**Data Science Project Report on Rice Leaf Disease**

**Project Team ID:** PTID-CDS-AUG-23-1607B

**Project ID:** PRCP- 1001- RiceLeaf disease detection

**BUSINESS CASE:- BASED ON THE GIVEN DATA WE NEED TO PREDICT THE MODEL WHICH CLASSIFY THE THREE MAJOR ATTACKING DISEASES OF RICE PLANTS LIKE LEAF BLAST, BACTERIAL BLIGHT AND BROWN SPOT.**

**Abstract:**

Rice is amongst the majorly cultivated crops in India and its leaf diseases can have a substantial impact on output and quality. The most important component is identifying rice leaf diseases, which have a direct impact on the economy and food security. Brown spot, Leaf Blast are the most frequently occurring rice leaf diseases. To resolve this issue, we have studied various machine learning and deep learning approaches for detecting the diseases on their leaves by calculating their accuracy, n to measure the performance. This study helps the farmers by detecting the diseases in rice leaves in order to get a healthy crop yield. The deep learning models perform well when compared with the machine learning methods.

**Methodology** The objective of this research is to build a deep learning model which will detect and classify the diseases in rice crop using the data base.In this research Convolutional Neural Network (CNN) is used to extract the features from the rice leaf images.

**Device Project In-to Multiple Steps:**

1. Data Collection
2. Load Data
3. Data Preprocessing
4. Data Augmentation
5. Visualise Images
6. Building Deep learning Model
7. Compile Model
8. Fit Model for Training
9. Model Evaluation
10. Prediction on Test Data

**Data Collection:**

The dataset received from the Datamites educational platform and provided in a JPG format.

**Loading data:**

Before load the Data we import some necessary libraries.

To Load Data, We can utilize the Split\_folders library, a convenient tool in Python, to create three subsets. to split the dataset into three subsets: the training set, the testing set, and the validation set. This division allows for training the model on one subset, evaluating its performance on another, and validating the final results on a separate portion of the data.

**The Dataset Has Three Classes Labelled As-**

* 'Bacterial leaf blight': 0,
* 'Brown spot': 1,
* 'Leaf smut': 2

**Bacterial Leaf Blight:**

**Bacterial leaf blight** is a common disease in plants caused by bacteria. It manifests as water-soaked lesions on leaves, which eventually turn brown. This disease affects various plants, including rice, causing damage to the crops.

**Brown Spot:**

**Brown spot** is another plant disease characterized by the appearance of small, dark brown lesions on leaves. It is caused by a fungus and can affect a wide range of plants, including rice. Brown spot disease can lead to reduced yield and quality of crops.

**Leaf Smut:**

**Leaf smut** is a fungal disease that affects the leaves of plants, causing them to develop black spore masses. This disease can severely impact plant growth and reproduction, leading to significant agricultural losses.

**Image Pre-Processing:-**

* Before any modeling or classification is performed on the images, several image pre-processing steps must be performed first Common pre-processing steps include:

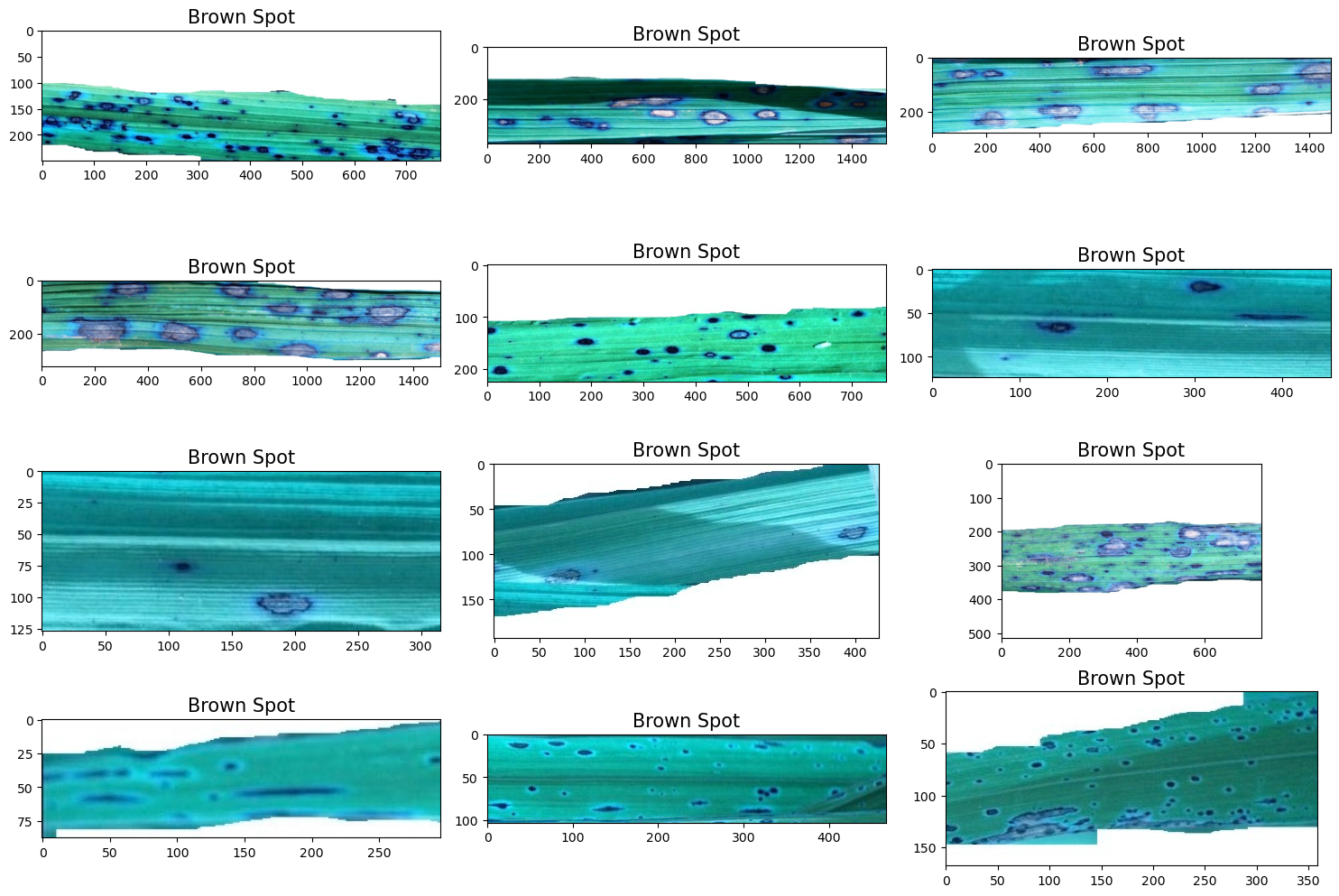
**Resizing**: Ensure all images have the same dimensions

### **Data Augmentation:**

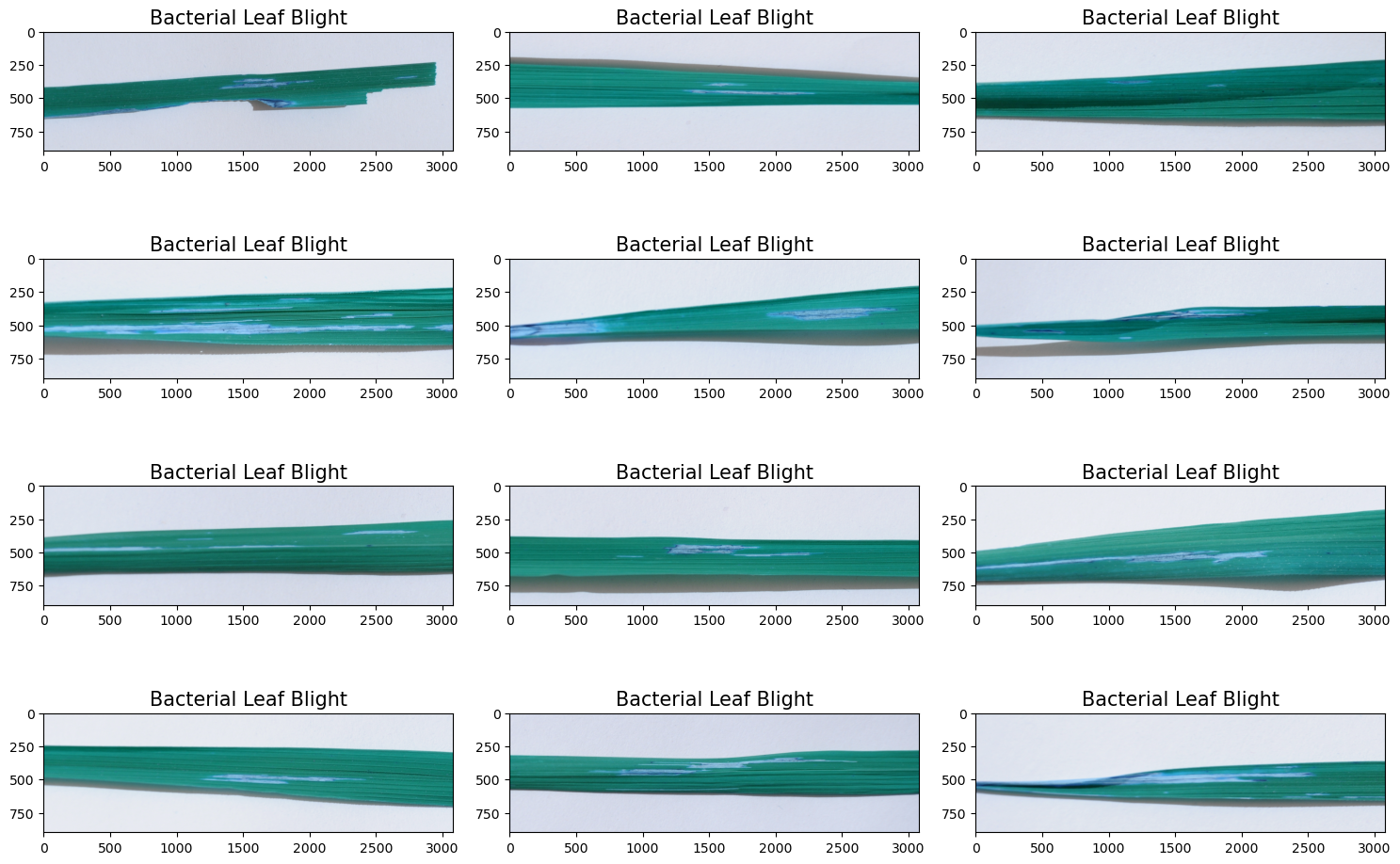
To account for different environments upon taking the picture, e.g. rotated images, shifted objects, zoomed in pictures, image augmentation is performed. The augmented images will be transformed versions of the dataset by randomly rotating the image up to 40∘, shifting the image vertically or horizontally up to 0.2 of the image length, shearing the image up to 20∘, and zooming in the image 0.2x.

