Leaf Diseases Identification and Classification of Self-Collected Dataset on Groundnut Crop using Progressive Convolutional Neural Network (PGCNN)

Anna Anbumozhi, Shanthini A
Department of Data Science and Business Systems
SRM Institute of Science and Technology, Kattankulathur, Chennai 603203

***Abstract*—A healthy crop is required to provide high-quality food for daily consumption. Crop leaf diseases have more influence on agronomic production and our country. Earlier, many scholars relied on traditional techniques to detect and classify leaf diseases. Furthermore, classification at an early stage is impossible when there are not enough experts and inadequate research facilities. As technology progresses into our day to day life, an Artificial Intelligence subset called Deep Learning (DL) models plays a vital role in the automatic identification of groundnut leaf diseases. The essential for controlling diseases that are spread to the healthy development of groundnut farming. Deep Learning can resolve the issues of finding leaf diseases early and effectively. Most of the researchers concentrate on detecting leaf diseases by doing research in Machine Learning (ML) approaches, which leads to low accuracy and high loss. To achieve better accuracy and decreases the loss in the DL model by identifying the leaf diseases of groundnut crops at an early stage, we propose the Progressive Groundnut Convolutional Neural Network (PGCNN) model. This paper mainly focuses on identifying and classifying groundnut leaf diseases with a self-collected dataset which is collected from the various climatic conditions around the village located nearby Pudukkottai district, Tamil Nadu, India. The common diseases that occurred in those areas were gathered namely early spot, late spot, rust, and rosette. Model Performance metrics analysis was done to evaluate the model performance and also compared with the various DL architectures like AlexNet, VGG11, VGG13, VGG16, and VGG19. The proposed models have achieved a training accuracy of 99.39% and a validation accuracy of 97.58%, continuing with an overall accuracy of 97.58%.**

Keywords*—*Leaf Diseases Identification (LDI); Progressive Groundnut Convolutional Neural Networks (PGCNN); Self-Collected Dataset; AlexNet; VGG Models

I. Introduction

Agriculture is a major source of revenue for the farmers in India. Despite the fact that agriculture no longer makes up the majority of the country's Gross Domestic Product (GDP) and that other industries' contributions have grown more quickly, agricultural productivity has increased. Timely identification of crop leaf diseases is critical in agricultural yield, quality management, and decision-making. [1] This research builds the model based on CNN architecture for groundnut leaf diseases with five categories which was common in Gujarat area with the progressive resizing using cross entropy loss function and achieved the results as 96.12% accuracy. The peanut, scientifically known as arachis hypogaea, it is a herbaceous yearly plant that belongs to the Fabaceae family and is cultivated for its oil, edible nuts, and nutritious snacks. The symptoms that are shown in the leaf diseases can be easily differentiated between the two infections based on their appearance, spot color, and shape. However, farmers must suffer as a result of inadequate agricultural income. Plants that produce peanuts are typically quite small and have slender stems and leaves that resemble feathers. The leaves are attached to the stalk in a manner that resembles a leaf and is arranged in pairs that alternate. [2] Describes the end-to-end Internet of Things (IoT) constructed system for groundnut leaf disease detection and castor oil plant leaf disease recognition. The vast majority of peanuts grown for commercial purposes are crushed up to extract their oil, which is then put to use in culinary endeavours. A pressed cake is produced as a by-product of oil extraction, and in addition to its use as animal feed, it is also put to use in the manufacturing of peanut flour. Raw kernels are frequently roasted into a snack food and consumed by roasted or boiled manner.

Here, the four types of diseased leaves were selected to experiment with disease identification: early spot, late spot, rust, and rosette. These four have been taken as unhealthy class for model creation in this paper. The above leaf disease images were gathered from the area namely Pudukkottai, Tamil Nadu, India in various climates. All the images are captured with a smartphone-branded VIVO V2029 model with a camera quality of 13 Mpx LED. Gathered images are transferred to the system for further processing, whether the leaf images were affected by diseases or not. To identify groundnut leaf diseases automatically we are proposing Progressive Groundnut Convolutional Neural Networks (PGCNN). The study [3] represents the idea for the recognition and classification on groundnut leaf diseases using the backpropagation procedure by the color transformation in leaves. Previously the traditional CNN architecture was used in many different research papers to identify leaf diseases but there no accurate identification was found. So, to address that problem we are developing the model with a progressive convolutional stack with customized pooling layers to reduce the size of the image and activation function as sigmoid for binary classification.

Convolutional neural networks (CNNs) have been demonstrated to perform well in a variety of image classification tasks. We apply CNNs to the problem of identifying and classifying peanut leaf diseases in this paper. Peanut is a key food crop in many regions of the world, and leaf diseases can drastically limit productivity. Early and accurate disease diagnosis is critical for effective disease control. However, due to a large number of diseases and the similarity of symptoms among them, diagnosing peanut leaf diseases is frequently difficult. CNNs have been demonstrated to be effective in a variety of image classification tasks. Our CNN model achieves 96.7% accuracy on a held-out test set, which is much greater than the accuracy of a baseline model that uses traditional hand-crafted features.

The rest of the paper is sectioned by the following pattern: Section II details about the existing methods that are used for identifying and classifying the leaf diseases as literature review with various peer review journals; Section III describes about the datasets which was collected from the agricultural field from Pudukkottai district and also shows the sample leaves of healthy and unhealthy leaves; Section IV discusses the details of flow diagram of dataset and the proposed PGCNN model for leaf disease identification and classification among the groundnut self-collected dataset; Section V explains about the implementation details of PGCNN model and the results were achieved, as well as the discussion that was pertinent to them are presented; Section VI shows the results and discussion about the PGCNN model and plots the graphs for training and validation accuracy; Section VII compares our proposed model to the existing CNN architecture to show the proposed model works well than the existing model; and Section VIII concludes the paper with briefs about the proposed model results with future enhancement.

II. Literature Review

A. Identification of Diseases

In recent days, the identification and Classification of groundnut crop leaf diseases are done automatically by ML and DL algorithms. Identification of healthy or diseased leaves on groundnut crops is based on the features extracted from the leaves as yellow or brown changes in color from the original green texture of normal groundnut leaves. After identifying the healthy and diseased leaves, it has to be classified whether it is infected or not. Some of the research papers were taken as reference papers for this work to attain the solution with a Deep Learning model for the evaluation of leaf diseases on groundnut crops.

