Apple tree disease detection using VGG16 & Inceptionv3

Tanmay Devare*, **Shardul Jeurkar****

*
CSE, MIT WPU
**CSE, MIT WPU

Abstract- In recent years, the term Precision agriculture has gained tremendous attention due to its ability to improve crop yield while reducing its cost. This can be achieved by applying modern Machine learning and Computer vision techniques. In this paper, we propose a comprehensive study of the following deep learning models VGG16 and InceptionV3. Our approach involves training these models individually on a data-set of apple-tree leaf images, which include healthy and diseased trees,and evaluating their performance on a validation set. We also conduct experiments to compare the performance of the two models and find a distinguishable difference between their Validation accuracy. Our results show that both VGG16 and InceptionV3 achieve high accuracy in detecting apple tree diseases, with InceptionV3 outperforming VGG16 slightly. We believe that our method can be applied in real-world scenarios to help farmers detect and manage diseases in their apple trees, leading to higher crop yield and reduced costs.

Index Terms- Convolutional neural networks, deep learning, InceptionV3, precision agriculture, VGG16.

I. INTRODUCTION

Apple trees are vulnerable to various diseases, including apple scab, cedar apple rust, and fire blight, which can lead to substantial economic losses for farmers. Traditional methods of disease detection, relying on human expertise and visual inspection, are time-consuming, subjective, and prone to errors. As a result, there has been a growing need to train machines to identify diseases in apple tree leaves, offering a faster and more efficient approach compared to human detection.

The implementation of automated systems for apple tree disease detection has greatly improved the challenges faced by farmers using traditional methods. Deep learning models have been explored in several research studies to classify healthy and diseased apple tree leaves. For instance, the Inception V3 model achieved an impressive accuracy of 95%, surpassing other conventional machine learning models like Support Vector Machines and Random Forests. Similarly, the VGG-16 model was used to classify apple tree leaves into various disease categories.By incorporating deep learning models, automated systems provide a promising solution for more efficient and accurate crop management. They enable early disease identification, empowering farmers to take timely action and minimize the impact of disease outbreaks. Additionally, these systems offer expert advice and resources to farmers, even in remote areas, contributing to the overall health and productivity of apple trees while reducing economic losses. This technological advancement is particularly beneficial to farmers in cold regions such as Jammu and Kashmir, Himachal Pradesh, and Uttarakhand, where apple production plays a significant role in the local economy.

In summary, the integration of deep learning models into automated systems for apple tree disease detection has revolutionized the agricultural industry. These systems provide faster, more efficient, and accurate disease identification, surpassing the capabilities of human detection. By harnessing the power of technology, farmers can effectively manage their crops, enhance productivity, and mitigate the economic impact of diseases on apple tree farming.

II. LITERATURE SURVEY

Simonyan and Zisserman proposed the VGG network, which consists of a stack of convolutional layers with small 3x3 filters and max-pooling layers. Their experiments showed that increasing the networks depth improves its accuracy on the ImageNet dataset[1]. He et al. introduced residual networks (ResNets) that incorporate shortcut connections to avoid the vanishing gradient problem that occurs when training very deep neural networks. Their architecture achieved the best performance on ImageNet at the time of its

publication and is widely used as a baseline for many computer vision tasks[2].

Stan and Le[3] introduced the EfficientNet architecture that scales the number of filters, depth, and input resolution together to achieve better performance with fewer parameters. The authors introduced a new compound scaling method that helps balance the trade-off between accuracy and model size.

