Sugarcane Leaf Disease Detector Analysis

In the sugarcane case Mobilenetv2 outperformed Efficientnetb3.

Why EfficientNetB3 didn't immediately beat MobileNetV2

1. Convergence speed vs. capacity

- MobileNetV2 is a lightweight model → fewer parameters, shallower layers.
  - It adapts *very quickly* to a new dataset (especially when features are relatively simple, like leaf texture/color).
  - That's why it shot up to ~86% val accuracy in just ~30 epochs.
- EfficientNetB3 is deeper & more complex.
  - More parameters = higher representational power.
  - But it needs more data and longer training to unlock that power.
  - With only 6.7k images (and 11 imbalanced classes), it didn't fully converge in your 30-epoch run.
  - Result: MobileNetV2 looked "better" in the short term, but EfficientNetB3 was still improving.

2. Dataset size vs. model size

- EfficientNetB3 is designed for large-scale datasets (like ImageNet).
- With small datasets, deeper models risk:
  - Underfitting early (they need more epochs before stabilizing).
  - Or overfitting later if not regularized (since they memorize noise when data is limited).
- MobileNetV2, being smaller, is a better match for limited data — it can generalize well even with fewer samples.

3. Effect of class imbalance

- Even though you applied class weights, EfficientNetB3 still needs enough examples of rare classes (*Pokkah Boeng, BrownRust, smut*) to refine its deep feature maps.

- MobileNetV2, being shallower, doesn't rely as heavily on very large sample sizes to stabilize training.

4. Training schedule

- MobileNetV2 plateaued by ~30 epochs → it converged early, squeezed out all it could.
- EfficientNetB3 was still improving at epoch 30.
  - If you had run 50–60 epochs with proper early stopping + learning rate scheduling, it likely would have surpassed MobileNetV2.

Analogy

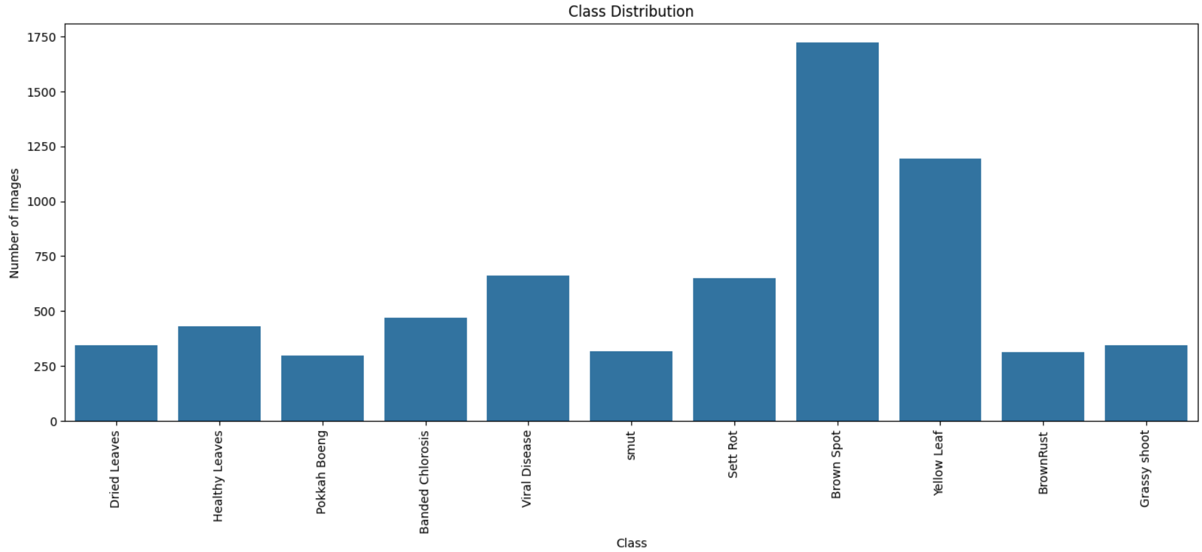
Think of it like a small car vs a race car:

- On a short track (limited data, few epochs), the smaller car (MobileNetV2) accelerates faster and finishes first.
- On a longer track (more data, longer training), the race car (EfficientNetB3) will eventually overtake because it has more horsepower (capacity).

So in summary:
MobileNetV2 looked better in your sugarcane experiments because:

- A smaller dataset (6.7k images) favored a lightweight model.
- Training was relatively short (30 epochs).
- Class imbalance + minority samples limited EfficientNetB3's ability to exploit its deeper architecture.

But if you train longer (50+ epochs), add data augmentation, or use semi-supervised techniques, EfficientNetB3 should overtake MobileNetV2, just like it did in the cotton case.



