International Journal of Electrical and Computer Engineering (IJECE)
Vol. 15, No. 3, June 2025, pp. 3127~3138
ISSN: 2088-8708, DOI: 10.11591/ijece.v15i3.pp3127-3138

Cucumber leaf disease identification in real-time via deep learning based algorithms

Md Mizanur Rahman[1], Mahimul Islam Nadim[1], Mahinur Akther[2], Ahad Ullah[3], Jakaria Ahmed[1], Muhammad Jalal Uddin Ahmed[1], Israt Jahan[1]
[1]Department of Computer Science and Engineering, Faculty of Science and Information Technology, Daffodil International University, Dhaka, Bangladesh
[2]Department of English, Faculty of Humanities and Social Science, Daffodil International University, Dhaka, Bangladesh
[3]Department of Civil Engineering, Faculty of Engineering, Daffodil International University, Dhaka, Bangladesh

Article Info

Article history:

Received Jul 31, 2024
Revised Feb 6, 2025
Accepted Mar 4, 2025

Keywords:

Cucumber disease
Cucumber leaf
Deep learning
Disease recognition
EfficientNetB4
Image processing

ABSTRACT

Cucumber is a cash crop in Bangladesh as it is a side dish grown commercially in cultivable lands year-round. The early prediction of disease-prone crops could save grooming time and minimize losses. The conventional method of examining leaves just through observation of the human eye could only detect the diseases at an advanced stage without a concrete decision of which disease it might be and regular inspection is labour intensive, inaccurate and often unreliable. This study evaluates machine learning-based image analysis for classifying healthy and diseased cucumber leaves by training deep learning models to detect and identify observable traits. Total 1,629 images use as primary dataset and all the data collected from the cucumber field of Bangladesh. To fulfill this purpose, convolutional neural network (CNN), InceptionV3, and EfficientNetB4 are the models implemented in this paper to improve the classification of objects. The dataset was optimized by pre-processing techniques and the leaves are classified into four categories, namely angular leaf spot, downy mildew, powdery mildew, and good leaf. The EfficienNetB4 model achieved the highest train and test accuracy respectively 95% and 87%. A comparative examination of the available models was conducted in this paper to reach a solid decision.

Corresponding Author:

Md Mizanur Rahman
Department of Computer Science and Engineering, Faculty of Science and Information Technology,
Daffodil International University
Daffodil Smart City, Birulia, Savar, Dhaka-1216
Email: mizanurrahman.cse@diu.edu.bd

1. INTRODUCTION

To improve crop health and raise agricultural output, it is imperative to diagnose plant diseases correctly and promptly. A crucial indication of early success in agriculture depends upon the farmer's ability to early detection of diseases which could substantially assist in the prevention of further progression of the disease. Hence, it is imperative that not only there are measures in place for the early detection of such ailments in crops but also detection of the particular disease at play so that the parties concerned may be able to take steps to cure the particular ailment which would determine the wider success of the crop. Machine learning (ML) and image processing highlight the importance of early disease identification in cucumbers. [1]. Autonomous device using convolutional neural network (CNN), artificial neural network (ANN) and others model achieves accurate detection, suggesting further research for real-time capabilities [2].

The dataset used in this paper is primary data. We take pictures of cucumber leaves at various seasons in very remote areas of Bangladesh. Four distinct classes are included in the data set: angular leaf spot, powdery mildew, downy mildew, and good leaf. An experienced agricultural officer has confirmed each class and images. We attempt to develop a suitable detection system based on prior research by achieving a satisfactory accuracy and making it more stable and high-quality than prior research. Since Bangladesh is an agricultural nation, it can play a significant role in the agricultural sector's future growth and requires a scientific approach to implementing artificial intelligence (AI) [3].

On leaf disease there are several relatable study is existing specially in cucumber leaf, we try to make an overview of this relevant study. Like Mia *et al.* try to detect timely and treatment of these diseases are critical for minimizing losses and implementing timely management strategies. Cucumber leaf diseases are classified into seven categories and training images in ML methods using K-means_clustering and random_forest provides an accuracy of 89.93%. Additionally, the study addresses the performance of InceptionV3, MobileNetV2, and visual geometry group 16 (VGG16) [4]. Krithika and her colleagues implemented a computer-aided approaches for early disease classification in salad cucumber leaves, with an emphasis on smart organic farming difficulties in India were addressed by combining support vector machine (SVM) and K-Means_clustering enabling the existing trend of machine learning-based image processing for precision agriculture. The system possesses practical viability for widespread agricultural benefit as it has demonstrated efficacy in identifying diseases in crops such as paddy leaves and cotton [5]. Zhang *et al.* try to shows - Tackles important issues in real-world plant disease identification with internet of things (IoT). The suggested approach, which combines two-stage segmentation, rotation, translation, augmented reality-generative adversarial network (AR-GAN) based augmentation, and dilated inception convolutional neural network (DICNN), outperforms existing techniques with remarkable average identification accuracy is 90.67% on raw field diseased leaf. This is achieved using deep convolutional neural networks (CNNs) and a small sample size. A major step towards the actual application of advanced technology in agricultural IoT systems and precision agriculture, the work is in line with the emerging trend of using CNNs and generative adversarial networks (GANs) for strong plant disease recognition [6]. The fusion-based method achieves 94.30% Training accuracy in leaf disease identification and their dataset was not as much as healthy and this research done by colleagues of Kainat *et al.* [7]. Another cucumber leaf disease severity assessment by wang C and his colleagues as Two-stage model and achieves 93.27% segmentation accuracy [8]. The overcomes the difficulties of manual diagnosis by a tuned CNN for the determination of diseases after the creation of a new dataset and the dataset class Varity is less and validation of this data is not as impactful. The authors show that it outperforms pre-trained models with a high recognition accuracy [9]. Li *et al.* [10] addresses the insufficient training samples and imperfect measurement settings by implementing the extended collaborative representation (ECR) model for cucumber leaf disease classification with a 94.7% diagnostic accuracy.