In [4] researcher has done the experiment to identify the Maize leaf diseases based on the feature enhancement and designed a neural network with the base as Alexnet architecture. DMS-Robust Alexnet architecture was proposed with dilated and multi-scale convolution to improve the quality of feature extraction. To avoid overfitting batch normalization is used, to improve the convergence and accuracy researcher used the activation function as PRelu, and Optimizer as Adabound and achieved 98.62% with DMS-Robust Alexnet.

In [5] author proposed a model with a deep learning approach called the Inception model and Rainbow concatenation (INAR-SSD) model which is based on Improved Convolutional Neural Networks for apple disease detection. The major five types of leaf diseases were selected and implemented in the INAR-SSD model with a recognition performance of 78.80% in Apple Leaf Disease Dataset (ALDD).

The research [6] proposed the Leaf-GAN model to create four different types of leaf diseases for grape crops to identify the healthy or diseased images and fed the generated images to training. There were total of 4062 images before using Leaf GAN, after using augmentation techniques 8124 images are generated to reduce the overfitting problem. Upon CNN models Xception attains an accuracy of 98.70% on testing the datasets.

The study [7] proposed a fine-grained GAN method for identification of grape leaf spot disease to improve the training images to get good performance accuracy and also faster R-CNN is integrated with the above method. With the proposed method higher accuracy was achieved with Resnet-50 as 96.27%.

The study [8] achieved 98.75% and 96.25% of accuracy by implementing the pretrained CNN model called AlexNet and GoolgleNet models. This model performed better than the traditional pattern recognition techniques. Here the researcher used five-fold cross-validation approach and crop selected for the plant disease identification is soyabean.

The author [9] proposed Restructured Residual Dense Network (RRDN) for identification of tomato leaf diseases with datasets of AI challenger 2018. With residual and dense networks, which minimize the amount of parameters to increase computation accuracy, the result was 95% with top-1 average accuracy.

In [10] the study achieved top-1 level accuracy 94.33% by experimenting Deep Convolutional Generative Adversarial Network (DCGAN) to improve the tomato leaf disease recognition that increases the generalization ability. This network improves the performance, decreases the dataset collection cost and also establishes the diversity of generalization of DCGAN recognition models.

Also, [11] Proposed Deep Convolutional Neural Networks with multiclass classifier for detecting common rice crop diseases with the base model of AlexNet by SGD optimizer on the learning rate of 0.0001 and achieved 91.23% accuracy.

B. Classification of Leaf Diseases

The research [12] proposed the pretrained model of modified InceptionResNet-V2 (MIR-V2) based on CNN by transfer learning to identify the illness of tomato leaves. This model is trained with both the public and custom dataset of 7 types of diseases. Model achieved the accuracy of 98.92% and F1-Score is 97.94%.

The research work in [13] proposed new model which is based on recognition and classification of groundnut crop leaf diseases. ICS algorithm is used for segmenting the leaves which are affected by diseases. Then MSO algorithm has been used for multiclass feature extraction and MO-DNN algorithm is used for disease classification of multi-classes. At last, GLD-HML was proposed for analyzing the benchmark datasets to show the performance metrics achieved better than existing.

The authors in [14] used two pre-trained models namely EfficientNetB0 and DenseNet121 for feature extraction from the corn leaves to identify and classify the diseases. The results achieved by new model is 98.56% and compared with two existing model called ResNet152 and InceptionV3 achieved accuracy of achieved accuracy of 7% and 96.26% respectively.

In [15] researcher proposed the Deep Convolutional Neural Network (DCNN) which is based on the evolved concept of transfer learning. Here three types of optimizers were verified with the new model for tomato leaf diseases to identify which optimizer (Adam, SGD and RMSprop) works well for the model. The experimental results showed that the transfer learning model attained the good accuracy with Adam optimizers.

The study [16] utilized the DCNN model for determination and classification of groundnut crop leaves with different types of diseased leaves. This paper achieves accuracy of 95.28% with the optimizer of stochastics gradient with momentum method in DCNN model. Overall accuracy for the proposed model delivers 99.88%.

The research [17] proposed the new model for classifying the groundnut crops leaf as healthy and unhealthy with CNN algorithm with the image compression techniques as DCT, DFT and DWT. The results show that the overall computational time was reduced by comparing the CNN with ResNet50 architecture.

The author [18] created the method for precise detection and classification of groundnut leaf diseases, the method which is proposed by the author, algorithms named H2K which combines the strengths of the three-concept called Harris corner detector, the HOG, and the K-Nearest Neighbor classifier.

In [19] the researcher developed the system for the detection and classification of nutrient deficiency in groundnut leaf especially on nitrogen level. DCNN was implemented for the dataset to achieve the high accuracy and also achieved 95% for training, 92% for validation result.

III. Groundnut Datasets

In the section we are going to describe about the groundnut dataset where we have collected from Pudukkottai district, Tamil Nadu, India. Here the detailed description of groundnut leaves and total number of leaves collected with various diseases occurred in the leaf with healthy leaf as on category. [20] explains about the disease-free plant growth which produces more productivity with ELM algorithm with normalization using the benchmark dataset and obtains the optimal learning and better generalization. The dataset is collected with the five types of diseases including healthy leaves to identify whether the leaves are affected by disease or not. In this paper, the commonly affected diseases are selected for the identification of the leaf diseases process. Because these are the diseases that were affected in the area where the dataset was collected. Details and stages of leaf disease that occurred in groundnut crops are briefed as follows;

A. Dataset Description

First Early Leaf spot is a common disease that reduces yield significantly and can be found anywhere the groundnut is grown. This encompasses countries and regions such as the United States of America, Australia, Fiji, the Solomon Islands, Tonga and in Indian states such as Gujarat, Andra Pradesh, Karnataka, Tamil Nadu, and Maharashtra. It is estimated that the leaf spot causes a reduction in yields of at least fifty percent in countries located in the Pacific. Approximately one month after sowing, the first infection of early leaf spot leaf disease will take place. The symptoms manifest themselves on the leaf as a reddish-brown large spot that is not perfectly circular and is surrounded by a yellow halo. The lesion on the lower surface is a light brown color.

Secondly, we took Late Leaf Spot types for the identification of diseases. This type will have infections at the beginning approximately seven weeks after sowing and appear as mostly circular dark brown small spots without a yellow halo. Carbon black color will display in lower surface lesions caused by late leaf spots.

