

## Plant disease prediction using convolutional neural network

Hema M S<sup>1</sup>, Niteesha Sharma<sup>2</sup>, Y Sowjanya<sup>3</sup> Ch. Santoshini<sup>4</sup>,  
R Sri Durga<sup>5</sup>, V. Akhila<sup>6</sup>

<sup>1,2,3</sup>Department of Information Technology Anurag University, Hyderabad, India

<sup>4,5,6</sup>Student, Department of Information Technology Anurag University, Hyderabad, India

Email: nitishasharma@cvsr.ac.in, sowjanyait@cvsr.ac.in, Santoshinich@gmail.com,

Durga2008@gmail.com, Akhilav22@gmail.com

Correspondence Author: hemait@anurag.edu.in

*Received September 2, 2021; Revised October 4, 2021; Accepted November 5, 2021*

### Abstract

Every year India losses the significant amount of annual crop yield due to unidentified plant diseases. The traditional method of disease detection is manual examination by either farmers or experts, which may be time-consuming and inaccurate. It is proving infeasible for many small and medium-sized farms around the world. To mitigate this issue, computer aided disease recognition model is proposed. It uses leaf image classification with the help of deep convolutional networks. In this paper, VGG16 and Resnet34 CNN was proposed to detect the plant disease. It has three processing steps namely feature extraction, downsizing image and classification. In CNN, the convolutional layer extracts the feature from plant image. The pooling layer downsizing the image. The disease classification was done in dense layer. The proposed model can recognize 38 differing types of plant diseases out of 14 different plants with the power to differentiate plant leaves from their surroundings. The performance of VGG16 and Resnet34 was compared. The accuracy, sensitivity and specificity was taken as performance Metrix. It helps to give personalized recommendations to the farmers based on soil features, temperature and humidity.

**Keywords:** Plant disease, Convolutional neural network, VGG16, Resnet34, leaf image.

### 1. INTRODUCTION

Agriculture plays vital role in Indian economy. Agriculture contributes nearly 17-18% country GDP. Due to various factors, sometime yield of the agriculture is not as good as expected. The agriculture yield is affected by nature as well as disease. The challenges in agricultural field are seed selection, irrigation techniques, soil defects, weather condition and water treatment [1]. The virus, fungi and bacteria causes the plant disease. It is another challenge in the agriculture field. If the plant diseases are not correctly and timely identified, then it will lead to heavy loss in the yield of

crops. The identification of plant disease needs experts and infrastructure. It is not available in all places. The computer aided identification plant disease identification helps the experts to find the disease accurately and timely manner. The machine techniques are used to identify, classify and predict the plant's disease. The traditional machine techniques such as Support Vector Machine, Decision Tree classification, Random forest, Naïve Bayes are used for plant disease classification and prediction. Nowadays convolutional neural networks are used for plant disease classification and prediction [2-4].

The main contribution of the article is to predict the plant disease using convolutional neural network. The CNN extracts the features from plant and it identifies the disease of the plant.

The article is organized as follows, it has four section apart from introduction section. Literature analysis is presented in section 2. In section 3 describes the proposed methodology. The results and discussion presented in section 4. The conclusion drawn in section 5.

## **2. RELATED WORKS**

A review was presented for plant disease prediction using machine learning techniques [1]. An automatic plant disease prediction using image processing approach was proposed [2,5]. The authors proposed CNN based methodology to predict plant disease [4].

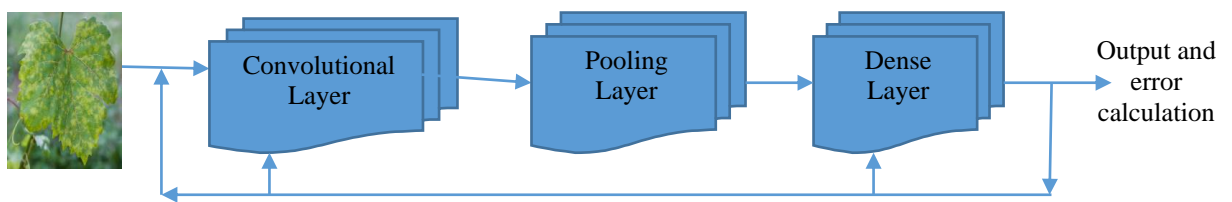
The study has been conducted to detect and predict the plant disease at the early stage using hyperspectral image analysis techniques [5]. The author proposed to predict the Parkinson disease using machine learning techniques [6]. The review is conducted to discover the plant resistance genes using machine learning techniques [7]. 3D convolutional neural network was proposed for plant disease identification [8]. The hybrid approach was proposed to identify plant's leaf disease using convolutional neural network and encoders [9]. The query processing in federated database is proposed [10]. The review was presented to detect and classify plant disease using deep learning approaches [11]. A novel two stage neural network architecture was proposed to classify the plant disease. The traditional augmentation and state of the art style generative adversarial networks were used to augment the images in the database [12]. The Random Forest classifier was used to distinguish disease plant and healthy plant. The leaf of the plant is used for classification [13]. Image processing based study was presented for plant disease detection [14]. The convolutional neural network was proposed to identify plant disease severity. The apple black rot image from plant village database was used for experimentation [15]. The CNN was proposed to identify plant disease. It has five steps namely image acquisition, image preprocessing, image segmentation, feature extraction and classification [16]. The grapes leaf disease identification is done using CNN approach. The grapes leaf images were taken from national research centre [17]. Multiclass support vector machine algorithm was implemented to identify disease in tomato plant leaves [18].

