https://doi.org/10.33451/florafauna.v28i1pp77-85
FLORA AND FAUNA
ISSN 2456 - 9364 (Online)
2022 Vol. 28 No. 1 PP 77-85
ISSN 0971 - 6920 (Print)

Plants Disease Detection Using Deep Learning Plant Disease CNN Models for Image Segmentation

***Abhinav Agarwal, D.K.Agarwal[1], T.K. Sharma[2], Vijay Kumar Yadav[3] and Mukesh Srivastava[1]**

*Department of Computer Science,
Singhania University,
JHUNJHUNU (RAJASTHAN), INDIA
[1]Department of Chemistry,
[2]Department of Botany,
[3]Department of Zoology,
Bipin Bihari College,
JHANSI (U.P.) INDIA
*Corresponding Author
Email: abhinavkiot2410@gmail.com

Received **:** 18.02.2022; ***Revised*** **:** 10.03.2022; ***Accepted*** : 28.03.2022

ABSTRACT

For increased agricultural productivity, early diagnosis and management of plants illness are critical. The texture, color, and spots of a diseased plant's leaves may be used to distinguish it from a healthy one. Observing leaves a traditional way requires a certain level of experience. Farmers that lack experience and resources might benefit from the creation of plant illness detection utilizing Deep Learning methods. Deep Learning methods are used in this study to categorise the various plant illnesses. Due to its great success in image-based classifications, the plant illness convolutional neural network (CNN) construction is deployed. When compared to manual observation of leaves, the Deep Learning models are quicker and more precise. The Plant Pathology 2020 FGVC7 data set is used to train the PD CNN framework in this study. The model's 95 percent accuracy is the best among them.

Figures : 09 References : 13 Tables : 03
KEY WORDS : Convolutional Neural Networks, Data Augmentation, Deep Learning, Plant disease

Introduction

The diagnosis of plant diseases is a major difficulty in agriculture. Signs may be seen on the leaves of some of the plants. To stop the transmission of illness, these leaf patterns might be used to detect it. Even experts have trouble seeing the symptoms of many of these plant illnesses since they are challenging to see with the naked eye[1]. There is a direct correlation between the accuracy of manual forecast and the person's expertise and understanding. Plant illnesses have been highlighted as a growing hazard to food security all around the world. As a result, the most critical stage in growing healthy crops is the diagnosis of plant illnesses. As a result of the wide diversity & similarities of plants in nature, it is impossible to categorize plants "with and without illness."[11]

Fig. 1: Healthy Leaves

Seventeen percent of India's GDP and 60 percent of the nation's total employment are attributable to agriculture. It is impossible to regulate the effects of climate changes on crop yields in several sections of the nation. Plant illnesses may, however, be controlled, resulting in a decrease in ill effect on yields. Several Plant illnesses impact agricultural yields[8]. Crops are suffering from a wide range of plant illnesses, which not only harm the economy but also influence the quality & quantity of food. Different techniques & procedures are being used to identify these plant illnesses. However, certain farmers

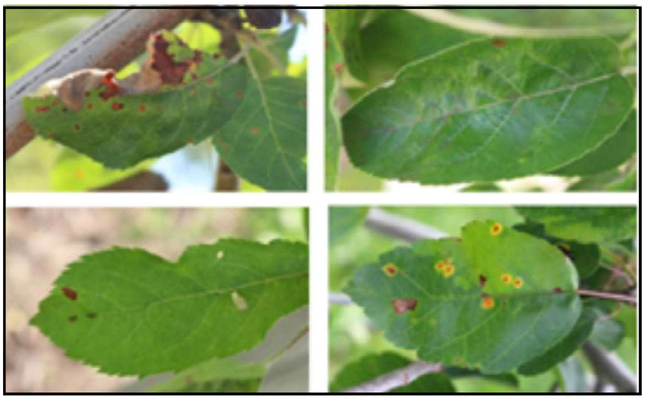

Fig. 2: Multiple plant disease

Fig. 3: Rust Leaves

can afford laboratories and the supervision of scientists. The vast majority of them cannot afford to hire professionals or invest in the equipment needed to identify and avoid the spreading of plant illnesses. Farmers also need to send a sample of the diseased plant to certified laboratories to determine it. Economically, these laboratories are not feasible, and the funds should be better spent on other initiatives that would increase the return. As a result, to reduce the cost of a lab and the expertise it requires to identify plant illnesses and increase yields for both farmers & users, we must use digital methods.[13]