Thirdly, Plants that have been infected with rosette will have the symptoms of a solid clump or gnome shoots, and each will have a tuft of small leaves arranged in a rosette pattern. Chlorosis and mosaic mottling are characteristics of the plant. The infected plants will continue to be underdeveloped and produce the flowers, with the immature flower only a small percentage of the groundnut pegs will mature into nuts, and there will be no seeds produced.

Last, Rust disease infects every part of the plant that is above ground. In most cases, the diseases were discovered at the time of plants at six weeks old approximately. On the backside of the groundnut leaves, tiny eruptions that range in color from brown to dusty chestnut are known as uredosori.

TABLE I. Types of Groundnut Diseases

Sl. No.	Disease Selected	Symptoms	Duration of disease attacks
1.	Early Leaf Spot	The reddish-brown spot upon the leaf is surrounded by yellow halo nature.	1 month from sowing the seed.
2.	Late Leaf Spot	Carbon black color that will occur under the surface of the leas.	7 weeks after sowing the seed.
3.	Rosette	Yellow color leaf with mottling of the foliage	2 or 3 weeks from sowing.
4.	Rust	Tiny pustules that range in color from brown to dusty	After completing 6 weeks
5.	Healthy	doesn’t change its natural green color.	Same as the nature of groundnut leaf color as green

The Table I shows about the groundnut leves diseases name with symptoms and describes about the time when the diseases will attack the leaf.

B. Early Leaf Spot (ELS)

- Fig. 1. shows the sample of early leaf spot diseases in groundnut crops are fungal diseases caused by the fungus Cercospora arachidicola. This disease can cause extensive leaf spots and defoliation in groundnut crops. About one month after sowing, this disease begins to affect the crop.

Fig. 1. Early leaf spot

- The disease affects all above-ground plant parts, especially leaves. The two infections' leaf symptoms differ in appearance, spot color, and shape. Both fungi damage the petiole, stem, and pegs. As infection spreads, lesions from both species merge and spotted leaves drop early. Severe infections impair nut quality and production.

C. Late Leaf Spot (LLS)

- Fig. 2. represents the LLS fungal disease that affects groundnut crops. This disease is triggered by the fungus Phaeoisariopsis personata and is characterized by the thick brown or black dots on the leaves. The spots can vary in size and shape, and they may be surrounded by a yellow halo. The disease can cause the leaves to drop off the plant, which can reduce the plant's ability to produce nuts.

Fig. 2. Late leaf spot

- The binomial name of the LLS is phaeoisariopsis personatum. Mycelium and haustoria are produced by the fungus. The symptoms of LLS are defined as Conidia are cylindrical or obclavate, short, hyaline to olive brown, usually straight or curved slightly with 1-9 septa, not constricted but mostly 3-4 septate. The favorable conditions for this disease to affect the leaf which are in high humidity for three days, if the temperature is as low as 20 degrees Celsius with droplets on the lower surface, heavy dosage of fertilizers like nitrogen, phosphorus, and magnesium deficiency in soil.

D. Rosette Diseased Leaf (RoDL)

- Fig. 3. depicted the RoDL are affected with the viruses types and lead with stunted growth and reduced yields in groundnut crops. These diseases are spread by aphids and can be controlled by using insecticides. The affected plants appear as compressed clumps with tufts of small leaves appeared on the leaves. Chlorosis and mosaic mottling can be seen on the plant.

Fig. 3. Rosette leaf

- The affected plants leaf won’t develop fully as they needed for the conversion of nuts from flowers, if they convert also, it won’t be good seed to harvest.

E. Rust Diseased Leaf (RuDL)

- The Fig. 4. shows the rust diseases in groundnut crops are caused by a fungus that affects the leaves of the plant. This can cause the leaves to turn yellow and eventually drop off. The fungus can also cause the nuts to become discolored and shrivelled. The scientific name of the leaf disease called rust is Puccinia arachidis.

Fig. 4. Rust leaf

- The disease attacks the plant's entire aerial structure. Rust leaf disease is discovered when the groundnut crops are grown till the week of 6, then this disease will get affect and won’t get good maturity of the pods. On the lower surface of the leaves, little brown to uredosori will appear. The above disease can only infect the leaf if the following circumstances are met: the humidity value above 95% high, significant rainfall, and a low temperature (20-25 degrees Celsius).

F. Healthy Leaf (HL)

- A healthy leaf in a groundnut crop is one that is free from disease, pests, and deformities. It should be a deep green color and be able to photosynthesize efficiently. The below Fig. 5. represents the healthy leaf of groundnut crop. The peanut, scientifically known as arachis hypogaea, is an herbaceous annual plant that belongs to the Fabaceae family and is cultivated for its oil and edible nuts.

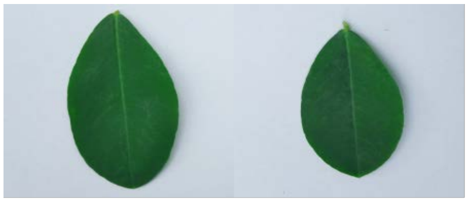


Fig. 5. Healthy groundnut leaf

- Plants that produce peanuts are typically rather tiny and have slender stems and leaves that resemble feathers. The leaves are attached to the stalk in a manner that resembles a leaf and is organized in alternate pairs. The peanut plant can produce yellow, orange, cream, or white flowers. These blooms give rise to 'pegs', which are distinctive floral structures that are pushed down into the soil to facilitate the growth of pods.

IV. Flow Process of Dataset

The Fig. 6. provides a visual representation of the comprehensive real-time detection process. The first step in developing the GLDD is to collect images of both diseased and healthy groundnuts from an actual groundnut field. The initial GLDD is then subjected to several data augmentation procedures, during which it is manually annotated and expanded. The dataset was categorized into is split into two parts: healthy and unhealthy. In these two classes we have specified for training as 80% and testing as 20%. The findings of the detection include information on healthy and unhealthy classes. Groundnut Dataset Images were collected with the VIVO 2029 smartphone camera having 13 Mpx resolution to identify the groundnut crop leaves which is affected by diseases or not. This paper contains two categories healthy and diseased. In the diseased category we collected four types of leaves which are as follows with the number of leaves collected.

