International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Deep Learning for Green Growth: Custom CNN Uncovers Cucumber Leaf Disease Secrets

Mohammad Tahmid Noor[1], B. M. Shahria Alam[2], Golam Kibria[3], Sidratul Moontaha[4], Fahad Ahammed[5]

[1,2,3,4,5]Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh

Abstract
There are many popular vegetables in the world, among them cucumber is one of the most famous and healthiest foods. It comes with multiple health benefits, especially in losing weight and in skincare. However, there can be multiple diseases that can attack the cucumber leaf and the cucumber itself. Cucumber leaf disease research plays a cardinal role in agricultural fields. Cucumber leaf disease is a special disease that harms its leaves and causes different types of diseases. In real fields, it is difficult to identify different diseases on the leaf due to the complex environment. It is also a very challenging and time-consuming process to identify leaf diseases. To overcome those barriers and detect the diseases we have used two learning methods namely DenseNet201 and VGG16 model. In this research, we have created a new dataset with four classes containing 2159 pictures with an accuracy rate of 98.97% from DenseNet201. The image recognition methods choose the image and find the disease's spot by breaking the image frame by frame. We have rescaled all the images while preprocessing where the image size is 224×224. The methods show great success in finding the disease spots.

Keywords: CNN, Cucumber, Disease, Machine learning, Image Classification

1. Introduction

Cucumber is very important vegetable for the human body. Cucumber contains almost 95% of water which make our body hydrated. Cucumber is a good source of fiber which aids digestion and prevents constipation. Cucumber plays a vital role in regulating blood pressure, they help with digestion, reduce blood sugar, help lose weight, reduce the risk of cancer, and so on. Because of its massive health and agricultural benefits, a report shows that in the year 2020, 90.35 million tons were the global yield, and 2.25 million hectares were the cucumber planting area in the global [1].

In the agricultural sector, cucumbers play a vital role in the whole world. Cucumbers are one of the most popular vegetables in the world and much more popular for farmers to crop. But its growth falls as different kinds of diseases like Powdery Mildew, Downy Mildew, Angular Leaf Spot, Bacterial Wilt, Belly Rot and anthracnose lesions attack the leaf [2]. To identify its diseases, farmers and doctors work in this fields. It is also proven that the identification of these diseases manually can consume vast amount of time, manpower, and money. Wrong identification and using wrong pesticides can be harmful to cucumber agriculture and can bring large amounts of production costs and huge losses to the farmers.

To relieve the farmers, several methods were introduced to identify the diseases such as probabilistic segmentation [3], pattern recognition, image classification, deep machine learning, etc. Diseases on the plant can be easily recognized by its color, spots, and texture extracted from the leaf. It is also very challenging to identify diseases as there are vast numbers of gardens, and a variety of diseases. There are several types of conditions are responsible for the different types of diseases. That's why different types of intelligence base approaches have been introduced to identify them.

It is necessary to have accurate and efficient disease detection for modern farming. There are different types of diseases like anthracnose lesions, downy mildew, and powdery mildew. For real time monitoring and data collection, the (IoT) Internet of Things has led us to some new opportunities [4]. It has a very magnificent part in the disease detection fields. It captures the images of the affected leaf and collects its data by implementing sensors, and cameras. In real-time, the information on the leaf can be transmitted and processed to identify the diseases. Deep learning methodologies shown significantly enhancement is due to its recent advancement [5]. Because of having different colors and appearances of varying leaf, it is difficult to identify plants from other developing problems. As a result, the general response has been poor.

There are several different ways in which researchers have attempted to address this problem. The traditional method with manual image analysis is established by segmenting healthy and diseased leaves from the characteristics like color, shape. But these approaches often grievously lacked in accuracy. Computers have been used in the analysis of images more recently. Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs) techniques have been used [6]. These techniques revolve around teaching computers to detect visual patterns that are relevant for specific diseases.

CNN models have proven effective in image processing and image analysis tasks. Nazar Hussain's work based CNN models, VGG and inception v3 deep learning models with using five types of cucumber disease from a private dataset and able to gain much higher accuracy of 96.5% [7]. Computer vision framework also undoubtedly performed pretty well. Sumair Aziz have done a work by using this method and able to gain 94% of accuracy from five different datasets of tomato leaf dataset [8]. A deep learning based idea which worked by Snehal Banarase utilizes symptom data of the apple leaf diseases with MobileNetV2 model [9]. They have done some magnificent worked with four different classes and got achievement of 99.36% accuracy rate [10].

One branch of AI, deep learning has shown the potential to achieve this. This technology is vastly used to classification and recognition [11]. Further, to improve accuracy, the combination of various CNN models like VGG19 and Inception v3 are used [12]. Another good thing is that the deep learning approaches have proven to be inexpensive. Based on the accuracy and error rate, the performances are evaluated. Researchers have presented a number of deep learning models where the error rates for AlexNet, a VGG deep model, and GoogleNet were 15.3%, 7.3%, and 5.1%, respectively [7].