We cannot control the impact of climate-changing on agricultural production throughout the country. It is important to know the many plant diseases that might affect agricultural productivity. As a result of a variety of plant ailments, the economy is suffering and the quality and quantity of food are being affected. There are a variety of tools and processes being employed to diagnose these plant diseases. Then there are a few farmers who can afford labs and the supervision of experts. They are unable to pay experts or purchase the necessary equipment to diagnose and prevent the development of plant diseases. A sick plant sample must be sent to a recognized laboratory for examination to confirm the diagnosis. These labs are not financially viable, and the funding should instead be used for other projects that would provide a greater return. Because of this, we must adopt digital approaches to lower the costs of a lab and the knowledge required to diagnose plant ailments and enhance yields for both farmers and consumers.

Deep Learning approaches and state-of-the-art algorithms are employed in this research to discover plant illnesses that may assist those in agriculture who want to solve the current issue in this field.

Related Work

Various research works have been conducted over the years for identifying plant diseases using leaves, thus the most recent methodologies pertinent to this subject are discussed in this section.

Using four CNN algorithms, a novel plant disease detecting method has been created[4]. To experiment, researchers utilized an open-source database containing 36258 photos divided up into 10 plant species and 61 categories of leaf types in good condition and those with illness. Two datasets of 36258 photos were created, containing 31718 pictures for the training set and 4540 for the validation set. Four different types of CNN architectures were used, included Inception, Resnet, Inception Resnet, and Densenet. An 87 percent accuracy rate was reached utilizing the stacking strategy, which is a substantial increase above the single CNN model's results. It seems the use of a mixture of CNN models combined with stacking technique may be an effective strategy that may be applied to realistic cultivation circumstances as an advanced plant illness warning tool.

Applying a computing approach on Deep Learning systems[9] that are dependent on artificial neural networks, this enables for earlier identification of plant illnesses, utilizing convolutional neural networks (CNNs) acquainted with a few of the renowned designs, especially "ResNet" design, using an enhanced dataset including photos of healthy and sick leaves (every leaf is carefully chopped

Fig. 4: Scab Leaves

Model: "sequential"

Layer (type)	**Output Shape**	**Param #**
Conv2d (Conv2D) max_pooling2d (Max Pooling2D)	(None, 196, 196, 32) (None, 98, 98, 32)	896 0
batch_normalization (BatchNo)	(None, 98, 98, 32)	128
conv2d 1 (Conv2D)	(None, 98, 98, 64)	18496
max_pooling2d 1 (Max Pooling2)	(None, 49, 49, 64)	0
batch_normalization_1 (Batch) conv2d_2 (Conv2D)	(None, 49, 49, 64) (None, 49, 49, 128)	256 73856
max_pooling2d_2 (Max Pooling2)	(None, 24, 24, 128)	0
batch normalization 2 (Batch)	(None, 24, 24, 128)	512
conv2d_3 (Conv2D) max_pooling2d_3 (MaxPooling2)	(None, 24, 24, 256) (None, 12, 12, 256)	295168 0
batch normalization 3 (Batch)	(None, 12, 12, 256)	1024
conv2d_4 (Conv2D)	(None, 12, 12, 384)	885120
max pooling2d 4 (Max Pooling 2)	(None, 6, 6, 384)	0
batch_normalization_4 (Batch) conv2d_5 (Conv2D)	(None, 6, 6, 512)	1769984
max_pooling2d_5 (MaxPooling2)	(None, 3, 3, 512)	0
batch_normalization_5 (Batch)	(None, 3, 3, 512)	2048
flatten (Flatten)	(None, 4608)	0
dense (Dense)	(None, 1024)	4719616
dropout (Dropout)	(None, 1024)	0
dense 1 (Dense)	(None, 4)	4100
Total params: 7,772,740 Trainable params: 7,769,988 None-trainable params: 2,752		

Fig. 5: CNN layered architecture

TABLE-1: Training Parameters

Parameter	Value
Epochs	100
Batch size	32
Learning rate	.0001
Optimizer	ADAM

and placed on a consistent backdrop) with acceptable accuracy rates in the study. When applied to a variety of object identification situations, our Deep Learning approach has shown exceptional efficiency. The framework performs its function by categorizing photos into two groups: those that are free of illness and those that are not (*i.e.* diseased). By the findings, the created system outperforms advanced detection systems in terms of detecting performances. The implementations under Anaconda 2019.10 are used to contrast their efficiency at the end of the process.