There is multiple research done on finding disease and others issue in crop plants like cucumber but here is applying model and approach is on directly cucumber not any leaves. After analyzing we find Hussain *et al.* shows The issue in the detection of automated disease classification in cucumbers needs to be handled in a time-efficient manner-which is essential to maintaining agricultural productivity-it is tackled in this work. The suggested architecture combines VGG19, InceptionV3, deep learning, and a unique parallel maximum correlation fusion (PMCF) technique to reach an impressive 96.5% accuracy. When compared to more current methods, it emphasizes the value of sophisticated computational methods for illness diagnosis. we investigate existing techniques such as feature selection, dilated convolution, SVM, and sparse representation [11]. Kan *et al.* presents briefly that The study in tackles the critical problem of identifying cucumber diseases, which is essential to maintaining agriculture. With an Entropy-ELM-based architecture and deep learning, the study attains an impressive 98.48% accuracy. Comparisons with current methods highlight the superiority of the suggested method. Previous studies that use techniques such as feature selection (Entropy ELM), pre-trained models (VGG19, Inception V3), and data augmentation are presented. There are noteworthy initiatives that use a variety of models (DeepLabV3+, U-Net, EfficientNet, methodologies (whale optimization algorithm (WOA)-based feature selection), and methods (SHSB saliency, probability distribution-based entropy). By combining these methods, this article contributes and highlights the value of sophisticated computational methods for disease classification and classification in precision agriculture [12]. The body of research highlights a shift towards reliable, automated techniques, such as global pooling dilated convolutional neural network (GPDCNN), which hold the promise of better crop disease classification, especially with irregularly diseased leaf images' complexity [13]. A hybrid framework approach by Zhang *et al.* achieves 93.50% accuracy in cucumber disease classification, emphasizing real-time applications. Sparse representation and clustering achieve 85.7% recognition, addressing automated crop disease classification challenges. But it focuses on Cucumber not the leaves [14].

Regarding time of previous literature study we found that relatable study on others plants or crops are implemented and in cucumber sector it use in different angle like focusing in a particular disease or different image segmentation. Jasim and Al-Tuwaijari mates employs a unique dataset CNN achieves a remarkable accuracy (98.29% for training and 98.029% for testing) with a focus on tomatoes, peppers, and potatoes Plant [15]. Ozguven uses the faster region-based convolutional neural network (faster R-CNN) model, with object classification technique and acquired a respectable 94.86% accurate classification rate but this article focus only for mildew disease [16]. A CNN-based segmentation achieves 95.08% accuracy in powdery mildew assessment and here not any other segmentation or class implemented by Lin *et al.* [17]. Pawar and his colleagues applied Image processing technique on their dataset and ANN model achieve 80.45% accuracy in early crop illness diagnosis [18]. Cap *et al.* shows a diffferent deep learning system and achieves 65.8% precision in wide-angle plant disease diagnostics which is quite a different role [19]. Table 1 presents a close analysis of relevant works, which aids in making our study stand out. We also attempt to narrow the gap and make a significant addition to this field.

Table 1. Analytical view of previous literature

Authors	Studies	Dataset	Accuracy	Remarks
Mia *et al.* [4]	K-means clustering and random forest (RF)	Cucumber leaf disease	89%	Others Implemented model of InceptionV3, MobileNetV2, and VGG16 gives poor performance
Zhang *et al.* [6]	DICNN	Cucumber leaf disease	90.67%	Small and less categories data
Kainat *et al.* [7]	Fusion based method	Cucumber leaf disease	94.30%	Methodological studies is quite different
Hussain *et al.* [11]	VGG19, Inception V3	Cucumber vegetable data	93.79%	This paper not focused on leaf disease
Kan *et al.* [12]	Entropy-ELM-based architecture and deep learning	Cucumber disease	98%	This try to disease of Cucumber directly not the leaves diseases.
Zhang *et al.* [14]	Hybrid framework approach	Cucumber disease classification	85.7%	The main focus is introduced a hybrid framework and the dataset is from open source.
Jasim *et al.* [15]	CNN	Plant leaf disease	98%	Tomato, peppers and potatoes leaf disease
Power *et al.* [18]	ANN	Early crop disease	80.45%	Focusing on early crop illness diagnosis
Cap *et al.* [19]	Deep learning	Plant diseases	65.8%	Approach a deep learning model for wide-angle plant disease diagnostics

We attempt to close all of the previous research gaps in this study, including the absence of pure dataset resources, the use of contemporary preprocessing techniques, and the use of the most recent deep learning models. Our final objective is to use our primary data to develop an automated system and reduce significant limitations. Below is a list of our objectives, which truly represent our main contribution:

- Based on the perspective of Bangladesh cucumber fields, we have created a healthy dataset of cucumber leaf disease that includes four classes.
- In order to provide a healthier and more useful dataset for the researchers, we used pre-processing techniques.
- We examine prior research and compare it to ours, finding that our accuracy is higher.
- We use a methodical approach to putting in place an automated system for direct users.

2. METHOD

Bangladesh is an agricultural nation, and the quality of its soil makes it ideal for growing a wide range of products, including vegetables. More than $10 million in business opportunities exist in the cucumber field. We attempt to do highly influential research in order to directly benefit our rural users through the application of cutting-edge deep learning models and lessen their losses through early disease classification knowledge. We described our suggested system and implementation process in this part.

2.1. Proposed methodology

Our goal in this research project is to develop an ideal method for highly accurate and confidently diagnosing cucumber disease. We started by creating the ideal dataset and processing it to match the deep learning model. Our methodology flowchart is displayed step-by-step in Figure 1. It is evident that we used a variety of deep learning models and pre-processing methods after gathering the data. We finally achieved the results we had anticipated. The following subsections go into detail about each phase.

