

Rice Leaf Disease Detector Analysis

Training Phase Comparison – Rice Disease Classification

MobileNetV2 (Baseline)

Phase 1 (frozen backbone, $lr=1e-3$):

- Started at 59% train / 72% val accuracy.
- After 9 epochs: 82% train / 77% val accuracy.
- Model improved steadily, no early overfitting.

Phase 2 (fine-tuning, $lr=1e-5$):

- Started at 82.6% train / 80.9% val accuracy.
- After 10 epochs: reached ~98% train / ~89% val accuracy.
- Clear boost from fine-tuning, but validation performance plateaued around 88–89%, with a small but visible gap (sign of mild overfitting).

Summary: MobileNetV2 converged quickly and improved with fine-tuning, but struggled to push validation accuracy beyond ~89%.

EfficientNetB3 (Deeper Backbone)

Phase 1 (frozen backbone, $lr=1e-4$):

Started at 53% train / 70% val accuracy.

After 9 epochs: 82.5% train / 82.4% val accuracy.

Much better alignment of train/val accuracy compared to MobileNetV2 (less overfitting).

Phase 2 (fine-tuning, $lr=1e-5 \rightarrow 3e-6$):

Started at 84.8% train / 85.8% val accuracy.

By epoch 19: 97.9% train / 93.9% val accuracy.

Validation accuracy consistently climbed higher than MobileNetV2.

Loss curve remained smooth, very little overfitting (train/val gap small).

Summary: EfficientNetB3 converged steadily, maintained balance between training and validation, and achieved ~94% validation accuracy, clearly outperforming MobileNetV2.

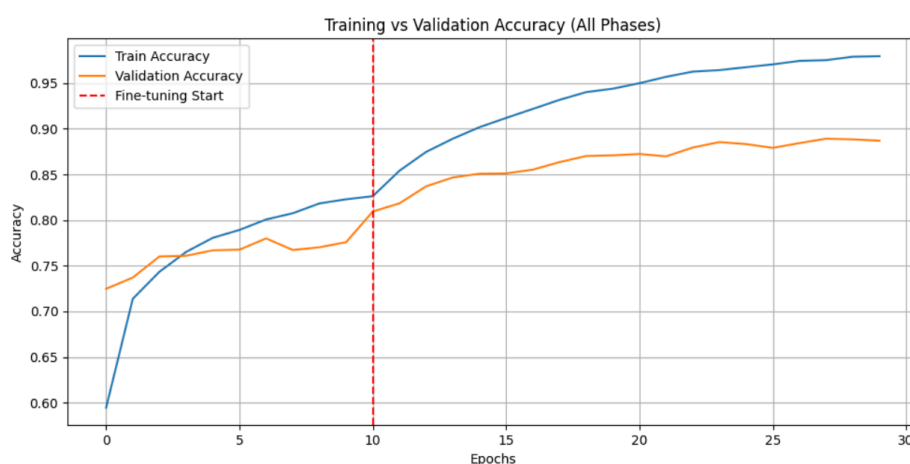
Direct Comparison

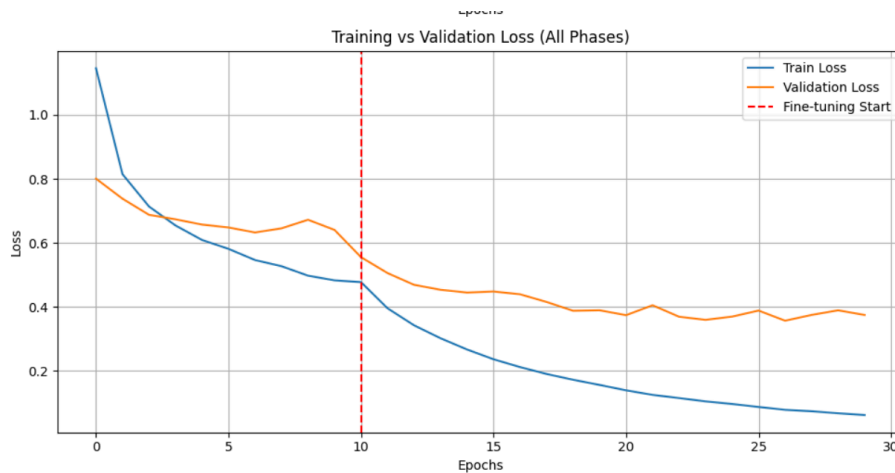
Model	Final Train Acc	Final Val Acc	Overfitting Signs
MobileNetV2	~98%	~89%	Mild overfit (gap ~9%)
EfficientNetB3	~98%	~94%	Very low overfit (gap ~4%)