**Visualise Images**

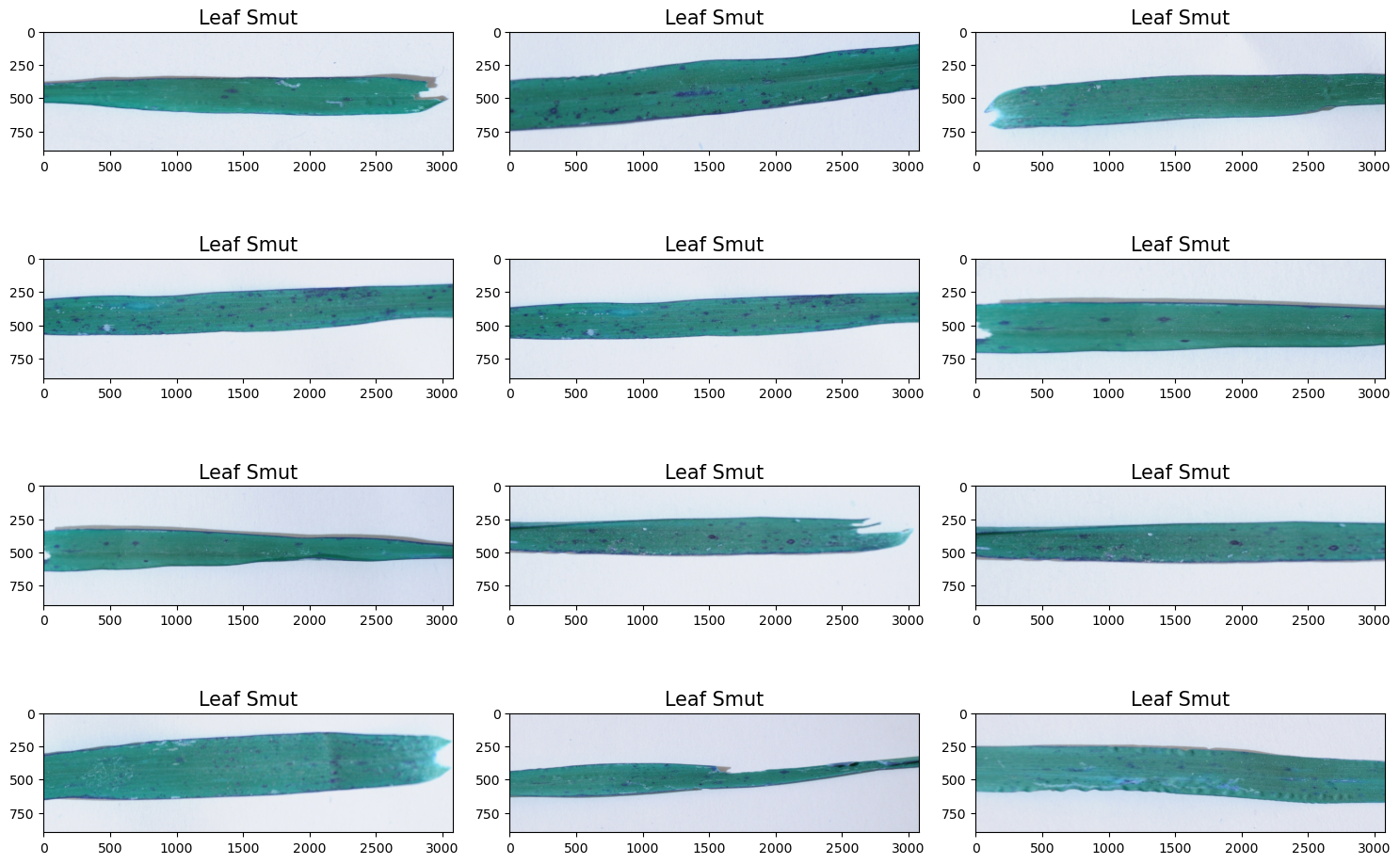
**Brown spot Images**



**Bacterial Leaf Blight**



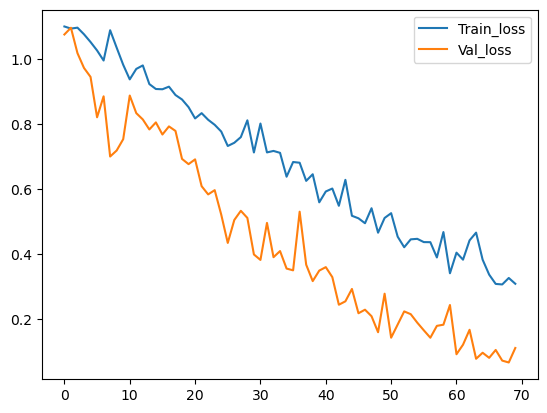
**Leaf Smut**



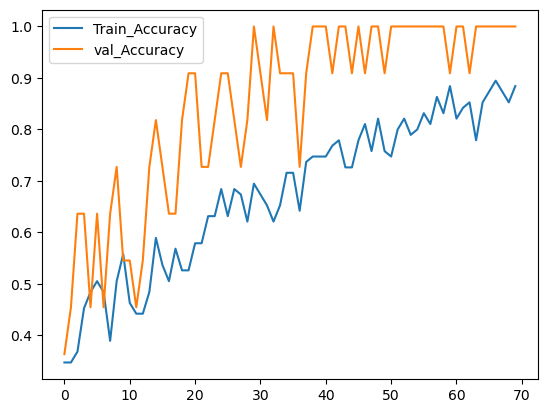
**Model Building**

* First, we initiate the construction of a Convolutional Neural Network (CNN) model, incorporating three essential types of layers: convolutional layers for feature extraction, pooling layers for down-sampling and dimensionality reduction, and fully connected layers for classification.
* Following the model's architectural design, a graphical representation is generated using the VisualKeras library, providing an insightful layered view of the network with a legend for enhanced interpretability. This visualization aids in understanding the connectivity and flow of information within the model, complementing the textual description.
* In Model Summary Subsequently, a comprehensive summary of the model is obtained.
* In Model compile method, key parameters such as the loss function (set to 'categorical\_crossentropy' for this classification task), the optimizer (utilizing 'adam' for gradient descent optimization), and evaluation metrics (including 'accuracy') are specified.
* compiled model is ready for training, and the fitting process is initiated using the fit\_generator method. This involves feeding the training dataset (train\_set) to the model for a specified number of epochs (set to 70 in this instance), with validation data (val\_set) used to assess the model's performance during training. This step represents the core process of updating the model's weights based on the training data, progressively improving its ability to make accurate predictions.
* Plotting the training accuracy and validation accuracy and training loass and validation loss.

PLOTTING THE TRAINING ACCURACY AND VALIDATION ACCURACY:



TRAINING LOSS AND VALIDATION LOSS:



* summarizes the model's performance, indicating the training and test losses, as well as the corresponding accuracies.

|  |  |  |
| --- | --- | --- |
|  | **Train** | **Test** |
| **Loss** | **0.305303** | **0.343684** |
| **Accuracy** | **0.852632** | **0.923077** |

**NOTE-** CNN model is very well perfrom on testing side, but in training side model is slightly perfrom low and overfitting occurs.

* Save the model
* Prediction on test data

## 

## Conclusion and Future Work In this project, we have proposed a custom CNN-based model that can classifyﬁve common rice leaf diseases. Our model istrained to recognize the rice leaf diseases in diﬀerent image backgrounds andcapture conditions. Our model achieves 92% accuracy on independent testimages.

**Thank You**