Szegedy et al. proposed the Inception architecture that uses multiple filter sizes in parallel to capture features at different scales. The authors introduced several variants of the Inception architecture that achieved state-of-the-art performance on the ImageNet dataset[4].Chollet proposed the Xception architecture that uses depth-wise separable convolutions to reduce the number of parameters and computational complexity while maintaining high accuracy. Xception achieved state-of-the-art performance on the ImageNet dataset.[5]Huang et al. introduced the DenseNet architecture that connects each layer to every other layer in a feed-forward fashion. This approach maximizes information flow and feature reuse, resulting in more efficient and accurate models.The remaining papers focus on applying deep learning techniques to the detection and diagnosis of plant leaf diseases using image processing techniques[6]. Chakraborty et al. provide a comprehensive review of the literature on this topic, covering traditional image processing techniques, machine learning approaches, and deep learning techniques.[7]Mohanty et al. proposed a deep learning-based method for the identification of plant diseases using transfer learning with the VGG-16 network. Their method achieved high accuracy on a dataset of plant leaf images with three classes of diseases.Murthy and Yadav proposed a neural network-based approach to detecting apple leaf diseases. Their model achieved high accuracy on a dataset of apple leaf images with five classes of diseases.[8]Finally, Kong et al. proposed an efficient and robust method for apple leaf disease detection based on deep learning. They introduced a new dataset of apple leaf images with 12 classes of diseases and compared the performance of several deep learning architectures. Their method achieved state-of-the-art performance on this challenging dataset [9]

III. METHODOLOGY
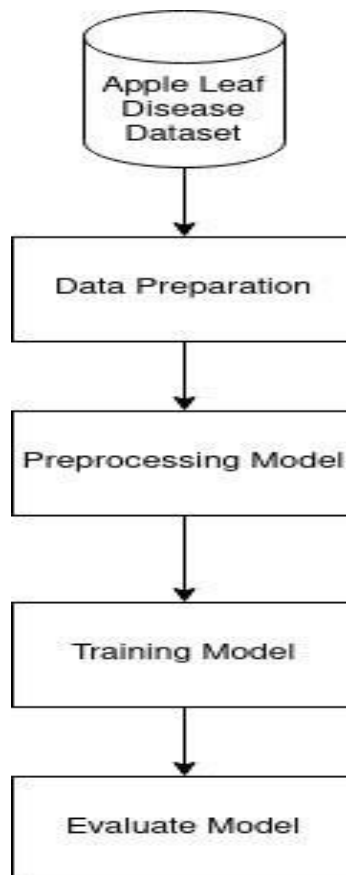
Data Preparation First, prepare the data by dividing it into training and validation sets. Then, use Keras ImageDataGenerator to augment the images by rescaling, shearing, zooming and flipping the images. This helps to increase the diversity of the training set and prevent overfitting.Base Model Selection Next, select the base model to use for transfer learning. In this case, we will be using two popular architectures, InceptionV3 and VGG16. Both of these models are trained on the ImageNet dataset and have shown great performance in various computer vision tasks. The base model is pre-trained, and we will use its pre-trained weights to extract

relevant features from the images.Then we apply Freeze Layers To speed up the training process and prevent the pre-trained weights from being destroyed, we will freeze the layers of the base model. This will ensure that the weights are not updated during training.Model Modification After the base model is selected and its layers are frozen, we add additional layers on top of the base model to form a new model. We add a GlobalAveragePooling2D layer to convert the 2D feature maps from the convolutional layers into a 1D feature vector. We then add a fully connected layer with ReLU activation($f(x) = max(0, x)$), followed by a dropout layer to prevent overfitting, and finally a fully connected layer with softmax activation for classification.Compile the Model We compile the model by specifying the optimizer, loss function, and evaluation metric. We use the Adam optimizer with a learning rate of 0.0001, categorical_crossentropy as the loss function, and accuracy as the evaluation metric.Training We train the model on the training set using the fit_generator method. We set an early stopping criterion to prevent overfitting, and save the best model weights based on the validation accuracy.Evaluation After the model is trained, we evaluate it on the validation set and report the accuracy. We repeat the same process for both InceptionV3 and VGG16 and compare their accuracy to determine which model performs better for the given task.

