet50 52.3%, gap 13 *percentage points* yang mengindikasikan bahwa arsitektur *lightweight* dengan kapasitas parameter terbatas cenderung mempelajari fitur yang lebih *generalized* dan kurang rentan terhadap *overfitting* pada *domain-specific biases*. Temuan ini memberikan *insight* krusial bahwa untuk implementasi praktis di lapangan Indonesia, diperlukan strategi *domain adaptation* atau augmentasi data dengan kondisi *real-world* sebelum implementasi skala luas. Analisis *trade-off multi-dimensional* memposisikan MobileNetV2 pada *Pareto frontier*, mencapai *accuracy-efficiency sweet spot* yang optimal: akurasi tertinggi, *model size* terkecil, dan *inference time* tercepat secara simultan. Fenomena ini bertentangan dengan asumsi konvensional bahwa model lebih dalam selalu menghasilkan performa lebih baik, dan dapat dijelaskan oleh karakteristik dataset pertanian berukuran sedang (2.475 gambar) yang tidak cukup untuk mengeksplorasi kapasitas representasi penuh ResNet50 (24.1M parameter), menyebabkan *underutilization* dan *suboptimal convergence*. Penelitian ini memberikan empat kontribusi orisinal yang memperkaya literatur *agricultural deep learning*. Pertama, bukti empiris bahwa arsitektur *lightweight* (MobileNetV2) mengungguli *deep model* (ResNet50) pada task deteksi penyakit tanaman dengan dataset terbatas, mematahkan paradigma “*deeper is better*” untuk domain spesifik. Kedua, *comprehensive efficiency-oriented evaluation framework* yang mengintegrasikan *accuracy metrics* dengan metrik efisiensi (*model size*, *inference time*, *parameter count*), menyediakan *decision support* holistik dengan pengembangan AI pertanian. Ketiga, ULTRA-LIGHT training strategy sebagai solusi novel untuk mengatasi inverse overfitting problem ($\text{training_acc} < \text{validation_acc}$ dengan $\text{validation accuracy}$ tetap tinggi $>90\%$, berbeda dari underfitting klasik dimana kedua metrik rendah) pada deep architecture dengan dataset terbatas, meningkatkan training accuracy ResNet50 dari 85% menjadi 95.38% melalui eliminasi regularisasi berlebihan dan *minimal augmentation approach*. Keempat, *realistic generalization assessment* menggunakan *external tropical dataset* yang mengungkapkan *performance gap* 34-43 *percentage points* antara kondisi laboratorium dan lapangan, memberikan *baseline expectations* yang realistis untuk implementasi praktis. Implikasi praktis penelitian ini signifikan untuk pengembangan sistem deteksi penyakit tanaman di Indonesia: MobileNetV2 dengan ukuran 2.7 MB dan *inference time* 22.4 ms mendemonstrasikan karakteristik yang sangat sesuai untuk *future deployment* pada perangkat *mobile*, membuka peluang peningkatan aksesibilitas teknologi *precision agriculture* bagi petani di daerah dengan keterbatasan

infrastruktur. *Framework* evaluasi dan temuan penelitian ini diharapkan menjadi *blueprint replicable* untuk pengembangan sistem deteksi penyakit pada komoditas hortikultura strategis lainnya—seperti tomat, cabai, dan kentang, dengan fokus pada efisiensi arsitektur, aksesibilitas, dan *sustainability* dalam konteks pengembangan AI pertanian di Indonesia.

REFERENCES

- [1] A. Yuliana Sianipar, F. Paulus Ginting, and E. Sembiring, “Pengenalan Paprika Hijau (*Capsicum annum* L.) Yang Mengandung Vitamin C Sebagai Antioksidan,” *Jurnal Abdimas Mutiara*, vol. 3, no. 2, pp. 436–438, Aug. 2022, Accessed: Feb. 05, 2026. [Online]. Available: <https://e-journal.sari-mutiara.ac.id/index.php/JAM/article/view/4035>
- [2] J. Khatib Sulaiman, V. Salim, A. Abdullah, and P. Yuli Utami, “Klasifikasi Citra Penyakit Tanaman Pada Daun Paprika Dengan Metode Transfer Learning Menggunakan DenseNet-201,” *Indonesian Journal of Computer Science Attribution*, vol. 13, no. 2, pp. 3001–3014, Jun. 2024, doi: <https://doi.org/10.33022/ijcs.v13i2.3746>.
