International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 5,2025
https://theaspd.com/index.php

Accuracy Assessment of Convolutional Neural Networks for Plant Disease Prediction with Multi-Stage Image Preprocessing

A. Geetha Devi[1], Dr. K. Thirupal Reddy[2], V. Dilip Kumar[3], Dr. M. Venkata Subbarao[4]
Dr. Yeswanth Kumar Alapati[5], J. Hymavathi[6], *Pamula Udayaraju[7]
[1]Department of ECE, Prasad V Potluri Siddhartha Institute of Technology, Andhra Pradesh, India.
[2,4]Department of CSE, Malla Reddy Engineering College for Women, Hyderabad, Telangana-India.
[3]Department of CSE, SRKR Engineering College, Bhimavaram, Andhra Pradesh, India.
[5]Department of IT, RVR & JC College of Engineering, Andhra Pradesh, India.
[6]Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India.
[7]Department of CSE, SRM University- AP, Amaravathi, Andhra Pradesh, India.
geetha.agd@gmail.com[1] , k.thirupalreddy2020@gmail.com[2] , dilipv510@gmail.com[3] , subbaraomudragada@gmail.com[4], alapatimail@gmail.com[5] , hymavathi.janaswami@gmail.com[6], *udayaraju.p@srmap.edu.in[7]

Abstract

Plant disease forecasting is among the advanced research areas targeted by most agricultural institutions and government bodies to enhance yields. Forecasting diseases is a very vital topic of international agriculture since crop health directly affects human health proportionately. Various plant organs, like root, stem, leaves, fruits, or a major proportion of the crop, are diagnosed to identify diseases, while various previous work employed images of plant leaves to identify disease. Previous works employed standard practices, optimization techniques, and machine learning methods for plant disease prediction but specifically for sets of leaf images. No work adopted a generalized approach to disease detection for any plant image data because the dataset was large, colored, aligned, and resolution varied. The goal of this paper is to adopt a Preprocessing Framework with a Convolution Neural Network (PF-CNN). It works on any plant leaf image. The preprocessing framework includes alignment, rotation, resizing, cropping, color transformation, and image enhancement. Final output images are passed through CNN for disease classification. As this work is the beginning of the research work, CNN is trained using pretrained images and checked against ground truth outputs to validate the disease class. Next, CNN is tested with test images and confirmed using validation images. Experiment output is cross-checked and compared with other analogous methods to analyse the performance of PF-CNN. Keywords: Preprocessing, CNN, Plant Disease Detection, Leaf Images, Tomato Leaves, Maize Leaves.

Keywords: *Preprocessing, CNN, Plant Disease Detection, Leaf Images, Tomato Leaves, Maize Leaves.*

INTRODUCTION

One of the branches of art and science is agriculture. It is cultivating plants, crops, fiber, biofuel, and livestock to increase the lifespan of living beings. Agriculture includes preparing plant and animal products for humans to use and market their distributions. It provides a high portion of global food and fabric materials. Agriculture influences the economic development of both developed and developing countries. Most developing countries have relied entirely on agriculture to increase the national economy. Developed countries have contributed a small portion of agricultural outcomes for federal income compared to developing countries. Especially in India, 70% of the people are involved in agriculture and farming [1]. Most low-income regions consider agriculture as the primary source of their income. It also creates many employment opportunities among people with a high percentage of income [2]. Agriculture fulfills the food and non-food product requirements globally. Some of the essential food items of the Asian continent are rice, vegetables, groundnuts, fruits, pulses, wheat, dairy products, maize, and sugarcane, and the non-food products are paper, rope, manufacturing materials, and raw materials of cosmetics [3]. Compared to past decades, the total landscape of the agricultural area is shriked, and the demand for crop production has increased worldwide [4].

Figure-1. Plant Disease Affects Leaves, Stem, and Fruits

Though various advanced technologies and cultural changes have occurred, living beings still require agricultural farm products because agrarian products are needed daily. Humans consume more agricultural products for construction, education, garments, cleaning, and food. Thus, the production rate of the crops needs to be increased. The healthiness of the plant influences human health directly. People need to intake more vitamins, proteins, enzymes, and other nutrients to lead a sustainable life [5], which requires cultivating crops with good health and more yield. Crop yielding measures the quantity of crops harvested in a particular area based on measurements like square feet, acres, miles, and hectares [6]. After measuring the area to cultivate, every farmer has to know the following steps to cultivate healthy crops: crop cultivation time and type, soil characteristics and proper establishment method, proper maintenance of water irrigation, weed management system, diseases management, harmless pesticide usage, climatic changes, and effective plant prevention method from early damage. The disease affects various parts of the plant, like stem, leaves, and fruits (Figure-1), whereas it initially affects and easy identification can be made only by diagnosing leaves. Before involving the fruits (yield), they can be protected by detection at the earlier stage through leaf image diagnosis. Diagnosing and detecting plant diseases from plant leaves is very easy and fast.

Plant disease management is one of the major processes to maintain successful crop production and economic ratio. Because plant diseases have significantly minimized the quality and quantity of crop production [7], they also reduce the food's nutrition level and health condition. Protozoa, insects, fungi, viruses, and bacteria are the major pathogenic organisms that cause plant diseases. The symptoms of plant diseases are evident in various parts of the crops, and they are easily analyzed from the leaves. The known and unknown plant diseases are identified based on the signs and symptoms. It is placed through visual microscopy, serology techniques, and media studies. However, these techniques require more time to analyze and require continuous monitoring. And it is more expensive to accurately diagnose the diseases at the early stage. So, automatic disease detection is necessary to reduce the time and cost of diagnosing plant diseases. Compared to the traditional methods, an automatic plant disease detection method is more beneficial for analyzing large crop areas [8].