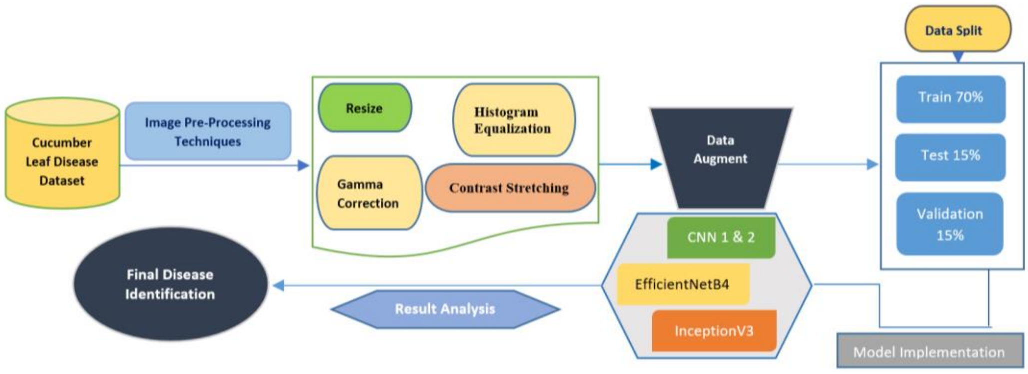

Figure 1. Proposed methodological diagram

2.2. Data acquisition procedure

A varied set of high-definition photos of cucumber leaves that covered a wider range of growth phases, in infected and healthy conditions along with varied lighting circumstances were taken into consideration as a dataset. We made sure that the dataset was representative of common cucumber leaf diseases that affect cucumbers. We accurately labelled the dataset with the presence of the particular diseases in consideration for them to be passed onto the supervised training phase. The collection of datasets on Cucumber diseases is a significant contribution of this research. The data collection for this research was done in two seasons to cover all the bases since Cucumbers grow and are affected by distinct weather characteristics The photos were collected via the Mobile Phone Samsung A50 which possesses the specifications of a 48 MP primary camera. Various cucumber fields farmers were visited for the collection of data. A total of 543 images were acquired out of which there are 130 pictures for Angular Leaf Spot, 122 images of Downy Mildew, 105 images of Powdery Mildew and 186 images of Good Leaf.

2.3. Data preprocessing

Following data collecting, we used a variety of data preprocessing techniques to improve the data's quality and make it more appropriate and useful for the model training phase. …

a. Resizing: all images were shaped in 256×256 pixel dimensions. Because when we collect data from the agricultural field the image sizes are not the same.
b. Histogram equalization: this is the technique of image improvement. Here all image's contrast and brightness are not the same. So we apply the histogram equalization technique for the same contrast and brightness. As a result, the low-contrast area of image becomes easier to detect [20].
c. Gamma correction: a digital image processing method called gamma correction modifies intensity values to alter the brightness and contrast of an image. It applies a nonlinear technique to the pixel values to accommodate changes in the way monitors display light. By altering the brightness and contrast of the images in our collection, it can enhance their visual appeal. Additionally, it's frequently used to correct photos that seem too bright or dark [21].
d. Contrast stretching: the stretching of contrast is an image enrichment technique that aims to increase the visibility of details by expanding the range of hues in a picture. This is often accomplished in an 8-bit grayscale image by linearly increasing the initial intensity values to encompass the full range. By intensifying contrast, this technique makes dark parts darker and bright areas brighter [21].
e. Augmentation: for picture augmentation in this study, we have used shearing, zoom horizontal flip, width shifting, height shifting, and rotation techniques. In Table 2 we present the scenario of raw & augmented images quantity for each class.
f. Splitting: the image dataset was augmented into training, testing and validation respectively by 70%, 15% and 15%.

Table 2. Dataset table

Dataset	Classes	Original images	Augmented
Cucumber leaf disease dataset	Angular leaf Spot	130	390
	Downy Mildew	122	366
	Powdery Mildew	105	315
	Good leaf	186	558
	Total	543	1629

Our raw photos underwent data pre-processing, which resulted in more precise and significant data. Following the application of all techniques, in particular A, B, C, and D, the data became more suited to the model that was used. Section 3.6 presents the accuracy of the raw data results, demonstrating the extent of quality improvement achieved through the application of post-processing procedures. After the entire process is finished, Figure 2 displays the visualization and injects the pre-proceeding components.

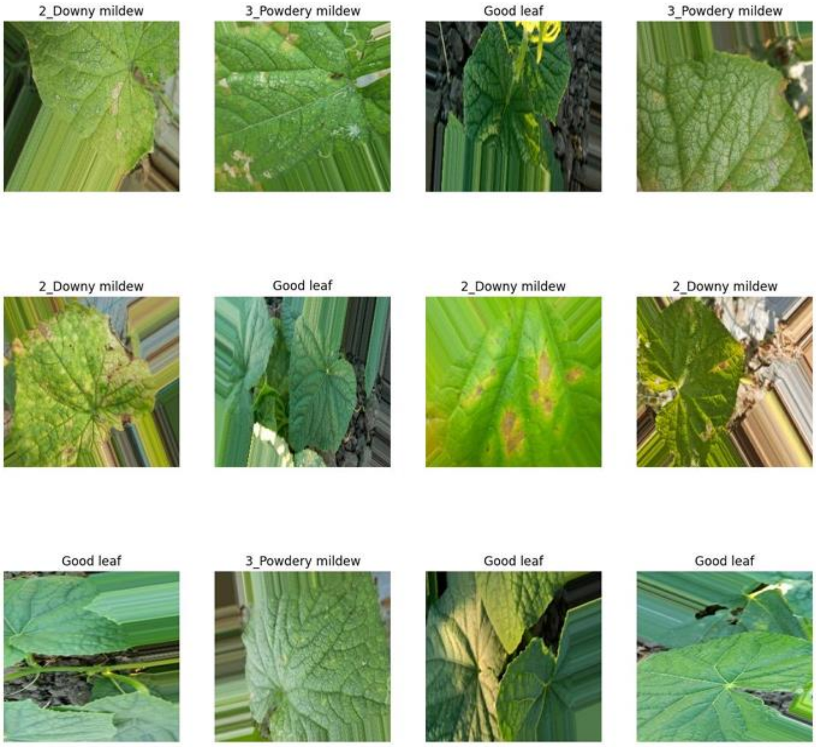

Figure 2. Sample of preprocessed images

2.4. Model implemetation

Our dataset implements three major significant models, and we anticipate excellent results. Our applied model is explained in detail below. …