A study[10] facilitated the detection of numerous illnesses in a diversity of plant species.Apple, corn, grapes, potato, sugarcane & tomato were all included in the system's scope of detection & recognition. Many plant diseases may be detected with the technique. The researcher was able to train Deep Learning frameworks to identify & distinguish plant illnesses and the absence of these illnesses utilizing 35,000 photos of healthy and diseased plant leaves. 95.5% accuracy has been reached by the modeled trained system, which is capable of registering 100 percent accuracy in the detection and recognition of plant variety & illness kind of diseased plant.

Picture processing & Machine Learning techniques were used in the study[7] of plant illness detection. Early and effective disease identification is critical to the quality & production of crops. Early detection and treatment of plant diseases may save money and avoid the need for unneeded medication. In this work, leaf photos from a variety of plant species were gathered & feature extraction was performed using the transfer learning approach. The suggested framework was 94% successful using a variety of machine learning techniques.

The convolutional autoencoder was used to develop an unsupervised feature learning system for the identification of plant illnesses[12]. Using a convolutional autoencoder provides two primary benefits. You don't need to create features manually, since the network can learn to generate discriminative characteristics. No labeling is necessary since this operation is unsupervised, therefore there's no need for it. Plant illnesses may be automatically detected using SVM-based classifiers that employ the output of an autoencoder. Autoencoders using additional hidden layers seem to be inferior to this approach.

Material and Methods

This section presents the dataset that was utilized in the research. Additionally, the part discusses the application of Deep Learning for the detection of damaged crops by analyzing the plant's leaves using this technology. The categorization of plant leaf photos as unhealthy or not, is carried out using deep learning utilizing CNN.

A. Dataset

This research relies on the Plant Pathology 2020 FGVC7 dataset. In this dataset, there are 1821 training photos, all of which have been given the same size of 196 × 196 pixels, such as apples with the scab, apples without scab, grapes with black rot, *etc.* This dataset is available in three separate formats: color, grayscale, & segmentation. There are plant-disease pairings in this dataset, which is split into 70% for training & 30% for testing. Leaf pictures from datasets, like healthy, numerous illness, scab, & rust-infested leaves, are shown in Fig. 1-4[6].

B. Image Preprocessing

As part of the picture assessment and interpretation, preprocessing is utilized. Because disease identification & diagnosis are becoming more important in agricultural research, the preprocessing of plant pictures is a difficulty that has to be overcome to ensure disease-free crops. At the most basic level of abstraction, one of the most important operations on picture data is preprocessing. To minimize distortion, preprocessing aims to improve the image's characteristics. Image preprocessing is an essential step in the picture analysis procedure because it ensures that the picture is clean

TABLE-2: Results of proposed PD CONV Model

Model	Loss	Training Acc	Validation Loss	Validation Acc
PDConv	0.1788	0.9546	1.1523	0.6967

