# Deep Learning-Based Approach for Disease Stage Classification of Sunflower Leaf

Original Scientific Paper

**Rupali Sarode***
Thadomal Shahani Engineering College, Computer Engineering Department,
Off Linking Road, Bandra(west), India
rupali.patil@thadomal.org

**Arti Deshpande**
Thadomal Shahani Engineering College, Computer Engineering Department,
Off Linking Road, Bandra(west), India
arti.deshpande@thadomal.org

*Corresponding author

***Abstract*** *– Accurate disease severity evaluation is crucial for managing the disease and yield loss. The classification of disease stages is essential for the estimation of disease severity. It takes extensive time for cultivators and botanical researchers to meticulously examine each leaf image and identify the disease stage to assess the severity of the disease at the field scale. Extracting the damaged leaf area is also achievable with image segmentation, although there are drawbacks such as threshold selection and lack of grayscale difference. Thus, deep learning has produced recent breakthroughs in various fields, such as high-resolution image synthesis, recognition, and categorization of images. In this work, the disease stages of two diseases (Alternaria leaf blight and Powdery Mildew) are classified using sunflower leaf images taken from sunflower farms in India (Marathwada State) during the Rabi season. With the help of botanists, images are labeled as three disease stage classes and one healthy stage as ground truth. A series of deep convolutional neural networks (Visual Geometry Group models with 16 and 19 neurons, respectively) with transfer learning and fine-tuning approach is trained, validated, and tested using stratified k-fold values four and five. The findings indicate that VGG16, with k-fold=5, gives the highest testing accuracy, which is 90.25%, with fine-tuning for Alternaria Leaf Blight. For VGG19 with kfold=5, the highest testing accuracy is 86.89% with fine-tuning for Powdery Mildew. Additionally, confidence interval calculation shows smaller intervals of 3% and 4% with a significance level of 95% for the VGG16 and VGG19 models, respectively.*

***Keywords***: *Convolution neural network, transfer learning, fine-tuning, multiclass classification, Alternaria leaf blight, Powdery Mildew*

*Received: August 24, 2024; Received in revised form: October 27, 2024; Accepted: November 12, 2024*

## 1. INTRODUCTION

The sunflower plant serves several important functions. It recycles nutrients and organic matter from the soil through its roots, produces oil from its seeds, provides feed for animals, acts as green manure when its leaves are used, and also produces flowers and honey [1]. In various regions of India, diseases like Powdery Mildew, Alternaria leaf blight, and Downy Mildew have impacted sunflower plants [2]. Growers can make informed decisions at the field level to protect plants using early disease prediction and forecasting. Predicting and forecasting diseases will be essential for safeguarding plants, enabling prompt action to prevent crop loss and enhance oilseed yield and production. This research will develop technologies to enhance the nutritional and medicinal benefits of sunflower oilseed crops, which are widely used as functional food [3]. Sunflower has various nutrients which are beneficial for humans as well as for animal health. It is rich in essential nutrients, including protein, fiber, unsaturated fats, copper, zinc, selenium, iron, and vitamins, especially vitamin E. Sunflower seed meal is commonly used as animal and pet feed due to its high content of sulfuric amino acids. It can also be used as a salted or roasted snack or used as cooking oil. [4] Sunflower seed production Vol. in India was 544 in the fiscal year 2013. It decreased to 228 in the fiscal year 2022, then rose to 250 in the fiscal year 2023, and finally increased to 279 in the same year. [5] In contemporary times, CNN is often used in agricultural research due to its powerful image feature processing capabilities. The most common applications of deep learning include plant and crop classification, which aids in robotic harvesting, pest control, yield forecasting, and disaster monitoring. At the field scale, manual diagnosis of plant diseases and their severity is time-consuming. [6] Deep learning has greatly enhanced computer vision by enabling computers to analyze, understand, and make intelligent decisions based on visual data.

Computer vision advances through the development of convolutional neural networks and deep learning algorithms [7]. Harm to the plant and its growth is based on the extent of the disease. Growers assess disease severity on a field scale by observing individual plant leaves. This manual process is expensive and takes a lot of time. As a result, by employing various new deep learning technologies, this lengthy procedure could become automated, allowing growers to make decisions earlier and, in less time, to protect plants.

## 2. RECENT STUDIES:

Numerous researchers have focused on agriculture to assist growers in identifying and classifying various crop diseases, as well as in diagnosing and forecasting them. Field observations can often be slow and inefficient. Deep learning approaches have been applied by multiple researchers to automate disease assessment for rice, apple, corn, and cotton crops. Zahraa Al Sahili et al. [8] provided an AgriNet dataset of 1.6 lakh images recorded from 19 distinct places and grouped into 423 plant types and illness classifications. Five ImageNet architectures were used to categorize plant species, diseases, pests, and weeds, including VGG16, VGG19, Inception-V3, InceptionResNet-V2, and Xception pre-trained networks. The VGG19 model provided the highest accuracy of 94%, whereas the InceptionV3 model achieved the minimum accuracy of 87%. To further enhance the AgriNet project, the authors suggested performing more complex data augmentation on its datasets. However, due to the limited number of agricultural public datasets, the research community is recommended to convert the data set's private information to publicly available at any moment.

Srinivasa Rao Dammavalam et al. [9] recommended employing Deep Learning models such as VGG16 and VGG19 to classify leaf images. From the Kaggle data sets [10], 14 crop leaf images were used and classified as healthy and infected images. Both models gave better results than others, such as ResNet50, DenseNet121, ResNet50V2, MobileNet, MobileNetV2, etc. The study's authors strongly advocate for including unique leaf disease classifications for each crop in future research.