a. CNN: Convolutional neural networks developed for object classification and localization in pictures. A feed-forward neural network underlies it. It also covers data processing using a grid-like layout for presenting data visually. This model has a big influence on the deep learning algorithm. The system consists of three main layers: convolutional, pooling, and completely linked. It is capable of data analysis and autonomous event prediction. This calculation provides answers to issues with Among other things, segmentation, classification, and computational imaging [22].
b. InceptionV3: InceptionV3 is the name of a 48-layer pre-trained model. Inception-v3, an enhanced GoogLeNet architecture, uses transfer learning to improve biomedical categorization [23]. This provides an architecture with several filters (convolutional) of varying sizes. Google developed Inception V3 for picture classification and object recognition [23].
c. EfficientNetB4: Compound scaling is used by the CNN architecture called EfficientNet-B4 to attain great accuracy and computational efficiency. It is a member of the EfficientNet family, which also includes versions B0 through B7. Accuracy rises with the number of models without appreciably raising the amount of parameters required for training. Models B1–B7 can be created using EfficientNetB4's compound scaling technique, which also optimizes FLOPs and increase efficiency of the parameters [24].

3. RESULTS AND DISCUSSION

We are able to detect with high accuracy by applying numerous algorithms and use our own dataset. Several measurement metrics demonstrated its quality and stability with evidence. We offer several visualizations, such as a confusion matrix, an accuracy and loss graph, each in its own part, as evidence.

3.1. Performance measurement metrics

Important criteria like accuracy, recall, specificity, and F1-score were taken into account in our study. These metrics enable a thorough evaluation of the models' ability to identify illnesses of cucumber leaves, providing insightful information for real-world agricultural applications.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

True positive, true negative, false positive, and false negative are the full forms of TP, TN, FP, and FN in this context [25].

3.2. Result discussion of preprocessed dataset

After using all of the deep learning models on our pre-processed data, we were able to attain extremely high accuracy. Our top model, EfficientNetB4, offers the highest training accuracy of 95.67% and the highest test accuracy of 87%. The precision and recall values are fairly equal, indicating that overfitting and underfitting are issues that are almost nonexistent. Table 3 provides a summary of the measurement metrics values for all four models.

Table 3. Image measurments score (precision, recall, F1-score)

Model	Class	Precision	Recall	F1-score
CNN-1	Angular leaf spot	0.74	0.74	0.71
	Downy mildew	0.69	0.56	0.67
	Powdery mildew	0.77	0.75	0.69
	Good leaf	0.99	0.99	0.99
CNN-2	Angular leaf spot	0.79	0.70	0.74
	Downy mildew	0.72	0.63	0.70
	Powdery mildew	0.69	0.84	0.76
	Good leaf	0.99	0.98	0.98
InceptionV3	Angular leaf spot	0.62	0.74	0.62
	Downy mildew	0.60	0.32	0.34
	Powdery mildew	0.61	0.82	0.63
	Good leaf	0.90	0.98	0.94
EfficientNetB4	Angular leaf spot	0.85	0.83	0.86
	Downy mildew	0.76	0.75	0.74
	Powdery mildew	0.77	1.00	0.86
	Good leaf	1.00	1.00	1.00

The average result summary of all applicable models' truthfulness as seen from the standpoint of various measurements is shown in Table 4. The efficienNetB4 achieved the maximum accuracy for both train and test data, as can be seen here. There other measurement’s matrix also be shown for better compares.

Table 4. Result summury of all implemented models

Model	Training accuracy (%)	Testing accuracy (%)	Precision (%)	Recall (%)	F1-score
CNN-1	84.21	77.34	79	77	77
CNN-2	88.76	83.59	85	84	84
InceptionV3	79.31	71.48	70	71	65
EfficientNetB4	95.67	86.62	88	86	86

3.3. Confusion metrix (CF-M) analysis

Figure 3 confusion matrix for the all applied model illustrates how stable our detection accuracy is. Figure 3 also demonstrates our result quality and dynamic situation. The CNN1 confusion view, which shows a decent TP rate for four classes, is presented in Figure 3(a). While the FP rate of the second class, Downey_Mildew, is quite high, CNN2 in Figure 3(b) helps to mitigate the lack of this class. However, there are some noisy views in Powdery_Mildew because the TP is significantly lower than the FN and FP rate. The CF matrix of a different model is shown in Figure 3(c), and the Inception v3 model has very little value in the Second Class. However, every single flaw and TP rate has been subtly fixed in our best model, EfficientNetB4, which is displaying extremely accurate and lively values from every angle, including TP, TN, FP, and FN. The CF-M of the EfficientNetB4 models is displayed in Figure 3(d), providing evidence of the high values of all the measurement metrics.