Traditional methods used multiple techniques to diagnose the plant images accurately. The automatic system regularly analyses the crop field area and generates the result in images using cameras and sensors. It used a sequence of steps, such as image pre-preprocessing, segmentation, and future extraction, for detecting and classifying plant diseases. Image preprocessing helps to improve the quality of the input image; it includes cropping, rotating, aligning, de-noising, and other preprocessing techniques. The segmentation process reduces the overall computational complexity and increases the prediction accuracy due to preprocessing. Threshold-based, edge-based, clustering-based, and region-based methods are standard segmentation methods used in the earlier research. Then, the feature extraction is done to extract the essential features from the segmented images. Finally, the classification process is applied to classify the plant diseases and severity levels. Machine learning [12-14], support vector machine [10], random forest [11], and fuzzy logic [9] methods were used to predict plant diseases in earlier days. Though these models have accurately classified the plant diseases, it takes more time to compute the data and can process only a limited amount.

In this paper, a Convolution Neural Network model is implemented to overcome these limitations of the existing models. The main focus of this paper is to evaluate the efficiency of the ANN model in classifying the preprocessed input leaf images. Image preprocessing involves image denoising, alignment, rotation, cropping, resizing, color conversion, and image enhancement. After these steps, the enhanced input image is fed into the CNN model, generating the final prediction output (see Figure-2). The following section discusses the earlier research work, proposed methodology, workflow, and the model's output with limitations. It helps to plan and design the proposed model of this paper.

LITERATURE SURVEY

This section discusses various earlier research and review work on plant disease detection. In recent times, many of the learning methods, such as AI, ML, DL, and transferring learning approaches, have been followed by agriculture-based researchers. In that sense, a comparative study is conducted on five ML classifiers, such as SVM, CNN, ANN, Fuzzy, and KNN, to demonstrate the model's efficiency in detecting plant diseases (U. Shruthi et al. (2019)). The analysis shows that most existing researchers have used the SVM model for disease prediction. But in terms of accuracy, the CNN model has detected more crop diseases than others. A machine learning approach is applied and experimented with to detect regular and disease-affected rice crops (K. Ahmed et al.(2019)). The affected crop images with the white background are taken as input and analyzed using different ML classifiers such as KNN, J48, NB, DT, and LR. Compared to other methods, the review results show that the DT algorithm has classified affected rice crop images with 97% accuracy. Supervised machine learning approaches, namely NB, KNN, DT, SVM, and RF, are used to detect the disease in the maize leaf (K. Panigrahi et al. (2020)). The result of each proposed model is compared with each other to find the optimal model for plant disease detection. The overall result of the comparison shows that the RF model outperformed the other models with 79.23% accuracy.

A detailed review is presented to define the performance of the deep learning model in classifying plant diseases (L. Li et al. (2021) and J. Liu and X. Wang (2021)). The current and future trends in plant disease detection are explained in detail. Various challenges in implementing DL and image techniques to detect plant diseases are discussed, and possible solutions are suggested. Along with this, multiple suggestions are provided to overcome the limitations of conventional methods. Overall, this review is considered a guideline for the current and upcoming research on detecting various types of plant diseases. The uses of different imaging techniques are reviewed, and the efficiency of those models is discussed in detail (V. Sing et al. (2020)). The review represents the efficiency of imaging techniques in classifying plant diseases. Current trends and challenges in plant disease detection using imaging techniques and computer vision are summarized. In addition, various earlier research works on plant disease detection are reviewed. The review shows that the major disease prediction techniques are SVM, KNN, DL, and K-mean clustering. An image processing and computational intelligence model is proposed to classify plant disease detection and classification (V.K. Vishnoi et al. (2021)). Most of the earlier research has used plant leaves to detect the type of diseases using soft computing and computer vision techniques. The advantages and disadvantages of those researchers are summarized in this research work. Common symptoms and different stages of plant diseases are discussed. And the efficiency of new feature extraction techniques is also elaborately briefed. The main focus of this research work is to provide a guide to the recent research to understand the efficiency of the image processing and computational intelligence model on plant disease detection.

A deep SVM and 11-CNN classifier of the transfer learning model is evaluated to detect the four types of rice plant diseases (P.K. Sethy et al. (2020)). The model's performance is evaluated using various performance metrics such as F1-score, sensitivity, specificity, FPR, training time, and accuracy. Once again, a statistical analysis is performed to select the optimal plant disease prediction classifier. The simulation result shows that, compared to the transfer learning model, the SVM has performed well with an F1-score of 0.9838. ML-based classifiers such as SVM, RF, and LR are proposed to classify different types of plant diseases (D. Das et al. (2020)). The results of the models are evaluated through various performance metrics. The obtained performance comparison result shows that the SVM model accurately classifies the different types of plant diseases.

The literature survey identified that the learning methods did not focus on preprocessing tasks on the input data. The experiment was carried out with a dataset of a size and only one or two kinds of dataset. Plant disease prediction through the machine learning algorithm discussed in the literature mainly involves processing data obtained through various sensors. It also provided minimal accuracy and failed to generalize for different applications. It also provided less accuracy for different plant diseases. The comparison between different classification algorithms like SVM, KNN, DT, NB, and RF algorithms has provided less accuracy compared to the regression-based models. It could also not process image-based datasets required to predict and identify diseases. Due to the lower accuracy, neural networks were considered for processing the obtained data, which utilized image processing techniques to improve prediction accuracy. Neural networks provided better accuracy in predicting plant diseases using the above-said methods. It was capable of processing image-based datasets and also IoT-based datasets. Among the various neural networks adopted for prediction, CNN algorithms provided better prediction accuracy and utilized image processing and enhancement to improve its accuracy. Hence, a plant disease prediction model using the CNN algorithm provides better results than the conventional machine learning models. Preprocessing Methods Several methods are involved in image preprocessing, and some of them used in this paper are explained here. Image Denoising Image denoising is one of the necessary techniques followed in