### 3. ORIGINALITY

The proposed methodology is to predict the plant disease from the plant leaf using a convolutional neural network. It extracts the features and predicts the disease from the plant leaf. It is able to predict 38 different diseases from a plant leaf.

### 4. SYSTEM DESIGN

The main objective of the proposed model is to predict the plant disease. The convolutional neural network is used to implement the proposed model. The processing steps of the convolutional neural network are shown in Figure 1.



**Figure1.** Processing steps in CNN

The convolutional neural network has four layers namely convolution layer, pooling layer, flatten layer and dense or fully connected layer.

#### 4.1 Convolutional layer

The important features are extracted from the leaf in the convolutional layer. The kernels or filters are used to extract the features from the image. The kernels are operated on the input image and the feature map is generated. The feature map is calculated using the formula shown in equation 1

$$F(n, m) = (i * k)[n, m] = \sum_x \sum_y k[x, y] * i[n - x, m - y] \quad (1)$$

where  $i$  is the input image and  $k$  is the kernel. The  $n$  and  $m$  are the row and column indexes respectively. The hyper parameters of the convolutional layer are kernels, strides.

#### 4.2 Pooling layer

The images are downsized in the pooling layer to reduce the computational complexity. There are two types of pooling namely max pooling and average pooling. The output of the convolutional layer is divided into  $N \times N$  sub matrices. The max pooling will take the maximum value from each submatrix. The average value of each submatrix is calculated in the average pooling. The number of convolutional and pooling layers is used based on the requirements.

### 4.3 Dense Layer

The output of the convolutional layer or pooling layer is given as input to the dense layer. The dense layer will accept only flattened output. The inputs are flattening in the flatten layer. One or more dense layer is used to predict the plant disease. The output of the final dense layer is the network output. The whole process is called forward propagation. The input is multiplied with weighted and added all the inputs. It is represented in equation 2.

$$O_1 = \sum_{j=1}^m w_j x_j \quad (2)$$

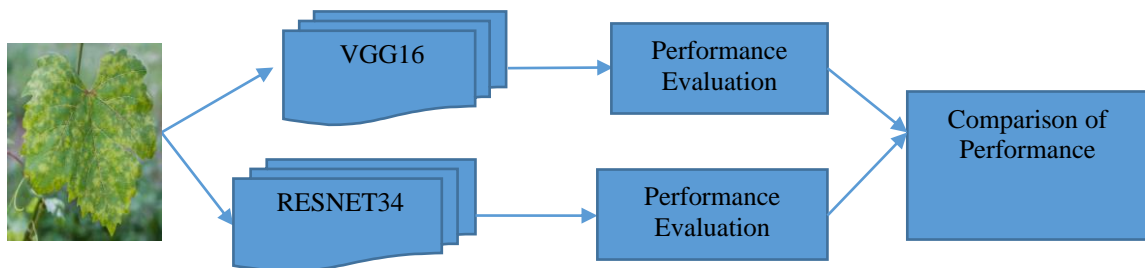
Where  $w$  is weight of each input and  $x$  is the feature input.

The output is given to activation function. It is represented in the equation 3

$$output = A(O_1) \quad (3)$$

Where  $A$  is the activation function. The convolutional layers and dense layers have activation function. The ReLU (Rectified Linear Unit) activation function is used in convolutional layers and softmax activation function is used in the last dense layer.

The performance of the network is not as good as expected. Then error is calculated and it is propagated back to dense layer and convolutional layer to adjust the weights. This process is called as backpropagation. After adjustments' weights, again forward and backward propagation is performed until the network reaches the expected performance.



**Figure 2.** Plant disease prediction steps

The VGG16 and Resnet 34 predefined Convolutional Neural Networks are proposed to predict the plant disease. The prediction steps of the proposed model is shown in figure 2. The VGG16 and Resnet34 are kind of convolutional neural network. Both are used for image classification. In VGG16, the 16 refers number layers in the network. VGG16 used small filters. The input image size is 224X224 RGB. It has nearly 138 million parameters.

