Plant Leaf Disease Classifier

1. Introduction

Plant disease detection through computer vision techniques has emerged as a promising solution to revolutionize crop management practices, offering automation and accuracy in identifying and classifying diseases affecting crops. Usually, the leaves of plants are the primary source for identifying plant diseases, and most of the symptoms of diseases may begin to appear on the leaves [1]. The problem of plant disease detection through leaves involves automating the identification and classification of plant diseases by analyzing images of plant leaves. Traditionally reliant on manual inspection, which is time-consuming and error-prone, leveraging Convolutional Neural Networks (CNNs) offers a faster alternative. CNNs are designed to efficiently process visual data, automatically learning relevant features, including embedded features, from raw pixel data, enabling accurate disease detection. Once trained, CNN-based models can rapidly analyze large volumes of images, streamlining crop management decisions and contributing to global food security efforts. Variability in disease symptoms across plant species, unpredictable image characteristics due to varied capture conditions, and the difficulty in delineating healthy and diseased regions present significant hurdles are some of the challenges faced in detecting plant diseases. Moreover, diseases may exhibit different characteristics depending on their stage of development and location on the plant, further complicating the detection process. Additionally, the presence of visually similar symptoms produced by different diseases necessitates robust discrimination methods [2]. To address these challenges, our research employs ResNet18, MobileNetV2, and InceptionV3 CNN architectures, each offering distinct advantages in terms of model complexity, computational efficiency, and performance. Through rigorous experimentation and evaluation, we aim to develop robust and scalable solutions for plant leaf disease detection, with the overarching goal of advancing agricultural practices and ensuring sustainable food production.

2. Possible Methodologies

For this project three datasets have been used. The first dataset, Plant Village contains 61k images of different plants like apple, potato, pepper etc and their diseases [3]. The second dataset, Corn Maize Leaf Disease, with 3500 images of corn leaves [4]. The third dataset is tomato leaf diseases containing 16k images [5]. The statistics of the data is listed in Table 1

All images in the three datasets were preprocessed by renaming them to JPG format, removing images with a brightness greater than 200, resizing them to 224x224 pixels, applying random flips, and normalizing them. Each dataset was divided into 80-10-10 ratio, as training, validation and test dataset respectively.

Dataset details			
Datasets	Plant Village	Corn Maize Leaf	Tomato Leaf
Image Size	256x256x3	256x256x3	300x300x3
Image Count	61,000	3,500	16,000
Classes	25	4	10

Table 1. Dataset Details

This project has 11 tasks: three CNN models trained from scratch with each dataset and two models trained using transfer learning. The three CNN architectures selected are Residual Neural Network (ResNet-18), MobileNetv2 and Inceptionv3. The Residual Network was chosen as the residual block structure employed in it is designed to enhance the depth of the model while mitigating issues related to model instability and accuracy loss associated with increasing network depth. This architecture allows the model to learn deeper characteristics of spectral curves more effectively due to its skip connections. These connections facilitate gradient flow during training, enabling the model to retain important information from earlier layers. [6]. The MobileNetV2 model was chosen due to its state-of-the-art performance across various tasks and benchmarks, along with its efficient architecture tailored for mobile devices [7]. The inverted residual structure, lightweight depthwise convolutions, and removal of non-linearities in narrow layers contribute to its effectiveness in balancing performance and computational efficiency. InceptionV3 is chosen for its efficient utilization of computational resources, demonstrated by its factorized convolutions and aggressive regularization techniques, leading to significant improvements in performance across various benchmarks [8]. Until this point, for the Corn Maize and Tomato Leaf dataset we have trained the data with all three architectures, and for Plant Village dataset, we have trained with ResNet-18 and MobileNetv2. The hyperparameters have been kept constant throughout all the 8 models. The learning rate for the training process is 0.001 and number of epochs is 10. SGD optimizer is used for its fast and computational approach [9]. All the models were evaluated based on accuracy for training, testing, and validation datasets, along with visualization for the confusion matrix.

3. Attempts at solving the problem

In analyzing the performance of various models across three distinct datasets with the above-mentioned configuration, notable trends emerge regarding their accuracy and generalization capabilities.