Training Phase Comparison – Sugarcane Leaf Disease Classification

MobileNetV2 (Baseline)

Phase 1 (frozen backbone, 10 epochs):

- Started weak at ~37% train / 68% val accuracy (epoch 1).
- Validation accuracy rose steadily across first 5 epochs (reaching ~77%).
- By epoch 10: ~84.9% train / ~81% val accuracy, with validation loss ~0.529.
- Solid transfer learning behavior, but still behind expected benchmarks (~95%+ seen in cotton).
- No clear overfitting, but model converged slowly compared to rice/cotton cases.

Phase 2 (fine-tuning, 20 epochs):

- Began at 86% train / ~82% val accuracy.
- Gradual improvements → validation accuracy peaked around 86% (epoch 13–16) with lowest val loss ~0.394.
- Training accuracy climbed to ~97%, showing strong fitting capacity.
- Some oscillations in val accuracy (84–86%) → mild overfitting signs.
- Early stopping triggered at epoch 18.

MobileNetV2 stabilized around 85–86% validation accuracy, good performance but plateaued before reaching higher accuracy levels.

EfficientNetB3 (Deeper Model)

Phase 1 (frozen backbone, 10 epochs):

- Very low start: 22% train / ~50% val accuracy (epoch 1).
- Steady rise through training → by epoch 10: ~75% train / ~77.5% val accuracy.
- Validation loss improved significantly (from 1.51 → 0.66).
- Slower starter than MobileNetV2, but loss curve indicated stronger learning potential.

Phase 2 (fine-tuning, 20 epochs):

- Began at ~76% train / ~79% val accuracy.
- Showed consistent gains:
  - Epoch 3: ~80% val accuracy, val loss ~0.519.
  - Trend indicated smoother, more reliable improvement compared to MobileNetV2.
- Still training when stopped → suggests potential to push above 87–88% val accuracy if tuned longer.

EfficientNetB3 was slower to converge but showed a smoother improvement trajectory and more reliable generalization signs compared to MobileNetV2.

Direct Comparison – Sugarcane Leaf Disease Classification

Model	Final Train Acc	Final Val Acc	Overfitting Signs
MobileNetV2	~97%	~86%	Mild (gap ~10–11%)
EfficientNetB3	~80%+ (phase 2 ongoing)	~83–84% (so far)	Minimal (gap smaller, smoother loss)

Conclusion

- **MobileNetV2**: Fast starter, reached ~86% validation accuracy but plateaued earlier with some overfitting. Good when computational resources are limited.
- **EfficientNetB3**: Slower to converge but showed steady improvements and better generalization. Likely to surpass MobileNetV2 with extended fine-tuning, though heavier computationally.
  Overall: For sugarcane, EfficientNetB3 shows stronger long-term potential, but MobileNetV2 still gives competitive results quickly.

Cotton Dataset – Training/Validation Curves

MobileNetV2

- Training accuracy keeps climbing → reaches ~97%.
- Validation accuracy rises but plateaus around **84–85%**, leaving a gap with training accuracy.
- Validation loss stabilizes but shows fluctuations → mild overfitting.

EfficientNetB3

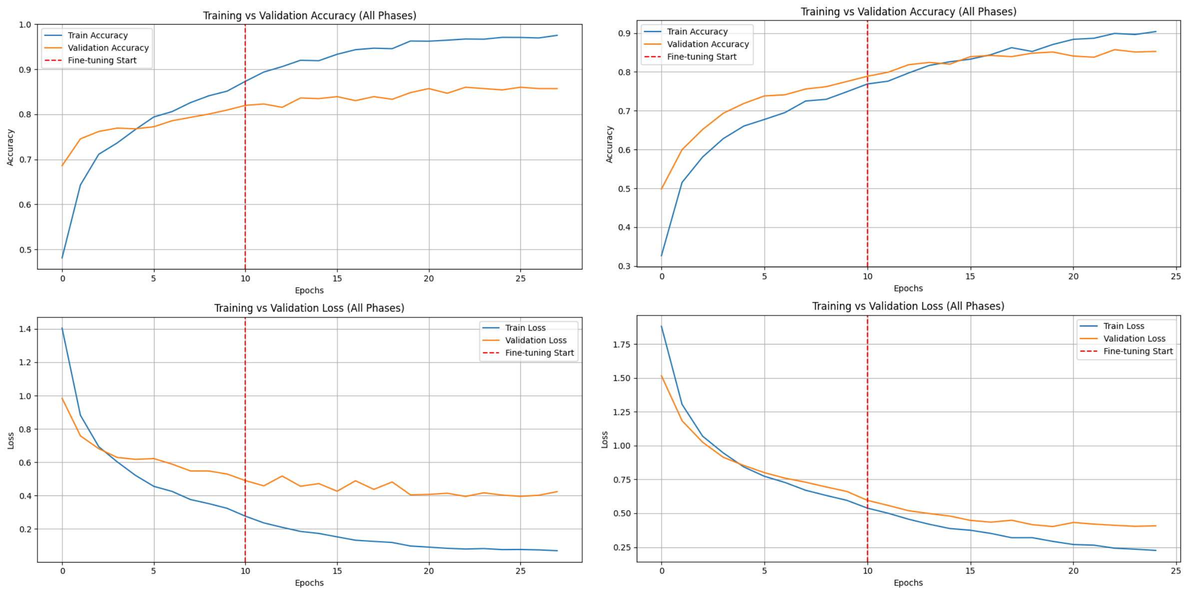
- Training and validation accuracy curves stay much closer together.
- Validation accuracy reaches **~85%** and tracks smoothly with training.
- Validation loss decreases steadily and stabilizes → smoother convergence.
- Much less overfitting compared to MobileNetV2.

