

# Plant Health Prediction and Monitoring Based on convolution Neural Network in North-East India

**Bhagwan Sahay Meena**

Department of Computer Science  
Assam University, Silchar  
Silchar, Cachar, India  
bsmeena.g1@gmail.com

**Abstract**— Agriculture is the main backbone of any society. In this modern era as the population continuously increases, resources for farming are also decreasing. If the health condition of the plant's determined at a regular interval, farmers can take action timely to improve the plant's health condition. Plant monitoring and plant health status at regular intervals may lead to better growth of foods. But the regular physical visit to the crop field to monitor plants and plant health is a critical task for a large crop field. To overcome this situation, we need to shift from a traditional cropping system to smart agriculture. Now, these days, a smart agriculture-based approach can use the internet of things, machine learning, and deep learning to predict the health condition of plants. In this paper, the internet of things (IoT) based smart agriculture system has been presented along with machine learning, deep learning, and transfer learning to monitor and health prediction of plants. The IoT-based system has been used to monitor the plant's surrounding parameters such as humidity, temperature, light intensity, and soil water moisture. The leaf images of plants have been used in deep learning (DL), machine learning (ML), and transfer learning (TL) to predict the health condition of plants. In this paper, the convolution neural network (CNN) based model has been proposed and it has been compared with the support vector machine (SVM), random forest (RF), VGG19, and mobilenet model. It has been concluded that the accuracy of the proposed CNN model is 81.5 %, which is the highest among SVM, RF, VGG19, and mobilenet..

**Keywords**- : Internet of things; Support Vector Machine; Random forest; Convolution Neural Network; VGG19; MobileNet

## I. INTRODUCTION

Agriculture is an art and a process of producing grains through a traditional cropping system. India is mainly an agriculturally based country. Agriculture plays a vital role in the Indian economy as it is the main source of food. Scientific research and good farming techniques can help to achieve the green revolution and attain a self-reliant status in better food production.

**Problem Definition:** Plant monitoring and plant health status at regular intervals may lead to better growth of plants. But the physical visit to the crop field for monitoring plants and plant health is a critical task for a large crop field. Moreover, disease prediction along with health monitoring in plants using new technology such as deep learning, and machine learning is a research area.

**Objective:** The IoT, machine learning, deep learning, and CNN-based work has been proposed to monitor and disease prediction of plants under a single umbrella. In this paper, CNN-based plant health prediction system have been previewed where the plant's surrounding humidity, the surrounding temperature, surrounding light, and plants soil moisture are targeted as these are some important factors of plant's growth and real-time data of the above-mentioned parameter will be sent to an IoT platform i.e. thingspeak for remote monitoring of plants. The deep learning-based CNN model also has been

proposed in this study that can predict the health status of a plant's leaf. To accomplish plant health prediction, different supervised techniques such as support vector machine (SVM), random forest (RF), and CNN model have been used in this paper. And it was also tested with VGG19 architecture and MobileNet architecture to predict the plant's health.

## II. LITERATURE REVIEW

IoT devices may be used for various applications such as user localization [1,2] home automation and to control of electric appliances in the office/home as per user movement [3,4,5]. The user/device localization and home automation have been done using IoT devices such as passive infrared sensors, radio frequency identification, light-dependent register, Bluetooth low energy, and smartphone. The result shows that the passive infrared sensor along with radio frequency identification-based IoT localization gives better accuracy of 93% than light dependent register [2]. Smartphone-based localization gives better accuracy than Bluetooth low energy-based localization [4].

In [6], Geetha et al. has proposed a model to predict diseases in tomato leaves and also proposed a machine learning methodology that focuses on generating an advanced and efficient system that makes processes of creating high-yield tomatoes much easier for farmers. The author aimed to detect the most common diseases that generally occur in a tomato

namely early blight, bacterial spot, and curl using different image processing techniques.