Le Yang et al. [11] suggested a model for identifying corn weeds using SE-VGG16, which was evaluated on an image dataset [12] of corn seedlings and weeds. Using a Canon PowerShot SX600 HS camera, 6000 photos from the dataset were taken. These photos include one corn seedling and four weed categories—bluegrass, Chenopodium album, Cirsium setosum, and sedge—each with 1200 photos. Squeeze-and-Excitation mechanism is used with the VGG16 model to concentrate an important part of images and it gave superior weed identification results than VGG16. To support the complete agricultural production process and increase agrarian efficiency in production, deep learning combined with agricultural output will be applied to other fields in the future.

Prabira Kumar Sethy et al. [13] used 11 CNN models with fine-tuning approaches and deep feature plus support vector machine (SVM). Four types of rice leaf diseases—bacterial blight, blast, brown spot, and tungro—were identified using 5932 on-site photos gathered from agricultural sites in Odisha, India's Sambalpur and Bargarh districts. To perform feature extraction, the mentioned deep learning models were applied: InceptionResNetV2, GoogleNet, AlexNet, VGG16, InceptionV3, VGG19, ResNet18, ResNet50, ResNet101, DenseNet201, and XceptionNet. A Support Vector Machine is used to classify the extracted features. The transfer learning approach was used again to identify the four diseases of rice leaf. Finally, the outcomes of transfer learning and feature extraction were assessed.

Ghazanfar Latif et.al [14] proposed detecting and classifying six distinct diseases: healthy, narrow brown spot, leaf scald, leaf blast, brown spot, and bacterial leaf blight using deep CNN transfer learning with the VGG19 model. The image dataset [15] used has 2167 photos classified based on the six diseases listed. The final evaluation uses accuracy, precision, and the F1 measure. This article proposed that the future scope is to acquire field-scale images using drone technology in conjunction with IoT technology and a deep learning approach in real-world circumstances.

## 3. DISEASE SEVERITY ESTIMATION:

Guan Wang et al. [16] mentioned that the degree of damage caused by a plant disease is determined by its severity. Usually, experienced professionals use visual inspection of plant tNo. to rate the extent of plant diseases. To estimate disease severity, based on disease extent, classes made as Healthy, Initial, Middle, and Last with the help of a botanist. Deep learning networks are trained and tested using such class-wise images and evaluated further to classify disease stages.

### 3.1 DATASET INFORMATION:

This article considers two sunflower leaf diseases: Alternaria Leaf Blight and Powdery Mildew. The images used are collected from India. A total of 378 sunflower leaf images were considered for this study. These images belong to two diseases, Alternaria leaf blight, and Powdery Mildew, with their disease stage, and 105 are healthy images. A smartphone camera with 64 Megapixels was used to take images of the Rabi season. With the aid of plant researchers, the disease's extent was determined and divided into three stages: the initial stage, which ranged from 0 to 30%. The middle stage ranged from 31% to 60%, and the last stage ranged from 61% to 100%. The following table shows the count of images for each disease and its stage.

**Table 1.** Disease Stage count per disease

| Sr. No | Disease Name | Disease Grade and Stage | Count of Original images |
|---|---|---|---|
| 1. | Powdery Mildew | 1-Stage1 | 64 |
| | | 2-Stage2 | 84 |
| | | 3-Stage3 | 38 |
| 2. | Alternaria Leaf Blight | 1-Stage1 | 34 |
| | | 2-Stage2 | 33 |
| | | 3-Stage3 | 20 |
| 3 | Disease Free | Healthy-Stage0 | 105 |
| | Total | | 378 |

### 3.2. DISEASE INFORMATION

**1.** Alternaria Leaf Blight:

The disease affects the stems, sepals, petals, and leaves, causing brown spots. The disease stage is shown in Fig 1 a), b), and c). The leaves have dark brown spots with a golden circle around them and a pale edge. Later on, the spots become larger and take on an irregular shape with concentric rings. Leaf fall and dryness are caused by bigger, uneven lesions formed when multiple spots come together [17].

(a) (b) (c)

**Fig.1.** Brown spots of Alternaria Leaf Blight, a) Stage1, b) Stage2, c) Stage3

**2.** Powdery Mildew:

On the leaves, the disease causes white, powdery growths. The disease stage is shown in Fig 2 a), b), and c). A white to grey mildew occurs on the upper surface of older leaves. As the plant ages, areas of white mildew can be seen with black pinhead-sized leaves. The impacted leaves lose color, curl, become chlorotic, and eventually die. [17]

(a) (b) (c)

**Fig. 2.** White powdery growth on Sunflower Leaves, a) Stage1, b) Stage2, c) Stage3

## 4. CONCEPTS AND METHODOLOGY USED

### 4.1. DATA PREPROCESSING

The ImageDataGenerator class from Kera's library was used to apply data augmentation, increasing the sample size. This procedure comprised a rotation range of 40, a shear and zoom factor of 0.2, and a brightness range of 0.5 to 1.5 with a horizontal flip. The desired image size is (224,224), and a total of images, including training and testing, were generated after the dataset augmentation [18].

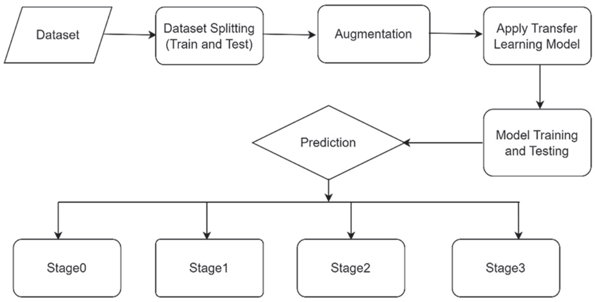



**Fig. 3.** Methodological block diagram

**Table 2.** Sample Size after Augmentation

| Stage | Number of images per class after Augmentation | |
|---|---|---|
| | Alternaria Leaf Blight | Powdery Mildew |
| Healthy-Stage0 | 610 | 610 |
| 1-Stage1 | 685 | 641 |
| 2-Stage2 | 672 | 732 |
| 3-Stage3 | 598 | 767 |
| Total | 2565 | 2750 |

### 4.2. DEEP LEARNING PROPOSAL:

**1.** Convolution Neural Network:

The foundation of Artificial Intelligence is Deep Learning. Deep Learning technologies have multiple uses in agriculture, including disease detection, disease identification and categorization, and disease severity assessment. The Convolution Neural Network, a component of Deep Learning architecture, is a multilayered, hierarchical network that functions nonlinearly and resembles the human brain. Convolution neural networks (CNN) have so far exhibited the greatest power in image classification.[19].