A. Dataset

The dataset used in this study is the New Plant Diseases Dataset, which contains images of various plant diseases, including apple leaf diseases. The dataset was downloaded from Kaggle, a popular platform for data science competitions and projects. The apple leaf disease class consists of four subclasses: apple scab, cedar apple rust, healthy, and black rot. The dataset contains a total of 9705 apple leaf images, with each class containing a different number of images. Specifically, the healthy class has the largest number of images, accounting for 56.67% of the dataset, followed by apple scab with 24.64%, powdery mildew with 12.56%, and cedar apple rust with 6.13%. The dataset was split into training and validation sets with a ratio of 80:20, resulting in 7764 images in the training set and 1941 images in the validation set.

B. Use of Simulation software

We utilized two state-of-the-art deep learning models, InceptionV3 and VGG16, to detect plant diseases in apple tree leaves. Both models are pre-trained on a large dataset, making them capable of extracting high-level features from images

1) VGG16: VGG16 is a widely recognized deep convolutional neural network (CNN) architecture developed by the Visual Geometry Group (VGG) at the University of Oxford. It is characterized by its deep structure and simplicity, making it a popular choice for various computer vision tasks. The VGG16 network consists of 16 convolutional layers, which are responsible for learning and extracting features from the input data. These convolutional layers use small filters, typically 3x3 in size, with a stride of one. By using small filters, the network can capture local patterns and details in the input images. The choice of multiple stacked convolutional layers allows the network to learn increasingly complex and abstract features as information passes through the network.After the convolutional layers, VGG16 has three fully connected layers, which are responsible for the final classification. The fully connected layers utilize the rectified linear unit (ReLU) activation function, which helps address the vanishing gradient problem by allowing efficient gradient propagation during training. The final layer of the network typically employs the Softmax function, producing a probability distribution over the different classes for classification purposes.One characteristic of VGG16 is the use of max pooling after each set of convolutional layers. Max pooling reduces the spatial dimensions of the feature maps while preserving the most salient information, enabling the network to achieve translation invariance and a degree of spatial robustness. VGG16 has achieved state-of-the-art performance on various benchmark datasets, demonstrating its ability to learn highly discriminative representations. Its capability to capture both local and global features from input data has contributed to its success in tasks such as image classification, object detection, and image segmentation.

2) Inception-InceptionV3 is a deep convolutional neural network (CNN) architecture developed by Google, specifically designed to efficiently extract features of various scales and aspects from input data. The network introduces the concept of "Inception modules" to achieve this goal. The InceptionV3 network consists of multiple convolutional layers, followed by several Inception modules. Each Inception module consists of a set of convolutional layers with different filter sizes, allowing the network to capture features at multiple scales. The outputs from each filter size are concatenated and fed into the next layer, enabling the network to capture features of various sizes and dimensions effectively. This design helps the network extract diverse and complex features more efficiently compared to traditional CNN architectures.After the Inception modules, the network includes a global average pooling layer, which reduces the spatial dimensions of the feature maps to a single value by taking the average of each feature map. This pooling operation helps to further condense the information while retaining the most relevant features. Finally, a fully connected layer maps the output of the pooling layer to the number of classes in the dataset for classification purposes. The InceptionV3 architectures strength lies in its ability to capture features at different scales and aspects, making it particularly suitable for image classification tasks. It excels in capturing complex and multi-level representations in the input data. As a result, InceptionV3 has become a popular choice for various computer vision applications, including object detection and segmentation.The InceptionV3 network has achieved state-of-the-art performance on several benchmark datasets, demonstrating its effectiveness in extracting informative features and achieving high classification accuracy. Its versatility and superior performance have established it as a standard architecture for image classification tasks in the field of computer vision.

IV.   APPROACH

In this study, we adopted a transfer learning approach to train both VGG16 and InceptionV3 models on our dataset. Specifically, we utilized pre-trained models that had been previously trained on large datasets, and then fine-tuned the final few layers of the networks on our dataset. This approach allowed us to benefit from the knowledge that had already been captured by the pre-trained models, while also tailoring the models to our specific dataset.