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 5,2025
https://theaspd.com/index.php

image processing applications to improve the quality of the input images. Denoising is critical in CAD systems for denoising images under different imaging modalities. Denoising is also challenging for computer vision applications to reach their goal of suppressing noise from noisy images. Different denoising methods are used according to the noise type and the level at which it affects the input image. For example, salt and pepper noise is removed by filtering methods, and wavelet methods remove Gaussian noise, and vice versa. The spectral noise removal method removes the mixture of noises, which involves multiple techniques to remove noises. Noise removal is a method of removing the random variations in the image pixel data. This paper removes Gaussian noise using a Gaussian filter, a low-pass filter that eliminates blurred regions and high-frequency components of the input image. A kernel matrix with an odd size is applied on each pixel of the RoI (Region of Interest) to obtain the denoised image. The mathematical form of the Gaussian function is:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (1)$$

(x, y) represents the coordinate values from the RoI, the Pi value is 3.13 and sigma represents the standard deviation. By using the above equation-(1), a 3 x 3 Gaussian matrix is created with the STD value of 1 appears as:

$$\frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \qquad (2)$$

Image Color Conversion

Image preprocessing includes color conversion called Color Transform, which converts the image color from RGB (Red, Green, Blue) into Grayscale, YUV, and HSV. It can also form back the transformed image to RGB color. From each pixel, the values of the primary colors (R, G, B) are obtained and applied to gamma expansion as:

$$C_{linear} = \begin{cases} \frac{C_{RGB}}{12.92} & C_{RGB} \le 0.04045 \\ \frac{(C_{RGB}+0.065)}{1.065} & C_{RGB} > 0.04045 \end{cases} \qquad (3)$$

Where C_{RGB} represents the values from 0 to 255, and C_{linear} represents the values from 0 to 1. The above equation shows the abstract mathematical model for color space, which helps to illustrate the image colors regarding intensity values. The color space uses a multi-dimensional coordinate system; depending on the application, it chooses the dimensions. The color conversion may be represented as a function:

$$y = f(x) \qquad (4)$$

Where the input image is x, and f(x) converts the color RGB(x) into grayscale as:

$$f(x) = 0.2989 * R + 0.5870 * G + 0.1140 * B \ (5)$$

and f(x) is the output.

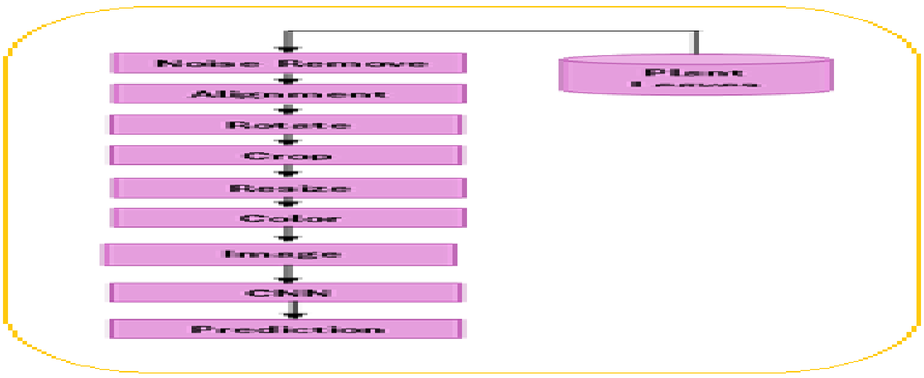

Figure-2. Proposed Workflow

Image Resizing

Resizing an image makes image interpolation when resizing an image from one size to another. The interpolation function obtained a new group of values using a function for the previous set of values given as input. Increasing or decreasing the number of pixels available in an image is essential. It is also called image scaling, which helps to reduce the computational complexity concerning number of pixels used in comparison. The interpolation is classified into nearest neighbor, linear, and polynomial interpolation, which are used in image resizing. In image resizing, the width and height of the image are scaled up or down into a new image. For example, in a bi-linear interpolation, it is assumed that $(\boldsymbol{x_1}, \boldsymbol{y})$ and $(\boldsymbol{x_2}, \boldsymbol{y})$ are the pixel coordinates, and their intensities are $\boldsymbol{I_1}$and $\boldsymbol{I_2}$ where $\boldsymbol{x_1} < \boldsymbol{x_2}$. The intensity of the unknown point is computed as:

$$Img_{new} = \frac{x_2 - x}{x_2 - x_1} * I_1 + \frac{x - x_1}{x_2 - x_1} * I_2$$

Where $x_1 \geq x \geq x_2$

Image Cropping

Image cropping is one of the transformation methods used to remove the unwanted portion of the image. Eliminating distractions, extra background, and other unwanted pieces in the image is vital. It provides a new image with a new dimension to preserve the area of the image required by the user. Three ways are used to crop the images, such as

Crop rectangle-based

Aspect Ratio-based

Pixel-based

Image Alignment

It is a process of aligning the overlaying images of the same manifestation under various conditions like viewpoints, illumination, sensors, and times. Image alignment consists of geometric-oriented evaluation and position-based evaluation. It helps to align the image position concerning the straight lines and planes. Three different alignment methods are followed:

Exact alignment

Alignment by aberration and

Alignment with the Bessel beam.

Image Rotation

Image rotation helps to rotate an image based on the axis or a center point. It can rotate the image up to 360° angle clockwise or anti-clockwise. It is a standard routing of image processing applications used in image alignment and mapping. It needs two images. The first one is the base image, where its angle is 0; based on this image, the second image's base angle concerning the x-axis increased or decreased clockwise or anti-clockwise, respectively.