**2.** Model Regeneration:

The deep learning model's potential is to acquire knowledge from a hierarchical representation of features effortlessly. The first layers of CNN-extracted features are always generic, while features at later layers are increasingly specialized. Hence, to perform classification based on required interest, the model needs to be regenerated by adding a new classifier as per the interest, and finally, the model needs to be fine-tuned based on three approaches:

1. Train the entire model learning from scratch.
2. Train some layers and freeze the other layers—The network's weights can be controlled by keeping more layers frozen for small datasets and training more layers for large datasets.

3. Freeze layers of the convolution model—Keep the convolution layer base model as it is and use its output to give to the classifier.

In the first two approaches, the learning rate must be carefully selected (usually smaller) to avoid knowledge loss. In the last approach, the pre-trained model can be used as it is and based on extracted features to classify required interest [20]. For this approach, the Adam optimizer is used with a learning rate of 0.0001.

**3.** Fine-tuning with Transfer Learning:

The transfer Learning process is based on the third approach, as mentioned above. In this process, first, select the pre-trained networks VGG [21] and InceptionV3 [22], which are available to use on Keras [23].

The second is to classify the required problem based on the size similarity matrix, as shown below in Fig 3 [20]. This size similarity matrix shows how the size of the dataset is related to the model's fine-tuning. Hence, there is a proper mapping based on the number of dataset images and the number of layers to be kept, not frozen. Finally, the third is to select and try various values of fine-tuning parameters based on the dataset size and pre-trained model dataset. So, for disease stage estimation, from Fig. 4, orange circled mapping is used to incorporate transfer learning with fine-tuning parameter of value=2; as per class, there are around 600 images.

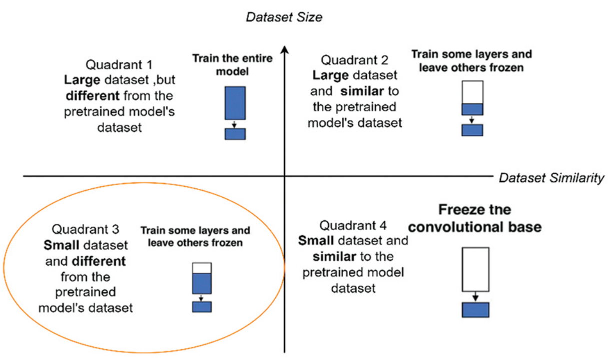



**Fig. 4.** Size-similarity matrix & decision map for fine-tuning pre-trained models

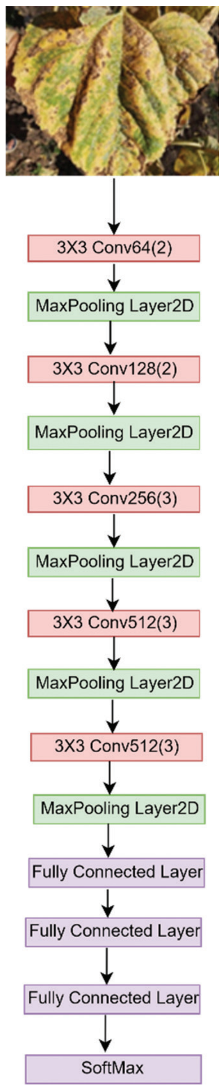
**4.** VGG16 and VGG19:

We employed two architectures, VGG16 and VGG19, for the fine-grained disease stage classification problem with minimal training data. Each architecture uses transfer learning by fine-tuning the higher layers of a pre-trained deep network to classify the disease stage.

The VGG architecture consists of three fully linked layers, a SoftMax activation function at the end, and Conv-1 Layer with 64 filters, Conv-2 with 128 filters, Conv-3 with 256 filters, and Conv 4 and Conv 5 with 512 filters. A max-pooling layer of a 2x2-pixel window with a stride of 2 follows each filter of size 3 × 3, a Rectified Linear Units (ReLu) activation, and all layers are followed by a dropout layer with a dropout ratio of 20% followed by the last convolutional layer. The final fully connected layer produces four outputs, one for each of the four classes. The SoftMax layer uses these outputs to determine the probability output. For both models, the input shape is 224 ×224 × 3, and the kernel size is 3 x 3 pixels. Fig 5 and 6 show the model architecture. The VGG16 model has 16 convolution layers, and we trained the model by freezing the first 12 layers for the VGG19 model, which has 19 convolution layers. So, both models are retrained by freezing 12 and 15 layers to obtain precise disease stage classification. (Refer to Table 3).

