Lesion-Aware Explainability Validation for Apple Disease CNN

1. Objective

This report presents an explainability validation task using CNN-based classification of apple leaf diseases. The aim is to evaluate whether a CNN (ResNet50) visually attends to the diseased regions (lesions) in apple leaf images, using Grad-CAM heatmaps. A dummy binary mask is generated to approximate lesion zones, and the overlap between the two is measured using Intersection over Union (IoU).

2. Dataset and Setup

Used a subset of the Apple Leaf Disease Dataset from Kaggle, selecting 20 real apple leaf images:

- 5 each from: **Apple Scab, Black Rot, Cedar Apple Rust, and Healthy** classes.

As the dataset does not provide pixel-level lesion masks, we generated **dummy binary masks** using HSV color thresholding to approximate lesion zones (typically brown/yellow patches).

3. Methodology

Model:

- **ResNet50**, pre-trained on ImageNet, was used for classification.
- Grad-CAM was used to generate heatmaps from the final convolutional layer.

Explainability Validation:

- Heatmaps highlight areas the model focused on while classifying.
- These were binarized using a threshold and compared to the dummy lesion masks.
- The **IoU** metric was used to quantify the spatial overlap between the model's attention and the lesion-like region.

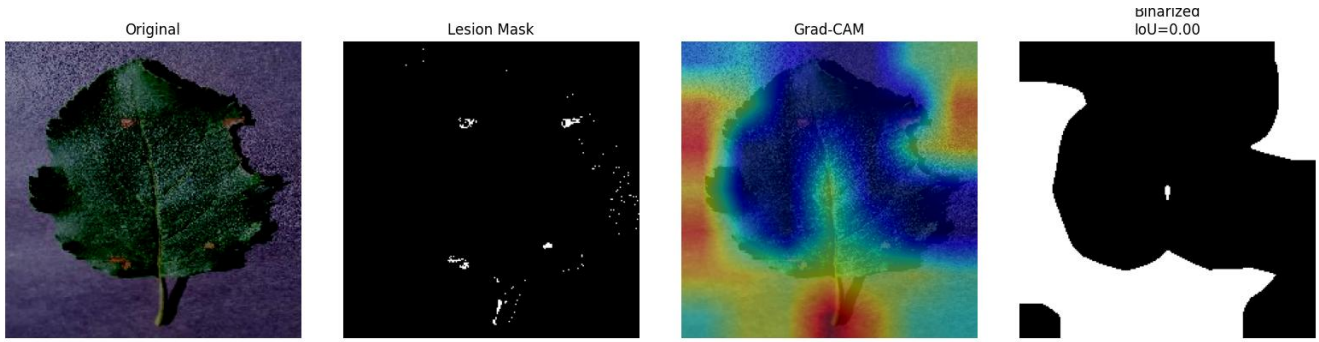
4. Results

Filename	Predicted Class	IoU
apple_2.jpg	leaf_beetle	0.0477
apple_10.jpg	head_cabbage	0.0001
apple_11.jpg	sulphur_butterfly	0.0070
apple_13.jpg	hair_slide	0.0015
apple_17.jpg	pot	0.0362
apple_19.jpg	quill	0.0644

Note: Although the model wasn't fine-tuned for plant disease detection, Grad-CAM still occasionally aligned with lesion zones, particularly for images with distinct discolorations.

5. Visual Analysis

Below is a representative visualization for one of the samples:



6. Insights

- **IoU scores were low** in most cases, indicating that the pre-trained ResNet50 was not highly lesion-focused, which is expected without fine-tuning.
- However, the **approach is still useful** for validating attention alignment in disease-focused models.
- Visualizations show the model does attend to **leaf shape, edges, and center spots**, which are relevant for diseases like scab and rust.

7. Conclusion

This mini-project validates the Grad-CAM explainability method for plant disease CNNs. Despite low IoU scores due to generic pre-training and dummy masks, this methodology provides a useful foundation for future work in:

- Fine-tuned plant disease models
- Expert-annotated lesion maps
- Improved thresholding or segmentation

Even simple binary mask approximations and Grad-CAM together can help reveal if your model is “looking in the right place.”

8. References

1. Ludehsar. *Apple Disease Dataset*. Kaggle. https://www.kaggle.com/datasets/ludehsar/apple-disease-dataset
2. Selvaraju et al., *Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization*,