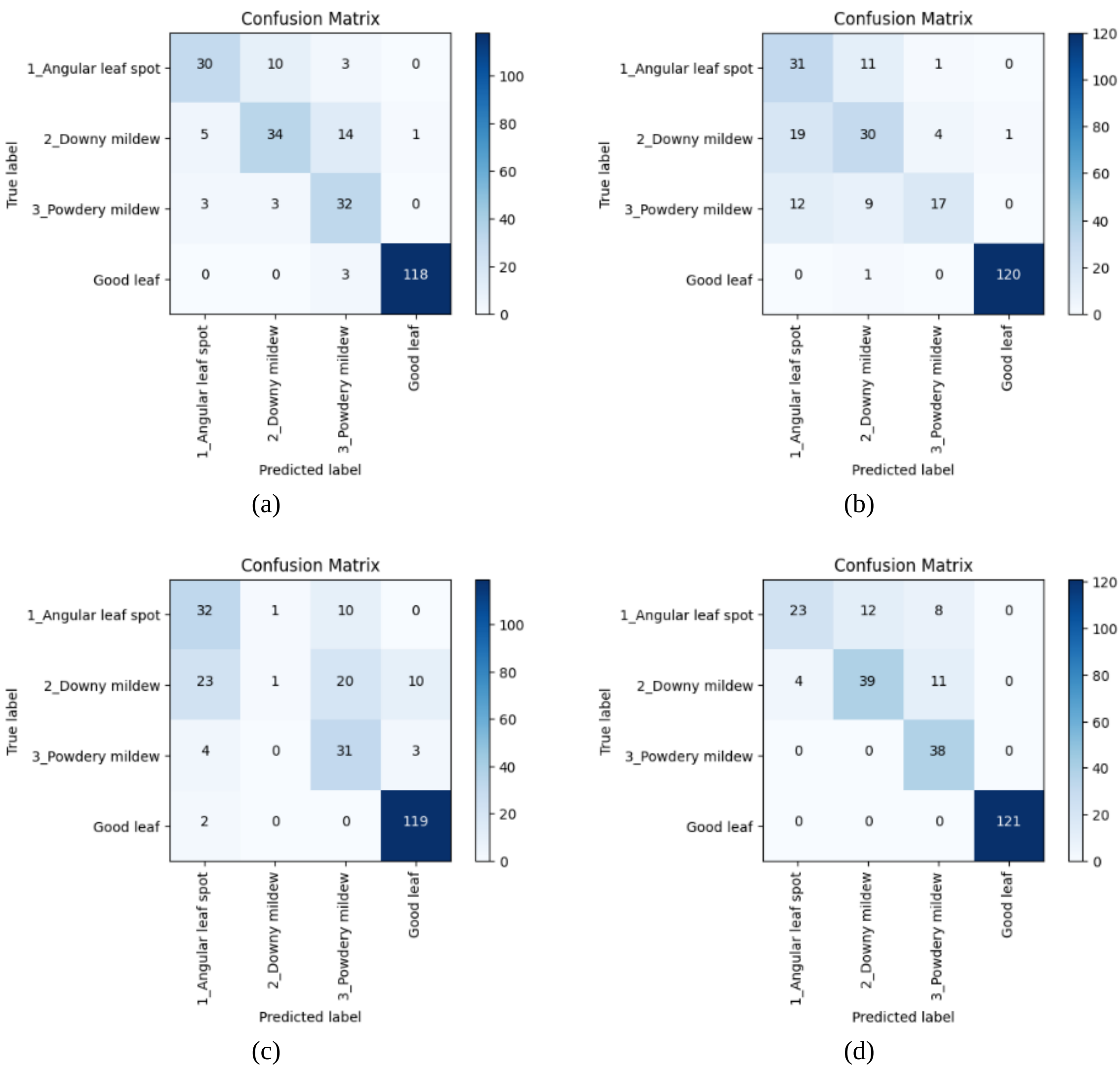

Figure 3. Confusion matrix of (a) CNN,1 (b) CNN2, (c) Inception V3, and (d) EfficientNetB4

3.4. Accuracy and loss graph analysis

Accuracy graph shows how well a piece of data is recognized based on its class during training and testing. Figure 4 shows the accuracy of four applied models together with a loss graph. It is also critical to understand whether the validation loss is little or large. We can conclude that the model is fairly good and satisfied if the loss is smaller. Accuracy and loss graph is clear presentation of model performance [22].

The CNN1 accuracy and loss graph is shown in Figure 4(a). The visual graph shows that when the model's epoch number increases, both its training and test accuracy increase. However, CNN2, as shown in Figure 4(b), performs less well than CNN1, and the main cause of this underfitting and overfitting is the model's in 10-15 number epoch. The accuracy graph of the InceptionV3 model is shown in Figure 4(c). In that case it is proved, while the accuracy throughout training and testing is generally consistent, there are occasional periods when it produces noise as a result of improper implementation procedures. However, EfficientNetB4 ultimately demonstrates its excellent accuracy and makes it evident that it has been quite steady from the start. The accuracy and loss graph of EfficientNetB4 in Figure 4(d) are far better than those of any other model, and our own dataset is well-fitted into this methodology. We guarantee that the data we provide to our algorithm yields a high degree of accuracy, and the graph provides compelling evidence of our findings. Note that for accuracy and loss graph X-Axis denotes the number of epoch and Y-Axis denotes the accuracy/loss level.