**Table 3.** Details of fine-tuning used

| Model | Frozen layer | Trainable Layer |
|---|---|---|
| VGG16 | 12 | 4 |
| VGG19 | 15 | 4 |



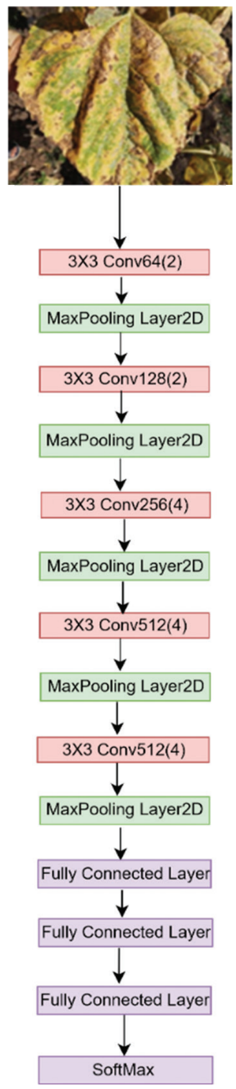




**Fig. 5.** VGG16 architecture **Fig. 6.** VGG19 architecture

## 5. EXPERIMENTAL PREPARATION AND ASSESSMENT

This work uses a dataset of sunflower images of two diseases. Google Colab is used to train, validate, and test phase of both models. As shown in Table 2, each class has uneven samples per class. Hence, this uneven sample distribution makes the model learn slowly, resulting in a biased model. Thus, to mitigate this, stratified K-fold cross-validation is employed. The samples are organized into K strata to provide nonoverlapping sets. The first strata from each class are then combined into the first fold, the second strata from each class into the second fold, and so on, to generate the stratified folds. By replicating the dataset's initial groupings, folds are created. After that, one-fold is used as the test set, and the remaining K-1 folds for training in each K-Fold Cross-Validation process iteration (Refer to Figs. 7 and 8).

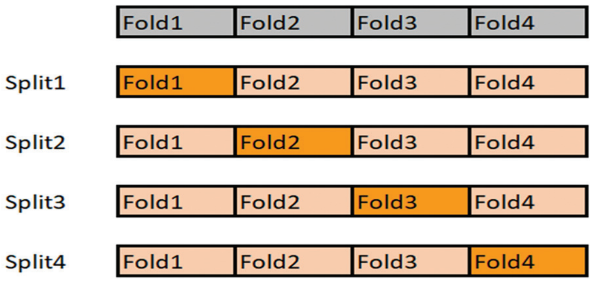



**Fig. 7.** K4 Cross Validation

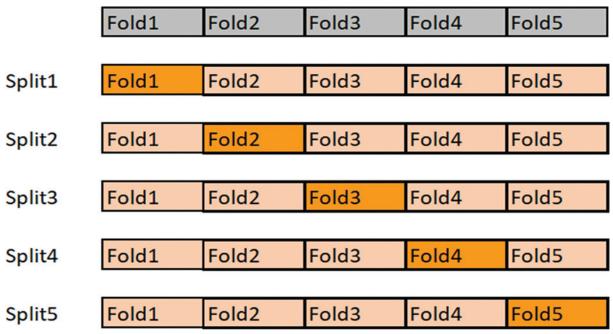



**Fig. 8.** K5 Cross Validation

Well-known Python machine-learning toolkit Scikit-Learn natively facilitates stratified K-Fold Cross-Validation. To reduce data balancing, stratified K-fold cross-validation ensures that samples from each class are used for testing and training the model. Using K=4 and 5 values, image samples are arranged into K strata in this work to construct non-overlapping sets. There will be 2360 images distributed across four classes since 590 image samples are kept for each class to have even samples. There are 236 images for testing and 2124 for training because the split is 10% for testing and 90% for training. Since this work uses K4 and K5 cross-validation, there are 531 and 425 images per fold, respectively. The training dataset was used to train the transfer learning model with a batch size 16. and the test dataset is used to assess it. Both models trained till epoch10 with fine-tuning =2 and without fine-tuning. TensorFlow, Kera's, and Sklearn Python libraries are combined in the models. The hyperparameters for both models are defined in Table 4. The network weights were iteratively adjusted using the Adam optimizer (with a learning rate of 0.0001) to underrate the loss function during the model construction based on training data [24]. A sparse categorical cross-entropy function with metric accuracy is utilized as the loss function because this task requires multiclass image classification. The loss function calculates the difference between the input label and the predicted result.

**Table. 4.** Hyperparameters used

| Metrics | Metrics Value |
|---|---|
| Batch size | 16 |
| Optimizer | Adam |
| Learning rate | 0.0001 |
| Criterion | Categorical cross Entropy Loss |

## 6. EXPERIMENTAL RESULTS

### 6.1. ALTERNARIA LEAF BLIGHT

Fig. 9 shows both model's training and testing accuracies per fold without fine-tuning for Alternaria Leaf Blight. The testing accuracy of VGG16 on the holdout dataset is greater than VGG19 after K5 cross-validation.

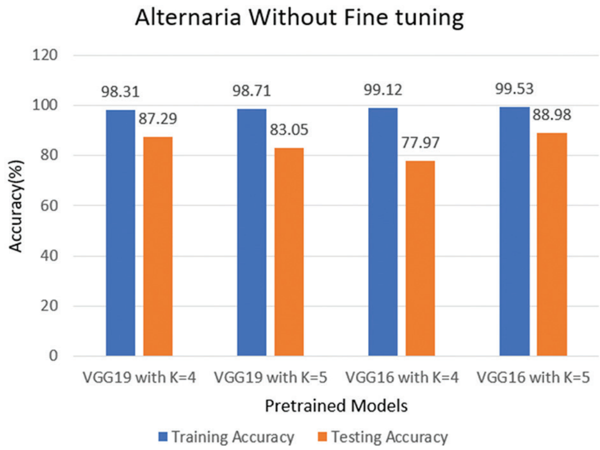



**Fig. 9.** Training and Testing accuracies for Alternaria Leaf Blight (without fine tuning)

The confusion matrix for the VGG16 model on the hold-out test set is presented in Table 5. In the initial stage, the model achieves a classification accuracy of 100% for K=5. The accuracy rates for the healthy and last stages are 93.22%, while the middle stage has a lower accuracy of 74.57%, making it more susceptible to misclassification.

