## **Title:** Tomato Disease Detection

## **Abstract**:

## Agriculture stands as a cornerstone of Bangladesh's economy, yet its growth is hindered by the pervasive threat of vegetable diseases. The plight of farmers grappling with the afflictions of vegetables and fruits is undeniable, resulting in substantial vegetable wastage and economic losses. To mitigate this challenge, we propose the implementation of a deep learning model, specifically a Convolutional Neural Network (CNN), to address the detection of diseases in tomato plants. By leveraging the power of artificial intelligence, our model swiftly identifies and classifies ailments afflicting tomato leaves, distinguishing between healthy foliage and diseased counterparts with an impressive accuracy of 97.40%. This innovative solution promises to revolutionize agricultural practices, empowering farmers with a rapid and precise means of disease diagnosis. By swiftly identifying and treating vegetable diseases, we envisage a future where farmers can preemptively safeguard their yields, minimizing losses and bolstering agricultural productivity in Bangladesh.

## **1. Introduction:**

Bangladesh, being an agricultural country, most of its economy depends on it. But vegetables can easily get affected by various diseases. Late treatment causes the loss of the vegetables.

2. Literature Review

2.1 The methodology utilizes image processing and machine learning, primarily support vector machines, for wheat leaf disease detection. It involves preprocessing, segmentation, feature extraction, and pattern recognition, alongside a review of relevant machine learning methods. This approach implements these methods for disease detection and is illustrated by a brief block diagram. It has problems like Bias or inconsistency due to different shapes and sizes of diseased leaf images. Subjectivity and potential errors in feature extraction

3. Methodology

3.1 Working Diagram: -

Classify the Images

Detect and show result

Build CNN Model

Prepare Images

Collect Images

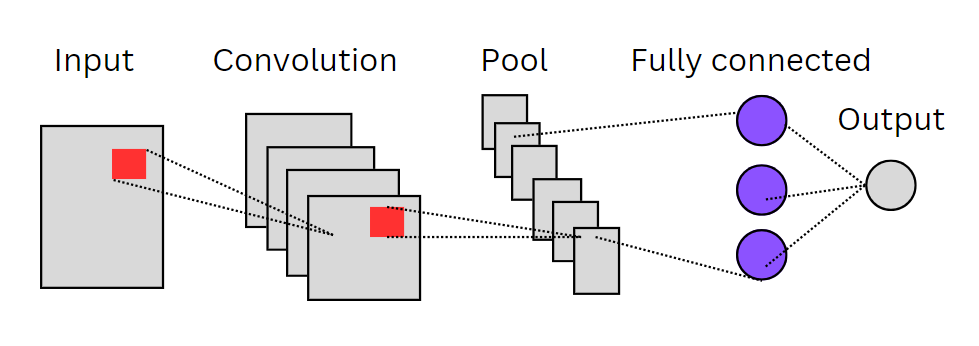
🡪 🡪 🡪 🡪

3.2 Dataset Collection

I have gathered a comprehensive dataset from Kaggle consisting of images representing eight different classes of tomato diseases. This dataset contains a total of 18,345 images, providing a diverse range of samples for each disease class. Each image in the dataset showcases the visual manifestations of various tomato diseases, including bacterial spot, early blight, late blight, leaf mold, Septoria leaf spot, spider mites, target spot, and yellow leaf curl virus. These images serve as valuable resources for researchers, agronomists, and machine learning practitioners interested in studying and addressing the challenges posed by tomato diseases. It is essential to acknowledge the sources and authors of these images to avoid plagiarism and ensure proper attribution of the valuable contributions made by the creators of this dataset.

3.3 Neural Network

 In the realm of image analysis, Convolutional Neural Networks (CNNs) reign supreme, adeptly navigating the intricacies of supervised learning while holding promise for unsupervised exploration. Central to CNNs' functionality are their convolution layers, which play a pivotal role in extracting crucial features from input data. As demonstrated in our previous code, these convolutional operations unfold in two-dimensional space, meticulously dissecting images through localized feature extraction. By applying filters across the input image, CNNs discern spatial hierarchies and subtle details essential for classification tasks. This nuanced approach empowers CNNs to distill complex visual information into meaningful representations, thereby enabling accurate image classification and analysis. Through the transformative capabilities of convolution layers, CNNs transcend conventional image processing, emerging as potent instruments for image understanding and exploration, heralding a new era of computational vision.

Where, 𝑦𝑗 𝑙 is the future map and 𝑓(𝑧𝑗 𝑙) is the activation function. Datasets are stored in 2 dimentional convolution operation in convolutional neural network (CNN).