Choudhary et al. [7], developed a deep convolutional neural network based on the recently developed Efficient CNN architecture, where the model was fine-tuned for healthy and different unhealthy classes of tomato leaves like bacterial infected and viral infected. The modified U-Net, Efficient Net B7, and Efficient Net B4 models have been used in binary classification, six class classification, and ten class classification of plant leaves to predict the diseases with an accuracy of 98.66 %, 99.95%, and 99.89% respectively.

Pavel et al. [8] have proposed an IoT device-based model which sends real-time environment data stored in the database along with the image of a plant's leaf that classifies various diseases using image processing and support vector machine techniques. The accuracy of the proposed system is 97%.

Patel et al. [9], has presented a IoT and machine learning based agriculture system that monitor humidity, ph, temperature, and soil moisture for precision farming and gives a recommendation of fertilizers using machine learning techniques. The proposed system gives an accuracy of 99.31%.

Doshi et al. [10], proposed an esp32 camera-based IoT monitoring system that monitors the temperature, humidity, soil moisture, visible light index, and UV index of sunlight using blynk application. Moreover, many sensors have been discussed that can be used in smart farming.

This paper [11], presented a detailed discussion of machine learning techniques has been discussed that can be used to detect plant diseases. The traditional and advanced deep learning methods that are associated with the data acquisition modalities, IoT, and satellite images are discussed and the last section also gives some focus on data fusion.

Disha et al. [12], discussed different issues in the traditional cropping system. They also discussed an IoT-based monitoring system that can be used to measure a plant's soil temperature and water consumption. They have also discussed different deep learning techniques such as CNN and recurrent neural network (RNN) that can be used to minimize agriculture problems.

Reshma et al. [13], highlighted different work in the field of agriculture along with new agriculture policies for traditional cropping systems that used artificial intelligence techniques and it also shows how machine learning and deep learning techniques can be combined with an IoT. It can minimize the current problem in agriculture by increasing the productivity of food, making maximum use of available land, and controlling the pest.

Dewangan has presented a smart agriculture system that comprises micro-controllers, sensors, and a consolidated water quality system [14]. In this work, the real-time data have been

monitored and the regression-based water toxicity detection system is used to detect the toxicity of the water.

In the literature review, first, the IoT-based existing works in the different domains have been discussed. Thereafter, the ML techniques, DL techniques, CNN-based techniques, and their various methods have been discussed to detect plant diseases. The IoT-based smart agriculture work was independent of ML, DL, and CNN-based agriculture work. The ML and DL-based architecture make data of plant leaves in six classes that are tested on SVM and RF algorithms with different parameters. The SVM gives a more accurate prediction than the RF algorithm. Thereafter, a CNN-based model has been proposed in this study. The CNN model has been tested on mobileNet and VGG19 architecture. The output layer of VGG19 and mobilenet cut down and set with 6 neurons as we have 6 classes. The procedure of cutting down the output layer of VGG19 and mobilenet has been discussed in section 3.4.

### III. PROPOSED METHODOLOGY

An IoT-based model has been proposed that monitors the plant's surrounding humidity, temperature, light intensity, and soil moisture/ soil health. The monitored data have been sent to IoT architecture. In this work, two types of plants are targeted namely spinal gourd and lady finger to monitor and health predictions. Samples of plants are being collected from the region of northeast India through field visits. Three types of techniques are used to design the health prediction model of plants namely ML [15], CNN [16], and TL (VGG19, MobileNet). Each technique has used many parameters that are being discussed in the next section.

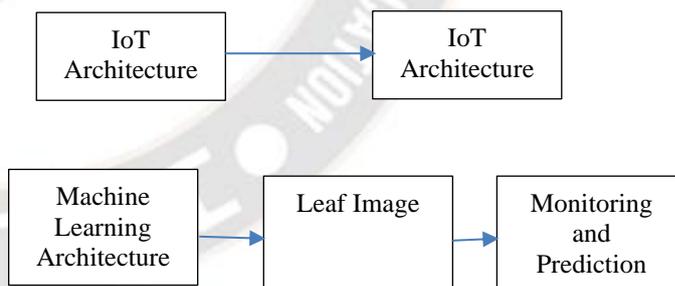


Figure 1. Proposed Architecture

As shown in figure 1, the proposed model consists of an IoT architecture and machine learning architecture. An IoT architecture deals with IoT data that are generated by IoT devices and the machine learning architecture deal with the leaf images of spinal gourd and lady's finger.