As we all know that these diseases can highly impact on crops, so accurate detection is crucial to get effective results. The method we are using such as deep learning and CNN models have shown some promising results in the identification of the diseases. After collecting the dataset, we create a custom CNN model for the affected cucumber leaf analysis. Then we divided the dataset into training, validation and testing sets. Using the model TensorFlow core we are going to build and run the CNN model. Then the performance of the trained model will be evaluated by using metrics like accuracy, precision and recall. The next part which is in Section II presents the literature review of related works. Research methodology, dataset, VGG16, DenseNet201 are narrated in Section III. Experimental results and comparison of performance of different previous works have been discussed in Section IV. Some concluding touch on and future works are given in Section V.

2. Related Works

We have gathered some related work for better understanding. These related works can lead us to a comparison with our proposed work. Shanwen Zhang and his team worked on a CNN model to identify cucumber leaf disease [17]. They segmented the disease leaf images, extracted shape and color and classifying disease leaf by sparse representation. By using this method they have gained 85.7% accuracy with 4 classes and able to manage seven major cucumber diseases. Juncheng Ma has done their work using the deep convolutional neural network method with 4 classes containing 14,208 images and gained an accuracy of 93.4% through DCNN. ReLu and softmax are used for the activation and to classify the disease [18].

Zia ur Rehman and his team have done their work using Deep learning and MASK RCNN method with 4 classes which contain around 770 images. They used a pre-trained CNN model extract the feature and gained an accuracy of 96.6% [19]. Jingyao Zhang has done their work using the Deep learning method with 3 classes. They used Inception convolutional neural network to identify accurately of the leaf diseases using generated training samples and gain an accuracy of 96.11% [4]. Xuebing Bai also worked on a cucumber leaf spot disease based on 129 images. An effective and robust segmentation was provided by their proposed method and their average segmentation error was only 0.12% [20]. Nusrat Sultana and her team worked on a data set based on 8 classes and 1280 raw images. They built an effective machine vision-based model for efficiently detecting the cucumber leaf diseases. They gain a total of 96% accuracy [21].

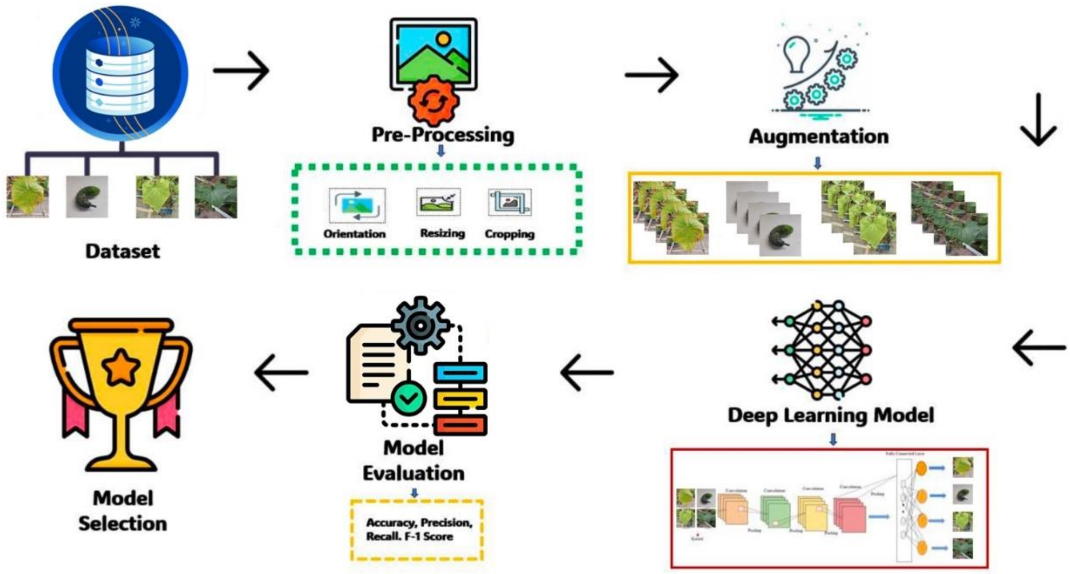
Researchers use deep learning based approaches so that the leaf disease detection can be more efficient. Zhang et al. proposed a method called deep learning using the dilated convolutional kernel to extend the extraction process [13]. To minimize overfitting probability, they use global pooling layer reducing the size of CNN parameters [14]. By using a multi-kernel for feature extraction and the ability to learn fuzzy rules were able achieve excellent performance in disease identification accuracy against database images. The deep feature is somehow obligatory for the selection step [15]. A study segmented the infected part of the cucumber leaf by a new Sharif saliency-based (SHSB) approach and extracted deep features from VGG19, and VGG16 [7]. A novel method used a Vese-Osher active contour model with fused feature maps generated at various levels on both layers. To fully automate the detection and classification processes new tools are required [16].