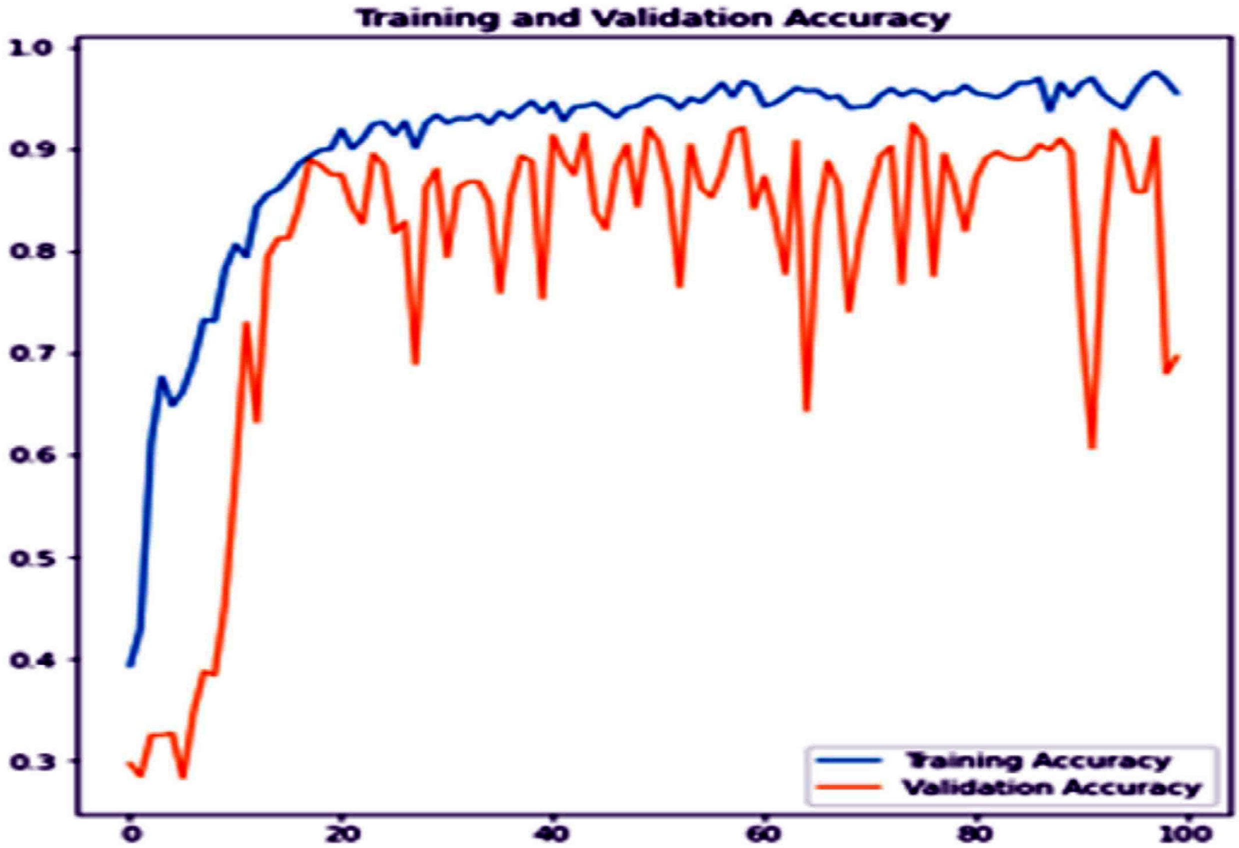

Fig. 6: Accuracy plots for training and validation of PDCNN

and ready for subsequent steps like segmentation, feature extraction, & categorization. Deep Learning uses picture preprocessing to improve the model's efficiency[2]. These downsized photos are used to train and test models using pictures that have been reduced to 196x196 pixels.

C. Data Augmentation

The quantity of data needed for illness detection on leaves utilizing deep learning is enormous, but obtaining it is a challenge. We thus decide to use a more efficient strategy of data augmentation to tackle this issue[5]. Increasing the "train" as a whole also assists to alleviate the issue of over-adjustment. For the optimal outcomes, the following settings were specified in the "Keras Image Data Generator" class: (a) Rotational Range of 40 (b) Zoom & Shear Range of 0.2 (c) True for horizontal flipping (d) Partitioning for Validation: 0.3, (e) Rescale:1/255.

D. Proposed Method

A CNN is a most often used approach for extracting useful information from large datasets. A bespoke CNN is built using 3 convolutional layers, followed by RELU activation and 3Max Pooling 2D layers (a max-pooling layer succeeding every convolutional layer). A Plant disease (PD) CNN is suggested for this purpose. (196*196) is the size of the picture that is sent into the network. Filters of various diameters are used. Every convolutional layer of the plant disease dataset has an instance of every class. The activation function is'relu' on

TABLE-3 : Result comparison between existing CNN and proposed Plant disease (PD) Conv model

Model	Loss	Training Acc	Validation Loss	Validation Acc
CNN	0.8985	0.6299	1.2327	0.4411
PD Conv	0.1788	0.9546	1.1523	0.6967