Table 2 shows the results obtained from each model corresponding to the dataset for training 10 epochs. It

Dataset	Model	Train Accuracy	Test Accuracy
Corn - Maize [4]	Resnet-18	0.83	0.83
	Inception V3	0.76	0.75
	MobileNetV2	0.81	0.80
Tomato-leaf [5]	Resnet-18	0.93	0.92
	Inception V3	0.48	0.49
	MobileNetV2	0.82	0.87
Plant leaf disease [3]	Resnet-18	0.97	0.97
	MobileNetV2	0.97	0.97

Table 2. Comparison of Model Accuracy on Different Datasets.

is seen that at the current state, the ResNet-18 performs well with all the three datasets leading to an accuracy of around 90% for all the datasets. This might be due to the fact that the number of learning parameters is high in the model compared to the others. ResNet-18 is able to capture intricate designs and patterns. The skip connections alleviate the vanishing gradient problem, enabling ResNet-18 to effectively learn hierarchical representations of the data. This makes ResNet-18 particularly well-suited for datasets with complex and diverse images, such as those depicting corn maize and tomato leaves. In these datasets, where the diseases may exhibit subtle variations and nuanced patterns, ResNet-18's capacity to capture fine-grained features enables it to achieve higher accuracies compared to other architectures. The second-best model would be MobileNetV2 which provides a reasonable accuracy of 80% in both train and unseen data. On the other hand, it is observed from the predictions that the Plant leaf Disease dataset performs quite similarly when trained with resnet and MobileNetv2 yielding an accuracy of almost 95%. InceptionV3 and MobileNetV2 are designed to be lighter and more computationally efficient, making them suitable for scenarios where resource constraints are a concern. However, these architectures may struggle with datasets that require deeper representations or finer feature extraction. InceptionV3's inception modules allow it to capture features at multiple scales, but its relatively shallower depth compared to ResNet18 might limit its ability to learn complex patterns effectively. Similarly, MobileNetV2's emphasis on efficiency through depth-wise separable convolutions may compromise its capability to capture fine details in datasets with intricate features. Figure 1 shows sample predictions made by the model(inceptionv3) for unseen data after training for corn. The train and the test accuracies of the models are plotted according to their respective datasets in Figure 2 and the confusion matrix are provided in the supplementary section. These initial findings highlight the importance of model selection based on the specific task and available resources. While ResNet-18 offers superior accuracy, its computational demands might be prohibitive for certain applications. On the other hand, MobilenetV2 presents a promising balance, demonstrating reasonable accuracy with a lightweight architecture. Further optimization techniques could potentially improve its performance. InceptionV3's inconsistent performance across datasets warrants further investigation. Hyperparameter tuning or data augmentation strategies might be necessary to enhance its generalizability.

Figure 1. Sample Predictions by InceptionV3 for Unseen Data (Corn)

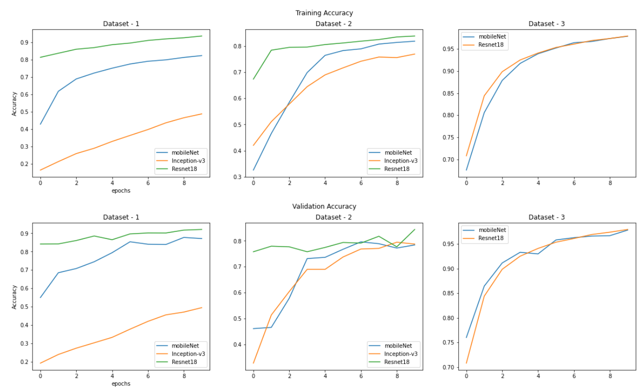

Figure 2. Plot on comparison of results.

Overall, the 8 models trained show stable and consistent performance across various evaluation metrics. This indicates that they effectively generalize to unseen data without resulting in either overfitting or underfitting, showcasing a balanced learning approach.

4. Future Steps

Furthermore, the hyperparameter tuning will be employed to optimize InceptionV3's performance on the Tomato Leaf disease dataset. This will involve adjusting learning rates, batch sizes, and regularization techniques to enhance pattern learning. Additionally, more experiments will be conducted to explore transfer learning for poorly performing architectures. Leveraging pre-trained models and fine-tuning specific plant disease datasets is projected to improve generalization and convergence speed. Finally, a comparative study will be conducted to evaluate different models, analyzing their strengths and weaknesses to identify the best model for each dataset based on various metrics, including accuracy, precision, recall, and F1-score.