M. Attique Khan proposed an automated framework using the deep learning method and best feature selection. They evaluated their proposed framework in each step and gained an accuracy of 98.4% [22]. Chunshan Wang and his team have done their work using DeepLabV3+ and U-Net on the dataset of 9 classes. They proposed to use a two-stage model which are DeepLabV3 and U-Net for cucumber leaf disease and gain an accuracy of 93.27% [18]. Abdul Rehman and his team conduct work using LTriDP on the dataset of cucumber. Their main aim was to design a model that is robust to noise in real-time. They gain an accuracy of 96.80% [23]. Supreetha S has done their work using the Deep learning method MobileNetV2 on the data set of 2 classes. They have proposed a lightweight transfer learning-based approach which proves an effective processing method. It enhances the leaf images, improves the classification performance and gains an accuracy of 92.16% [24].

3. Methodology

We described in detail about the proposed work here. Firstly, we previously trained the deep CNN model to extract deep features with the help of transfer learning. This is an important feature for recognizing the patterns. This feature mainly uses shape, color, point, and appearance to represent an object. A deep CNN model has many types of layers in which, the first layer is used to pass the input image to the convolutional layer to calculate the dot product of weights. The pooling and removal of the inactive neurons were done by the ReLu layer. To classify the features that are computed SoftMax layer is used. In Figure 1, shows the proposed methodology. Finally, one pre trained deep CNN model namely VGG16 is used to extract its features.

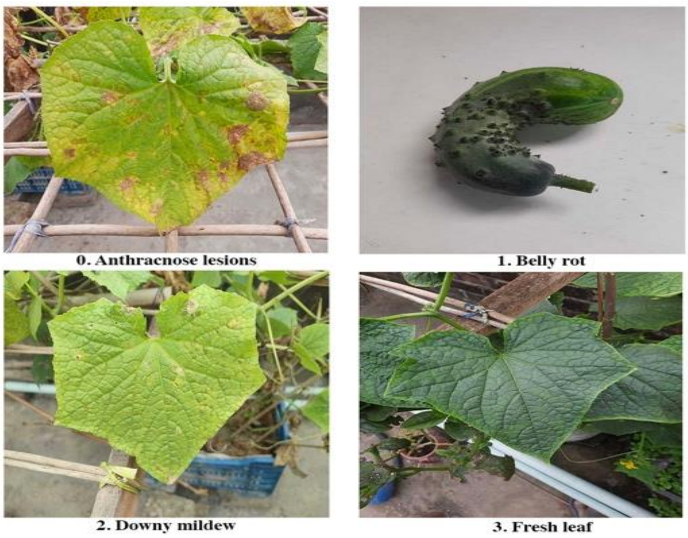
Figure 1: Methodology.



1. Dataset

We have used the proposed method for disease detection. The dataset has four diseases which are downy mildew (565 images), anthracnose lesions (536 images), Belly Rot (534 images), and fresh leaves (528 images). The proposed methods simulation was performed on MacBook Air M2 15 inch. We have used this dataset which was collected from Mendeley dataset for our proposed work [25]. We have shown a sample inputs of the dataset below in Figure 2.