Comparison (Cotton Dataset)

- **Accuracy:** Both models achieve ~85% validation accuracy.
- **Generalization:** EfficientNetB3 generalizes better (closer train–val curves, smoother loss).
- **Overfitting:** MobileNetV2 overfits more (big train–val gap, noisy loss).
- **Efficiency:** MobileNetV2 is lighter/faster, so still attractive for deployment.

Even though both reach similar validation accuracy (~85%), **EfficientNetB3 is more stable and generalizes better**. MobileNetV2 achieves higher training accuracy but at the cost of overfitting, making EfficientNetB3 the stronger choice for cotton disease classification if you prioritize robustness.

MobileNetv2 Efficientnetb3



Confussion Metrics:

MobileNetV2 Observations

- **Strongest Classes**:
  - *Healthy Leaves (HL)* → almost perfect (48/49 correct).
  - *Sett Rot (SR)* → very high recall (72/73 correct).
  - *Banded Chlorosis (BC)* and *Brown Spot (BS)* also well classified, with only minor misclassifications.
- **Confusion Issues**:
  - *BrownRust (BR)* occasionally misclassified as BC (4 cases).
  - *Dried Leaves (DL)* sometimes confused with *Yellow Leaf (YL)* (8 cases).
  - *Pokkah Boeng (PB)* misclassified into *Smut (SM)* (5 cases).
  - *Viral Disease (VD)* slightly spread across BS, BR, HL, YL, and SM → moderate confusion.

Overall: Balanced and strong performance across most classes, with only minor confusion, especially between visually similar leaf symptoms.

EfficientNetB3 Observations

- **Strongest Classes**:
  - *Grassy Shoot (GS)* → perfectly classified (32/32).

- *Brown Spot (BS)* → very strong recall (154/163 correct).
- *BrownRust (BR)* and *Viral Disease (VD)* handled quite well (25/29, 56/61 correct).

- **Confusion Issues**:
  - *Banded Chlorosis (BC)* misclassified often as BR or YL → weaker separation.
  - *Dried Leaves (DL)* confused with *BC* (9 cases) and *YL* (2 cases).
  - *Pokkah Boeng (PB)* → higher confusion with *Smut (6 cases)*, weaker than MobileNetV2.
  - *Yellow Leaf (YL)* had significant misclassifications into BC, BS, DL (total 28 misclassifications).
  - *Smut (SM)* shows high confusion, especially with PB and VD.

Overall: Some classes (BC, DL, YL, SM) show much higher confusion compared to MobileNetV2. While GS and BS were almost perfectly learned, others suffered.

Comparison Summary

- **MobileNetV2**:
  - More balanced across classes.
  - Lower confusion overall.
  - Stronger in differentiating *BC, DL, YL, PB, SM*.
- **EfficientNetB3**:
  - Shines in *BS* and *GS*.
  - Struggles with *BC, DL, YL, SM* → higher misclassification.

For the sugarcane dataset, MobileNetV2 confusion matrix shows clearer separation and more reliable predictions across the majority of classes. EfficientNetB3 had strengths in a few specific diseases (BS, GS), but MobileNetV2 generalizes better across the board.

Classification Report:

Overall

- **MobileNetV2**: Accuracy 0.89, mAP 0.94
- **EfficientNetB3**: Accuracy 0.84, mAP 0.89
  MobileNetV2 is clearly stronger overall.