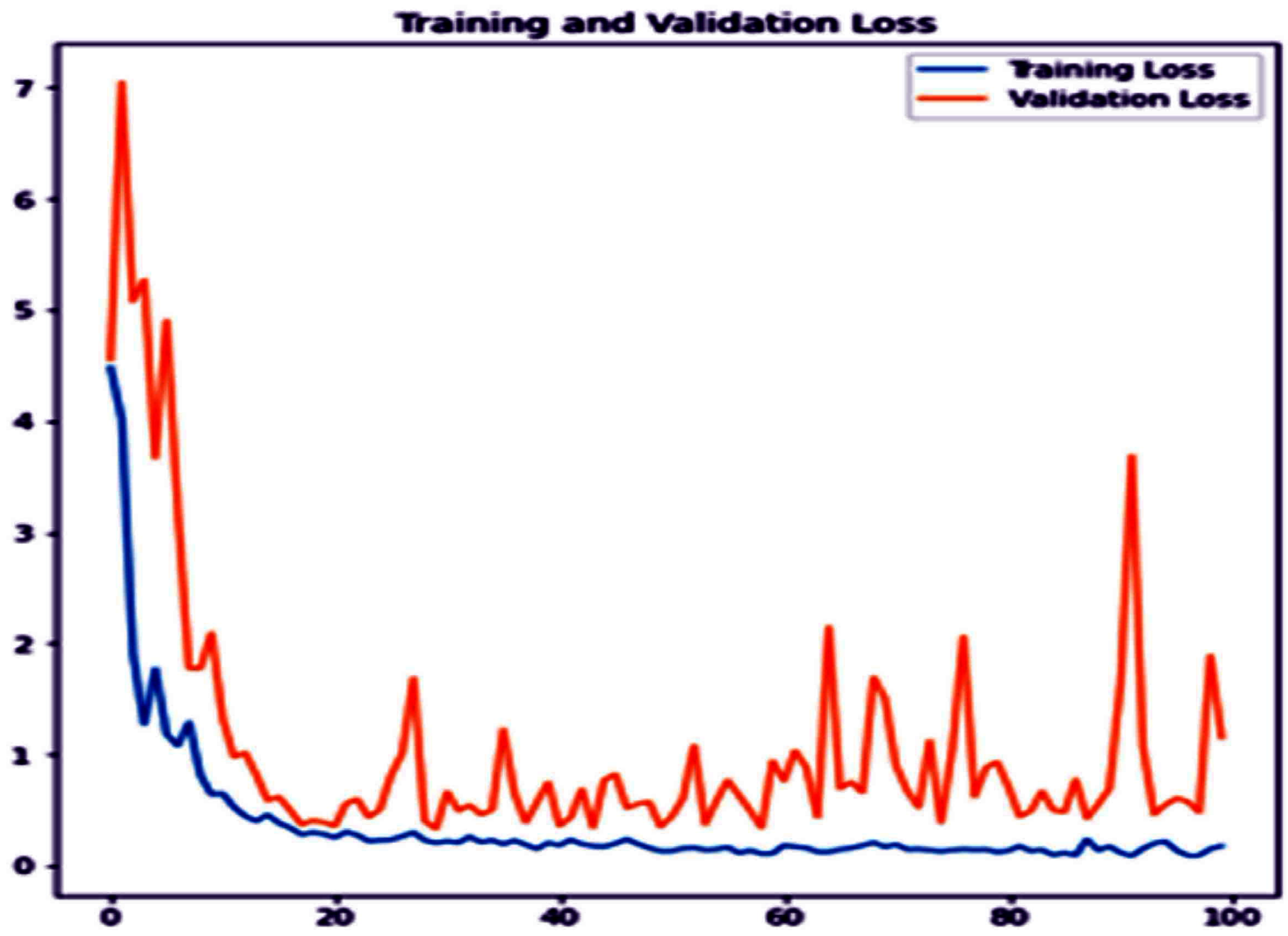

Fig. 7: Plots for validation and training loss of PDCNN

single, fully connected layers. Batch Normalization is also used to stabilize the model's learning procedure and to shorten the time it takes to train. If the prototype overfits during training, a dropout function may help it be more resilient on validation data. Among the model's 7,772,740 total parameters, 7,769,988 of them are trainable & 2,752 are non-trainable. Figure 5 depicts the network layer structure of CNNs.

Result and Discussion

After finishing the course, testing, & validating suggested models utilizing multiple datasets, we were able to give the required result. The following paragraphs provide further information about this particular model result.

Python, a simulation program that offers a runtime for deep learning and free access to GPU computers, was employed in this study[3]. Jupyter notebooks were configured on a project-by-project basis.