Figure 2: Sample inputs of the Cucumber leaf

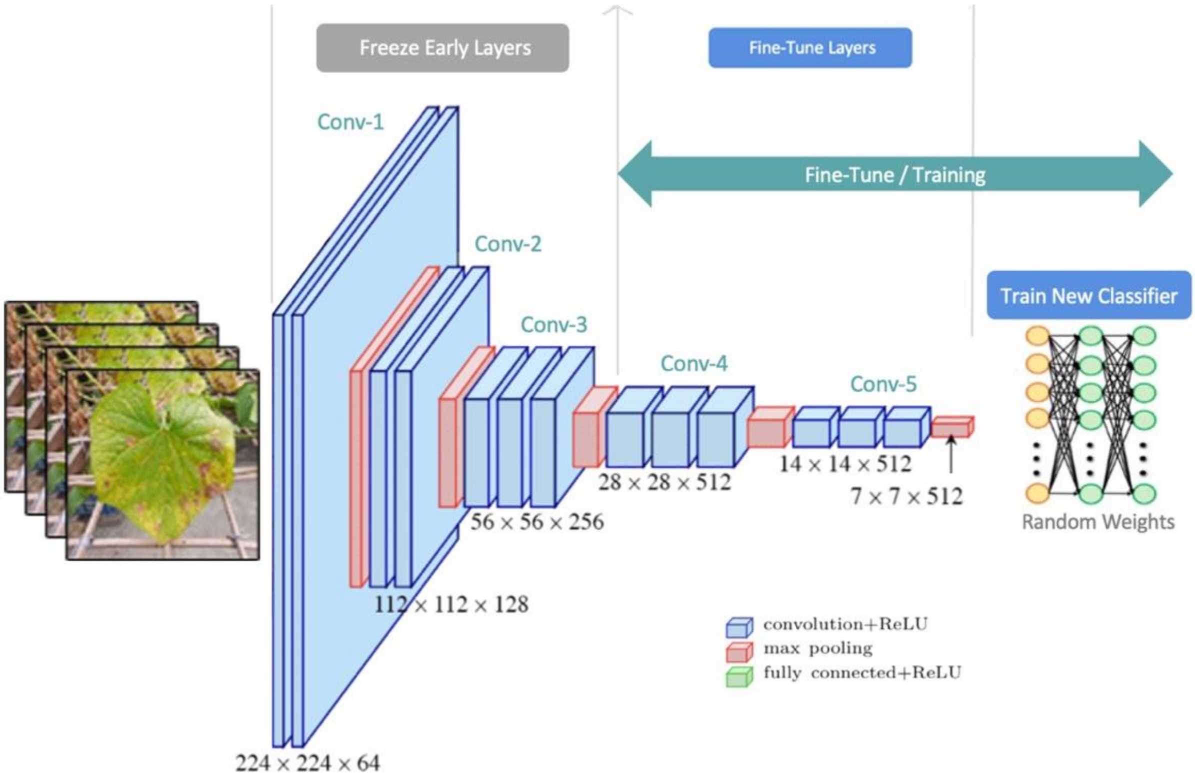


In the dataset image, 0 indicates anthracnose lesions diseases. It causes some spots in the leaf as little and then grow round yellow dots. 1 indicates Belly rot that shows the symptoms on the underside of a cucumber and produces water-soaked. 2 indicates Downy Mildew disease. The affected leaves turn into pale green or yellow on their peak. 3 indicating the fresh and healthy cucumber leaf which has a vibrant green color. The leaf is so fresh and healthy that gives a soothing feeling.

2. VGG-16

VGG16 is a very powerful convolutional neural network (CNN) that has performed impressive job in image classification. On the basis of a massive dataset, it is a pre-trained model. Using the pre trained model, we extract features and then fine tune the entire model. VGG16 mainly expands the depth of the network by joining more convolutional layers while using very small convolutional filters in all layers [13]. In Figure 3 the architecture of the VGG16 model is shown. We can say that VGG16 is a very powerful foundation for cucumber leaf disease detection. Table 1 showcases hyperparameter tuning for VGG16. Figure 4 shows the layer structure of VGG16.

Figure 3: Architecture of the model VGG16.



International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Figure 4: Structure of the model VGG16

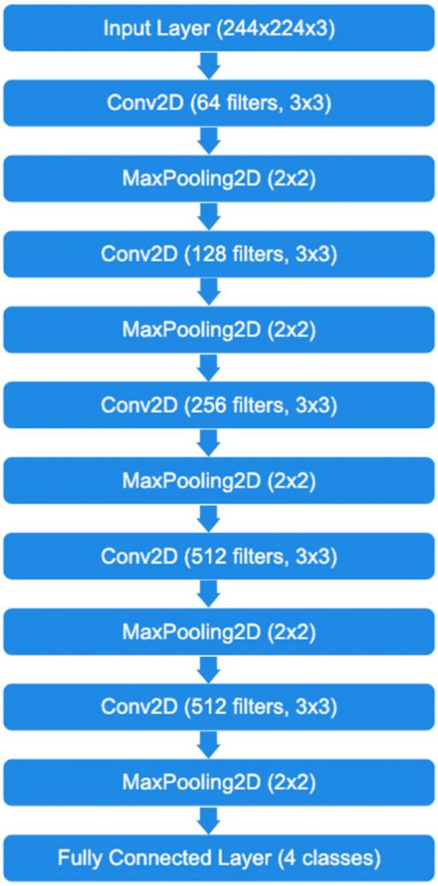


Table 1: Hyperparameter tuning for VGG16.

Batch size	32
Learning rate	0.0002
Optimizer	Adam
Loss function	categorical_crossentropy
Number of epochs	80
Patience	10

3. DenseNet201

DenseNet201 is basically a convolutional neural network that establishes connections between layers. DenseNet201 has 201 layers which include convolutional networks, pooling, and fully connected layers. It improves computational efficiency by utilizing bottleneck layers and transition layers and also reduces the number of parameters. DenseNet201 is a powerful tool for cucumber leaf disease detection. Figure 5 shows the architecture of DenseNet201. Here in Table 2, shows hyperparameter tuning for DenseNet201.

International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Figure 5: Architecture of the model DenseNet201.