**Table 5.** Confusion Matrix for Alternaria Leaf Blight for VGG16 at K5 cross-validation (without Fine-tuning)

| | | Predicted | | | |
|---|---|---|---|---|---|
| | | Stage 0 | Stage1 | Stage2 | Stage3 |
| Ground Truth | Stage 0 | **55** | 1 | 0 | 3 |
| | Stage1 | 0 | **59** | 0 | 0 |
| | Stage2 | 0 | 9 | **44** | 6 |
| | Stage3 | 1 | 3 | 0 | **55** |

Table 6 presents the confusion matrix for the VGG19 model tested on the hold-out dataset. All initial stages are accurately classified for K=4. The accuracy for the healthy stage is 91.52%, while the last stage achieves an accuracy of 86.44%. However, the middle stage has a lower accuracy of 71.18%, indicating it was frequently misclassified.

**Table 6.** Confusion Matrix for Alternaria Leaf Blight for VGG19 at K4 cross-validation without fine-tuning)

| | | Predicted | | | |
|---|---|---|---|---|---|
| | | Stage 0 | Stage1 | Stage2 | Stage3 |
| Ground Truth | Stage 0 | 54 | 0 | 0 | 5 |
| | Stage1 | 0 | 59 | 0 | 0 |
| | Stage2 | 0 | 9 | 42 | 8 |
| | Stage3 | 1 | 5 | 2 | 51 |

Fig. 10 shows training and testing accuracies per fold with fine-tuning for Alternaria Leaf Blight. The testing accuracy for VGG16 and VGG19 on the holdout dataset is improved with fine tuning=2

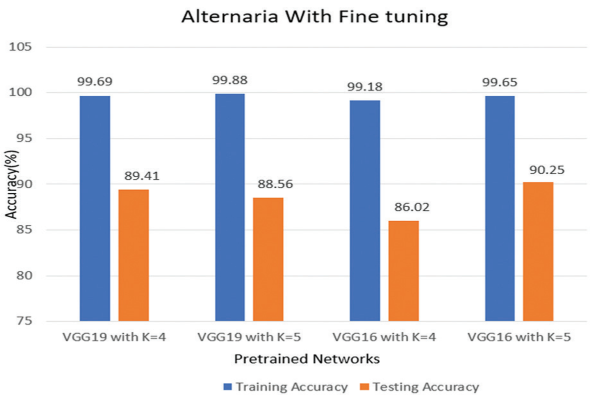



**Fig 10.** Training and Testing accuracies for Alternaria Leaf Blight with fine-tuning

The confusion matrix of the VGG16 model on the hold-out test set is shown in Table 7. The accuracy of the healthy stage is 86.44%, and the accuracy of the initial and last stages is 96.61%. The middle stage is not classified correctly, with an accuracy of 76.27%.

**Table 7.** Confusion Matrix for Alternaria Leaf Blight for VGG16 at K5 cross-validation (with fine-tuning)

| | | Predicted | | | |
|---|---|---|---|---|---|
| | | Stage 0 | Stage1 | Stage2 | Stage3 |
| Ground Truth | Stage 0 | 54 | 0 | 0 | 5 |
| | Stage1 | 0 | 59 | 0 | 0 |
| | Stage2 | 0 | 9 | 42 | 8 |
| | Stage3 | 1 | 5 | 2 | 51 |

The confusion matrix for the VGG19 model with K4 cross-validation on the hold-out data set is shown in Table 8. The Accuracy of the healthy stage is 72.88%. The initial stage is correctly classified at K=4. The Accuracies for the middle and last stages are 94.91% and 89.83% respectively.

**Table 8.** Confusion Matrix for Alternaria Leaf Blight for VGG19 at K4 cross-validation (with fine-tuning)

| | | Predicted | | | |
|---|---|---|---|---|---|
| | | Stage 0 | Stage1 | Stage2 | Stage3 |
| Ground Truth | Stage 0 | 43 | 0 | 3 | 13 |
| | Stage1 | 0 | 59 | 0 | 0 |
| | Stage2 | 0 | 0 | 56 | 3 |
| | Stage3 | 0 | 5 | 1 | 53 |

By the VGG19 model, stage 1 is more correctly classified than VGG16.

### 6.2. POWDERY MILDEW:

Both model's training and Testing accuracy plots (K4 and K5 cross-validation without fine-tuning) are shown in Fig 11. The plots show that the VGG19 model has the same accuracy for both K4 and K5 cross-validation. Also, VGG16 accuracy is increased to 74.15% after K5 cross-validation.

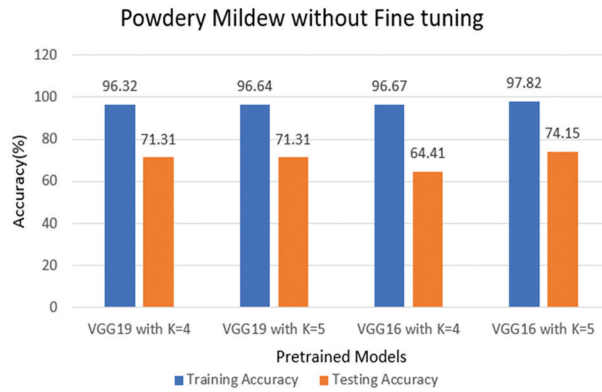



**Fig 11.** Training and Testing accuracies for Powdery Mildew (without fine tuning)

The confusion matrix of the VGG16 model on the hold-out test set is shown in Table 9. Accuracies of healthy and initial stage accuracies are 67.79% and 67.79%, respectively, and also the accuracy of the last stage is 96.61%, and the middle stage is 100% classified.