#### A. IoT Architecture

To build up this architecture, a system has been intended that comprises a composed way. As shown in figure 2, Three sensors namely DHT11, light-dependent register (LDR), and soil

moisture have been used in this work as IoT sensors. The sensors are chosen for both Surrounding and soil water moisture level estimation of plants. Sensor DHT11 measures the air temperature and humidity level of the surrounding. LDR Sensor is used to measure the surrounding light intensity. Additionally, the soil moisture sensor measures the water soil moisture level of plants. All these sensors send their data to Arduino and the

Arduino sends the received data to nodeMCU and the nodeMCU will send this data to thingspeak (raspberry pi).

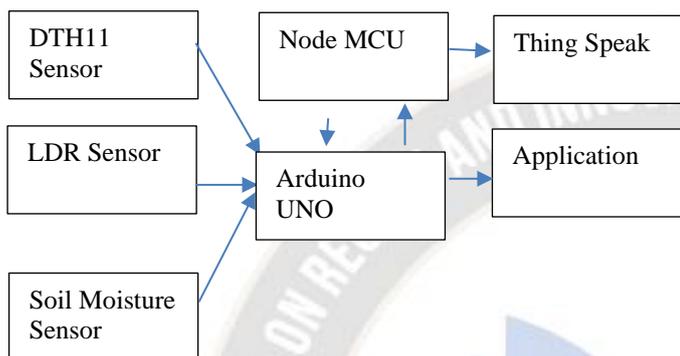


Figure 2. IoT Architecture

### B. Machine Learning (ML)

In crop diversification, for detecting spinal gourd and lady finger leaf health status, various samples of the mentioned leaves are collected from a local crop field via field visit. As shown in figure 3, Each image of leaf is resized into (224,224,3). At first, leaf images are converted into a pixel value in (1x150528) dimension each and then they are scaled into normalized form. While training the support vector classifier (SVC), it constructs hyperplanes in multi-dimensional space to separate the six classes. RF makes use of the many decision trees such that many base learners may come to a robust decision as compared to a single decision tree. In the algorithms, some hyperparameter tuning is done to get the best possible accuracy. SVM is best fitted with the kernel as RBF, gamma as scale, and C as 52. RF is hyper tuned with different n\_estimators with grid\_search\_CV. Each model is evaluated on some unused data. For training purposes, 80% of the data is used and the remaining 20% data is validated for testing.

### C. Convolution Neural Network (CNN)-

As the size of the data increases, the machine learning algorithm doesn't work well. To overcome such issues of machine learning, CNN-based model has been proposed in this study. CNN is specifically used to process pixel data with precise accuracy. The proposed architecture for CNN is given below in table 1. As shown in table 1, conv2d is performed in the input image with 32 filters having input shape of (224,224,3) with the activation function as relu, then max pooling operation is

performed with pool size (2,2) and Finally flattening into a 1-dimension array has been done. In the first and second hidden layers, 224 neurons are used with relu activation function. In between the second and third hidden layer, a dropout of 20% of neurons is done to avoid overfitting. In the third and fourth hidden layer, 128 and 32 neurons respectively are used with relu activation function along with the dropped out of 10% neurons during training. In the output layer, 6 neurons are used as the six possible outputs are there using softmax activation function. The model is trained and validated for 200 samples of leaves that are collected through physical visit of the crop field. whether leaves are healthy or infected of both plants (ladies' finger and spinal gourd) have been determined during the training and validation phase. To get better accuracy, the image augmentation technique is used during training that generates multiple augmented images for every single image. For the training purpose 80% of data is used and 20% for evaluating the model.