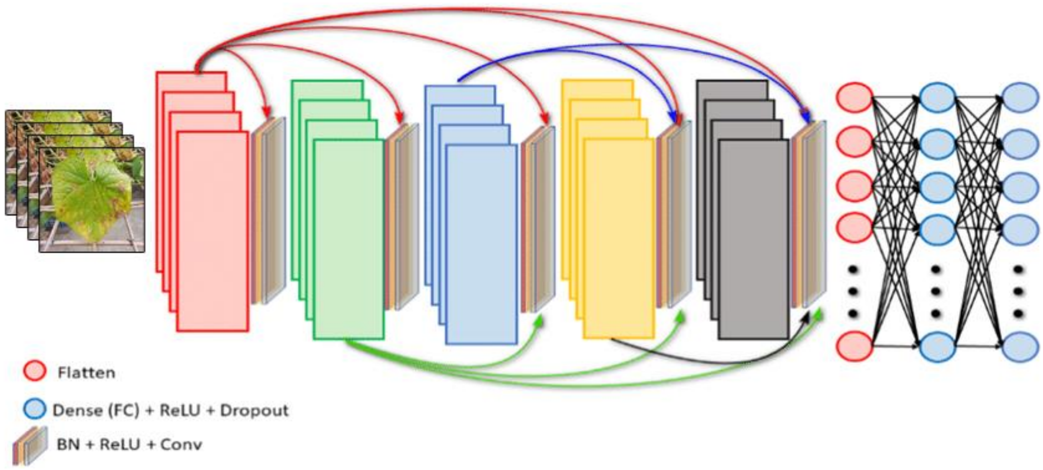
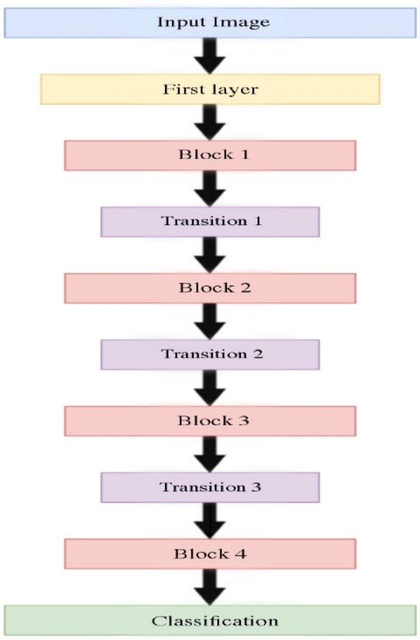


Table 2: Hyperparameter tuning for DenseNet201

Batch size	20
Learning rate	0.0002
Optimizer	Adam
Loss function	categorical_crossentropy
Number of epochs	36
Patience	10

An advanced deep learning model like DenseNet201 promises to enhance crop health and productivity in agricultural sectors. Figure 6 shows the layer structure of DenseNet201.

Figure 6: Structure of the model DenseNet201



4. Experimental Result and Analysis

The VGG16 model has successfully gained an accuracy of 98.46% in training and a validation accuracy of 98.44% succeeding in training for 80 epochs. 255 seconds per step has taken by the training phase and we applied 80 training epochs. The epochs are displayed in Figure 7. Classification-wise accuracy is shown in Table 3.

Figure 7: VGG16 model – quality metrics evaluation during training: (a) accuracy (b) loss.

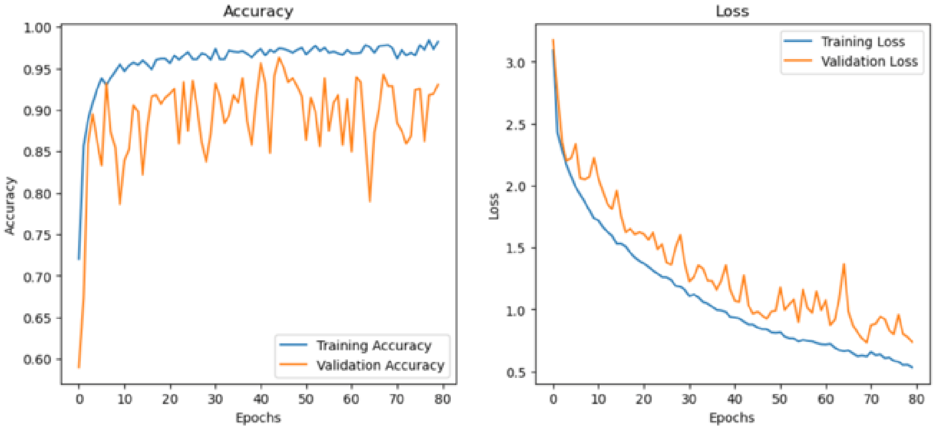


Table 3 Classification-wise accuracy table for VGG16

Classes	Precision	Recall	F1-score	Support	Accuracy	Weighted avg	Macro avg
Anthracnose lesions	0.99	0.99	0.99	192	0.98	0.98	0.98
Belly Rot	1.00	1.00	1.00	191			
Downy Mildew	0.99	0.96	0.97	202			
Fresh Leaf	0.95	0.99	0.97	189			

Figure 8 shows a confusion matrix of VGG16 model. For predicting cucumber leaf condition, the confusion matrix for the model VGG16 indicates a high accuracy. Some minor misclassifications are observed between anthracnose lesions and fresh leaves, indicating the model's strong performance but also leaving some room for potential improvement.