3.4 Building Model

In the architectural construction of the Convolutional Neural Network (CNN), a systematic assembly of convolutional layers orchestrates the process of feature extraction and hierarchical learning. Commencing with an initial pair of Conv2D layers, each furnished with 32 filters and employing a kernel size of 3x3, the network lays its foundational stones. Following this, a MaxPool2D layer intervenes, skillfully reducing spatial dimensions through a pooling size of 2x2 with a stride of 2. The subsequent layers unfold with a meticulous crescendo, amplifying the filter count progressively to 64, 128, 256, and finally 512. Infused with the ReLU activation function, each convolutional stratum engenders nonlinear transformations, thereby nurturing the cultivation of discriminative feature representations. Noteworthy is the strategic utilization of the 'same' padding parameter, ensuring congruence in output and input dimensions, thus preserving the spatial intricacies. Through this choreographed symphony of layers, the CNN orchestrates an intricate ballet of feature refinement, culminating in a nuanced understanding of input data, thereby enabling the discerning classification and analysis of images.  
FORMULA.

3.5 Compiling Model

When compiling the model, it's essential to consider parameters such as optimizer, loss function, and metrics. In this context, the "adam" optimizer is chosen for its effectiveness in adapting the learning rate throughout the training process. For the loss function, "categorical cross-entropy" is utilized, which is well-suited for multi-class classification tasks. Furthermore, the "accuracy" metric is employed to assess the model's performance on the validation set during training, providing an intuitive measure of classification accuracy.

3.6 Training and Testing the Model

The model is trained using the `fit()` function, with the testing dataset utilized as validation data. Within the `fit()` function, the number of epochs is specified, determining the number of cycles the model undergoes with the data. Once the training process is complete, the testing process ensues, verifying the efficacy of the trained CNN. This evaluation stage validates the performance of the model, ensuring its effectiveness in handling unseen data.

4. Result and Outcome:

4.1 Learning Rate:-

The learning rate graph visually represents how the weights of our network are adjusted based on the loss gradient. The relationship between weight, gradient, and learning rate is depicted as follows:

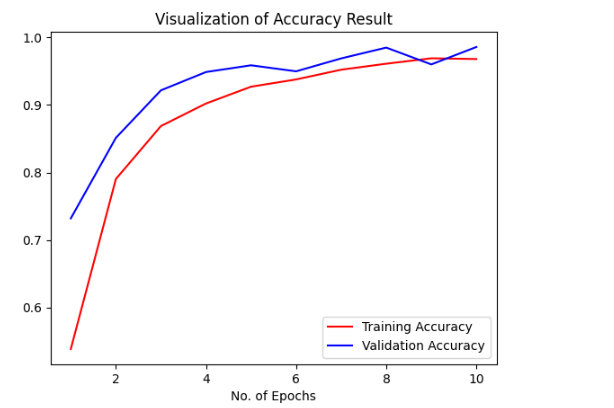
 New Weight = Existing Weight – Learning Rate × Gradient

Fig-4: Accuracy Result

Fig-4 illustrates the accuracy function graph of our model during training on datasets. Initially, the training accuracy is observed to be low, but over time, there is a notable increase in accuracy.

4.2 Confusion Matrix

The confusion matrix provides a visual representation of the performance of an algorithm by counting the values of correct and incorrect predictions [8]. The accuracy or classification rate formula, as well as the error of the model formula, are presented below:

Accuracy =

Error=1−Accuracy

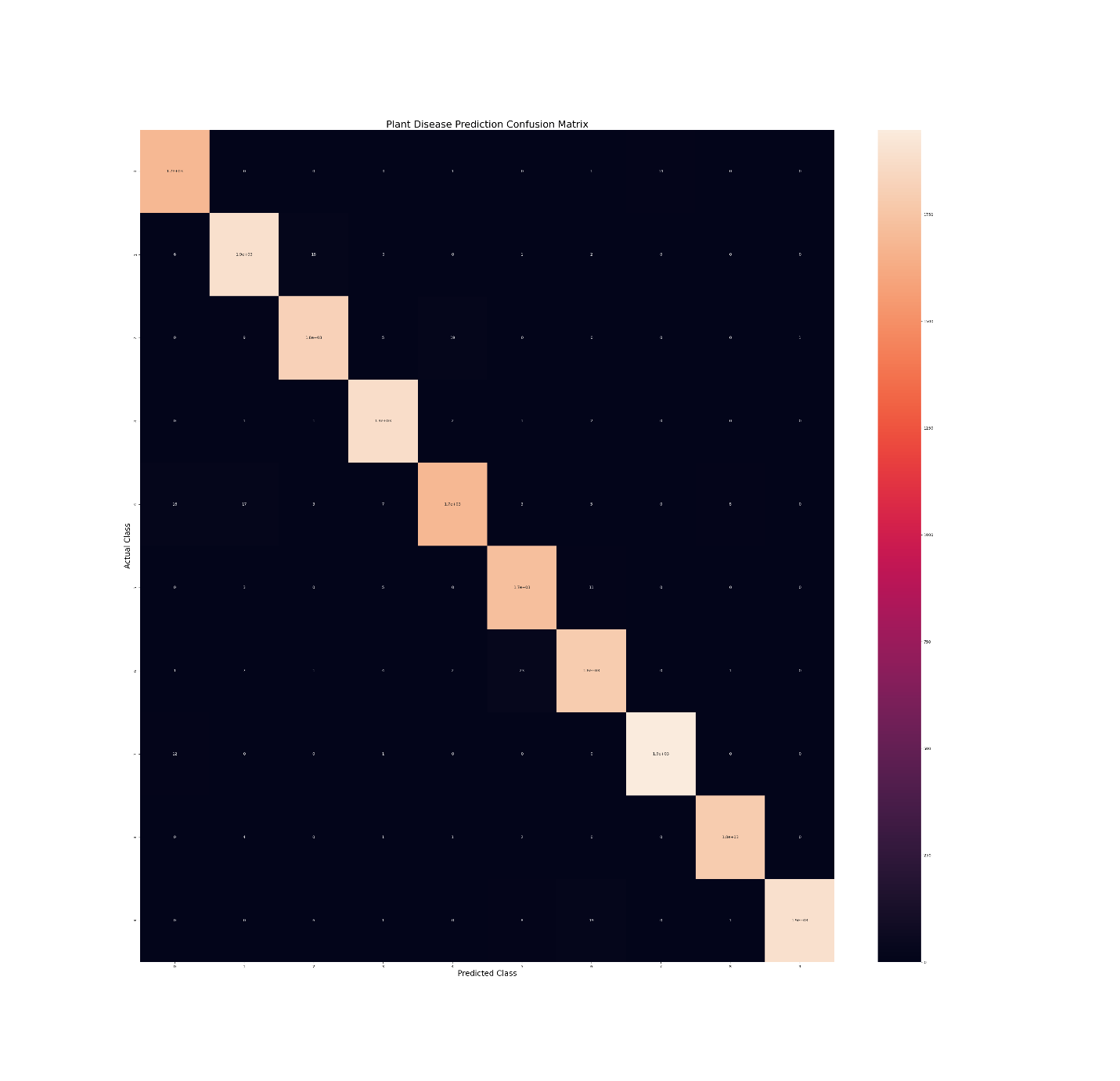


Fig-5: Confusion Matrix

Fig-5 presents the confusion matrix of our classification model. Notably, the diagonal values in each row are the highest, indicating that our algorithm achieves high accuracy.

4.3 Classification:

In the Jupyter Notebook, the entire process is implemented, starting with inputting various infected rice leaf images for classification. These images are selected from the dataset folder. Finally, the expected classified outcomes are obtained at the end of the process.

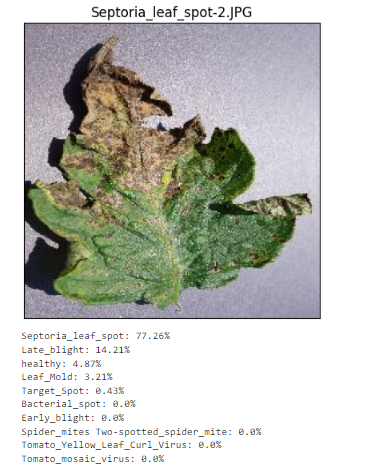
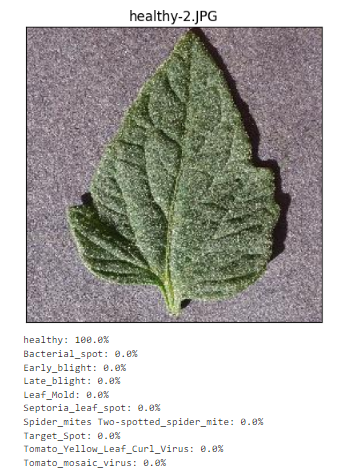
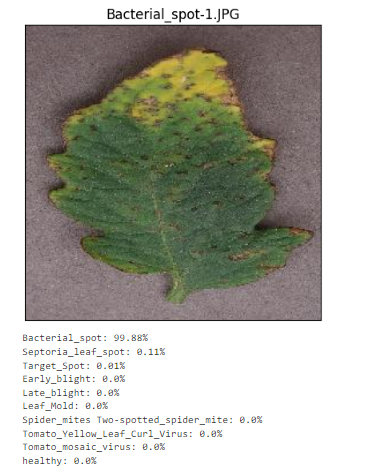
 Fig-6 Fig-7

Fig-8 Fig-9

5. CONCLUSIONS

In conclusion, our research presents a robust CNN-based approach for real-time tomato disease detection, offering valuable insights for agricultural practices. By leveraging advanced algorithms and extensive dataset analysis, we develop a highly accurate model capable of distinguishing between healthy and diseased tomato plants. Our findings underscore the significance of dataset quality and preprocessing in model performance. Looking ahead, our endeavor to explore memory-efficient CNN models reflects our commitment to continual improvement in plant disease classification. This work holds promise for revolutionizing agricultural practices and contributing to food security worldwide.

6. Reference