A. Analysis

After around 100 epochs of training, the loss curve converged around the x-axis for suggested CNN-based models (Fig. 6). Each batch of data was analyzed to determine the network's accuracy and loss score. Due to the massive quantity of training data, models' overall training time has increased. These prototypes were trained using parameters listed in Table-1. During the training process, an ADAM optimizer was used with a learning rate of 0.0001 and a batch size of 32. PDCNN trained from scratch acquired an accuracy of 95.4% and validation accuracy of 69.67% as (Fig. 6).

As seen in figure 6, several prototype training and validation accuracy graphs are provided. If contrasted to the other methods (PDCNN and so on.), CNN obtained the best accuracy.

Figure 7 demonstrates the plots used for training and validation loss of the Plant disease CNN. Training loss is recorded after every batch, whereas validating loss is measured after every epoch, hence on average, the training loss is calculated. 0.177 was the training loss whereas 1.152 was the validating loss.

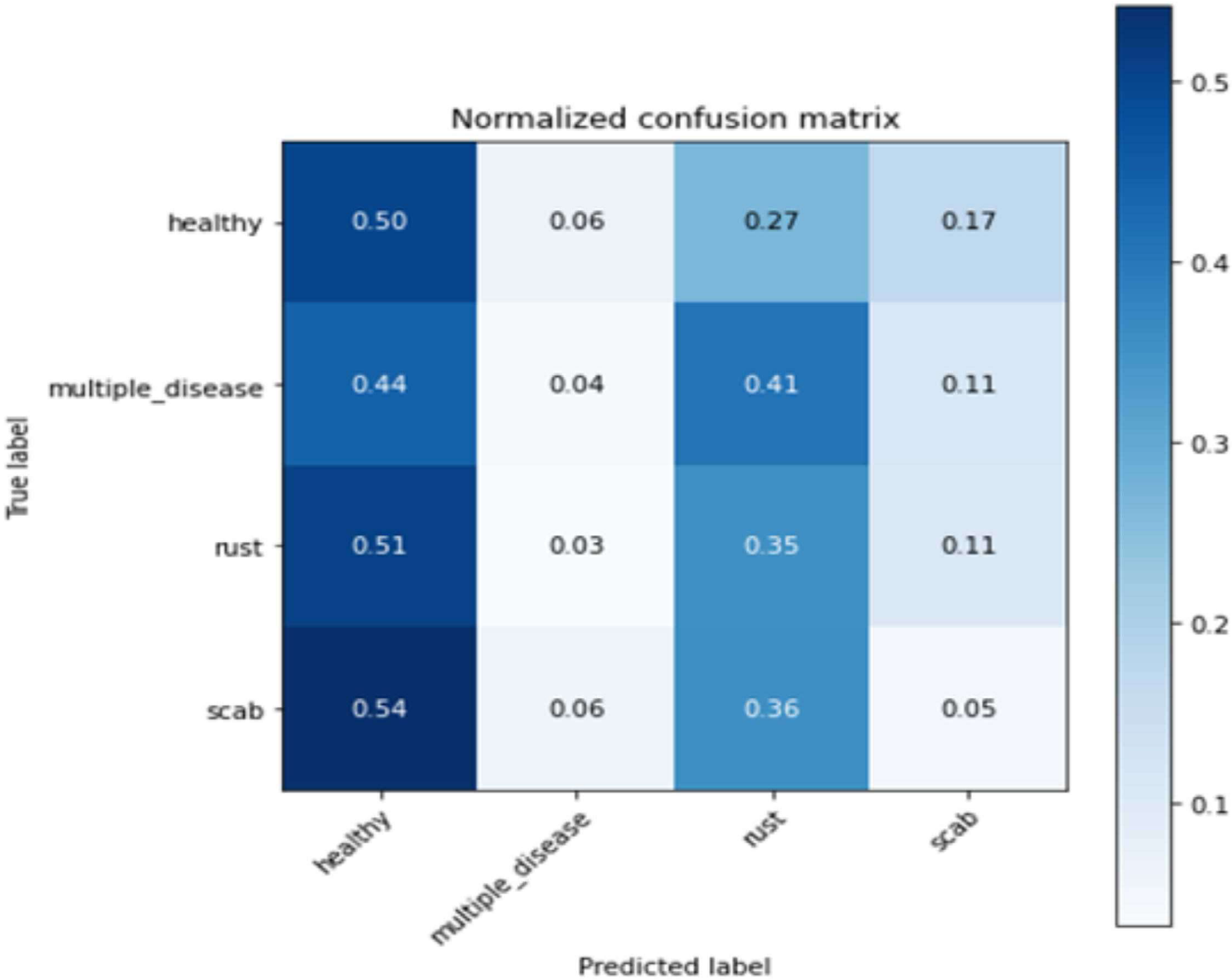

Fig. 8: Confusion Matrix

B. Confusion Matrix

There is a special error Table called a confusion or error matrix that shows the model's ability to perform well. Confusion matrix values that are placed diagonally tend to be more significant than those that are placed horizontally. Values placed on the diagonal in a (4x4) shape are displaying their highest efficiency compared to other positions, and this portion of data is maximized. The classification of images such as healthy, multiple diseases, rust, and scab leaves are presented by a confusion matrix. Fig.8. Shows confusion matrix.

C. Outcomes of Proposed Model

The results of our experiments show that after training for 100 epochs, our trained models have a training accuracy of 95.45%. After that, we'll use our data set using the "Data Augmentation" approach to increase the model's accuracy. Table 2 displays the outcomes of our deep learning method to plant illness detecting using the suggested PDconv models. Along with training accuracy and loss, as well as testing validity and testing loss, this displays the most effective outcomes obtained so far.

Comparative Performance

When every prototype was trained on the dataset for a distinct period of time and distinct epochs, the validation accuracy and validation loss for each model were varied as a result. Table 3 shows the Comparative performance of the existing CNN and the proposed PD Conv (Best results are highlighted in bold).

Figure 8 shows the comparison of performance parameters such as loss, training accuracy, validation loss and validation accuracy between the existing CNN model and the proposed PD Conv model that get better results better than the existing model.

Conclusion and Future Scope

One of the most significant aspects of deep learning is CNN. CNN frameworks have been used to detect plant