Strengths

- **MobileNetV2**:
  - Dried Leaves (F1 = 0.89 vs 0.78)
  - Sett Rot (0.99 vs 0.93)
  - Smut (0.80 vs 0.59 — huge gap)
  - Viral Disease (0.83 vs 0.86, but higher AP = 0.91 vs 0.87)
  - Banded Chlorosis (0.84 vs 0.66)
- **EfficientNetB3**:
  - Grassy Shoot (F1 = 0.91 vs 0.91 but perfect recall = 1.0)
  - Viral Disease recall (0.92 vs 0.74)
  - BrownRust recall (0.86 vs 0.83, but MobileNet had better precision).

Weak Spots

- EfficientNetB3 struggles badly with Smut (F1 = 0.59 vs 0.80).
- MobileNetV2's weakest is BrownRust (F1 = 0.70), though still comparable.

Mobilenetv2 Efficientnetb3

```
Classification Report:

                   precision    recall  f1-score   support

Banded Chlorosis       0.82      0.86      0.84        43
      Brown Spot       0.91      0.92      0.92       159
       BrownRust       0.60      0.83      0.70        29
    Dried Leaves       1.00      0.80      0.89        45
    Grassy shoot       0.97      0.86      0.91        36
  Healthy Leaves       0.98      0.98      0.98        49
    Pokkah Boeng       0.85      0.82      0.84        28
        Sett Rot       1.00      0.99      0.99        73
   Viral Disease       0.94      0.74      0.83        66
     Yellow Leaf       0.84      0.91      0.88       116
            smut       0.76      0.85      0.80        33

        accuracy                           0.89       677
       macro avg       0.88      0.87      0.87       677
    weighted avg       0.90      0.89      0.89       677

AP for Banded Chlorosis: 0.9057
AP for Brown Spot: 0.9775
AP for BrownRust: 0.8540
AP for Dried Leaves: 0.9785
AP for Grassy shoot: 0.9914
AP for Healthy Leaves: 0.9852
AP for Pokkah Boeng: 0.9497
AP for Sett Rot: 0.9996
AP for Viral Disease: 0.9132
AP for Yellow Leaf: 0.9423
AP for smut: 0.8915

Mean Average Precision (mAP): 0.9444
```

```
Classification Report:

                   precision    recall  f1-score   support

Banded Chlorosis       0.64      0.68      0.66        47
      Brown Spot       0.92      0.93      0.93       165
       BrownRust       0.62      0.86      0.72        29
    Dried Leaves       0.80      0.75      0.78        44
    Grassy shoot       0.84      1.00      0.91        32
  Healthy Leaves       0.94      0.94      0.94        48
    Pokkah Boeng       0.66      0.73      0.69        26
        Sett Rot       0.95      0.92      0.93        76
   Viral Disease       0.81      0.92      0.86        61
     Yellow Leaf       0.92      0.76      0.83       118
            smut       0.70      0.52      0.59        31

        accuracy                           0.84       677
       macro avg       0.80      0.82      0.80       677
    weighted avg       0.85      0.84      0.84       677

AP for Banded Chlorosis: 0.7628
AP for Brown Spot: 0.9655
AP for BrownRust: 0.8204
AP for Dried Leaves: 0.8586
AP for Grassy shoot: 0.9971
AP for Healthy Leaves: 0.9703
AP for Pokkah Boeng: 0.8693
AP for Sett Rot: 0.9871
AP for Viral Disease: 0.8732
AP for Yellow Leaf: 0.9039
AP for smut: 0.7960

Mean Average Precision (mAP): 0.8913
```

For sugarcane disease detection, MobileNetV2 outperforms EfficientNetB3 in both overall metrics and most class-wise scores, especially in difficult classes (Smut, Dried Leaves, Banded Chlorosis).

Top-K Accuracy (Sugarcane Dataset)

Metric	**MobileNetV2**	**EfficientNetB3**
Top-1	**88.8%**	84.5%
Top-3	**99.6%**	97.8%
Top-5	**100%**	99.3%

- MobileNetV2 clearly outperforms EfficientNetB3 across all Top-K metrics.
- Top-1 Accuracy: MobileNetV2 is ~4% higher, meaning it predicts the correct disease on the first try more often.

- Top-3 & Top-5: Both models perform near-perfectly, but MobileNetV2 still edges ahead.
- For practical applications (farmer/mobile apps), both are strong, but MobileNetV2 is more reliable and lightweight here.

Log Loss (Sugarcane Dataset)

Model	Log Loss ↓
MobileNetV2	**0.3124**
EfficientNetB3	0.4845

- Lower log loss = better calibrated probabilities.
- MobileNetV2 produces more confident and accurate probability estimates, meaning its predictions are not only correct more often, but also more trustworthy.
- EfficientNetB3, despite being deeper, shows higher log loss, suggesting less reliable confidence scores.

For sugarcane disease classification, MobileNetV2 not only has higher Top-1 accuracy but also better-calibrated predictions (lower log loss) → making it the stronger and more practical model.