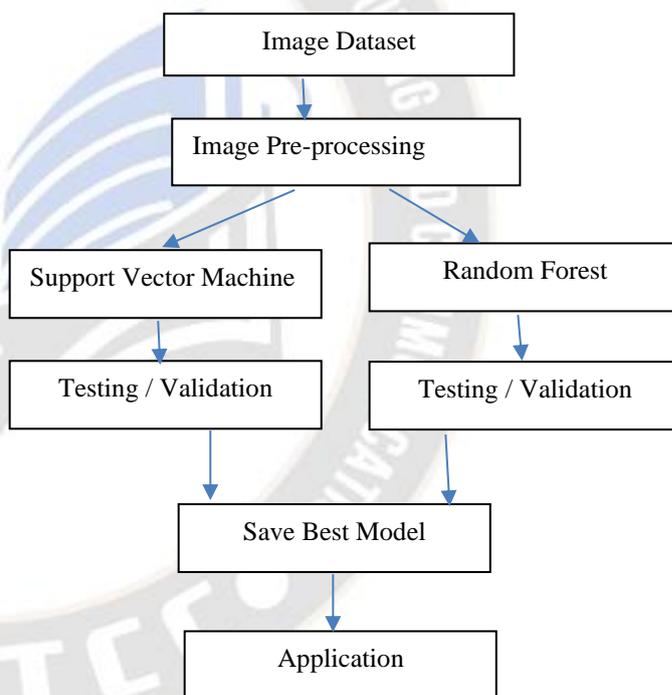


Figure 3. Machine Learning Architecture

### D. Model using Transfer Learning Techniques

In this proposed work, the leaf data have been divided into 6 classes as mentioned above. The VGG19 [17] and MobileNet [18] architecture is used for the classification problem of classes. VGG19 is a convolution neural network that has 19 layers. The pre-trained version of the network is used but the last layer is cut down and set with 6 neurons (we have 6 classes) with adamax as optimizer and loss function as Categorical\_crossentropy. MobileNet is also a prebuilt

convolution neural network such as VGG19 and consists of 53 layers. After cutting the output layer of MobileNet and adding an output layer of 6 neurons (as we have 6 classes) with adamax[19] as optimizer and loss function as categorical\_crossentropy, the images are trained with the pre-trained weight of imagenet. Here, 80% of the dataset is used for training and 20% dataset is used for testing.

#### IV. EXPERIMENTATIO AND ANALYSIS

At the outset, it is noted that Machine learning algorithms are outplayed to handle the classes like lightly infected and healthy of both plants. There is a high misclassification in predicting the two classes. Out of the two ML algorithms, SVM performed better after some parameter tuning than RF algorithm. The confusion matrix and the classification report of the Machine Learning algorithms are as below. The confusion matrix and classification report of SVM and RF are given in table 2 and table 3.

**Table 1.** This table gives the list of Parameters of the proposed CNN Model.

Model: "Sequential"		
Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 110, 110, 32)	2432
max_pooling2d (MaxPooling2D)	(None, 55, 55, 32)	0
flatten (Flatten)	(None, 96800)	0
dense (Dense)	(None, 224)	21683424
dense_1 (Dense)	(None, 224)	50400
dropout (Dropout)	(None, 224)	0
dropout_2 (Dense)	(None, 128)	28800
dropout_1 (Dropout)	(None, 128)	0
dense_3 (Dense)	(None, 32)	4128
dense_4 (Dense)	(None, 6)	198
Total params: 21, 7691 382		
Trainable params: 21, 769, 382		
Non-trainable params: 0		

**Table 2.** This table contain classification report of SVM

	precision	recall	f1-score	support
0	0.55	0.67	0.6	9
1	0.6	0.75	0.67	4
2	0.71	0.42	0.53	12
3	0.67	0.67	0.67	6
4	0.88	0.7	0.78	10
5	0	0	0	0
accuracy			0.61	41
macro avg	0.57	0.53	0.54	41
weighted avg	0.7	0.61	0.64	41

**Table 3.** This table contain classification report of RF.

	precision	recall	f1-score	support
0	0.82	0.6	0.69	15
1	0	0	0	0
2	0.67	0.55	0.6	11
3	0.17	0.5	0.25	2
4	0.87	0.65	0.74	20
5	0.25	1	0.4	1
accuracy			0.61	49
macro avg	0.46	0.55	0.45	49
weighted avg	0.77	0.61	0.67	49