Image Classification Using Convolution Neural Network

Nowadays, most image processing and recognition applications use deep neural network algorithms for automatic learning, feature extraction, and classification from high-dimensional data. Traditional algorithms use manual methods designed for feature extraction alone, which takes more time and is costly. Whenever data size increases, the computational complexity also increases. Deep learning algorithms are efficient and scalable for large-sized data analysis. The success of deep learning is its comprehensive performance in academics and industry, and it provides a more promising output than traditional models. One of the popular deep learning algorithms used mainly for image recognition is the Convolution Neural Network (CNN). CNN model has a sequence of defined layers to do predefined and customized tasks for data learning, feature extraction, dimensional reduction, and classification. The structure of the CNN model is shown in Figure-3.

Convolutional Neural Network

It is a complex model that comprises complex layers for predicting images. A CNN algorithm consists of convolutional, pooling, and fully connected layers. Every neuron in the convolutional layer is connected to all the neurons in the forthcoming layers. A pooling layer follows the convolutional layer and the fully connected layer. In the convolutional layer, the input data is convoluted into feature maps that carry the important features. Then, the pooling layer adopts max or average pooling to reduce the size of the data, which helps in the faster prediction of the results. In the convolution process, a convolution core slides through the images, extracting the features stored in the feature map. After the features are extracted, the pooling layer removes the unwanted features, and finally, they are passed to the fully connected layers, making the final predictions. This process is repeated over several iterations for multiple convolutional and pooling layers to improve the accuracy of the prediction process. Finally, the output result is obtained to predict the plant diseases in the images.

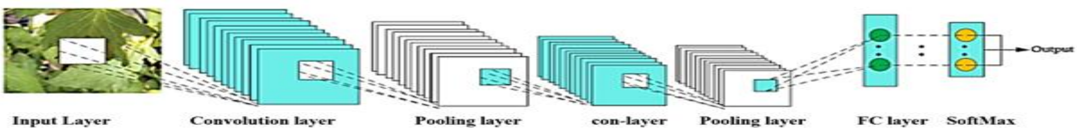

Figure-3. CNN Architecture

Experimental Setup

The above-discussed preprocessing methods and the CNN model are implemented in Python, and an experiment is conducted with a defined dataset taken from Kaggle and other resources. The implementation program is compiled and executed on Intel-Pentium-Core-i7, 7th gen, 3.0 Ghz processor, with NVIDIA GTX, 16 GB RAM, and 1TB HDD. The CNN model is created by training it with 80% of training data (a portion taken from the input dataset), testing it with 15% of test data, and validating it using 5% of random data. The CNN model is trained by collecting 80% of data from different datasets [22-25] available in the Kaggle source. A mixture of plant images is used in the training process to make the model a generalized model that can test any plant leaf image in the future. Input images are taken from Kaggle because they pertain to and are examined by more research works that can improve the training efficiency of the proposed CNN.

Implementation

The proposed model is implemented in TensorFlow, one of the widely used deep learning packages supported by Google. It is implemented in Google Colab with an open-source dataset from Kaggle. GPU acceleration is used for faster processing of the data. It provided effective processing of multiple images obtained from the dataset.

The architecture of the proposed CNN algorithm is shown in Table-1. The proposed algorithm comprises three convolutional layers, three max-pooling layers, and two dense layers. The batch normalization is used to normalize the input, and the convolutional layer accepts input as 248x248x3, with a bit size of 32. It is followed by a max pooling layer that reduces the number of parameters and the input size to 124x124x3, with the same bit size of 32. The next convolutional layer converts the input to 122x122x3 with 64-bit size, and after max_pooling is done with a resultant output 61x61x64, which is again continued with a convolutional and pooling combo. Finally, an output 29x29x128 is obtained and flattened to obtain 107648 parameters, passed through dense layers with 256 and 38 neurons.

Table-1. Proposed CNN Architecture

Layer (type)	Output Shape	Param #
batch_normalization_1 (Batch	(None, 250, 250, 3)	12
conv2d_3 (Conv2D)	(None, 248, 248, 32)	896
max_pooling2d_3 (MaxPooling2	(None, 124, 124, 32)	0
conv2d_4 (Conv2D)	(None, 122, 122, 64)	18496
max_pooling2d_4 (MaxPooling2	(None, 61, 61, 64)	0
conv2d_5 (Conv2D)	(None, 59, 59, 128)	73856
max_pooling2d_5 (MaxPooling2	(None, 29, 29, 128)	0
flatten_1 (Flatten)	(None, 107648)	0
dense_2 (Dense)	(None, 256)	27558144
dense_3 (Dense)	(None, 38)	9766
Total params: 27,661,170		
Traienable params: 27,661,164		
Non-ttrainable params: 6		

Dataset: The New plant disease dataset was obtained from Kaggle and is used for the plant disease prediction. The dataset consists of 87,000 RGB images, split for training and validation in the ratio of 80/20. A new directory with 33 images is used for the prediction process. The dataset is randomly augmented to improve the accuracy of the model. The dataset consists of affected leaf images from apples, blueberries, cherries, corn, grapes, peaches, oranges, peppers, potatoes, raspberries, soybeans, squash, strawberries, and tomatoes. Healthy and affected leaf images are classified for each plant, and 480 images are present for each field. Figure-4 shows some of the same leaf images labeled by their disease, improving the training accuracy and helping to evaluate the model.