References

[1] M. A. Ebrahimi, M. H. Khoshtaghaza, S. Minaei and B. Jamshidi, ”Vision-based pest detection based on SVM classification method”, Comput. Electron. Agricult., vol. 137, pp. 52-58, May 2017.

[2] Jayme Garcia Arnal Barbedo, A review on the main challenges in automatic plant disease identification based on visible range images, Biosystems Engineering, Volume 144, 2016, Pages 52-60, ISSN 1537-5110

[3] Plant Village Dataset from site https://data.mendeley.com/datasets/tywbtsjrjv/1 1

1

[4] Corn Maize leaf Disease Dataset from site https://www.kaggle.com/datasets/smaranjitghose/corn-or-maize-leaf-disease-dataset 1, 2

1, 2

[5] Tomato Leaf Disease Dataset from site https://www.kaggle.com/datasets/ashishmotwani/tomato 1, 2

[6] Zhao, Y., Zhang, X., Feng, W., Xu, J. Deep Learning Classification by ResNet-18 Based on the Real Spectral Dataset from Multispectral Remote Sensing Images. Remote Sens. 2022, 14, 4883. https://doi.org/10.3390/rs14194883 1

[7] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C. (2019). MobileNetV2: Inverted Residuals and Linear Bottlenecks. arXiv preprint arXiv:1801.04381. 1

[8] Christian S., Vincent V., Sergey I., Jonathon S., and Zbigniew W. (2015). Rethinking the Inception Architecture for Computer Vision. arXiv preprint arXiv:1512.00567. 1

[9] Hassan, E., Shams, M.Y., Hikal, N.A. et al. The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study. Multimed Tools Appl 82, 16591–16633 (2023). https://doi.org/10.1007/s11042-022-13820-0

5. Supplementary

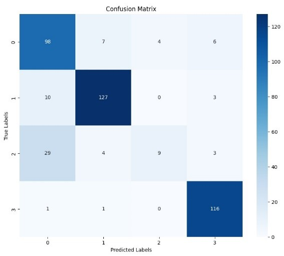

Figure 3. MobileNetV2 with Corn Maize dataset

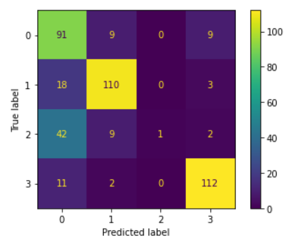

Figure 4. InceptionV3 with Corn Maize dataset

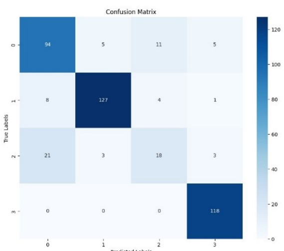

Figure 5. ResNet-18 with Corn Maize dataset

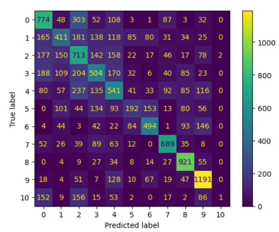

Figure 6. InceptionV3 with Tomato Leaf Disease dataset

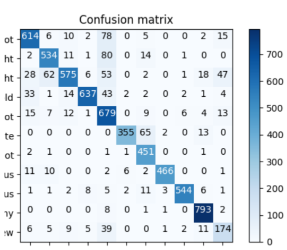

Figure 7. MobileNetV2 with Tomato Leaf Disease dataset

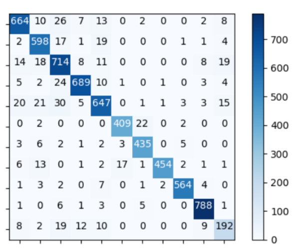

Figure 8. ResNet-18 with Tomato Leaf Disease dataset

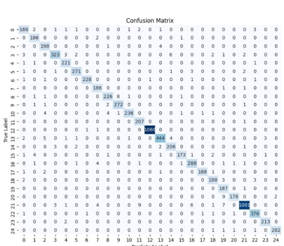

Figure 9. MobileNetV2 with Plant Village dataset

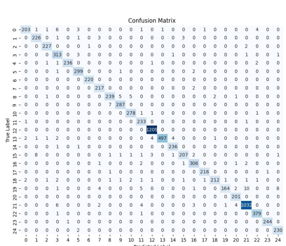

Figure 10. ResNet-18 with Plant Leaf dataset