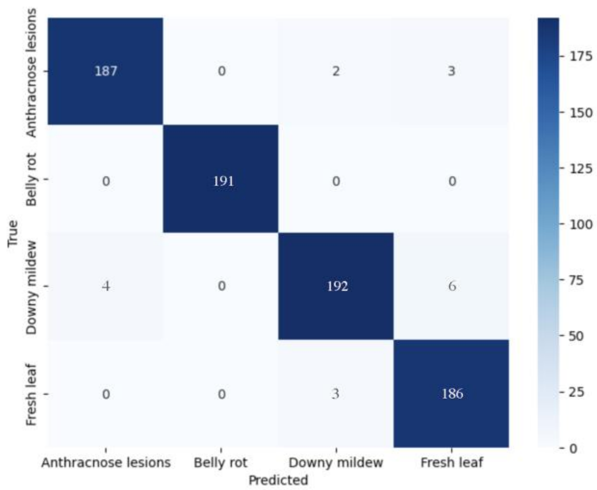


Figure 8 Confusion matrix of VGG16 model.

The DenseNet201 model has successfully gained accuracy in the training period of 99.95% and validation accuracy of 98.97% succeeding training for 35 epochs. 135 seconds per step has taken by the training phase and we applied 35 training epochs. The epochs are displayed in Figure 9. And the classification-wise accuracy is shown in Table 4.

Figure 9: DenseNet201 model – quality metrics evaluation during training: (a) accuracy (b) loss.

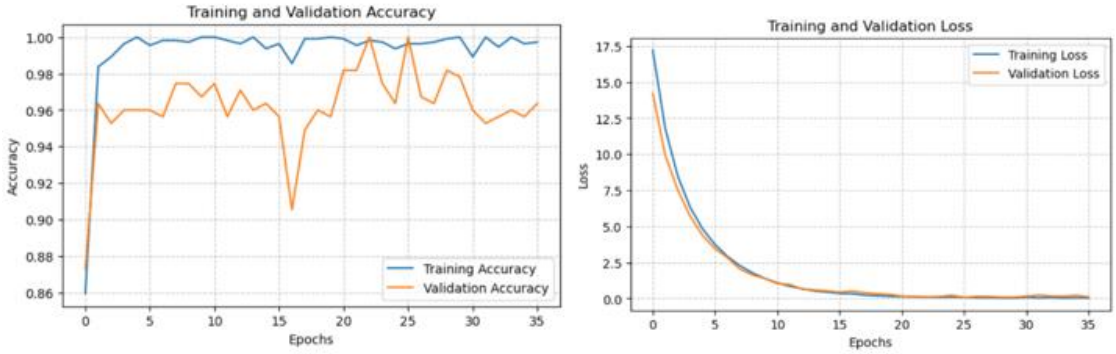
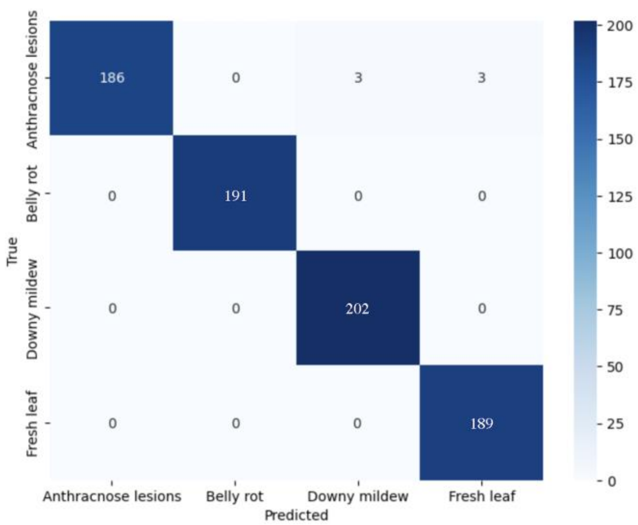


Table 4: Classification wise accuracy table for DenseNet201

Classes	Precision	Recall	F1-score	Support	Accuracy	Weighted avg	Macro avg
Anthracnose lesions	0.97	0.99	0.98	192	0.99	0.99	0.99
Belly Rot	1.00	1.00	1.00	191			
Downy Mildew	1.00	0.97	0.98	202			
Fresh Leaf	0.98	1.00	0.99	189			

Figure 10 shows a confusion matrix for the DenseNet201 model. This model demonstrates exceptional accuracy across all four categories with nearly all the predictions corrected. This model shows minimal confusion, particularly between anthracnose lesions and fresh leaves. This model holds a very strong capability to distinguish between different leaf diseases and healthy leaves.

Figure 10: Confusion matrix of DenseNet201 model.