A. Dataset

Most of the dataset used for model building are taken from publicly available dataset. The study [21] also uses both the self-collected dataset and PlantVillage dataset for pepper leaf disease detection. Groundnut leaves are collected from the various fields located near the village of Pudukkottai district and also from various websites through google search engine. We have collected the five categories of leaves which is segregated with healthy and unhealthy leaves of groundnut crop. Groundnut leaf disease patterns change with the season and with other elements like humidity, temperature, a lack of NPK, and wetness.

Table II details about the dataset with 619 images of unhealthy and healthy groundnut leaves obtained, which correspond to five categories: Healthy Groundnut Leaf, ELS, LLS, rosette, and rust. These are the five leaf diseases which occurred commonly on the groundnut leaves that were chosen for two reasons. First, these common diseases can be seen with the help of images which were collected by camera itself. Furthermore, they are responsible for its significant yield decrease in groundnut crop cultivation. The above five categories of groundnut leaves are the most common diseases affected in the fields where we collected the images.

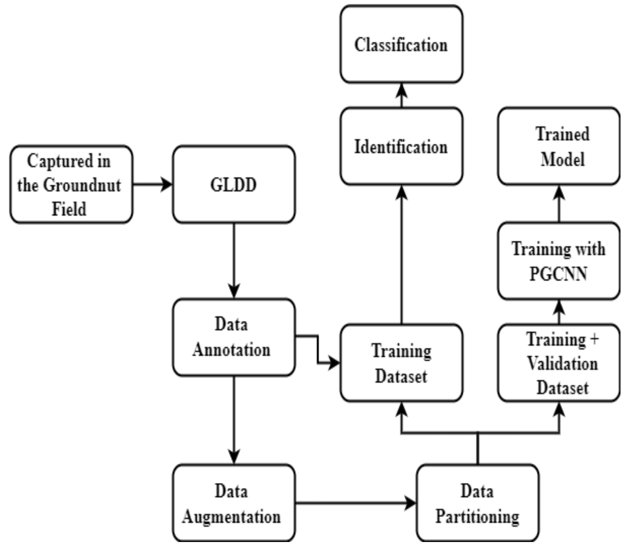


Fig. 6. Process flow of real-time detection of groundnut leaf diseases

TABLE II. Description of Groundnut Dataset

Sl. No	Name of the Diseases	Total No. of Images
1.	Healthy	190
2.	Unhealthy	429
Total		**619**

B. Progressive Groundnut Convolutional Neural Netwrok (PGCNN)

1) Input layer: In this paper, the image size of 1380x1380x3 is given to the input layer of our model to detect whether the image is healthy or diseased. Then the images will get rescaled to 224x224x3 during the model training and testing to reduce the computational time. Here we are feeding single images of four different types of diseases with one healthy leaf category of images with no background to convolve the images with filters to identify the patterns of leaf images.

2) Convolution Layer (Convnet + ReLU)

- From various types of convolution layers available in Keras API, we selected the Conv2D layer for this PGCNN model development because Conv2D layers are implemented over the spatial convolution over images. Arguments used for the Conv2D method are filters, kernel size, strides, padding, activation, and kernel initializers.
- Rescaled images of input image size 224x224x3 are convolved with 64 filters of size 3x3 kernel with strides value of (2, 2). The activation function used to take the value of feature map with "ReLU" as activation function for positive and negative values. While in convolution operation, various filters are used to detect the edges, patterns, etc., for extracting the features from the input images.
- In this paper we have used five conv2D layer to get good feature map. Operation performed by the

convolution layer uses the following equation in default, which had number of filters used in the convolution layer, filter size, padding and stride to convolve through the input image.

$$C_{out} = \frac{C_{in}+2P-F}{S} + 1 \quad (1)$$

- C_{out} represents the values which we can feed into next layer, which the input of image size 224x224x3 had been fed into the layer. C_{in} insists the size of the input image in convolution layer, P represents padding, F denotes the number of filters used in the convolution layer.

3) Pooling: The process of reducing the spatial dimensions (i.e., height and width of the input image) for making the computation faster by decreasing the training parameters. The pool size of the MaxPool2D is taken as (2, 2). There are 3 kinds of pooling to down-samples the input in pooling operations, they are Max Pooling, Average pooling, and Global Pooling. In this research paper, we use maxpooling2D operation to down-samples the input size of the groundnut leaf images. Max pooling chooses the most element from the feature map's filtered region. After max-pooling, the output is a feature map with the most prominent features of the preceding map.

4) Fully Connected Layer (FC): The final Pooling or Convolutional Layers output is flattened to convert the matrix form into one-dimension vector and those are fed into the input of the fully connected layer. Fig. 7 explains fully connected layer is the layer which applies the linear transformation and has the input connection to every neuron into the output layer of another layer.

$$Z_{ij}(y) = f\ \sum_{k=1}^{n_H} w_{ij} y_k + w_{io} \quad (2)$$

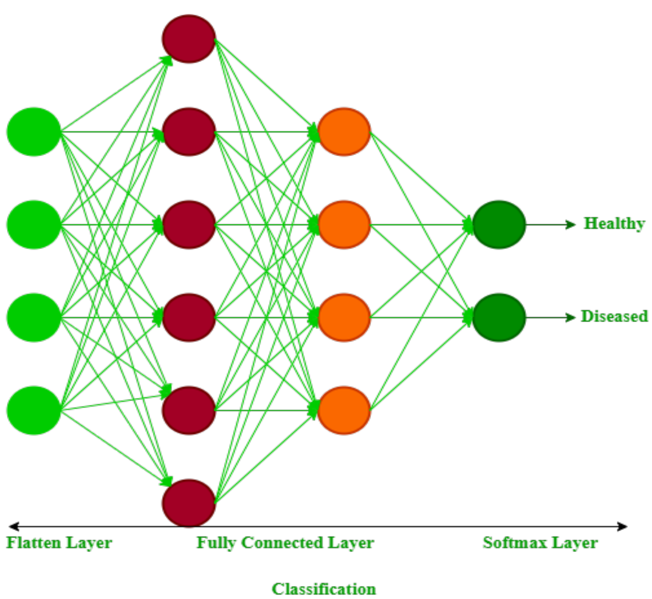


Fig. 7. FC for classification of groundnut leaf

The fully connected input layer (i.e., called flatten) is the conversion process from 2-dimensional array into a single dimensions array or linear vector. The output of the single vector is fed as the input for fully connected layer.

