CSC173 - Intelligent Systems
Dataset Description

Ramel Cary B. Jamen
ramelcary.jamen@g.msuiit.edu.ph

November 6, 2024

Background

Rice is a staple in Filipino cuisine and is deeply ingrained in the cultural identity and traditions of the Philippines. It is an essential food source, present at every meal and playing a vital role in daily life. Beyond its culinary significance, rice represents a symbol of unity and tradition for the Filipino people, and it is a part of celebrations, rituals, and community gatherings.

Problem Statement

Rice plants are susceptible to a wide array of diseases, including fungal, bacterial, and viral infections. These diseases pose severe threats to agricultural productivity, as they can significantly reduce rice yield, impair quality, and sometimes lead to total crop loss. For farmers, particularly those with limited resources, these issues can be devastating as they may lack access to costly disease-resistant seeds or chemical treatments. Furthermore, the warm, humid, and rainy climate of the Philippines fosters an environment conducive to the spread of these diseases, further compounding the risks. Therefore, timely and accurate identification of these diseases is essential to safeguard rice crops and ensure food security.

Structure of the Dataset

The dataset is collected as part of an initiative by Omdena's Local Chapter to create a Rice Disease Classifier using open-source data and computer vision. This dataset is designed to support the development of an automated solution to classify various rice diseases. Each image in the dataset represents a specific rice plant affected by one of the identified diseases, with the following key components:

- **Image Format**: To ensure consistency and facilitate model training, all images have been resized to a standard dimension of 224 x 224 pixels. This standardized format aids in the processing and analysis of images across diverse deep-learning models.
- **Labeling**: Each image is labeled according to the disease it represents, falling under one of the three categories: fungal, bacterial, or viral. The diseases themselves are divided into 13 distinct classes, making the dataset multi-class.
  - **Fungal**: Rice Blast, Sheath Blight, Brown Spot, Narrow Brown Spot, Sheath Rot, Stem Rot, Bakanae, Rice False Smut
  - **Bacterial**: Bacterial Leaf Blight, Bacterial Leaf Streak
  - **Viral**: Tungro Virus, Ragged Stunt Virus, Grassy Stunt Virus
- **Image Distribution**: The dataset is designed to have a balanced number of images across each disease class to minimize class imbalance, ensuring that models are not biased toward any one class.
- **Data Augmentation**: To enhance the model's generalization ability and improve performance in real-world scenarios, data augmentation techniques such as rotations, flips, and color adjustments has been applied to make it more robust.

Fungal Diseases

Fungal diseases are the most prevalent and impact various parts of the rice plant, including the leaf blade, sheath, stem, node, and panicle. The fungal diseases identified in the dataset include:

Figure 1: Rice Blast

Figure 2: Sheath Blight

Figure 3: Brown Spot

Figure 4: Narrow Brown Spot

Figure 5: Sheath Rot

Figure 6: Stem Rot

Figure 7: Bakanae

Figure 8: Rice False Smut

Bacterial Diseases

Bacterial diseases primarily affect the leaf blade. These infections can severely damage plant health, especially under humid conditions. The bacterial diseases in the dataset include:

Figure 9: Bacterial Leaf Blight

Figure 10: Bacterial Leaf Streak

Viral Diseases

Viral infections affect the leaf blade and sheath. These diseases spread rapidly and can be challenging to control without timely intervention. The viral diseases present in the dataset include:

Figure 11: Tungro Virus

Figure 12: Ragged Stunt Virus

Figure 13: Grassy Stunt Virus

These diseases exhibit distinct visual symptoms on affected rice plants, and this dataset captures these unique features to assist in accurate disease classification.

Solution

This dataset serves as a foundation for training deep-learning models to identify and classify rice diseases based on image data. By leveraging computer vision techniques, the aim is to develop a classifier capable of distinguishing between fungal, bacterial, and viral infections. Such a tool would enable early detection, allowing farmers to implement timely interventions and mitigate the spread of disease, thereby improving crop resilience and ensuring stable rice production.

The dataset provides valuable insights that support the advancement of technology for sustainable agriculture, ultimately contributing to the resilience and security of rice cultivation in the Philippines.