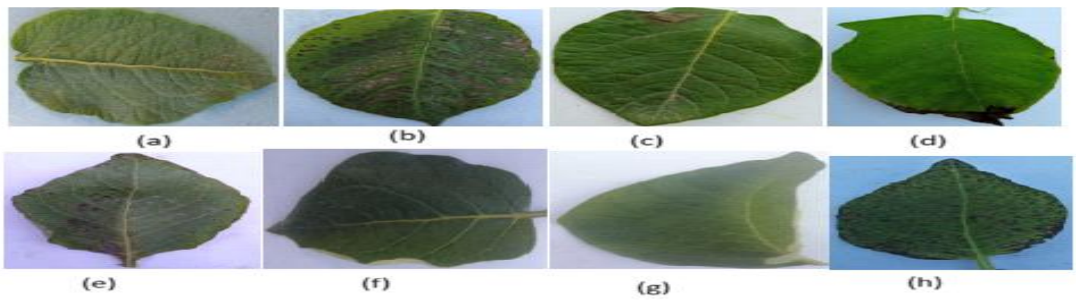

Figure-4. Sample Input Images, (a). Healthy, (b). Early-Blight, (c). Late Blight, (d). Late Blight, (e). Early Blight, (f). Healthy, (g). Healthy, (h). Early Blight

Evaluation Metrics It depends on the applications and study; however, for evaluating the plant disease prediction model, some of the evaluation metrics are widely considered; they are precision, recall, F1-score, P-average, accuracy, and loss value, which are shown as follows,

$$\boldsymbol{Precision = \frac{TP}{TP + FP}.100\% Recall = \frac{TP}{TP + FN}.100\%}$$

The mAP is usually used to obtain the average accuracy of the model. It can be seen that Precision and Recall are used to predict performance, and P_average is obtained to get average accuracy. With a better mAP, the prediction accuracy is high. The F1-score also shows the performance of the model.

$$\boldsymbol{P_{average} = \sum_{j=1}^{N(class)} Precision\,(j).Recall(j).100\%}$$

$$\boldsymbol{mAP = \frac{P_{average}}{N}(class)}$$

$$\boldsymbol{F1 = \frac{2Precision.Recall}{Precision + Recall}.100\%}$$

The performance of the proposed model is evaluated by calculating the performance metrics explained above. The model's performance is verified in terms of image preprocessing and classification. The output obtained from the proposed model is verified at each preprocessing stage and shown in Figures 5 to 17.