The SVM gives an accuracy of 70% and the RF gives an accuracy of 64%. These two algorithms do not work well for class-1 and class-5 classification. Between the two pertained models, 53 layered mobileNet architecture performs better than the 19 layered VGG19 architecture. MobileNet architecture gives a training accuracy of 95% and a testing accuracy of 75.5% which is an overfitting issue in terms of accuracy. VGG19 performs the worst and giving a training accuracy of 70% and testing accuracy of 50%. The graph of loss in training Vs validation, as well as the accuracy of training Vs validation, are shown in figure 4(a, b) and figure 5(a, b).

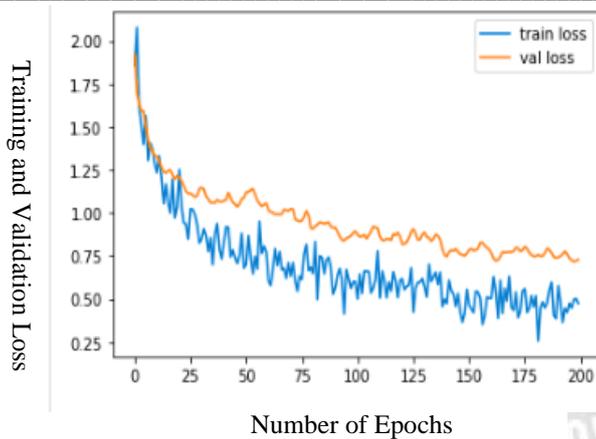


Figure 4 a. Training and Validation loss in MobileNet

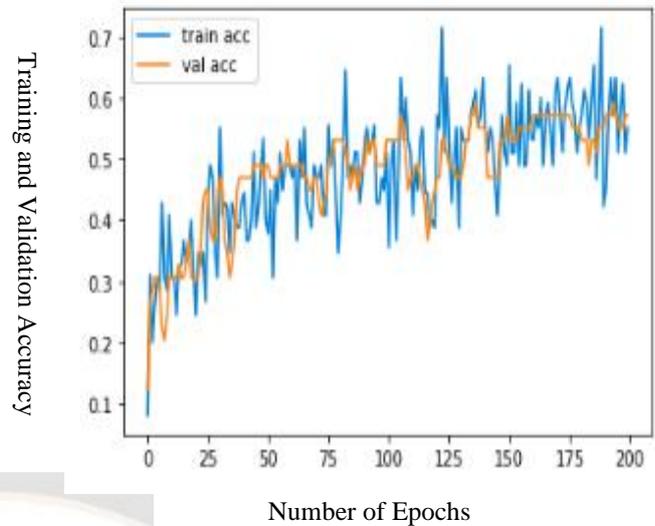


Figure 5 b. Training and Validation accuracy in VGG 19

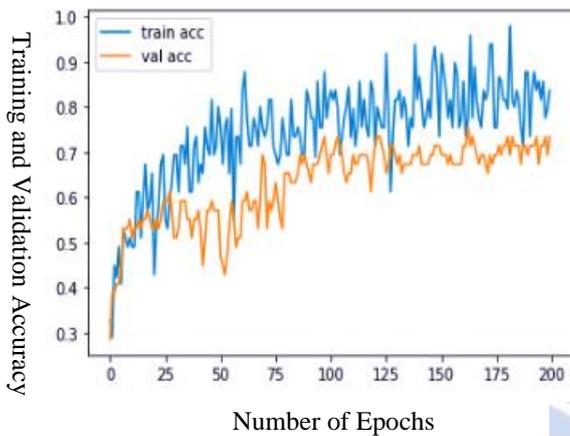


Figure 4b. Training and Validation accuracy in MobileNet

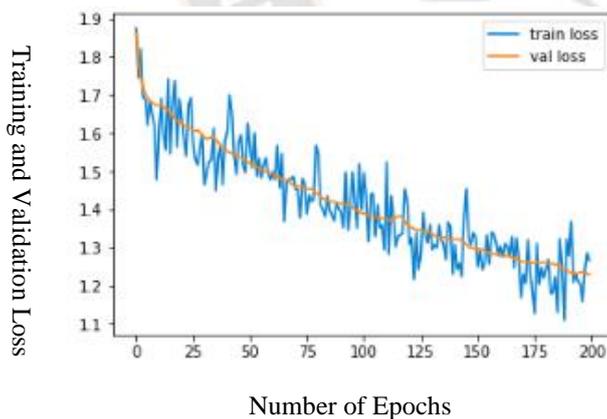


Figure 5 a. Training and Validation loss in VGG19.

In case of the MobileNet architecture, there is little over fitting issue in loss value and accuracy. In VGG19 architecture, there is no overfitting issue in respect of loss and accuracy but the model is underfitted and gives poor results. The confusion matrix of MobileNet is shown in table 4 and classification report of MobileNet is given in table 5. The is shown in table 6 and classification report of VGG19 is given in table 7.

Table 4. Confusion matrix of MobileNet.

Ladies finger healthy	9	1	1	0	0	0
Ladies finger infected	0	2	2	0	0	0
Ladies finger lightly infected	3	0	6	0	0	0
Spinal gourd healthy	0	0	0	4	2	0
Spinal gourd infected	0	0	0	0	14	1