5) Dense (or) output layer: Leaky ReLU, Sigmoid, Tanh, ReLU, Softmax, and other activation functions can be used in the final layer for the output layer based on the research problem. This paper focuses primarily on the healthy or diseased category. Because we developed the model for binary classification, the sigmoid is the activation function used at the final layer of the Progressive Groundnut Convolutional Neural Network (PGCNN) to determine whether the leaf is healthy or diseased.

6) Dropout layer: This is one of the main characteristics of CNN. This layer is used to set the input value as 0 to skip some of the unwanted neurons to continue to next layer and also it prevents overfitting problem. Here we used 30% of neurons to be dropped out for fed the neurons into next level.

V. Implementation Details

A. Groundnut Leaf Diseases Identification using PGCNN

Leaf diseases is a significant threat to agriculture, but we are facing the difficulty for the identification and classification of leaf diseases due to the lack of infrastructure development. To minimize the infrastructure cost and maximize the yield of groundnut cultivation artificial intelligence technology has evolved into farming methods also. DL is one of the subsection of Artificial Intelligence (AI), is causing a revolution in agriculture by replacing traditional methods with more efficient methods that assist farmers in the identification of leaf diseases with convolutional neural networks. The proposed PGCNN model experiments the leaf disease identification with two classes called healthy and diseased with five kinds of validation of the self-collected dataset as;

TABLE III. Validation Test with Various Ratio Percentages of Self-Collected Dataset

Sl. No.	Dataset Validation %	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1.	90-10%	98.74	03.99	91.94	0.2692
2.	**80-20%**	**99.19**	**03.23**	**91.94**	**0.3227**
3.	75-25%	98.70	03.49	92.90	0.2855
4.	70-30%	97.69	05.72	93.01	0.1745
5.	60-40%	97.30	08.77	90.32	0.4919

The Table III and Fig. 8 shows the various ranges of dataset splitting for training and testing (Validation) accuracy for an imbalanced dataset of groundnut leaf images. After plotting the table, we conclude that the dataset separation with ranges of 80% training and 20% testing works well and achieves a high of 99.19% training accuracy and 91.94% of testing accuracy than the other splitting.

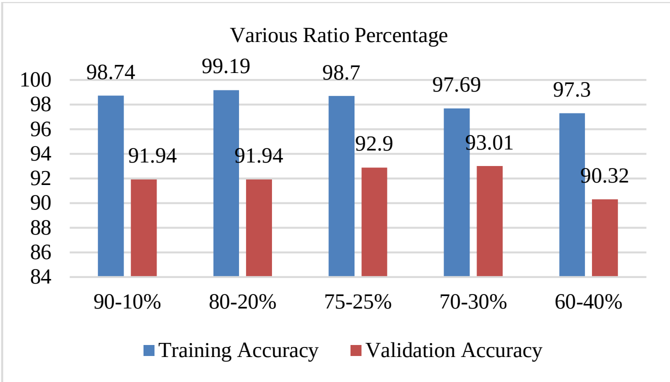


Fig. 8. Categories of ratio percentage

B. *Optimizers of PGCNN*

We have implemented five types of optimizers to reduce the loss value for our dataset which is self-prepared and validated with an 80% - 20% proportion. RMSprop, Adam, Adamax, Adagrad, and SGD were selected because these are similar to Adam optimizers with momentum.

TABLE IV. VARIOUS TYPES OF OPTIMIZERS TESTED IN PGCNN MODEL

Sl. No.	Optimizers	Training Accuracy	Validation Accuracy
1.	RMSprop	97.37	95.97
2.	**Adam**	**98.17**	**97.58**
3.	Adamax	96.56	91.13
4.	Adagrad	71.86	68.55
5.	SGD	78.34	76.61

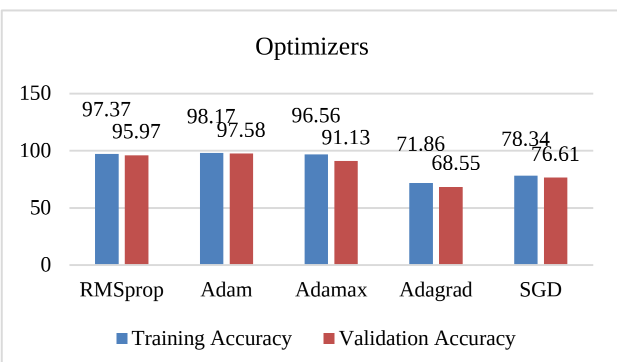


Fig. 9. Various optimizers

From the above Table V and Fig. 9 we achieved high accuracy in the Adam optimizer for training and validation with all the optimizers. The algorithm used for optimizing the PGCNN model is Adam achieves the good results in faster manner. Adam is named as adaptive moment estimation to converge the model in smooth way. This optimizer utilizes the two optimizer benefits, they are AdaGrad and RMSProp.

C. *PGCNN Model Flow Diagram*

The Fig. 10 shows the flow diagram of PGCNN model with sequence of layers which executes the model to achieve good performance. First the input image with the size of 224x224x3 has been fed into the classifier called sequential to arrange the layers stacked one by one to form a entire network.

There are five convolutional layers constructed to convolve the image to extract the feature map as 16, 32, 64, 128 etc. First Conv2D layer with 64 filter to specify the depth of the filters and 3x3 dimensions of kernel/filter represents the height and depth of the matrix with padding same value, following the activation function as "ReLU" of non-linear function. Then to reduce the dimensionality of images we apply max_pooling of 2x2 function. This reduces the size of the images with half the value from the previous layer input. By this way we are repeating five times of the convolution operation with filter of 64 on 3x3 size kernel by applying "relu" activation function with the conclusion of convolution is max_pooling layer.