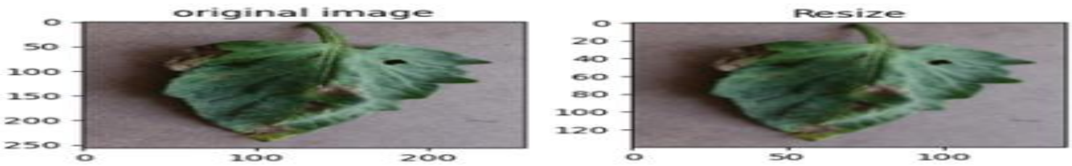

Figure-5. Resizing Based on The Features

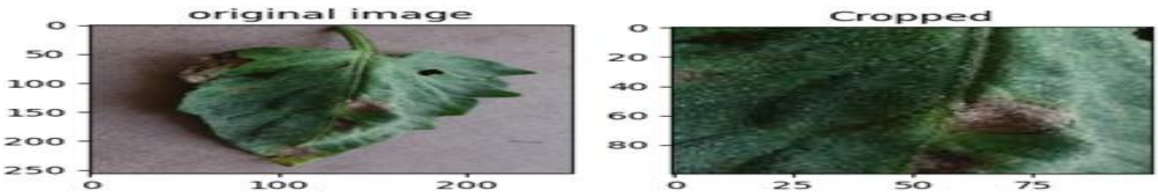

Figure-6. Image Cropped Towards the Affected Regions

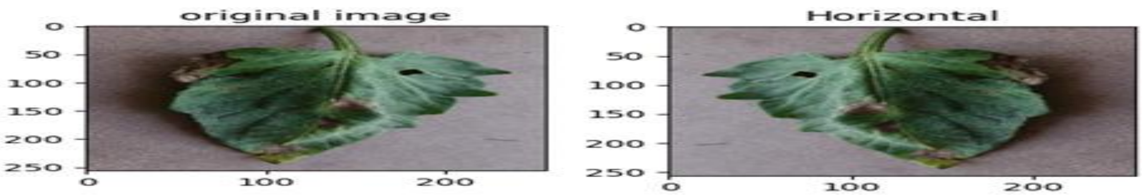

Figure-7. Horizontal Flipping of The Image

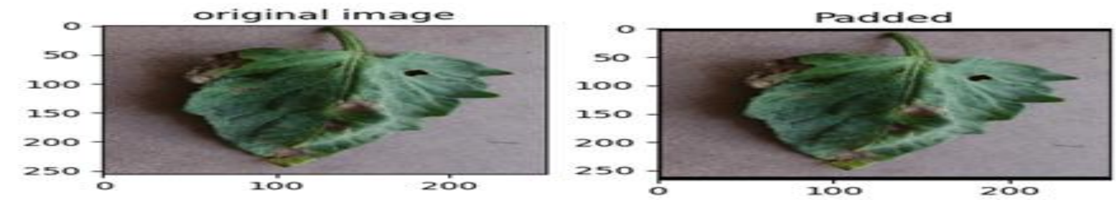

Figure-8. Padding The Images to Remove Unwanted Pixels

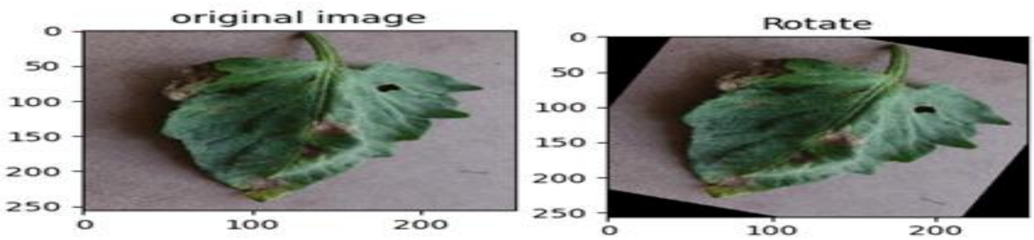

Figure-9. Rotating For Data Augmentation

Figure-10. Distorted Image

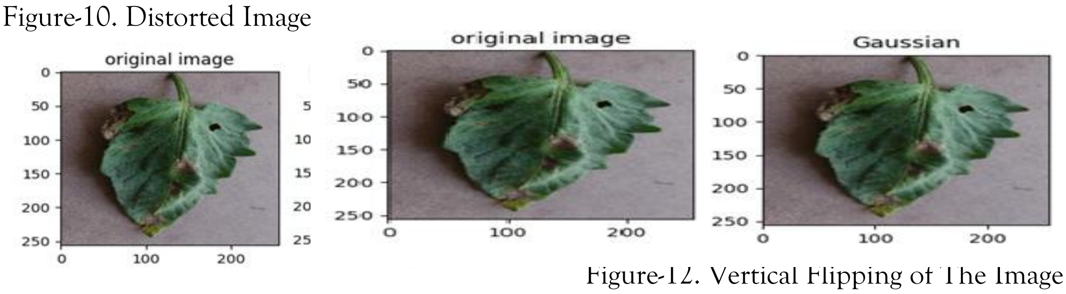

Figure-12. Vertical Flipping of The Image

Figure-13. Random Adjustment of The Brightness

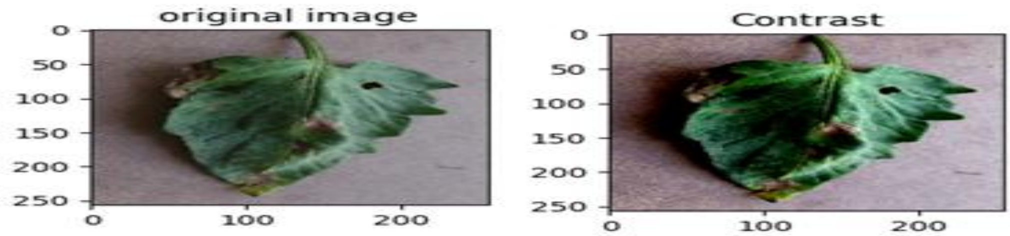

Figure-14. Contrast Enhancement of The Image

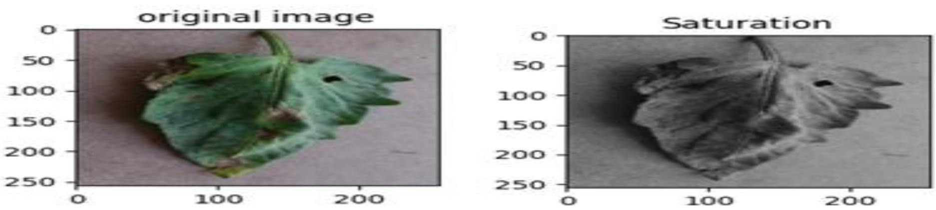

Figure-15. Saturated Image

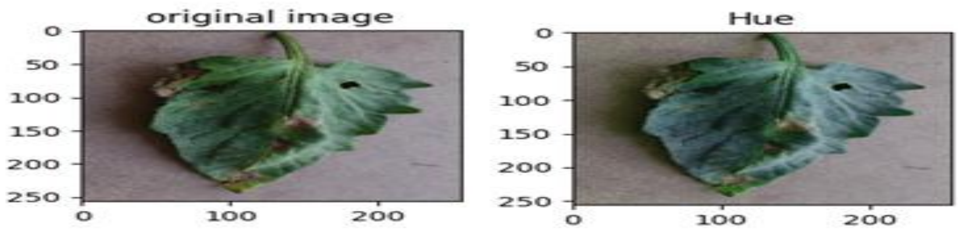

Figure-16. Adding Hue to The Original Image

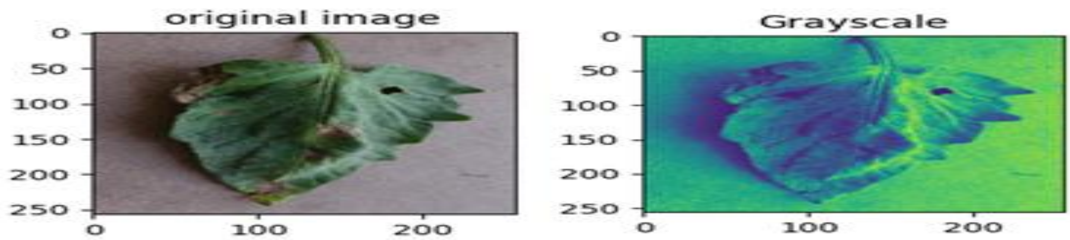

Figure-17. Gray Scale of The Image

The images are resized to emphasize important portions of the image. During the resizing of the image, details were lost, which can be seen in Figure-2. Figure-2 to 14 show the different preprocessing techniques applied to the images. Different features are highlighted during each process for the deep learning model to predict. It helps them better predict difficult-to-predict diseases. Most of these techniques are adopted in the data augmentation to improve the accuracy of the models.