Spinal gourd lightly infected	0	0	0	0	2	2
	Ladies finger healthy					

Table 5. Classification report of MobileNet

	precision	recall	f1-score	support
Ladies finger healthy	0.75	0.82	0.78	11
Ladies finger infected	0.67	0.5	0.57	4
Ladies Finger lightly infected	0.67	0.67	0.67	9
Spinal Gourd healthy	1	0.67	0.8	6
Spinal Gourd infected	0.78	0.93	0.85	15
Spinal Gourd Lightly infected	0.67	0.5	0.57	4
Accuracy			0.76	49
Macro avg	0.75	0.68	0.71	49
Weighted avg	0.76	0.76	0.75	49

Table 6. Confusion Matrix of VGG19

Ladies finger healthy	9	0	0	0	0	0
Ladies finger infected	4	0	0	0	0	0
Ladies finger lightly infected	3	0	0	0	2	0
Spinal gourd healthy	0	0	0	1	5	0
Spinal gourd infected	0	0	0	0	15	0
Spinal gourd lightly infected	0	0	0	0	3	1
Ladies finger healthy	Ladies finger healthy	Ladies finger healthy	Ladies finger healthy	Ladies finger healthy	Ladies finger healthy	Ladies finger healthy

Table 7. Classification report of VGG19

	precision	recall	f1-score	support
Ladies finger healthy	0.5	1	0.67	11
Ladies finger infected	0	0	0	4
Ladies Finger lightly infected	0	0	0	9
Spinal Gourd healthy	1	0.17	0.29	6
Spinal Gourd infected	0.6	1	0.75	15
Spinal Gourd Lightly infected	1	0.25	0.4	4
Accuracy			0.57	49
Macro avg	0.52	0.4	0.35	49
Weighted avg	0.5	0.57	0.45	49

The 5 layered convolutions neural network (CNN) gives training accuracy of 84% and testing accuracy of 81.5%. The model is best fitted for 200 epochs. Although there is a little overfitting [20] in terms of loss, the model is neither over-fitted nor under-fitted in terms of accuracy. The graph of loss in training Vs validation, as well as accuracy training Vs validation for the CNN, is shown in figure 6(a, b).