**Table 9.** Confusion Matrix for Powdery Mildew for VGG16 at K5 cross-validation (without fine-tuning)

| | | Predicted | | | |
|---|---|---|---|---|---|
| | | Stage 0 | Stage1 | Stage2 | Stage3 |
| Ground Truth | Stage 0 | 40 | 0 | 18 | 1 |
| | Stage1 | 0 | 40 | 19 | 0 |
| | Stage2 | 0 | 0 | 59 | 0 |
| | Stage3 | 0 | 0 | 2 | 57 |

Table 10 shows the confusion matrix for the VGG19 model on the hold-out data set. The Accuracy of the healthy stage is 91.52%. The initial stage is classified with an accuracy of 67.79%. The Accuracies for the middle and last stages are 98.30% and 79.66%, respectively.

**Table 10.** Confusion Matrix for Powdery Mildew for VGG19 at K5 cross-validation (without Fine-tuning)

| | | Predicted | | | |
|---|---|---|---|---|---|
| | | Stage 0 | Stage1 | Stage2 | Stage3 |
| Ground Truth | Stage 0 | 54 | 0 | 3 | 4 |
| | Stage1 | 5 | 40 | 15 | 1 |
| | Stage2 | 0 | 0 | 58 | 3 |
| | Stage3 | 1 | 0 | 13 | 47 |

VGG19 outperformed disease stage classification for the initial and middle stages with accuracy of 91.52% and 98.30% than VGG16.

From Fig. 12, we can observe that both model's testing accuracy is improved with fine tuning of 2 than without model fine-tuning (Refer Fig. 11). Also, after K5 cross-validation, the accuracy of both models is enhanced.

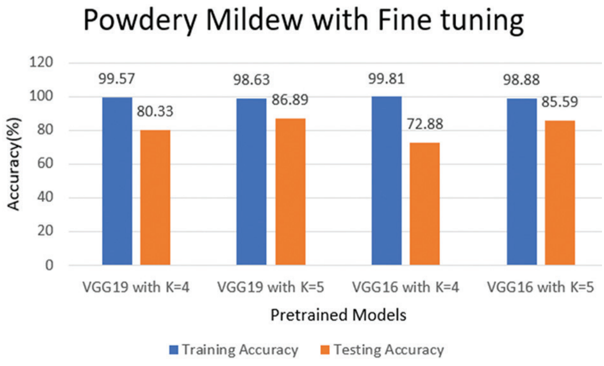



**Fig 12.** Training and Testing accuracies for Powdery Mildew (with fine tuning)

Table 11 shows the confusion matrix for the VGG16 model on the hold-out data set. The Accuracy of the healthy stage and last stage is 91.52%. The initial and middle stages are classified with an accuracy of 62.71% and 96.61%, respectively.

**Table 11.** Confusion Matrix for Powdery Mildew for VGG16 at K5 cross-validation (with Fine-tuning)

| | | Predicted | | | |
|---|---|---|---|---|---|
| | | Stage 0 | Stage1 | Stage2 | Stage3 |
| Ground Truth | Stage 0 | **54** | 1 | 2 | 2 |
| | Stage1 | 1 | **37** | 21 | 0 |
| | Stage2 | 1 | 0 | **57** | 1 |
| | Stage3 | 1 | 0 | 4 | **54** |

Table 12 shows the confusion matrix for the VGG19 model on the hold-out data set at K=5. The accuracy of the healthy and initial stages is 89.83% and 81.35%, whereas the middle and last stages are classified as 94.91% and 93.22%, respectively.

**Table 12.** Confusion Matrix for Powdery Mildew for VGG19 at K4 Cross-validation (with Fine-tuning)

| | | Predicted | | | |
|---|---|---|---|---|---|
| | | Stage 0 | Stage1 | Stage2 | Stage3 |
| Ground Truth | Stage 0 | **53** | 2 | 2 | 4 |
| | Stage1 | 3 | **48** | 9 | 1 |
| | Stage2 | 0 | 1 | **56** | 4 |
| | Stage3 | 0 | 0 | 6 | **55** |

Accuracies obtained by models are statistically witnessed by calculating confidence intervals. The confidence interval tells us how much precise the estimate values are calculated. This work estimates testing accuracy for each K fold with and without an acceptable tuning approach. This calculation shows accuracy is likely to come in which range. Gaussian distribution [25][26] of proportion helps to calculate the interval's radius. As this work is based on multiclass image classification, the radius of the interval [26] is calculated as per equation1 shown below:

$$interval = z * sqrt((accuracy * (1 - accuracy)) / n) \quad (1)$$

Where the interval is the radius of the confidence interval, accuracy is the estimate(testing accuracy of the model is used) that is to be witnessed, z is the number of standard deviations from the Gaussian distribution. z value is used as 1.96 with a 95% significance level for calculating the confidence interval, and n is the total number of samples. In this case, n is 236, which is the size of the out dataset. Table 13 shows the calculated confidence interval values with a range of testing accuracy.