Resnet34 has 34 layers. Each layer has number of blocks. Each block has operations. The Resnet34 takes residuals of each and every layer and

these residuals are used in the subsequent layers. The optimization in Resnet34 is easy. The performance of the VGG16 and Resnet34 is compared.

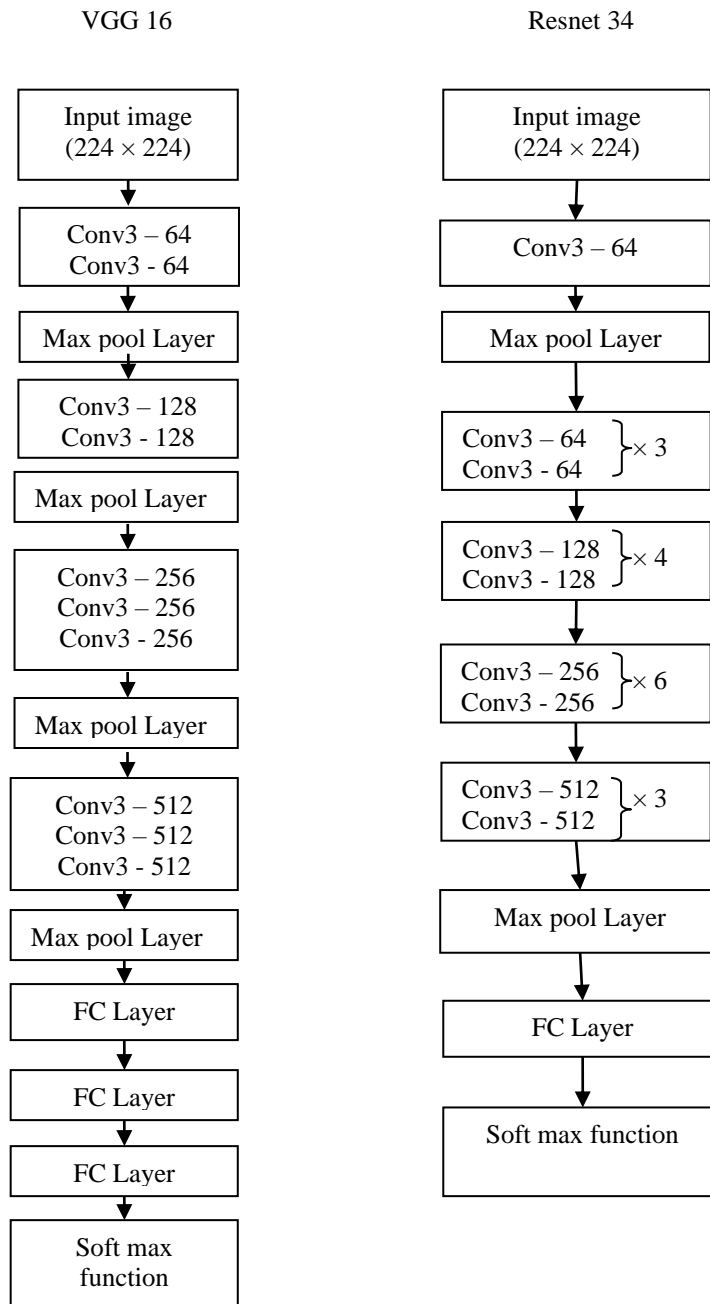
## 5. EXPERIMENT AND ANALYSIS

The plant leaf is given as input for disease prediction. The 150000 instances were taken from Kaggle database for experimentation. The leaves are taken from 14 different plants. The proposed model can recognize 38 differing types of plant diseases. Python language is used for implementation. The proposed CNN is implemented using Python. The transfer learning architectures such as VGG16 and Resnet34 were used for implementation. The filter size was  $3 \times 3$  in VGG16 and resnet34 architectures. The size of the filter used in VGG 16 and resnet 34 ranges from 64 to 512. The number of filters used in VGG 16 and resnet 34 was shown table 1.

**Table 1.** Filters in VGG 16 and Resnet 34

|           | Number Filters | Number of Chennal | Number Filters |
|-----------|----------------|-------------------|----------------|
| VGG 16    | 64             | 2                 | $3 \times 3$   |
|           | 128            | 2                 | $3 \times 3$   |
|           | 256            | 3                 | $3 \times 3$   |
|           | 512            | 6                 | $3 \times 3$   |
| Resnet 34 | 64             | 3                 | $3 \times 3$   |
|           | 128            | 4                 | $3 \times 3$   |
|           | 256            | 6                 | $3 \times 3$   |
|           | 512            | 3                 | $3 \times 3$   |

The maximum number filters used in CNN is not fixed. The maximum number filters are fixed based on the complexity of the problem. The configuration of VGG 16 and Resnet34 is given in figure 3.



