

# Plant Disease Classification using Handcrafted Features

**## Objective**

The goal of this project is to classify images of plant leaves into multiple categories — specifically:

- Healthy

- Multiple Diseases

- Rust

- Scab

This was done \*\*without using deep learning or transfer learning\*\*. Instead, we used \*\*traditional machine learning techniques\*\* with \*\*handcrafted image features\*\* (such as color histograms and texture descriptors like HOG).

---

**## Dataset**

The dataset used is from the [Plant Pathology 2020 - FGVC7](https://www.kaggle.com/c/plant-pathology-2020-fgvc7) Kaggle competition. The dataset consists of high-resolution images of apple leaves labeled according to the disease present.

---

**## Feature Extraction**

The features were extracted using:

- \*\*Color Histograms\*\*: To capture color distribution in the leaves

- \*\*HOG (Histogram of Oriented Gradients)\*\*: To capture shape and texture

- Combined feature vectors were used as input to the ML models.

---

**## Models Trained**

The following traditional machine learning classifiers were trained:

- \*\*Support Vector Machine (SVM)\*\*

- \*\*Random Forest\*\*

- \*\*Gradient Boosting Classifier\*\*

---

**## Best Model**

| Model | Accuracy |

|---------------------|----------|

| SVM | 53% |

| Random Forest | 62% |

| \*\*Gradient Boosting\*\* | \*\*61%\*\* (Best f1-score and balanced performance)

The \*\*Gradient Boosting Classifier\*\* was found to be the most consistent performer across classes, especially for detecting \*\*rust\*\* and \*\*scab\*\*.

---

**## Limitations**

- The model has \*\*difficulty detecting the 'multiple\_diseases' class\*\*, likely due to the small number of training samples for this class.

- Handcrafted features might not capture all complex patterns compared to deep learning methods.

- Performance may improve with feature engineering and ensemble techniques.