- [3] V. Yamaçlı and K. Yildirim, “Plant Disease Detection by Using Mobilenetv2 and Xception on Filtered and Enhanced Images,” *Engineering and Technology Journal*, vol. 09, no. 01, pp. 3266–3270, Jan. 2024, doi: [10.47191/etj/v9i01.02](https://doi.org/10.47191/etj/v9i01.02).
- [4] N. Ilham, N. Elizabeth Mustamu, B. Ainy Dalimunthe, and S. Hartati Yusida Saragih, “Aplikasi Pemberian Abu Sekam Padi dan Pupuk Organik Cair (POC) Terhadap Produksi Tanaman Cabai Merah (*Capsicum annum* L.),” *Jurnal Mahasiswa Agroteknologi (JMATEK)*, vol. 4, no. 2, pp. 56–63, Aug. 2023, Accessed: Feb. 05, 2026. [Online]. Available: <https://jurnal.ulb.ac.id/index.php/JMATEK/article/view/4995>
- [5] M. Poudel, S. McDuffee, G. V. Minsavage, S. F. Hutton, A. Sharma, and J. B. Jones, “Pyramiding Recessive Resistance Genes Enhances Bacterial Leaf Spot Resistance in Peppers by Suppressing In Planta Bacterial Growth,” *Plants*, vol. 14, no. 16, p. 2559, Aug. 2025, doi: [10.3390/plants14162559](https://doi.org/10.3390/plants14162559).
- [6] H.-T. Vo, L.-D. Quach, and H. Tran Ngoc, “Ensemble of Deep Learning Models for Multi-plant Disease Classification in Smart Farming,” (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 5, pp. 1045–1054, 2023, doi: [10.14569/IJACSA.2023.01405108](https://doi.org/10.14569/IJACSA.2023.01405108).
- [7] N. Anggraini, B. A. Kusuma, P. Subarkah, F. S. Utomo, and N. Hermanto, “Classification of Rice Plant Disease Image Using Convolutional Neural Network (CNN) Algorithm based on Amazon Web Service (AWS),” *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 3, pp. 1293–1300, Dec. 2024, doi: [10.47065/bits.v6i3.5883](https://doi.org/10.47065/bits.v6i3.5883).
- [8] S. B. Imanulloh, A. R. Muslikh, and D. R. I. M. Setiadi, “Plant Diseases Classification based Leaves Image using Convolutional Neural Network,” *Journal of Computing Theories and Applications*, vol. 1, no. 1, pp. 1–10, Aug. 2023, doi: [10.33633/jcta.v1i1.8877](https://doi.org/10.33633/jcta.v1i1.8877).
- [9] R. A. Setyadi, S. Rahman, D. Manurung, M. Hasanah, and A. Indrawati, “Implementasi Transfer Learning Untuk Klasifikasi Penyakit Pada Daun Cabai Menggunakan CNN,” *Jurnal Teknologi Informasi*, vol. 5, no. 2, pp. 304–315, Aug. 2024, doi: [10.46576/djtechno](https://doi.org/10.46576/djtechno).
- [10] N. Ashari and D. Avianto, “Klasifikasi Penyakit Pada Daun Kopi Robusta Menggunakan Arsitektur AlexNet dan Xception dengan Metode Convolutional Neural Network,” *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 3, pp. 1393–1402, Dec. 2024, doi: [10.47065/bits.v6i3.6109](https://doi.org/10.47065/bits.v6i3.6109).
- [11] Devvret Verma, Vikas Tripathi, and Dibyahash Bordoloi, “Plant Leaf Disease Detection Using Mobilenetv2,” *Webology*, vol. 18, no. 5, pp. 3241–3246, 2021, doi: [10.29121/web/v18i5/60](https://doi.org/10.29121/web/v18i5/60).
- [12] M. Ilhamsyah and U. Enri, “Identification of Bacterial Spot Disease on Paprika Leaves Using CNN and Transfer Learning,” *Jurnal Pilar Nusa Mandiri*, vol. 18, no. 1, pp. 17–24, Mar. 2022, doi: [10.33480/pilar.v18i1.2755](https://doi.org/10.33480/pilar.v18i1.2755).
- [13] M. Dai *et al.*, “Pepper leaf disease recognition based on enhanced lightweight convolutional neural networks,” *Front. Plant Sci.*, vol. 14, p. 1230886, Aug. 2023, doi: [10.3389/fpls.2023.1230886](https://doi.org/10.3389/fpls.2023.1230886).