Comparison between the two models

The training time between the two models VGG16 was taken 5 hour 0.9 minutes and the DenseNet201 was taken 1 hour 66 minutes. The accuracy of training and testing methods are compared in Table 5. The two models have gained an excellent accuracy of over 90%. The training and the test accuracy of all models have different numbers of training epochs.

Table 5: Accuracy comparison between the two models

Algorithm	VGG16	DenseNet201
Training accuracy	98.46	99.95
Test accuracy	98.44	98.97

To showcase our work and to prove that our method was more effective than the others we gathered some comparisons with some previous works shown in Table 6. We have successfully achieved more accuracy and computational time which was quite competitive with the others. Our proposed method has brought 98% accuracy by using Deep feature fusion with WOA (Whale Optimization Algorithm) optimization. On the other hand, Shanwen Zhang has gained an accuracy of 85.7% [17], Juncheng Ma has gained an accuracy of 93.4% [18], Nazar Hussain has gained an accuracy of 96.5% [7], Zia ur Rehman has gained an accuracy of 96.6% [19], Jingyao Zhang has gained accuracy of 96.11% respectively [4].

Table 6: Comparison between the proposed model and previous works

Author	**Method**	**Accuracy (%)**
Shanwen Zhang [17]	Sparse representation classification	85.7
Juncheng Ma [18]	Deep convolutional neural network	93.4
Nazar Hussain [7]	Best feature selection	96.5
Zia ur Rehman [19]	Deep learning and MASK RCNN	96.6
Jingyao Zhang [4]	Deep learning	96.11
Proposed model	CNN VGG16	98.44
	DenseNet201	**98.97**

5. Conclusion

Cucumbers are the common diet because of their refreshing taste and nutritional value. Economic importance, widespread consumption, and health benefits make it a significant agricultural crop. Cucumbers are susceptible to various diseases especially affecting the leaves. To identify the diseases perfectly many methods were introduced. Among the methods, we have used DenseNet201 and CNN VGG16 models achieving a remarkable accuracy of 98.97% and 98.44% respectively. In the revolution of plant disease identification, this high level of accuracy indicates the potential of deep learning techniques. While achieving outstanding accuracy in both models, DenseNet201 showed us slight advantages in the computational accuracy which makes it more suitable for real-time application. In conclusion, DenseNet201 and CNN VGG16 models seem a powerful tool for detecting cucumber leaf disease. Implementing both models offers significant advancement in agriculture sectors and more sustainable farming practices. In the future, we look forward to working more on additional deep learning models to amplify their accuracy and robustness. For practical and on field applications, we can implement real time disease detection on edge devices. We can explore the transfer learning methods with using larger agricultural datasets to improve the disease detection.

List of References

1. Chunshan Wang, Pengfei Du, Huarui Wu, Jiuxi Li, Chunjiang Zhao, Huaji Zhu,"A cucumber leaf disease severity classification method based on the fusion of DeepLabV3+ and U-Net". Comput. Electron. Agric. 2021, 189, 106373.
2. Muhammad Attique Khan, Abdullah Alqahtani, Aimal Khan, Shtwai Alsuba, Adel Binbusayyis, M Munawwar Iqbal Ch, Hwan-Seung Yong, Jaehyuk Cha, "Cucumber Leaf Diseases Recognition Using Multi Level Deep Entropy-ELM Feature Selection" Applied Sciences ,Volume 12, Issue 2 10.3390/app12020593
3. Akram, T.; Sharif, M.; Alhaisoni, M.; Saba, T.; Nawaz, N. A probabilistic segmentation and entropy-rank correlation-based feature selection approach for the recognition of fruit diseases. EURASIP J. Image Video Process. 2021, 2021, 1–28.
4. Jingyao Zhang, Yuan Rao, Chao Man, Zhaohui Jiang and Shaowen Li1, "Identification of cucumber leaf diseases using deep learning and small sample size for agricultural Internet of Things" International Journal of Distributed Sensor Networks, vol. 17(4), pages 15501477211, April.
5. Xiaohui Chen, Dongyuan Shi, Hengwei Zhang, José Antonio Sánchez Pérez, Xinting Yang, Ming Li "Early diagnosis of greenhouse cucumber downy mildew in seedling stage using chlorophyll fluorescence imaging technology", Biosystems engineering, Volume 242, June 2024, Pages 107-122.
6. L. Kaur and V. J. D. M. Laxmi, Detection of unhealthy region of plant leaves using neural network, vol. 1, no. 05, pp. 34-42, 2016.
7. Nazar Hussain, Muhammad Attique Khan, Usman Tariq, Seifedine Kadry, Muhammad Asfand E. Yar, Almetwally M. Mostafa, Abeer Ali Alnuaim6 and Shafiq Ahmad, "Multiclass Cucumber Leaf Diseases Recognition Using Best Feature Selection" CMC Journals, Vol. 70, No.2,2022
8. S. Aziz, M. Bashir, O. Mughal, M. U. Khan and A. Khan, "Image Pattern Classification for Plant Disease Identification using Local Tri-directional Features", 2019 IEEE 10th Annual Information Technology Electronics and Mobile Communication Conference (IEMCON), pp. 0973-0978, 2019.
9. M. Shyamala Devi, D. Umanandhini, Sakineti Aayush Kumar, "Seven Convolutional Nadam Optimized Deep Convolutional Neural Network based Cucumber Leaf Disease Prediction", 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp.1-6, 2023.
10. Snehal Banarase, Suresh Shirbahadurkar, "The Orchard Guard: Deep Learning powered apple leaf disease detection with MobileNetV2 model" Volume 12 No. 4 (2024)
11. J. G. J. B. e. Barbedo, Factors influencing the use of deep learning for plant disease recognition, vol. 172, pp. 84-91, 2018.
12. Afsana Mimi, Sayeda Fatema Tuj Zohura, Muhammad Ibrahim, Riddho Ridwanul Haque, Omar Farrok, Taskeed Jabid, Md Sawkat Ali,"Identifying Selected Diseases of Leaves using Deep Learning and Transfer Learning Models", April 2023Machine Graphics and Vision 32(1):55-71, DOI:10.22630/MGV.2023.32.1.3
13. S. Zhang, W. Huang and C. Zhang, "Three-channel convolutional neural networks for vegetable leaf disease recognition," Cognitive Systems Research, vol. 53, pp. 31–41, 2019.