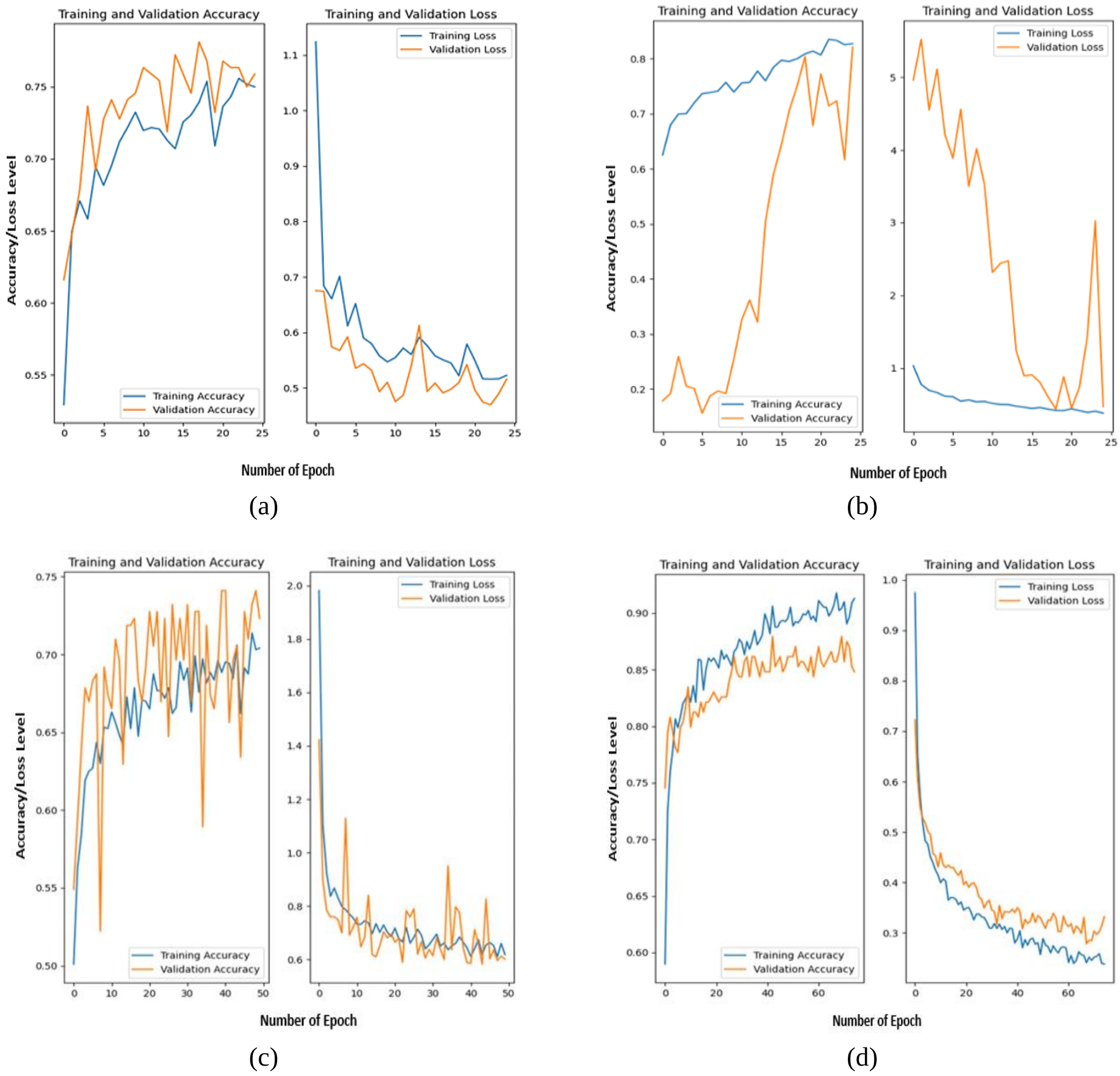

Figure 4. Accuracy and loss graph of (a) CNN1, (b) CNN2, (c) Inception V3, and (d) EfficientNetB4

3.5. Confidience level analysis

Another result analysis measure that truly expresses the confidences of the detection models for the specific data is the confidence level value [26]. We measure each model's confidence level (0–1) in our preprocessed dataset. Every model's accuracy level is incredibly confident, often approaching 100%. EfficientNetB4, our most accurate model, is shown in Figure 5 with a sample of detected images along with its confidence value, actual class, and forecast class. The confidence value at 95% is in the range of o.9 to 1. Thus, it is simple to state that the models' accuracy is quite respectable.

3.6. Result discussion of raw dataset

It is crucial to understand that we use the raw data in the same models only after resizing the image to 256 by 256 before processing the data for separate models. We were receiving extremely low accuracy in both the train and test cases at that time. Table 5 makes evident how extremely bad the precision, recall, and F1-score value is. We observe that in raw data, our most accurate model, EfficientNetB4, provides relatively little score. due to the extremely noisy raw data and the image's incompatibility with the model's algorithm. Because of this, we only obtained 56% accuracy in the train data and 53% accuracy in the test data, along with recall and an F1-score of 44.66% and 48.23, respectively. We can state that the data set is perfected by our pre-processing method, which also yields excellent accuracy and stability. This can be applied in a big sector for rural users who are directly involved in the cucumber farming industry.

Figure 5. Confidence level of EfficientNetB4 model in predicting

Table 5. Result summury of all implemented models for raw data

Model	Training accuracy (%)	Testing accuracy (%)	Precision (%)	Recall (%)	F1-score
CNN-1	51.11	49.44	49	46	46
CNN-2	53.46	52.32	45	44	44
InceptionV3	49.31	51.48	40	41	45
EfficientNetB4	56.12	53.05	48	48	44

4. CONCLUSION

We created our own dataset from Bangladesh's rural areas for our research, and we used customized CNN models, CNN1 and CNN2, as well as InceptionV3 and EfficientNetB4 models. It is demonstrated that by comparing the raw and processed photos, we obtain a satisfactory result, indicating that our pre-processing technique improves the data quality and accuracy over raw data. In agricultural countries like Bangladesh, India, and others, where cucumbers are a common crop with a wide variety of uses and a profitable market, it is going to have an immense effect on society. It will have a greater impact and be more beneficial to the development of our cultivation if we can concentrate on more disease classes and build a sizable dataset for numerous crops or vegetable leaves and plant disease in one frame. We are going to provide a system that will be useful to farmers and cultivators since it is an automated system that can identify disease at an early stage, aid minimize loss, and quickly identify a cure. Finally, we can state that our suggested approach, applied picture pre-processing techniques, and application of deep machine learning can be a great research in contemporary cultivation. Our implemented model, specifically EfficientNetB4, can be crucial in detecting or predicting the categorization of image disease.

