Technical Report: Plant Disease Classification & Decision Pipeline

Task 2: Core ML Implementation

1. Data Preprocessing

To ensure the model performs reliably under real-world field conditions, the following preprocessing pipeline was implemented:

- **Resizing (224x224):** Standardized input dimensions to match the requirements of the MobileNetV3 architecture.
- **Data Augmentation:**
  - **RandomHorizontalFlip:** Simulates various drone approach angles.
  - **ColorJitter (Brightness/Contrast):** Critical for normalizing variations in field lighting caused by cloud cover or time of day.
- **Normalization:** Applied ImageNet statistics to leverage pre-trained weights effectively.
- **Dataset Balancing:** The training set was strictly capped at 5,000 images and balanced (1,100 healthy vs. 1,100 diseased) to prevent class bias.

2. Model Architecture

The system utilizes **MobileNetV3-Small** for the following reasons:

- **Efficiency:** The model size is only **5.9 MB**, making it highly portable for local storage on farm hardware.
- **Speed:** Inference time is approximately **11.7 ms per image** on a standard CPU, enabling real-time processing without a GPU.
- **Offline Capability:** The lightweight nature of the model ensures it can run entirely on a laptop in fields with no internet connectivity.

3. Performance Metrics

The model was evaluated on a held-out test set (330 images) with an emphasis on recall:

- **Accuracy:** 96.67%
- **Precision:** 97.53%
- **Recall (Critical): 95.76%** — Chosen as the primary metric because missing a disease (False Negative) is more detrimental than a false alarm.
- **False Negative Rate:** 4.24%

4. Failure Case Analysis

A qualitative review revealed that failures primarily occur in:

- **Subtle Early Symptoms:** In cases like *Tomato Target Spot*, early lesions occupy a very small pixel area, leading the model to predict "Healthy".

- **Unusual Presentation:** Highly localized infections (e.g., *Cherry Powdery Mildew*) can be missed if the specific visual cues are partially occluded or resemble natural leaf variation.

Task 3: From Prediction to Decision

1. Why Model Confidence is Insufficient

In the provided scenario (12% diseased at 70% confidence), the raw confidence score is an unreliable standalone metric:

- **Calibration Issues:** A 70% confidence score does not always equate to a 70% probability of being correct; models can be overconfident in novel environments.
- **Context Blindness:** The model does not see external risk factors like humidity, wind, or the specific speed at which a particular pathogen spreads.

2. Recommended Action: "Verify then Treat"

Given the **Moderate Risk** (12% infection) and **Moderate Confidence** (70%), the system recommends a targeted manual verification:

- **Step 1:** The farmer should visit the specific flagged coordinates (roughly 12% of the field).
- **Step 2:** Collect 20–30 physical samples for visual confirmation.
- Step 3: Apply targeted treatment only to confirmed areas.
  This approach respects the farmer's limited budget by preventing expensive blanket chemical applications based on uncertain data.

3. Cost Asymmetry Analysis

The decision logic is biased toward catching disease because of the vast difference in error costs:

- **False Negative (Missed Disease):** ~$200 per plant in crop loss and secondary spread.
- **False Positive (False Alarm):** ~$15 per plant in wasted fungicide and labor.
- **The 13:1 Ratio:** Because missing a disease is **13 times more costly** than a false alarm, the system is tuned to "cry wolf" rather than remain silent during an outbreak.

4. Additional Signals for Future Improvement

To move from a proof-of-concept to a production-grade system, the following signals should be integrated:

- **7-Day Weather Forecast:** High humidity and rain forecasts should automatically lower the threshold for recommending treatment.
- **Pathogen Type:** Knowing if a disease is fungal (fast-spreading) vs. viral affects the urgency of the intervention.
- **Spatial Patterns:** Clustered infections are more likely to be real outbreaks, whereas scattered pixels might indicate sensor noise or shadows.