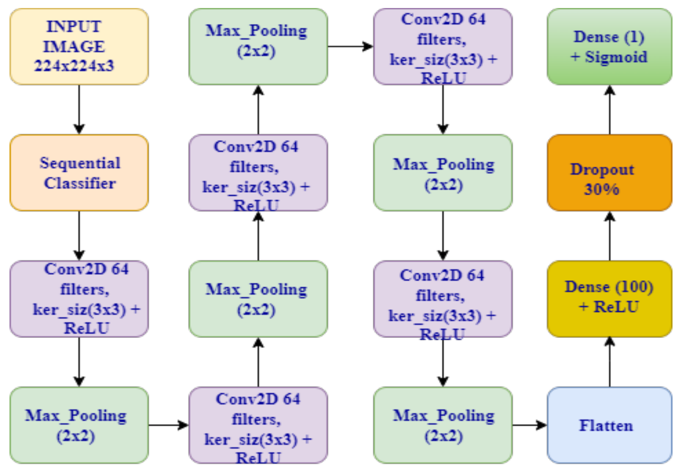


Fig. 10. PGCNN model flow diagram

TABLE V. VARIOUS RATIO OF PGCNN MODEL FOR TRAINING AND VALIDATION ACCURACY

Sl. No.	Sample Images	Training Ratio (%)	Validation Ratio (%)	Training Accuracy	Validation Accuracy
1	619	90	10	0.9910	0.9677
2		**80**	**20**	**0.9939**	**0.9758**
3		70	30	0.9954	0.9140
4		60	40	1.0000	0.8992

After completing the convolution operation we are applying the flatten() class to convert the two dimension image into one dimensional vector. With 100 nodes of neurons we are having the dense layer as the previous layer of output layer to spread the probability of values from the fully connected layer by the activation function as "ReLU". To improve the performance of PGCNN model and reduce the overfitting we added dropout function with 30% neurons to be dropped out randomly by selecting the important neuron and some of the neurons are ignored during training process. Dropout() is one of the type of regularization in deep learning models.

Finally, we concluded with dense layer with single neuron for binary classification and the activation function of "sigmoid" is used for the classification of PGCNN model as healthy or unhealthy images. We used sigmoid activation function because this function supports the binary classification for PGCNN model.

VI. RESULTS

The model creation and executions were performed in TensorFlow framework. Dataset has been prepared by self-collected with the samples of total 619 groundnut leaves as two categories such as Unhealthy and Healthy options. The leaves were collected from the village near by Pudukkottai town, Tamilnadu at different times with different groundnut crops.

The Table VI, Fig.11 and Fig. 12 explains about the results for the various ratios for the proposed PGCNN model for the groundnut leaves disease identification and classification of healthy and unhealthy category. We have taken four different ratios as 90-10%, 80-20%, 70-30% and 60-40% respectively. From these compilations we have achieved the good training accuracy and validation accuracy in 80-20% ratio. So, we freeze that ratio for further evaluation for PGCNN model. With the minimum number of sample images, the proposed model achieved 99.39% as the accuracy of training the model and 97.58% as the accuracy of validation of PGCNN model. The 90-10% ratio also achieves good training accuracy as 99.10% and validation accuracy as 96.77% with the training loss of 01.04%, validation loss as 18.79%. Next ratio tested with 70-30%, it achieves the training accuracy of 99.54% and validation accuracy of 91.40% with 01.11% of training loss and 90.05% of validation loss. At last we have tested with the ration of 60-40% which achieves the training accuracy of 100% and 89.92% of validation accuracy with the training loss of 00.41% and 39.59% validation loss.

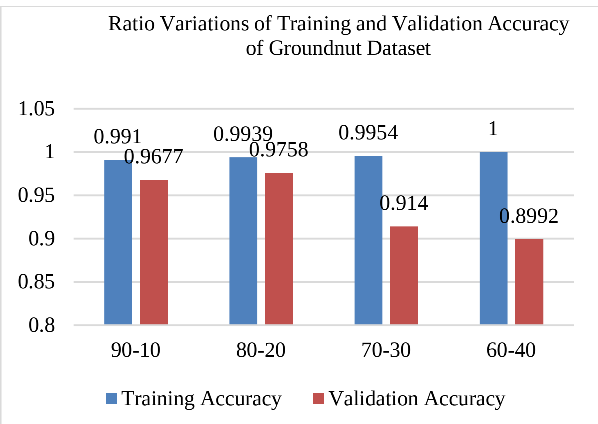


Fig. 11. Model training and validation accuracy graph

TABLE VI. VARIOUS RATIO OF PGCNN MODEL FOR TRAINING AND VALIDATION LOSS

Sl. No	Sample Images	Training Ratio (%)	Validation Ratio (%)	Training Loss	Validation Loss
1.	619	90	10	0.0194	0.1979
2.		**80**	**20**	**0.0148**	**0.1806**
3.		70	30	0.0111	0.9005
4.		60	40	0.0041	0.3959

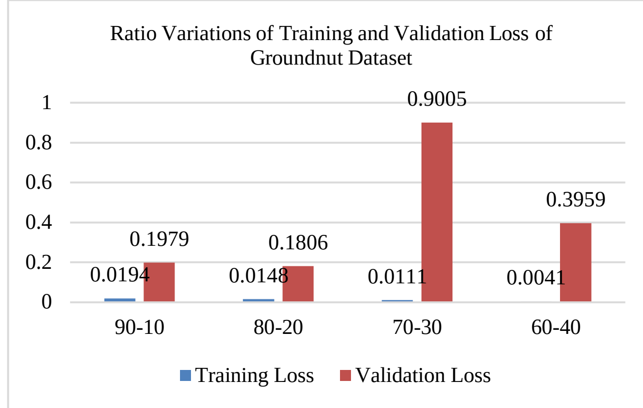


Fig. 12. Model training and validation loss

From the above details we have frozen the ratio for our PGCNN model as 80-20% for further evaluation of predicting the image as healthy or unhealthy. Because it ended up with good accuracy on both training and validation with less training and validation loss when compared to the other ratio results.

The Fig. 13. explains that the identification of groundnut leaves as healthy or unhealthy by the classes we defined in the dataset, it classifies the category by class names if it falls on healthy or unhealthy classes. And also it shows that the label defined as 1 for unhealthy class and 0 for healthy class.