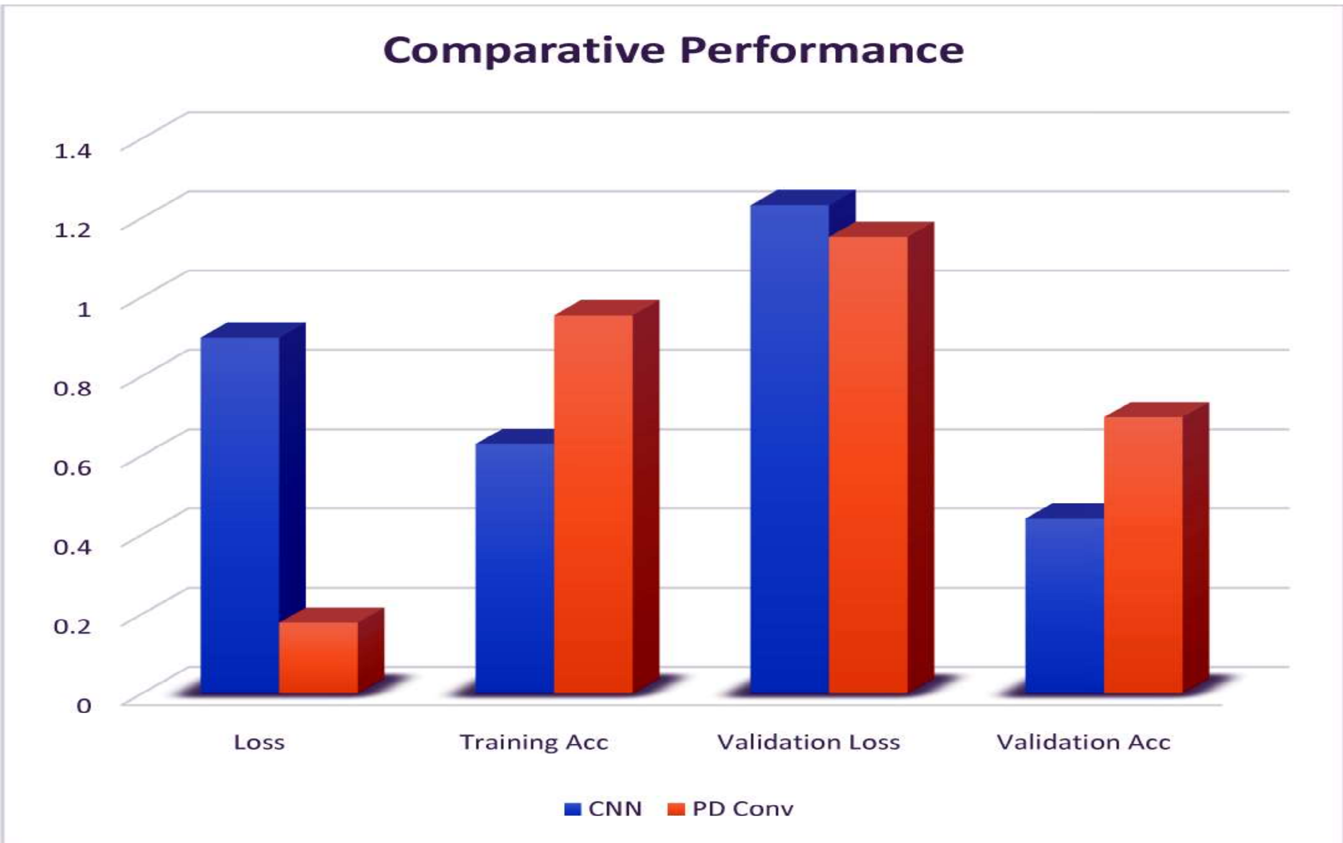

Fig. 9: Comparison bar Graph of performance parameters for Existing CNN and Proposed PD Conv model

diseases because of their capacity to extract features. Plant leaves were used as a test case to demonstrate the usefulness of the suggested method for illness detecting, with promising precision findings. Accurately classifying plant leaves as healthy, scab, rust and infected was the purpose of this work. Our research shows that a CNN model can be trained. A rise of 62.9% to 95.4% in prototype efficiency training accuracy on independent data and enhanced outcomes compared to a current prototype, like loss, training accuracy, validation loss & validation accuracy, can be seen when the same framework is trained using the (CNN) instead of training utilizing plant disease (PD-Conv).

The dataset will be expanded, and the number of classes will be increased in future work. Another future project will be the incorporation of the models into a site or software that will assist farmers and pathologists in the identification of various illnesses employing their mobile cameras. Improved datasets will be made accessible in the future, and the pre-processing of pictures before models training in CNN will become more important for achieving superior real-world performance.

References

1. Atila U, Ucar M, Akyol K, Ucar E. Plant leaf disease classification using efficient net deep learning model. *Ecol. Inform.* 2021; doi: 10.1016/ecoinf.2020.101182.

2. Barbedo JGA. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Comput. Electron. Agric.* 2018; doi: 10.1016/j.compag.2018.08.013. 10.1016/j.compag.2020.105220