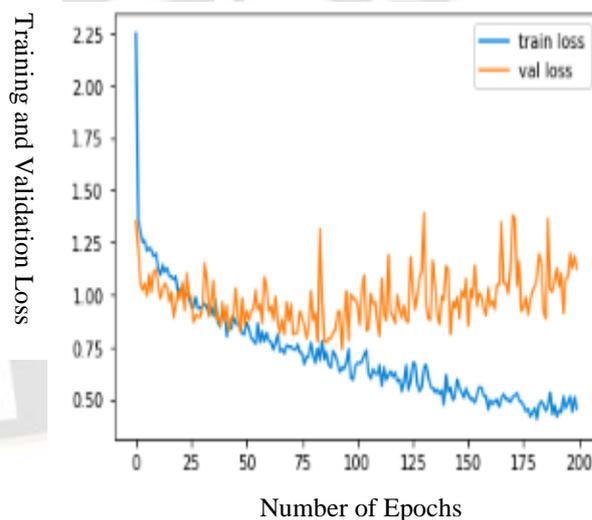


Figure 6 a. Training and Validation loss in CNN model

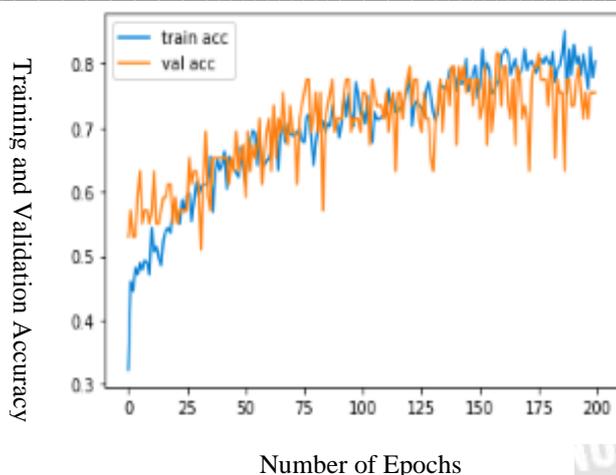


Figure 6 b. Training and Validation accuracy in CNN model

During the implementation, it has been observed that after 90 epochs, there is slight overfitting in the loss graph as shown in figure 6(a), but concerning training and validation accuracy, there is no overfitting issue as shown in figure 6(b).

## V. CONCLUSION

The proposed work is implemented using different techniques along with recently developed VGG19 and MobileNet architecture. Among the three types of techniques i.e. ML, CNN, and TF, the proposed convolution neural network performs better in terms of accuracy. ML algorithms perform poorly as the no of classes increases, ML algorithms were not able to classify lightly infected and infected classes with good precision. The SVM algorithm with some parameter tuning gives an accuracy of 70%. The RF algorithm gives an accuracy of 66%. Among the three ML techniques, the SVM performs better. The VGG19 architecture gives an accuracy of only 50% for the given dataset and MobileNet gives an accuracy of 75.5%. The CNN gives an accuracy of 81.5% which is the best fit for these types of leaf images. Finally, it's worth noting that the given approaches that have been discussed in the paper replace the existing solution for plant disease diagnosis, but light weighted CNN architecture that has been designed can be incorporated with these types of classes for plant health prediction.

## VI. FUTURE SCOPE

As far as only IoT devices are concern for plant diseases and health monitoring, Still, many IoT devices can provide more information about the plant such as biomass, CO<sub>2</sub>, chlorophyll, etc. That information is sufficient for plant disease detection and health monitoring but these IoT devices are costly. Therefore, the plant leaves are being used along with ML, TL, and CNN for plant disease detection in this study. Moreover, instead of 6

leaf classes, more classes could have been formed as the plants surrounding data increased.