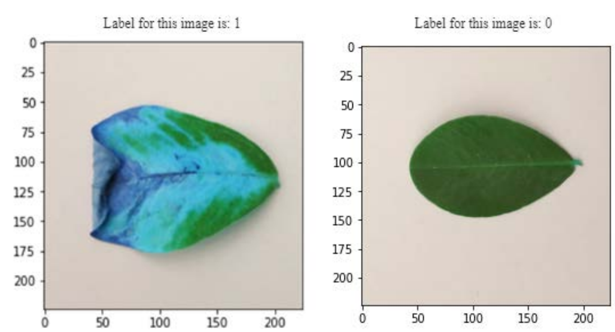
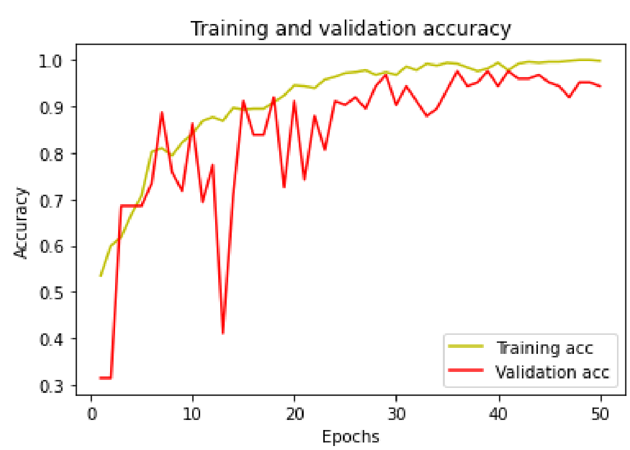
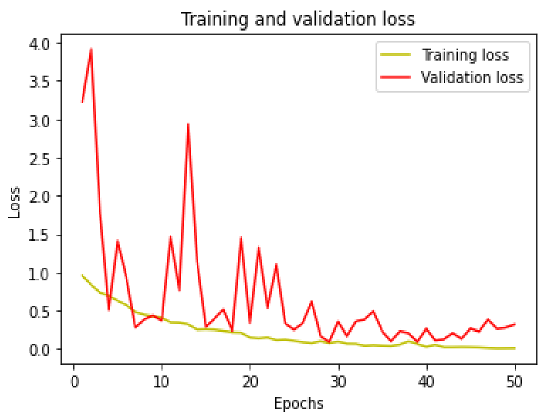


Fig. 13. Identification and classification of groundnut leaf as unhealthy or healthy



(a)



(b)

Fig. 14. (a). PGCNN model training and validation accuracy, (b). PGCNN model training and validation loss

The Fig. 14(a) and Fig. 14(b) shows that the graph for the proposed model of PGCNN training, validation of accuracy and loss against the groundnut dataset which was collected from the field located nearby the Pudukkottai district. The curve indicates how the model fits into the accuracy of both training and validation data. The training loss indicates that the training data learns well and it decreases over time to fits the curve as good fit. The validation loss jumps up and down to fit into the curve for model that learns the new data to fit the curve in good fit. The above model learns all data in a good manner to fit the curve of loss that decreases in certain time. Likewise, the training and validation accuracy curve also fits in good manner by gradual increase in the curve.

TABLE VII. COMPARATIVE ANALYSIS

Sl. No	CNN Models	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	**PGCNN**	**0.9939**	**0.0194**	**0.9758**	**0.1806**
2	AlexNet	0.9737	0.1039	0.7419	0.1148
3	VGG11	0.6949	0.6160	0.6855	0.6229
4	VGG13	0.6949	0.6159	0.6855	0.6228
5	VGG16	0.6949	0.6161	0.6229	0.6855
6	VGG19	0.6949	0.6161	0.6855	0.6229

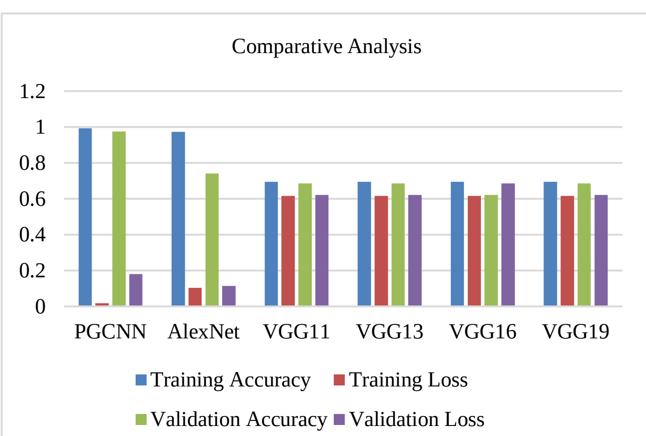


Fig. 15. Various model comparison on proposed PGCNN model

The Table VII and Fig. 15 shows that the comparative analysis of six CNN architecture including the proposed PGCNN model. It shows that PGCNN gives good accuracy and reduced loss for training and validation process. Here we have taken AlexNet, VGG11, 13, 16 and 19 models. The reason behind the model selection like VGG series is that we have built the model based on this architecture. By comparing the obtained results of proposed model with the pretrained models like VGG architectures. The proposed model has achieved the good results as compared to VGG model architecture.

VII. DISCUSSION

This research paper discuses about the groundnut leaf diseases identification and classification as whether the leaf is healthy or unhealthy. For experimenting this dataset with appropriate base model we have gone through various paper as related works to concluded our model. With the use of literature review we confirmed to develop PGCNN model which we used CNN architecture as our baseline to implement the model. After implementing the model, we have compared with the pretrained CNN architecture for performance comparisons. We achieved the good performance accuracy compared to the pretrained models which has limited layers. The proposed model achieved the training accuracy of 99.39% and validation accuracy of 97.58%. Such that the training loss as 0.01% and validation loss as 0.18% respectively. These results achieved with 619 total images for the dataset which was collected from the fields. In future we are planning to collect more images for large dataset and also, we will be preparing our dataset to make as a benchmark dataset for publicly available for the researcher to use and explore it. In next section we concluded our research work that have done with the dataset by PGCNN model.

VIII. CONCLUSION

This paper was implemented with the dataset which was collected from the real-life scenario of field located nearby Pudukkottai, Tamil Nadu, India. Groundnut dataset has two category of groundnut leaves namely healthy and unhealthy for identification of diseases from the leaves. The detailed description of the groundnut dataset has been listed under the Section IV. Here we focused to identify and classify the dataset with two classes. Five different CNN architecture were designed and executed to compare the proposed model whether the PGCNN model performs well when compared to the pretrained architectures like VGG11, 13,16, etc. we used performance metrics as accuracy for all the architectures, in that PGCNN achieves good accuracy for both training and validation process. The proposed PGCNN training accuracy was 99.39% and validation accuracy was 97.58%. The model loss also reduced as compared to other models as training loss 0.0194 and validation loss as 0.1806.