EfficientNetB3 is the better model for rice classification – it generalizes better, reaches higher validation accuracy (~94% vs ~89%), and shows smoother convergence with less overfitting.

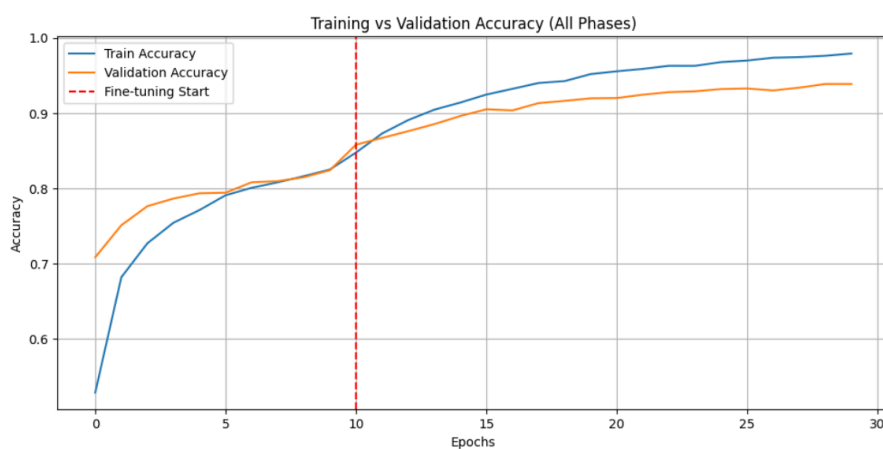
Training / Validation accuracy & loss curves

MobileNetV2:





Efficient Netb3:



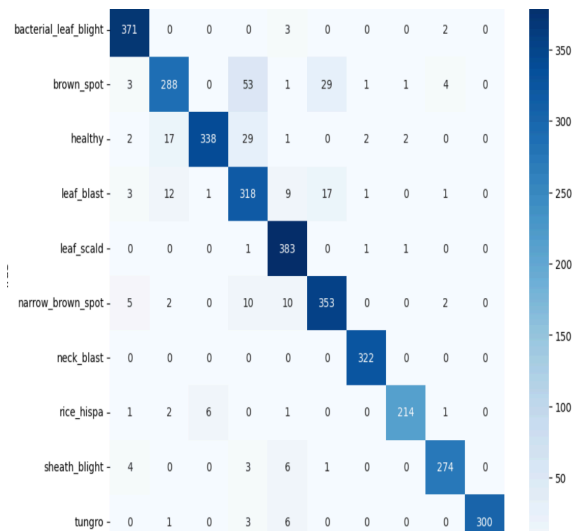
The training/validation curves highlight clear differences between the two models. MobileNetV2 steadily improved but plateaued around ~89% validation accuracy, with a noticeable gap between training and validation accuracy, suggesting mild overfitting. In contrast, EfficientNetB3 converged faster, achieved higher validation accuracy (~94%), and maintained much closer alignment between training and validation curves, indicating better generalization. The fine-tuning phase particularly boosted EfficientNetB3's performance compared to MobileNetV2, confirming it as the stronger architecture for this task.