3. Foundation PS. Python programming langwage. Python.org, [online]. Available: https://www.python.org/.

4. Guan X. A novel method of plant leaf disease detection based on deep learning and convolutional neural network. 2021; doi: 10.1119/icsp50882.2021.9408806.

5. Hawkins DM. The problem of overfitting. *Journal of Chemical Information and Computer Sciences.* 2004; doi: 10.1021/ci0342472.

6. Kaggle.com.plant-pathology-2020. https://www.kaggle.com/c/plant-pathology-2020-fgvc7.
7. Korkut UB, Gokturk OB, Yildiz O. Detection of plant diseases by machine learning. 2018; doi: 10.1109/siu.2018.8404692.
8. Lee SH, Goeau H, Bonnet P, Joly A. New perspectives on plant disease characterization based on deep learning. Comput. Electron. *Agric.* 2020; doi:
9. Marzougui F, Elleuch M, Kherallah M. A deep CNN approach for plant disease detection. 2020; doi: 10.1109/acit50332.2020.9300072.
10. Militante SV, Gerardo BD, Dionisio NV. Plant leaf detection and disease recognition using deep learning. 2019; doi: 10.1109/ecice47484.2019.8942686.
11. Mohanty SP, Hughes DP, Salathe M. Using deep learning for image-based plant disease detection. *Front. Plant Sci.* 2016; doi: 10.3389/fpls.2016.01419.
12. Pardede HF, Suryawati E, Sustika R, Zilvan V. Unsupervised convolutional autoencoder-based feature learning for automatic detection of plant diseases. 2019; doi: 10.1109/ic3ina.2018.8629518.
13. Sladojevic S, Arsenovic M, Anderla A, Culibrk D, Stefanovic D. Deep neural networks based recognition of plant diseases by leaf image classification. *Comput. Intell. Neurosci.* 2016; doi: 10.1155/2016/3289801.