In future we will be focusing the dataset on multi-class classification to classify the leaves by their diseased names and incorporate it into the prediction-based solutions. We will be focusing on more data to be collected on various seasons and measures the performance.

REFERENCES

[1] R. M. Rakholia, J. H. Tailor, J. R. Saini, J. Kaur, and H. Pahuja, "Groundnuts Leaf Disease Recognition using Neural Network with Progressive Resizing," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 6, pp. 83–88, 2022, doi: 10.14569/IJACSA.2022.0130611.

[2] S. S. Koshy, V. S. Sunnam, P. Rajgarhia, K. Chinnusamy, D. P. Ravulapalli, and S. Chunduri, "Application of the internet of things (IoT) for smart farming: a case study on groundnut and castor pest and disease forewarning," CSI Trans. ICT, vol. 6, no. 3–4, pp. 311–318, 2018, doi: 10.1007/s40012-018-0213-0.

[3] M. Ramakrishnan and A. N. A. Sahaya, "Groundnut leaf disease detection and classification by using back probagation algorithm," 2015 Int. Conf. Commun. Signal Process. ICCSP 2015, no. 7092512506, pp. 964–968, 2015, doi: 10.1109/ICCSP.2015.7322641.

[4] M. Lv, G. Zhou, M. He, A. Chen, W. Zhang, and Y. Hu, "Maize Leaf Disease Identification Based on Feature Enhancement and DMS-Robust Alexnet," IEEE Access, vol. 8, pp. 57952–57966, 2020, doi: 10.1109/ACCESS.2020.2982443.

[5] P. Jiang, Y. Chen, B. Liu, D. He, and C. Liang, "Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks," IEEE Access, vol. 7, pp. 59069–59080, 2019, doi: 10.1109/ACCESS.2019.2914929.

[6] B. Liu, C. Tan, S. Li, J. He, and H. Wang, "A Data Augmentation Method Based on Generative Adversarial Networks for Grape Leaf Disease Identification," IEEE Access, vol. 8, pp. 102188–102198, 2020, doi: 10.1109/ACCESS.2020.2998839.

[7] C. Zhou, Z. Zhang, S. Zhou, J. Xing, Q. Wu, and J. Song, "Grape leaf spot identification under limited samples by fine grained-GAN," IEEE Access, vol. 9, pp. 100480–100489, 2021, doi: 10.1109/ACCESS.2021.3097050.

[8] S. B. Jadhav, V. R. Udupi, and S. B. Patil, "Identification of plant diseases using convolutional neural networks," Int. J. Inf. Technol., vol. 13, no. 6, pp. 2461–2470, 2021, doi: 10.1007/s41870-020-00437-5.

[9] C. Zhou, S. Zhou, J. Xing, and J. Song, "Tomato Leaf Disease Identification by Restructured Deep Residual Dense Network," IEEE Access, vol. 9, pp. 28822–28831, 2021, doi: 10.1109/ACCESS.2021.3058947.

[10] Q. Wu, Y. Chen, and J. Meng, "Dcgan-based data augmentation for tomato leaf disease identification," IEEE Access, vol. 8, pp. 98716–98728, 2020, doi: 10.1109/ACCESS.2020.2997001.

[11] R. R. Atole and D. Park, "A Multiclass Deep Convolutional Neural Network Classifier for Detection of Common Rice Plant Anomalies," no. February, 2018, doi: 10.14569/IJACSA.2018.090109.

[12] P. Kaur, S. Harnal, V. Gautam, M. P. Singh, and S. P. Singh, "A novel transfer deep learning method for detection and classification of plant leaf disease," J. Ambient Intell. Humaniz. Comput., no. 0123456789, 2022, doi: 10.1007/s12652-022-04331-9.

[13] Suresh and K. Seetharaman, "Real-time automatic detection and classification of groundnut leaf disease using hybrid machine learning techniques," Multimed. Tools Appl., 2022, doi: 10.1007/s11042-022-12893-1.

[14] H. Amin, A. Darwish, A. E. Hassanien, and M. Soliman, "End-to-End Deep Learning Model for Corn Leaf Disease Classification," IEEE Access, vol. 10, pp. 31103–31115, 2022, doi: 10.1109/ACCESS.2022.3159678.

[15] R. Thangaraj, S. Anandamurugan, and V. K. Kaliappan, "Automated tomato leaf disease classification using transfer learning-based deep convolution neural network," Journal of Plant Diseases and Protection, vol. 128, no. 1. pp. 73–86, 2021. doi: 10.1007/s41348-020-00403-0.

[16] M. P. Vaishnnave, K. Suganya Devi, and P. Ganeshkumar, "Automatic method for classification of groundnut diseases using deep convolutional neural network," Soft Comput., vol. 24, no. 21, pp. 16347–16360, 2020, doi: 10.1007/s00500-020-04946-0.

[17] S. Muthukumaran, P. Geetha, E. Ramaraj, and C. Head, "Leaf Spot Disease Image Classification for Groundnut Crop using Deep Convolutional Neural Network," Turkish J. Comput. Math. Educ., vol. 12, no. 14, pp. 3166–3179, 2021.

[18] K. Suganya Devi, P. Srinivasan, and S. Bandhopadhyay, "H2K – A robust and optimum approach for detection and classification of groundnut leaf diseases," Comput. Electron. Agric., vol. 178, no. October, p. 105749, 2020, doi: 10.1016/j.compag.2020.105749.

[19] M. Janani and R. Jebakumar, "Detection and classification of groundnut leaf nutrient level extraction in RGB images," Adv. Eng. Softw., vol. 175, no. October 2022, p. 103320, 2023, doi: 10.1016/j.advengsoft.2022.103320.

[20] R. Dwivedi, T. Dutta, and Y. C. Hu, "A Leaf Disease Detection Mechanism Based on L1-Norm Minimization Extreme Learning Machine," IEEE Geosci. Remote Sens. Lett., vol. 19, 2022, doi: 10.1109/LGRS.2021.3110287.

[21] M. Ahmad, M. Abdullah, H. Moon, and D. Han, "Plant Disease Detection in Imbalanced Datasets Using Efficient Convolutional Neural Networks with Stepwise Transfer Learning," IEEE Access, vol. 9, pp. 140565–140580, 2021, doi: 10.1109/ACCESS.2021.3119655.