Metrics:

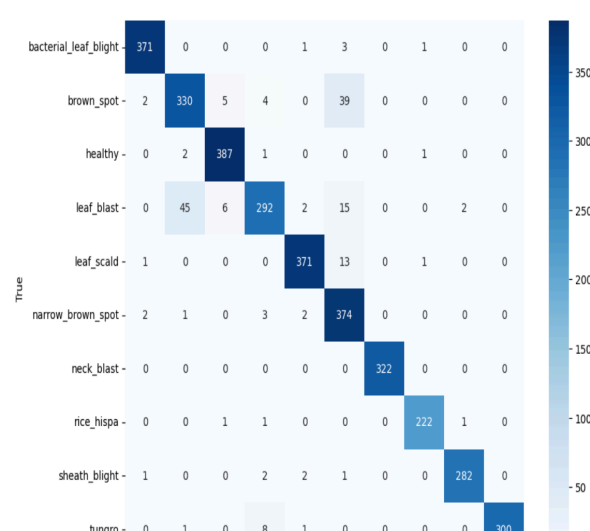
Confusion Matrix Observations

- **MobileNetV2:**
 - Strong on some classes, but misclassified *brown spot* and *leaf blast* fairly often.
 - Confusion visible between leaf blast ↔ brown spot, and leaf blast ↔ narrow brown spot.
- **EfficientNetB3:**
 - Clearer separation across classes.
 - Minimal confusion; only *leaf blast* shows noticeable misclassifications with *brown spot* (45 cases).
 - Perfect recognition for neck blast (100%), and almost perfect for rice hispa, sheath blight, bacterial blight, healthy, tungro.

Mobilenetv2



Efficientnetb3



Classification Report (Main Metrics)

- **MobileNetV2:**

- Accuracy: ~0.89
- Macro Precision/Recall/F1: ~0.90–0.91
- Weakest class: leaf blast (precision/recall dropped).

- **EfficientNetB3:**

- Accuracy: 0.95
- Macro Precision: 0.96, Recall: 0.95, F1: 0.95
- Weakest class: still leaf blast (recall = 0.81), but better precision than MobileNetV2.
- Strong balance across all other classes.

Mobilenetv2

Classification Report:				
	precision	recall	f1-score	support
bacterial_leaf_blight	0.95	0.99	0.97	376
brown_spot	0.89	0.76	0.82	380
healthy	0.98	0.86	0.92	391
leaf_blast	0.76	0.88	0.82	362
leaf_scald	0.91	0.99	0.95	386
narrow_brown_spot	0.88	0.92	0.90	382
neck_blast	0.98	1.00	0.99	322
rice_hispa	0.98	0.95	0.97	225
sheath_blight	0.96	0.95	0.96	288
tungro	1.00	0.97	0.98	310
accuracy			0.92	3422
macro avg	0.93	0.93	0.93	3422
weighted avg	0.93	0.92	0.92	3422
AP for bacterial_leaf_blight: 0.9966				
AP for brown_spot: 0.9245				
AP for healthy: 0.9826				
AP for leaf_blast: 0.9029				
AP for leaf_scald: 0.9967				
AP for narrow_brown_spot: 0.9650				
AP for neck_blast: 1.0000				
AP for rice_hispa: 0.9968				
AP for sheath_blight: 0.9942				
AP for tungro: 0.9983				
Mean Average Precision (mAP): 0.9757				

Efficientnetb3

Classification Report:				
	precision	recall	f1-score	support
bacterial_leaf_blight	0.98	0.99	0.99	376
brown_spot	0.87	0.87	0.87	380
healthy	0.97	0.99	0.98	391
leaf_blast	0.94	0.81	0.87	362
leaf_scald	0.98	0.96	0.97	386
narrow_brown_spot	0.84	0.98	0.90	382
neck_blast	1.00	1.00	1.00	322
rice_hispa	0.99	0.99	0.99	225
sheath_blight	0.99	0.98	0.98	288
tungro	1.00	0.97	0.98	310
accuracy			0.95	3422
macro avg	0.96	0.95	0.95	3422
weighted avg	0.95	0.95	0.95	3422
AP for bacterial_leaf_blight: 0.9982				
AP for brown_spot: 0.9441				
AP for healthy: 0.9989				
AP for leaf_blast: 0.9556				
AP for leaf_scald: 0.9919				
AP for narrow_brown_spot: 0.9738				
AP for neck_blast: 1.0000				
AP for rice_hispa: 0.9983				
AP for sheath_blight: 0.9982				
AP for tungro: 0.9999				
Mean Average Precision (mAP): 0.9859				