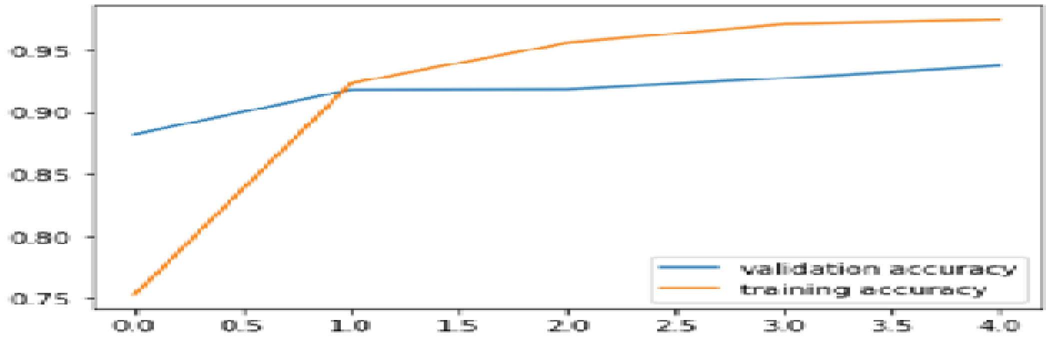

Figure-18. Training And Validation Accuracy of The Model

Figure-18 shows the training and validation accuracy of the CNN algorithm. The graph shows that the training accuracy is higher than the validation accuracy. During the initial training, the model's performance is below 90%, but after certain epochs, the model's accuracy stabilizes. The proposed model is evaluated with the validation dataset, during which the model provides a stable accuracy.

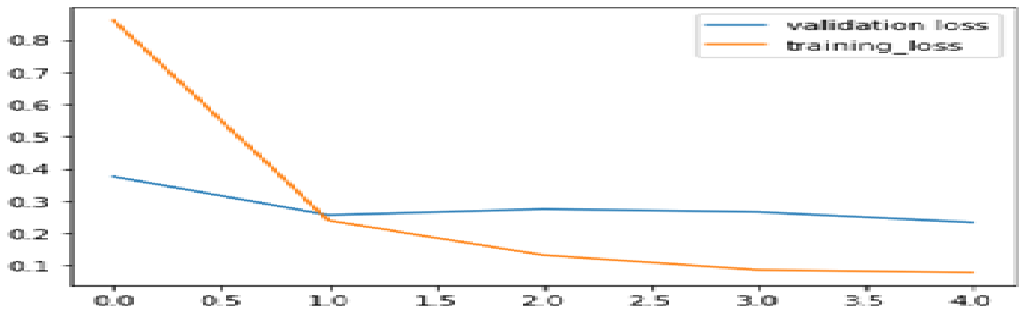

Figure-19. Training And Validation Loss of The Model

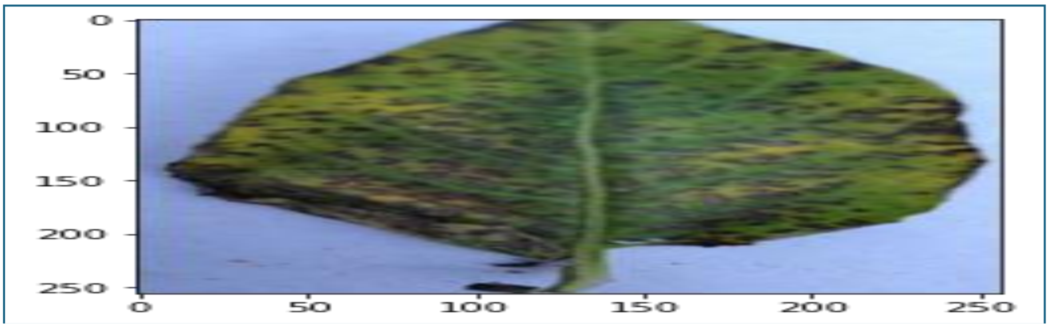

Figure-20. Test Result for Single Leaf

The training and validation loss of the model are shown in Figure-19. Loss of the model should be minimal for achieving a reliable prediction model. The proposed model provides very little loss value during training and validation. The validation loss is higher than the training loss, which is widely seen in most models. Figure-20 shows the test output obtained for a single leaf image. The actual class of the image is Early_Blight, where the prediction output obtained from the proposed CNN is also Early-Blight. Figure-21 shows the classification output obtained for validation data randomly given as input in the experiment.