ACKNOWLEDGEMENTS

We would like to sincerely thank Daffodil International University's Department of CSE and Computer and Programming Club (CPC) for their crucial support and direction during this research. Their encouragement and resources were instrumental in the successful completion of this study.

FUNDING INFORMATION

The study process that we carried out using our own funds and the article processing fee was made possible by a research grant from Daffodil International University. The Grant No is DIU/DoR/24/F&A/34.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Md Mizanur Rahman	✓	✓	✓	✓	✓	✓			✓		✓		✓	✓
Mahimul Islam Nadim		✓						✓	✓					
Mahinur Akther		✓				✓		✓	✓	✓	✓	✓		
Ahad Ullah	✓		✓	✓			✓			✓	✓		✓	
Jakaria Ahmed					✓			✓						
Muhammad Jalal Uddin Ahmed			✓		✓			✓			✓			
Israt Jahan							✓			✓		✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Derived data supporting the findings of this study are available from the corresponding author, MMR, on request.

REFERENCES

[1] R. Manavalan, “Cucumber leaves diseases detection through computational approaches: A review,” *Asian Journal of Research in Biosciences*, vol. 3, no. 2, pp. 100–110, 2021.

[2] K. Asefpour Vakilian and J. Massah, “An artificial neural network approach to identify fungal diseases of cucumber (Cucumis sativus L.) plants using digital image processing,” *Archives of Phytopathology and Plant Protection*, vol. 46, no. 13, pp. 1580–1588, Aug. 2013, doi: 10.1080/03235408.2013.772321.

[3] M. A. Islam, M. M. Rahman, A. A. Shohan, R. A. Pulok, S. Akter, and M. T. Ahmed, “Deep neural network-based approach to identify the paddy leaf disease using ResNet50-V2,” in *2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques, EASCT 2023*, Oct. 2023, pp. 1–6. doi: 10.1109/EASCT59475.2023.10393586.

[4] M. J. Mia, S. K. Maria, S. S. Taki, and A. A. Biswas, “Cucumber disease recognition using machine learning and transfer learning,” *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 6, pp. 3432–3443, 2021, doi: 10.11591/eei.v10i6.3096.

[5] P. Krithika and S. Veni, “Leaf disease detection on cucumber leaves using multiclass support vector machine,” in *Proceedings of the 2017 International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2017*, Mar. 2017, pp. 1276–1281. doi: 10.1109/WiSPNET.2017.8299969.

[6] J. Zhang, Y. Rao, C. Man, Z. Jiang, and S. Li, “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, no. 4, Apr. 2021, doi: 10.1177/15501477211007407.

[7] J. Kainat, S. S. Ullah, F. S. Alharithi, R. Alroobaea, S. Hussain, and S. Nazir, “Blended features classification of leaf-based cucumber disease using image processing techniques,” *Complexity*, vol. 2021, no. 1, Jan. 2021, doi: 10.1155/2021/9736179.

[8] C. Wang, P. Du, H. Wu, J. Li, C. Zhao, and H. Zhu, “A cucumber leaf disease severity classification method based on the fusion of DeepLabV3+ and U-Net,” *Computers and Electronics in Agriculture*, vol. 189, p. 106373, Oct. 2021, doi: 10.1016/j.compag.2021.106373.

[9] S. M. Omer, K. Z. Ghafoor, and S. K. Askar, "An intelligent system for cucumber leaf disease diagnosis based on the tuned convolutional neural network algorithm," *Mobile Information Systems*, vol. 2022, pp. 1–16, 2022, doi: 10.1155/2022/8909121.

[10] Y. Li, Z. Luo, F. Wang, and Y. Wang, "Hyperspectral leaf image-based cucumber disease recognition using the extended collaborative representation model," *Sensors (Switzerland)*, vol. 20, no. 14, p. 4045, Jul. 2020, doi: 10.3390/s20144045.

[11] N. Hussain *et al.*, "Multiclass cucumber leaf diseases recognition using best feature selection," *Computers, Materials and Continua*, vol. 70, no. 2, pp. 3281–3294, 2022, doi: 10.32604/cmc.2022.019036.

[12] M. A. Khan *et al.*, "Cucumber leaf diseases recognition using multi-level deep entropy-ELM feature selection," *Applied Sciences (Switzerland)*, vol. 12, no. 2, p. 593, Jan. 2022, doi: 10.3390/app12020593.

[13] S. Zhang, S. Zhang, C. Zhang, X. Wang, and Y. Shi, "Cucumber leaf disease identification with global pooling dilated convolutional neural network," *Computers and Electronics in Agriculture*, vol. 162, pp. 422–430, Jul. 2019, doi: 10.1016/j.compag.2019.03.012.

[14] 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, Mar. 2017, doi: 10.1016/j.compag.2017.01.014.

[15] M. A. Jasim and J. M. Al-Tuwaijari, "Plant leaf diseases detection and classification using image processing and deep learning techniques," in *Proceedings of the 2020 International Conference on Computer Science and Software Engineering, CSASE 2020*, Apr. 2020, pp. 259–265. doi: 10.1109/CSASE48920.2020.9142097.

[16] M. M. Ozguven, "Deep learning algorithms for automatic detection and classification of mildew disease in cucumber," *Fresenius Environmental Bulletin*, vol. 29, no. July, pp. 7081–7087, 2020.

[17] K. Lin, L. Gong, Y. Huang, C. Liu, and J. Pan, "Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network," *Frontiers in Plant Science*, vol. 10, Feb. 2019, doi: 10.3389/fpls.2019.00155.