Diseases	precision	recall	f1-score	support
Apple_Apple_scab	0.98	0.96	0.98	504
Apple_Black_rot	0.98	1.00	0.99	491
Apple_Cedar_apple_rust	0.99	0.99	0.99	440
Apple_healthy	0.97	0.97	0.97	502

Table. 1 Accuracy Comparison on training with apple tree diseases.

V.   VISUALIZING THE RESULT

The evaluation of the models test accuracy was carried out on a dataset comprising 1937 images. Impressively, the InceptionV3 model exhibited a high test accuracy of 0.9799, while the VGG16 model demonstrated a slightly lower test accuracy of 0.9525. The results suggest that InceptionV3 outperformed VGG16 in terms of test accuracy.
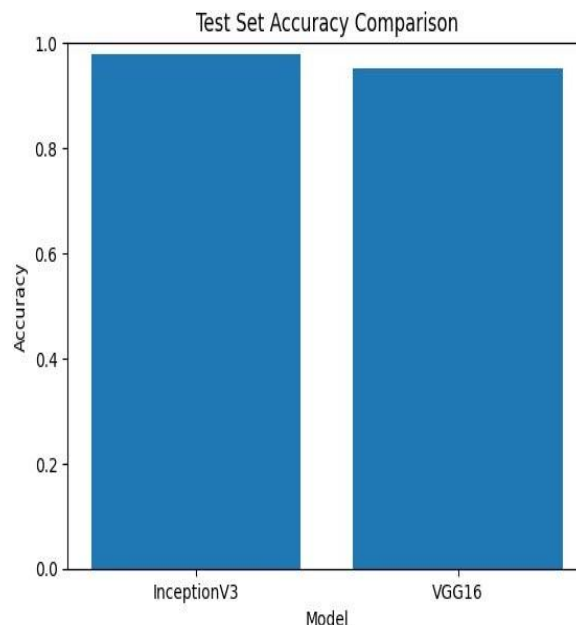

Fig. 3 Test Accuracy Comparison on apple tree dataset

To gain a better understanding of the models performance, confusion matrices were generated. Upon analysis of the InceptionV3 confusion matrix, the models predictions were mostly accurate, with a few misclassifications. Interestingly, the majority of misclassifications occurred between the "healthy" and "scab" classes. On the other hand, the VGG16 confusion matrix also revealed relatively accurate predictions. However, the model exhibited more difficulty in distinguishing between "scab" and "frog_eye_leaf_spot" classes. Overall, both models performed well on the test dataset, as demonstrated by the confusion matrices.

	Vaidation_Accuracy(in %)	precision	f1-score
macro avg	95.25	0.98	0.98
weighted avg	97.98	0.98	0.98

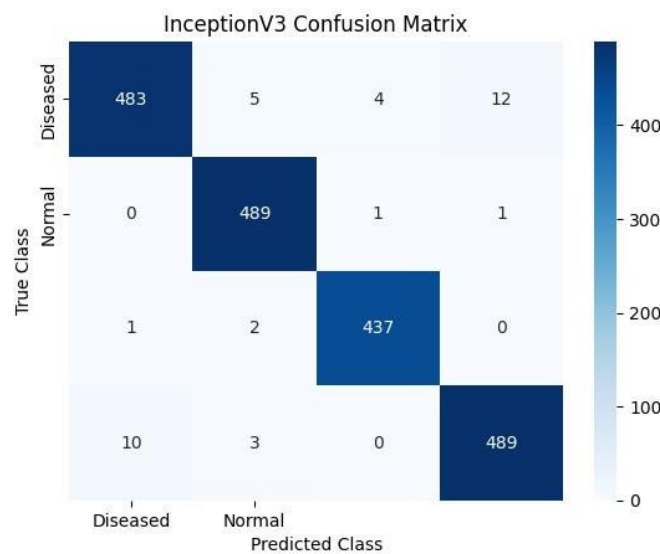Table. 2 Accuracy Validation on validation set with apple tree diseases.