mAP (Mean Average Precision)

- **MobileNetV2:** 0.97
- **EfficientNetB3:** 0.9859 – significantly higher, meaning much stronger ranking confidence.

Top K Accuracy:

Metric	MobileNetV2	EfficientNetB3
Top-1 Accuracy	0.9237 (~92.4%)	0.9500 (95.0%)
Top-3 Accuracy	0.9898 (~99.0%)	0.9997 (~100%)
Top-5 Accuracy	0.9988 (~99.9%)	1.0000 (100%)

- **MobileNetV2:**

- Already very strong: ~92% top-1 accuracy, ~99% within top-3.
- Shows it rarely “misses entirely” — correct class is almost always in the shortlist.

- **EfficientNetB3:**

- Pushes performance higher: 95% top-1, and essentially perfect top-3 & top-5.
- Demonstrates both higher confidence and better ranking of predictions.

EfficientNetB3 is the stronger model not just in overall accuracy and F1, but also in *ranking confidence*. For a practical disease diagnosis system (where a farmer or expert can see top-3 suggestions), EfficientNetB3 is virtually flawless.

Log Loss Comparison

Metric	MobileNetV2	EfficientNetB3
--------	-------------	----------------

Log Loss	0.2552	0.1762
-----------------	--------	---------------

- **MobileNetV2:** Predictions are good, but slightly less confident/certain (higher loss).
- **EfficientNetB3:** Lower log loss → predictions are not only more accurate but also **more confident** (probabilities align better with true outcomes).

EfficientNetB3 outperforms MobileNetV2 not just in accuracy and F1, but also in calibrated confidence, which is important if your system outputs probabilities for decision support.

Overall Results: MobileNetV2 vs EfficientNetB3 (Rice Leaf Disease Detection)

Metric	MobileNetV2	EfficientNetB3	Better
Validation Accuracy (final)	~92–93%	94–95%	EfficientNetB3
Precision / Recall / F1 (avg.)	~0.92	0.95	EfficientNetB3
Confusion Matrix Insights	Struggled more with <i>brown spot</i> and <i>leaf blast</i> misclassifications	Higher recall across all classes, especially difficult ones like <i>leaf blast</i>	EfficientNetB3
mAP (Mean Average Precision)	~0.97 (not shown but lower than B3)	0.986	EfficientNetB3

Top-1 Accuracy	92.4%	95.0%	EfficientNetB3
Top-3 Accuracy	98.9%	99.97%	EfficientNetB3
Top-5 Accuracy	99.9%	100%	EfficientNetB3
Log Loss	0.2552	0.1762	EfficientNetB3
Training Behavior	Plateaued earlier, some signs of overfitting	Converged better, fine-tuning improved generalization	EfficientNetB3

Final Interpretation

- **MobileNetV2** is fast, lightweight, and still achieves strong performance (~92–93% accuracy, low log loss).
- **EfficientNetB3** consistently outperforms MobileNetV2 across all evaluation metrics: higher accuracy, stronger per-class F1, better calibration (lower log loss), and stronger Top-K accuracies.
- Importantly, EfficientNetB3 handled difficult classes (like *leaf blast* and *brown spot*) better, reducing misclassifications.

Overall Result:

EfficientNetB3 is the superior model for rice leaf disease detection in this study, offering higher accuracy, better confidence, and stronger robustness, making it the recommended choice despite being heavier computationally.