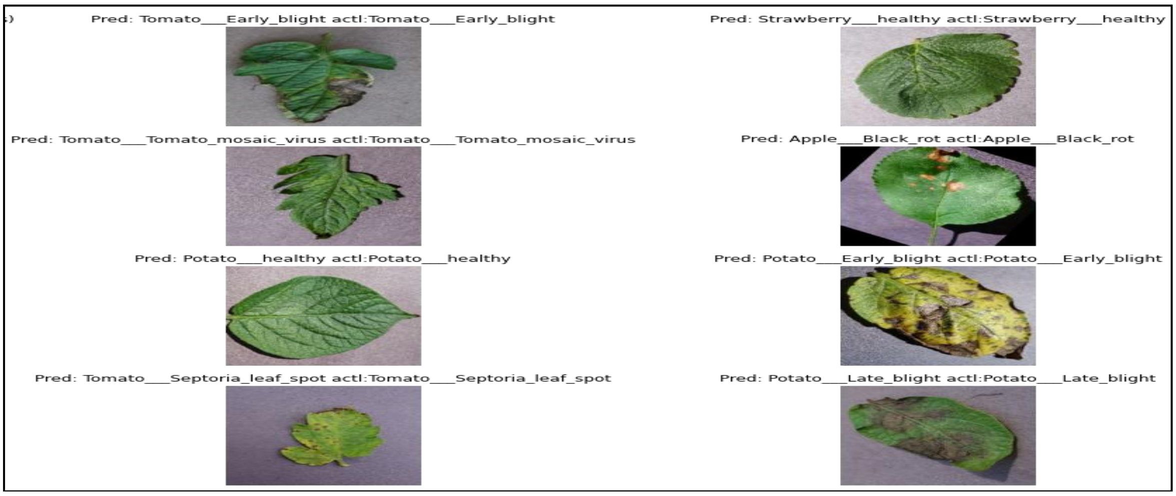

Figure-21. CNN Classification For Validation Data

The output values from the experiment are calculated, proving that the proposed PF-CNN outperforms the others. Compared to other methods, the PF-CNN obtained higher accuracy than other models for different datasets. The performance comparison in terms of accuracy is given in table-5. From the comparison, it is found that the proposed PF-CNN model outperforms others. The performance of the proposed method is evaluated by comparing its output with similar learning methods and given in Table-2. From the comparison, it is identified that the proposed PF-CNN outperforms the others. The proposed model obtained 99.04% accuracy, which is higher than that of others.

Table-2. Performance Comparison

Method	Accuracy
Wang et al.(2017), VGG	90.4%
Atole and Part (2018), AlexNet	91.23%
Coulibaly et al. (2019), VGG-16	95.0%
Hasin Rehana et al. (2023), VGG-16	96.17%
PF-CNN	99.04%

CONCLUSION

The main objective of this work is to explain the need for data preprocessing, which improves the data quality and further processing. Hence, this paper implements a Preprocessing Framework with a Convolution Neural Network (PF-CNN) for preprocessing the images using various preprocessing tasks, such as image alignment, rotation, resizing, cropping, color conversion, and enhancement. The final output is input to the proposed CNN for disease classification. Since this work is the initial stage of the research work, the CNN is trained with pretrained images and compared with the ground truth outputs to confirm the disease class. Then, CNN is tested with test images and validated with validation images. The experiment output is verified and compared with other similar methods to evaluate the performance of PF-CNN. The PF-CNN obtained 99% accuracy for the single image dataset and 98% for the mixed image dataset. In the future, the dataset size will be increased, and preprocessing should be done automatically by the deep learning model to evaluate its performance.

REFERENCES

1. Al-Hiary, H., Bani-Ahmad, S., Reyalat, M., Braik, M., & Alrahamneh, Z. (2011). Fast and accurate detection and classification of plant diseases. International Journal of Computer Applications, 17(1), 31-38.
2. Singh, V., & Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. Information processing in Agriculture, 4(1), 41-49.
3. Sahu, S. K., & Pandey, M. (2023). An optimal hybrid multiclass SVM for plant leaf disease detection using spatial Fuzzy C-Means model. Expert Systems with Applications, 214, 118989.
4. Acharya, S., Kar, T., Samal, U. C., & Patra, P. K. (2023). Performance Comparison between SVM and LS-SVM for Rice Leaf Disease Detection. EAI Endorsed Transactions on Scalable Information Systems, 10(6).
5. Saleem, M., Atta, B. M., Ali, Z., & Bilal, M. (2020). Laser-induced fluorescence spectroscopy for early disease detection in grapefruit plants. Photochemical & Photobiological Sciences, 19, 713-721.

6. Shruthi, U., Nagaveni, V., & Raghavendra, B. K. (2019, March). A review on machine learning classification techniques for plant disease detection. In 2019, the 5th International Conference on Advanced Computing & Communication Systems (ICACCS) (pp. 281-284). IEEE.

7. Ahmed, K., Shahidi, T. R., Alam, S. M. I., & Momen, S. (2019, December). Rice leaf disease detection using machine learning techniques. In 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI) (pp. 1-5). IEEE.

8. Panigrahi, K. P., Das, H., Sahoo, A. K., & Moharana, S. C. (2020). Maize leaf disease detection and classification using machine learning algorithms. In Progress in Computing, Analytics and Networking: Proceedings of ICCAN 2019 (pp. 659-669). Springer Singapore.

9. Li, L., Zhang, S., & Wang, B. (2021). Plant disease detection and classification by deep learning—a review. IEEE Access, 9, 56683-56698.

10. Liu, J., & Wang, X. (2021). Plant diseases and pests detection based on deep learning: a review. Plant Methods, 17, 1-18.

11. Singh, V., Sharma, N., & Singh, S. (2020). A review of imaging techniques for plant disease detection. Artificial Intelligence in Agriculture, 4, 229-242.

12. Vishnoi, V. K., Kumar, K., & Kumar, B. (2021). Plant disease detection using computational intelligence and image processing. Journal of Plant Diseases and Protection, 128, 19-53.

13. Das, D., Singh, M., Mohanty, S. S., & Chakravarty, S. (2020, July). Leaf disease detection using support vector machine. In 2020 International Conference on Communication and Signal Processing (ICCSP) (pp. 1036-1040). IEEE.

14. Sethy, P. K., Barpanda, N. K., Rath, A. K., & Behera, S. K. (2020). Deep feature-based rice leaf disease identification using a support vector machine. Computers and Electronics in Agriculture, 175, 105527.

15. Mayuri, A., Rashmi, V., Das, B., Mandal, A., & Neelakandan, S. (2023, June). Artificial Neural Network (ANN) With Chan-Vese (CV) Algorithm-Based Plant Disease Detection and Classification. In 2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC) (pp. 1-6). IEEE.

16. Wang, G., Sun, Y., Wang, J., 2017. Automatic image-based plant disease severity estimation using deep learning. Comput. Intell. Neurosci. 2017, 1–8.

17. Atole, R.R., Park, D., 2018. A multiclass deep convolutional neural network classifier for detecting common rice plant anomalies. International Journal of Advanced Computer Science and Applications, https:// doi.org/10.14569/IJACSA.2018.090109.