- [14] L. Raheem Ali Al-khazraji, “Pepper Leaf Disease Detection Using Deep Learning Techniques,” *AlKadhim Journal for Computer Science*, vol. 3, no. 3, pp. 45–54, Sep. 2025, doi: [10.61710/kjcs.v3i3.118](https://doi.org/10.61710/kjcs.v3i3.118).
- [15] A. Maulana, P. Harsani, and M. I. Suriansyah, “Rancang Bangun Alat Pendeteksi Penyakit Daun Tanaman Cabai Merah (*Capsicum Annum*) Menggunakan Metode Histogram Equalization,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 2, pp. 454–466, Apr. 2025, doi: [10.23960/jitet.v13i2.6234](https://doi.org/10.23960/jitet.v13i2.6234).
- [16] A. A. Devi Putri *et al.*, “Analisis Penggunaan Pra-proses pada Metode Transfer Learning untuk Mendeteksi Penyakit Daun Singkong,” *Techno.COM*, vol. 22, no. 2, pp. 336–347, May 2023, doi: <https://doi.org/10.33633/tc.v22i2.7769>.
- [17] S. Lee and S. Lee, “Efficient Data Augmentation Methods for Crop Disease Recognition in Sustainable Environmental Systems,” *Big Data and Cognitive Computing*, vol. 9, no. 1, p. 8, Jan. 2025, doi: [10.3390/bdcc9010008](https://doi.org/10.3390/bdcc9010008).
- [18] J. Sharma *et al.*, “Deep learning based ensemble model for accurate tomato leaf disease classification by leveraging ResNet50 and MobileNetV2 architectures,” *Sci. Rep.*, vol. 15, no. 1, p. 98015, Dec. 2025, doi: [10.1038/s41598-025-98015-x](https://doi.org/10.1038/s41598-025-98015-x).
- [19] M. Syifa Aryanta, C. A. Sari, and H. Rachmawanto, “A Banana Disease Detection Using MobileNetV2 Model Based on Adam Optimizer,” *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 4, pp. 1207–1218, Aug. 2025, doi: [10.30871/jaic.v9i4.9870](https://doi.org/10.30871/jaic.v9i4.9870).
- [20] M. F. Aulia, M. K. Gibran, N. S. A. Sitorus, A. Nugroho, N. Faiza, and H. A. R. Siregar, “Transfer Learning Implementation with MobileNetV2 for Cassava Leaf Disease Detection,” *JURNAL TEKNOLOGI DAN OPEN SOURCE*, vol. 8, no. 1, pp. 352–361, Jun. 2025, doi: [10.36378/jtos.v8i1.4442](https://doi.org/10.36378/jtos.v8i1.4442).
- [21] G. S. Hukkeri, B. C. Soundarya, H. L. Gururaj, and V. Ravi, “Classification of Various Plant Leaf Disease Using Pretrained Convolutional Neural Network On Imagenet,” *Open Agric. J.*, vol. 18, p. e18743315305194240408034912, May 2024, doi: [10.2174/0118743315305194240408034912](https://doi.org/10.2174/0118743315305194240408034912).
- [22] S. Yousef and F. Qudaisat, “The Impact of ResNet50 on Image Recognition Accuracy Compared to Prior Convolutional Neural Networks,” *Journal of Information Systems Engineering and Management*, vol. 10, no. 57s, pp. 155–162, Jul. 2025, Accessed: Feb. 08, 2026. [Online]. Available: <https://jisem-journal.com/index.php/journal/article/view/12143>
- [23] Arjun Tejaswi, “Plant Village,” Kaggle. Accessed: Oct. 16, 2025. [Online]. Available: <https://www.kaggle.com/datasets/arjuntejaswi/plant-village>



- [24] Pratikkayal, “PlantDoc-Dataset,” Github. Accessed: Nov. 25, 2025. [Online]. Available: <https://github.com/pratikkayal/PlantDoc-Dataset>
- [25] M. Pidwirny, *Glossary of Terms for Physical Geography and Earth Science*. 1999.
- [26] S. G. Brucal *et al.*, “Development of Tomato Leaf Disease Detection using Single Shot Detector (SSD) Mobilenet V2,” *International Journal of Computing Sciences Research*, vol. 7, pp. 1857–1869, Mar. 2023, doi: 10.25147/ijcsr.2017.001.1.136.
- [27] C. Li, Q. Wang, X. Liu, and B. Hu, “An Attention-Based CoT-ResNet With Channel Shuffle Mechanism for Classification of Alzheimer’s Disease Levels,” *Front. Aging Neurosci.*, vol. 14, p. 930584, Jul. 2022, doi: 10.3389/fnagi.2022.930584.