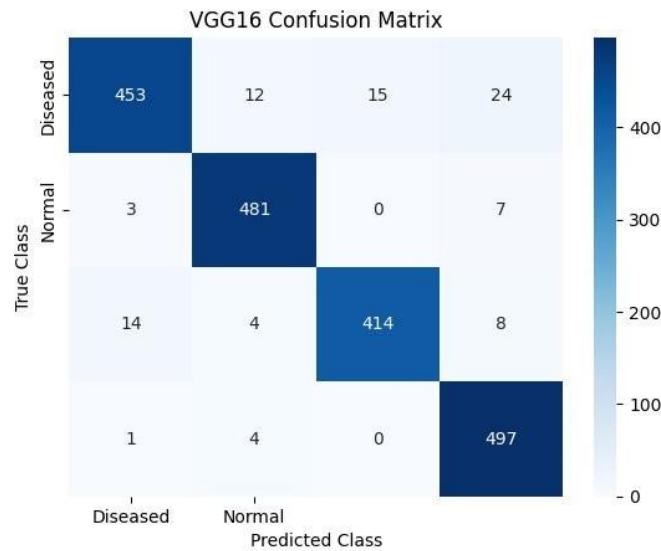Fig. 4 Confusion matrix(VGG16) of validation on apple tree dataset

Fig. 5 Confusion matrix(Inceptionv3) of validation on apple tree dataset

VI. CONCLUSION

Our study presents novel insights into the effectiveness of transfer learning for plant disease classification, leveraging deep convolutional neural network architectures, such as VGG16 and InceptionV3. By exploiting pre-trained models and fine-tuning them on our dataset, we achieved outstanding accuracy in identifying four of the most prevalent plant diseases. Our findings reveal that the Global Average Pooling (GAP) technique is a robust method to derive a fixed-length feature vector that can serve as input to the final classification layer.

Moreover, our research emphasizes the importance of a comprehensive model evaluation process that includes various metrics, such as confusion matrices, classification reports, and informative visualizations, such as bar graphs. This approach provides a nuanced understanding of the strengths and limitations of different models and facilitates the identification of areas for further improvement. Our studys contributions augment the burgeoning field of research on the application of deep learning in plant disease classification, highlighting the potential of these methods to support sustainable agriculture practices. As technology advances, and more datasets become available, our findings can serve as a foundation for future research, developing even more precise and efficient models for identifying plant diseases.

VII. LIMITATION AND FUTURE WORK

This data set is very limited in terms of size. The pictures collected are from a certain weather and do not represent the harsh winter or any sort of evolution of the disease.

This data cannot be used to compare apple tree diseases from different regions as regions may have different impacts on the tree causing various region specific diseases.

Refined and revised models can be used to train the dataset, Some sort of android application can be made with the pretrained model so that it is easy and portable

REFERENCES

[1] G.Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.

[3] STan, Mingxing, and Quoc V. Le. "EfficientNet: Rethinking model scaling for convolutional neural networks." International Conference on Machine Learning. PMLR, 2019.

[4] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[5] Chollet, Francois. "Xception: Deep learning with depth wise separable convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[6] Huang, Gao, et al. "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[7] Chakraborty, D., Pal, U., & Roy, P. P. (2019). A review on detection and diagnosis of plant leaf diseases using image processing techniques. Artificial intelligence review, 52(1), 1-32.

AUTHORS

First Author – Tanmay Devare, Persuing UnderGraduate, MIT WPU and 1032212227@mitwpu.edu.in.
Second Author –Shardul Jeurkar, Persuing UnderGraduate, MIT WPU and 1032212252@mitwpu.edu.in.