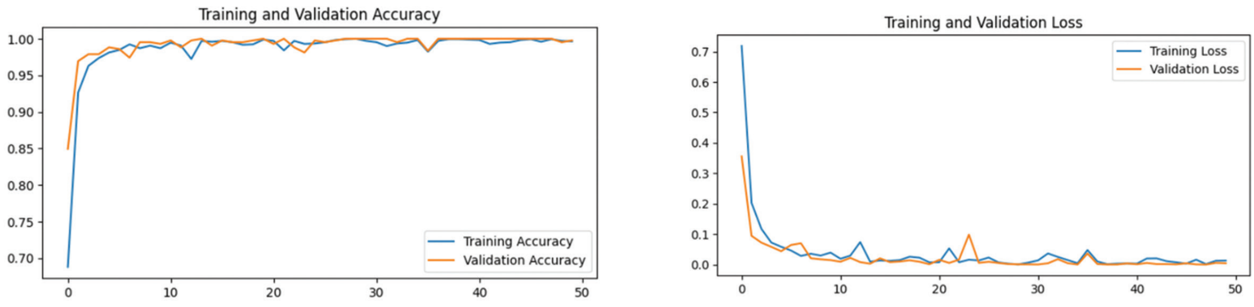
**Table 13.** Calculation of Confidence Intervals

| Model | Kfold | Accuracy (%) | Calculated Confidence Interval at 95% Significance Level | Range |
|---|---|---|---|---|
| **Alternaria Leaf Blight Without Fine Tuning** | | | | |
| VGG16 | K=4 | 77.97 | 5% | 72% to 82% |
| | K=5 | 88.98 | 4% | 84% to 92% |
| VGG19 | K=4 | 83.05 | 4% | 79% to 87% |
| | K=5 | 87.29 | 4% | 83% to 91% |
| **Alternaria Leaf Blight with fine-tuning** | | | | |
| VGG16 | K=4 | 86.02 | 4% | 82% to 90% |
| | **K=5** | **90.25** | **3%** | **87% to 93%** |
| VGG19 | K=4 | 88.56 | 4% | 84% to 92% |
| | K=5 | 89.41 | 4% | 85% to 93% |
| **Powdery Mildew Without Fine Tuning** | | | | |
| VGG16 | K=4 | 64.41 | 6% | 58% to 70% |
| | K=5 | 74.15 | 5% | 69% to 79% |
| VGG19 | K=4 | 71.31 | 5% | 66% to 76% |
| | K=5 | 71.31 | 5% | 66% to 76% |
| **Powdery Mildew with Fine Tuning** | | | | |
| VGG16 | K=4 | 72.88 | 5% | 67% to 77% |
| | K=5 | 85.59 | 4% | 81% to 89% |
| VGG19 | K=4 | 80.33 | 5% | 75% to 85% |
| | **K=5** | **86.89** | **4%** | **82% to 90%** |

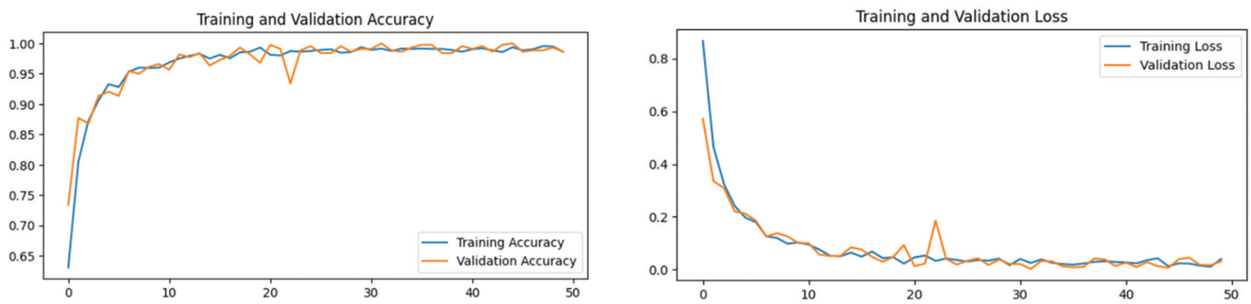
From the above table, a claim can be made that both models with K5 cross-validation by fine-tuning approach give smaller confidence intervals shown with yellow color highlighted form, which indicates the testing accuracy is more precise.

In this work, VGG16 and VGG19 models with K-fold (4 and 5) cross-validation with fine-tuning and without fine-tuning are applied to both diseases. But the confidence interval shows that VGG16 (K5 cross-validation) with finetuning = 2 for Alternaria leaf blight and VGG19(K5 cross-validation) with fine tuning=2 for Powdery Mildew giving smaller confidence interval and hence model accuracy 90.25% and 86.89% is claimed to be suitable for selected sample size respectively. Thus, Fig13(a) shows the accuracy and loss curves for models with precise testing accuracy. Models are trained till epoch 10. On the x-axis, 50 values are shown, which is epoch 10 per iteration, and the y-axis shows accuracy. Loss and accuracy curves are closer to each other. Fig

13(b) has higher training loss for initial iterations, and later, it decreases and flattens at the later iteration. Similarly, for both models' validation loss is also higher at the start of iterations, and later, it decreases. This indicates that the Powdery Mildew VGG19 model with K5 cross-validation, fine-tuned with 2 layers, and the Alternaria Leaf Blight VGG16 model with K5 cross-validation, fine-tuned with 2 layers, are good fit models.




(a)




(b)

**Fig. 13.** a) Accuracy and Loss curves for VGG16 after K5 cross validation(Alternaria Leaf Blight)
b) Accuracy and Loss curves for VGG19 after K5 cross validation(Powdery Mildew)

## 7. CONCLUSION

This study recommends a deep learning approach with fine tuning for performing disease stage classification of Alternaria leaf blight and Powdery Mildew diseases. This work creates a path for calculating plant disease severity. Both pre-trained models are trained with and without a fine-tuning approach based on training samples. Stratified K-fold validation is applied as samples per class are not uniformly distributed. The results showed that VGG16 and VGG19 both performed well after fine-tuning the parameter set to 2. VGG16 and VGG19 models gave more precise testing accuracy at 90.25% and 86.89% for Alternaria Leaf Blight and Powdery mildew, respectively, with small confidence intervals after K5 cross-validation. Both models have undergone K4 and K5 cross-validation; the result shows that with K5 cross-validation, the confidence interval is more minor. Hence K5 cross-validation is best suited for VGG16 and VGG19 models, demonstrating that the deep learning approach is the most favorable technique for disease stage classification based on plant disease severity estimation.