[18] P. Pawar, V. Turkar, and P. Patil, "Cucumber disease detection using artificial neural network," in *Proceedings of the International Conference on Inventive Computation Technologies, ICICT 2016*, Aug. 2016, vol. 2016, pp. 1–5. doi: 10.1109/INVENTIVE.2016.7830151.

[19] Q. H. Cap, K. Suwa, E. Fujita, S. Kagiwada, H. Uga, and H. Iyatomi, "An end-to-end practical plant disease diagnosis system for wide-angle cucumber images," *International Journal of Engineering and Technology(UAE)*, vol. 7, no. 4, pp. 106–111, Oct. 2018, doi: 10.14419/ijet.v7i4.11.20784.

[20] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," *IEEE Transactions on Consumer Electronics*, vol. 45, no. 1, pp. 68–75, Feb. 1999, doi: 10.1109/30.754419.

[21] A. Sohel *et al.*, "Sunflower disease identification using deep learning: A data-driven approach," in *2023 26th International Conference on Computer and Information Technology, ICCIT 2023*, 2023, pp. 1–6. doi: 10.1109/ICCIT60459.2023.10441285.

[22] A. Haque, R. A. Pulok, M. M. Rahman, S. Akter, N. Khan, and S. Haque, "Recognition of Bangladeshi sign language (BdSL) words using deep convolutional neural networks (DCNNs)," *Emerging Science Journal*, vol. 7, no. 6, pp. 2183–2201, Dec. 2023, doi: 10.28991/ESJ-2023-07-06-019.

[23] A. M. Hassan, M. B. El-Mashade, and A. Aboshosha, "Deep learning for cancer tumor classification using transfer learning and feature concatenation," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 6, pp. 6736–6743, Dec. 2022, doi: 10.11591/ijece.v12i6.pp6736-6743.

[24] V. K. Pratap and N. Suresh Kumar, "High-precision multiclass classification of chili leaf disease through customized EffecientNetB4 from chili leaf images," *Smart Agricultural Technology*, vol. 5, p. 100295, Oct. 2023, doi: 10.1016/j.atech.2023.100295.

[25] M. U. Hasan, F. N. Khan, H. Jahan, Z. Zaman, and M. M. Rahman, "Predicting pancreatic cancer survival and staging: a machine learning approach," in *2024 IEEE Conference on Computing Applications and Systems, COMPAS 2024*, Sep. 2024, pp. 1–5. doi: 10.1109/COMPAS60761.2024.10797176.

[26] C. Marois, D. Lafreniere, B. Macintosh, and R. Doyon, "Confidence level and sensitivity limits in high-contrast imaging," *The Astrophysical Journal*, vol. 673, no. 1, pp. 647–656, Jan. 2008, doi: 10.1086/523839.

BIOGRAPHIES OF AUTHORS

Md Mizanur Rahman is working as a lecturer, Department of CSE in Daffodil International University. He received his B.Sc. degree in computer science and engineering from Daffodil International University, Dhaka, Bangladesh in 2022 and Complete Master's Degree in Information Technology from Institute of Information Technology, University of Dhaka in 2023. His area of research interest includes data mining, machine learning, deep learning and computer vision. He is also working on developing course content for data mining and algorithms in online platform Like International Online University (IoU) and Go edu. He can be contacted at email: mizanurrahman.cse@diu.edu.bd.

Mahimul Islam Nadim is a lecturer in the Department of Computer Science and Engineering (CSE) at Daffodil International University (DIU) in Dhaka, Bangladesh. He holds a Bachelor of Science (B.Sc.) and Master of Science (M.Sc.) in computer science and engineering, both earned from Daffodil International University (DIU). Known for his dedication to student success, Mr. Nadim employs engaging teaching methods that foster a stimulating learning environment. His research interests lie in the exciting fields of deep learning, data mining, and machine learning. His email: nadim.cse@diu.edu.bd

Mahinur Akther is a lecturer in the English Department at Daffodil International University, where she has contributed to the academic community for the past 2.5 years. Holding a bachelor of arts (B.A.) in english from the University of Dhaka and a Master of Arts (M.A.) in both linguistics and social sciences from the same institution. Ms. Akther's research primarily focuses on the influence of social factors in various aspects of language learning and uses. mahinurakther.eng@diu.edu.bd

Ahad Ullah is a senior lecturer at Daffodil International University, specializing in Civil and Geotechnical Engineering. He is currently pursuing his M.Sc. in Geotechnical Engineering at the prestigious Bangladesh University of Engineering and Technology (BUET). With almost eight years of teaching experience, including a previous role as a lecturer at Port City International University in Chattogram, Ahad Ullah is known for his engaging teaching methods and dedication to student development. His research interests focus on soil mechanics, foundation engineering, and sustainable construction practices. He has contributed to several national and international journals and actively participates in conferences and workshops. ahad.ce@diu.edu.bd

Jakaria Ahmed is a student of Computer Science and Engineering Department at Daffodil International University. He is very passionate in research specially in data mining, machine learning and deep learning. He is very energetic and have thrust in exploring new research idea. His coding sense is quite good. He can be conducted in jakaria15-2215@diu.edu.bd.

Muhammad Jalal Uddin Ahmed is a student of Computer Science and Engineering Department at Daffodil International University. He is very passionate in research specially in data mining, machine learning and deep learning. He is very energetic and have thrust in exploring new research idea. His coding sense is quite good. He can be conducted in uddinjalal270@gmail.com.

Israt Jahan is a dedicated and passionate university lecturer (senior scale) of the Department of Computer Science and Engineering at Daffodil International University with a diverse academic background and a wealth of experience in teaching and education. Her interested research fields are machine learning, deep learning, digital image processing, and computer vision. She continues to make meaningful contributions to the academic community, leaving an indelible mark as a dedicated educator, researcher, and advocate for the transformative power of education. She can be contacted at email: isratjahan.cse@diu.edu.bd.