## REFERENCES

- [1]. Bhagwan sahay meena; K. Hemachandran. An Experimental Approach for Understanding Internet of Things (IoT) for Corporate Model and Various Platforms. *IJITEE*, 2019, 8,7, 228-240.
- [2]. Bhagwan Sahay Meena; K Hemachandran. Device-based Real-time Single User Indoor Localization Using Internet of Things. *RACSC*, 2021. doi: 10.2174/2666255814666210712114445.
- [3]. Bhagwan Sahay Meena; Ramin Uddin Laskar; K Hemachandran. Indoor Localization-Based Office Automation System Using IoT Devices. *Intelligent Computing in Engineering 2020*, Springer, 1125, 199-210.
- [4]. Bhagwan Sahay Meena; Sujoy Deb; K Hemachandran. Impact of heterogeneous IoT devices for indoor localization using RSSI. *Intelligent Computing in Engineering*, 2020, Springer, 1125, 187-198.
- [5]. Bhagwan sahay meena. Device based multi-user tracking system using internet of things, *IJRITCC*, 2022, (Article in press).
- [6]. G Geetha; S Samundeswari; G saranya; K Meenakshi; M Nithya. Plant Leaf Disease Classification and Detection System Using Machine Learning. *Journal of Physics: Conference Series*, 2020, 1712.
- [7]. Muhammad Chowdhury; Tawsifur Rahman; Amith Khadkar; Mohammad Arselene Ayari; AftabUlmahkhan; Muhammad Salman Khan; Nasser Al-Emadi; Mamum Bin Reaz; MohammadTariqul Islam; Sawar Hamid MdAli. Automatic and Reliable Leaf Disease Detection Using Deep Learning technique. *AgriEngineering 2021*, 3, 294-312.
- [8]. Monirul Islam Pavel; Syed Mohammad Kamruzzaman; SadmanSakibHasan; SaifurRahmanSabuj. An IoT Based Plant Health Monitoring System Implementing Image Processing. *IEEE 4th International Conference on Computer and Communication Systems*, singapor, 02 sep. 2019.
- [9]. Sameer Patel; Mittal B.Jain; Sarvesh Pai; Sagar Korde. Smart agriculture using IoT and Machine Learning. *IRJET 2021*, 08 (04), 4533-4536.
- [10]. Josh Doshi; Tirthkumar Patel; Santosh kumar Bharti. Smart Farming using IoT, a solution for optimally monitoring farming conditions. *Procedia Computer Science*, 160,746-751, 2019.
- [11]. Maryam Ouhami; Adel Hafiane; Youssef Es-Saady; Mohamed El Hajji; Raphael Canals. Computer Vision, IoT and Data Fusion for Crop Disease Detection Using Machine Learning: A Survey and Ongoing Research. *Remote Sens.* 2021 13(13), 2486.
- [12]. Disha Garg; Manshaf Alam. Deep Learning and IoT for Agriculture Applications. Book chapter - Internet of things (IoT), S.M.A.R.T, Springer, 273-284,2020. DOI:10.1007/978-3-030-37468-6\_14.
- [13]. R Juhi Reshma; Anitha S Pillai. Impact of Machine Learning and internet of things in Agriculture, The state of the art.

International Conference on Soft Computing and Pattern Recognition, 2018. DOI:10.1007/978-3-319-60618-7\_59.

- [14]. Amal Dewangan. Application of IoT And Machine Learning in Agriculture. *IJERT* 2020, 09 (07).
- [15]. Mario Bkassiny; Sudharman K. Jayaweera. A Survey on Machine-Learning Techniques in Cognitive Radios. *IEEE communications surveys and tutorials* 2013, 15 ,3.
- [16]. Keiron O'Shea; Ryan Nash. An Introduction to Convolutional Neural Networks, 2015.
- [17]. Muhammad Mateen; Junhao Wen; Nasrullah; Sun Song; Zhouping Huang. Fundus Image Classification Using VGG-19 Architecture with PCA and SVD. *Symmetry* 2018, 11, 1.
- [18]. Wei Wang; Yutao Li; Ting Zou; Xin Wang; Jieyu You; Yanhong Luo. A novel image classification via Dense-MobileNet model. *Mobile Information Systems*, 2020. doi.org/10.1155/2020/7602384.
- [19]. Ricardo Llugsi; Samira El Yacoubi; Allyx Fontaine; Pablo Lupera. Comparison between Adam, AdaMax and AdamW optimizers to implement a Weather Forecast based on Neural Networks for the Andean city of Quito. IEEE fifth ecudor technical chapters meeting, 13 November ,2021, Cuenca, Ecuador. 10.1109/ETCM53643.2021.9590681.
- [20]. Xue Ying. An overview of over fitting and its solution. *Journal of Physics: Conference series* 2018,2, 1168.