In future work, more image samples at different stages of both diseases will be collected with hyperspectral imaging with more powerful sensors, improving the model's performance by training with minutiae details of affected leaves. It can give better testing on unseen data. Also, the framework can be proposed to capture field intensity using advanced drone cameras.

## 8. ACKNOWLEDGEMENT:


This research could not have been completed without the invaluable support of Dr. R.D. Prasad, who serves as the Principal Scientist of Plant Pathology at ICAR-Indian Institute of Oilseeds Research. I extend my sincere appreciation to Dr. Maharudra Ghodke, who previously led the Latur Oilseed Research Centre, as well as to Santosh Waghmare and Sangita Aradwad, who hold the positions of senior and junior plant pathologists, respectively.


## 9. REFERENCES:


[1] Y. devi Puraikalan, M. Scott, "Sunflower Seeds (Helianthus Annuus) and Health Benefits: A Review", Recent Progress in Nutrition, Vol. 3, No. 3, 2023, pp. 1-5.

[2] R. Shivani, "Diseases of Sunflower: Necrosis, Leaf Blight, Mildew and Other Diseases", https://www.biologydiscussion.com/ (accessed: 2023)

[3] S. Guo, Y. Ge, K. Na Jom, "A review of phytochemistry, metabolite changes, and medicinal uses of the common sunflower seed and sprouts (Helianthus annuus L.)", Chemistry Central Journal, Vol. 11, 2017, p. 95.

[4] M. Alagawany, M. R. Farag, M. E. Abd El-Hack, K. Dhama, "The practical application of sunflower meal in poultry nutrition", Advances in Animal and Veterinary Sciences, Vol. 3, 2015, pp. 634-648.

[5] "India: sunflower seeds production Vol. 2023 | Statista", https://www.statista.com/statistics/769814/india-sunflower-seeds-production-Vol./ (accessed: 2023)

[6] J. Chai, H. Zeng, A. Li, W. T. E. Ngai, "Deep learning in computer vision: A critical review of emerging techniques and application scenarios", Machine Learning with Applications, Vol. 6, 2021.

[7] M. A. Hajam, T. Arif, A. M. Ud Din Khanday, M. Neshat, "An Effective Ensemble Convolutional Learning Model with Fine-Tuning for Medicinal Plant Leaf Identification", Information, Vol. 14, No. 11, 2023, p. 618.

[8] Z. Al Sahili, M. Awad, "The power of transfer learning in agricultural applications: AgriNet", Frontiers in Plant Science, Vol. 13, 2022.

[9] S. R. Dammavalam et al. "Leaf Image Classification with the Aid of Transfer Learning: A Deep Learning Approach", Current Chinese Computer Science, Vol. 1, 2020.

[10] S. Bhattarai, "New Plant Diseases Dataset (2019)", https://www.kaggle.com/vipoooool/new-plant-diseases-dataset (accessed: 2024)

[11] L. Yang, S. Xu, X.Y. Yu, H. B. Long, H. Zhang, Y. W. Zhu, "A new model based on improved VGG16 for corn weed identification", Frontiers in Plant Science, Vol. 14, 2023, p. 1205151.

[12] P. Lameski, "weed-datasets (2020)", https://gitee.com/Monster7/weed-datase/tree/master/ (accessed: 2024)

[13] P. K. Sethy, N. K. Barpanda, A. K. Rath, S. K. Behera, "Deep feature based rice leaf disease identification using support vector machine", Computers and Electronics in Agriculture, Vol. 175, 2020.

[14] G. Latif, S. E. Abdelhamid, R. E. Mallouhy, J. Alghazo, Z. A. Kazimi, "Deep Learning Utilization in Agriculture: Detection of Rice Plant Diseases Using an Improved CNN Model", Plants, Vol. 11, No. 17, 2022, p. 2230.

[15] A. Fiqri, "Rice Leafs 5 diseases (2021)", Kaggle, https://www.kaggle.com/datasets/adefiqri12/riceleafsv3 (accessed: 2024)

[16] G. Wang, Y. Sun, J. Wang, "Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning", Computational Intelligence and Neuroscience, Vol. 2017, 2017.

[17] R. Venkatachalam, K. Ilamurugu, "Sunflower: Index: Diseases of Sunflower", Development e-Courses for BSc Agriculture, Tamilnadu Agriculture University, http://www.eagri.org/, (accessed: 2024)

[18] Keras, "ImageDataGenerator | TensorFlow API Version", https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image/ImageDataGenerator (accessed: 2023)

[19] V. S. Magomadov, "Deep learning and its role in smart agriculture", Journal of Physics: Conference Series, Vol. 1399, No. 4, 2019.

[20] P. Marcelino, "Transfer learning from pre-trained models | Size-Similarity matrix and decision map for fine-tuning pre-trained models", https://towardsdatascience.com/transfer-learning-from-pre-trained-models (accessed: 2024)

[21] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv:1409.1556, 2014.

[22] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, "Rethinking the inception architecture for computer vision", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27-30 June 2016, pp. 2818–2826.

[23] Keras, "Keras 3 API documentation / Keras Applications", https://keras.io/api/applications/ (accessed: 2023)

[24] D. P. Kingma, J. Ba, "Adam: A method for stochastic optimization", Proceedings of the 3rd International Conference for Learning Representations, San Diego, CA, USA, 2015.

[25] A. Hazra, "Using the confidence interval confidently", Journal of Thoracic Disease, Vol. 9, No. 10, 2017, pp. 4124-4129.

[26] J. Brownlee, "Confidence Intervals for Machine Learning | Interval for classification accuracy", https://machinelearningmastery.com/confidence-intervals-for-machine-learning/ (accessed: 2024)