14. Varun Kumar, Deepak Banerjee, Deepak Upadhyay, Mukesh Singh, Kanegonda Ravi Chythanya, "CNN-Random Forest Hybrid Model for Improved Ginger Leaf Disease Classification", 2024 International Conference on E-mobility, Power Control and Smart Systems (ICEMPS), pp.1-6, 2024.
15. Akram, T.; Sharif, M.; Alhaisoni, M.; Saba, T.; Nawaz, N. A probabilistic segmentation and entropy-rank correlation-based feature selection approach for the recognition of fruit diseases. EURASIP J. Image Video Process. 2021, 2021, 1–28.
16. Iqbal, Z.; Sharif, M.; Shah, J.H.; ur Rehman, M.H.; Javed, K. An automated detection and classification of citrus plant diseases using image processing techniques: A review. Comput. Electron. Agric. 2018, 153, 12–32.
17. S. Zhang, X. Wu, Z. You and L. Zhang, "Leaf image based cucumber disease recognition using sparse representation classification," Computers and Electronics in Agriculture, vol. 134, pp. 135–141, 2017.
18. J. Ma, K. Du, F. Zheng, L. Zhang and Z. Sun, "A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network," Computers and Electronics in Agriculture, vol. 154, pp. 18–24, 2018.
19. Z. u. Rehman, M. A. Khan, F. Ahmed, R. Damaševicius and W. Nisar, "Recognizing apple leaf ˇ diseases using a novel parallel real-time processing framework based on MASK RCNN and transfer learning: An application for smart agriculture," IET Image Processing, vol. 1, pp. 1–23, 2021.
20. Xuebing Bai, Xinxing Li, Zetian Fu, Xiongjie Lv and Lingxian Zhang, "A fuzzy clustering segmentation method based on neighborhood grayscale information for defining cucumber leaf spot disease images" Computers and Electronics in Agriculture Volume 136, 15 April 2017, Pages 157-165.
21. Nusrat Sultana, Sumaita Binte Shorif, Morium Akter, and Mohammad Shorif Uddin, "A dataset for successful recognition of cucumber diseases" Data Brief, v.49; 2023 Aug. PMC10338290 Sharif, M., et al.: Detection and classification of citrus diseases in agri-culture based on optimized weighted segmentation and feature selection.Comput. Electron. Agric. 150, 220–234, (2018).
22. Akram, T.; Sharif, M.; Javed, K.; Raza, M.; Saba, T. An automated system for cucumber leaf diseased spot detection and classification using improved saliency method and deep features selection. Multimed. Tools Appl. 2020, 79, 18627–18656.
23. Abdul Rehman, Zain Tariq, Shoaib ul din Memon, Ahmed Zaib, Muhammad Umar Khan, Sumair Aziz," Cucumber Leaf Disease Classification using Local Tri-directional Patterns and Haralick Features" ICAI52203.2021.9445237.
24. Supreetha S, S H Manjula, "Identification of Downy Mildew and Leaf Miner Diseases in Hybrid Cucumber", 2024 International Conference on Green Energy, Computing and Sustainable Technology (GECOST), pp.240-244, 2024.
25. Rashid, Mohammad Rifat Ahmmad; Tarin, Taslima Khan ; Kamara, Raka ; Mou, Mahbuba Yasmin ; Rabbi, Sheikh Fajlay ; Hasan, Mahamudul (2024), "Plant Leaf Freshness and Disease Detection Dataset From Bangladesh", Mendeley Data, V3, doi: 10.17632/n67gctmjyj.3