**Figure 3.** VGG16 and Resnet34 configuration

The data is divided into training data and validation data. 80% of the data is taken for training and 20% of the data is taken for validation. The Cross entropy error function is used for error calculation. It is shown equation 4.

$$E(X_i, T_i) = - \sum_{k=1}^n T_{ik} * \log(p_{ik}) \quad (4)$$

Where  $T_{ik}$  is the target vector.  $p_{ik}$  is the probability of  $i$ th element in  $k^{\text{th}}$  class. Accuracy, sensitivity and specificity was used as performance metrics to assess the performance of the proposed approach. using The accuracy is calculated using the equation 4

$$\text{Accuracy} = \frac{(\text{True}_{\text{positive}} + \text{True}_{\text{negative}})}{(\text{True}_{\text{positive}} + \text{True}_{\text{negative}} + \text{False}_{\text{positive}} + \text{False}_{\text{negative}})} \quad (4)$$

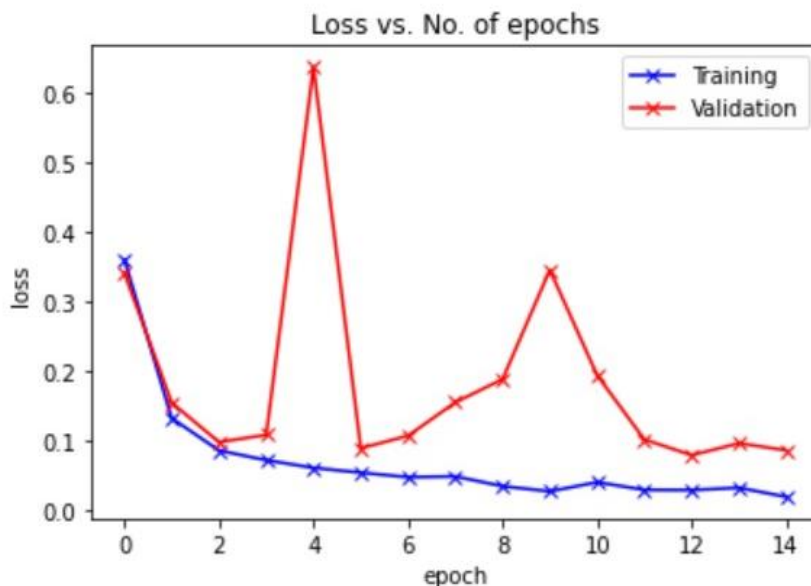
The sensitivity is calculated using the formula shown in equation 5

$$\text{Sensitivity} = \frac{\text{True}_{\text{positive}}}{(\text{True}_{\text{positive}} + \text{False}_{\text{negative}})} \quad (5)$$

The specificity is calculated using the formula shown in equation 6

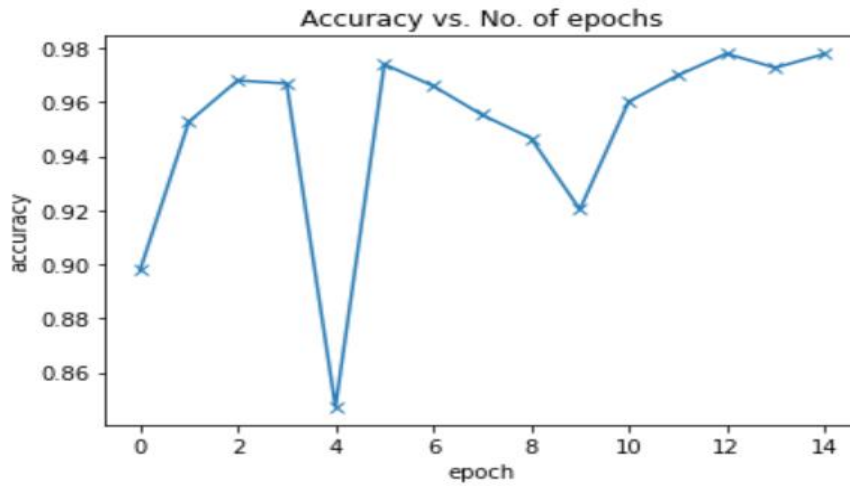
$$\text{Specificity} = \frac{\text{True}_{\text{negative}}}{(\text{True}_{\text{negative}} + \text{False}_{\text{positive}})} \quad (6)$$

The error rate of the Resnet34 Model is shown in figure 3. The error rate of the training data is reduced uniformly when the number epochs increased. But, the validation data error rate is not uniformly reduced because of uncertainty of the data.



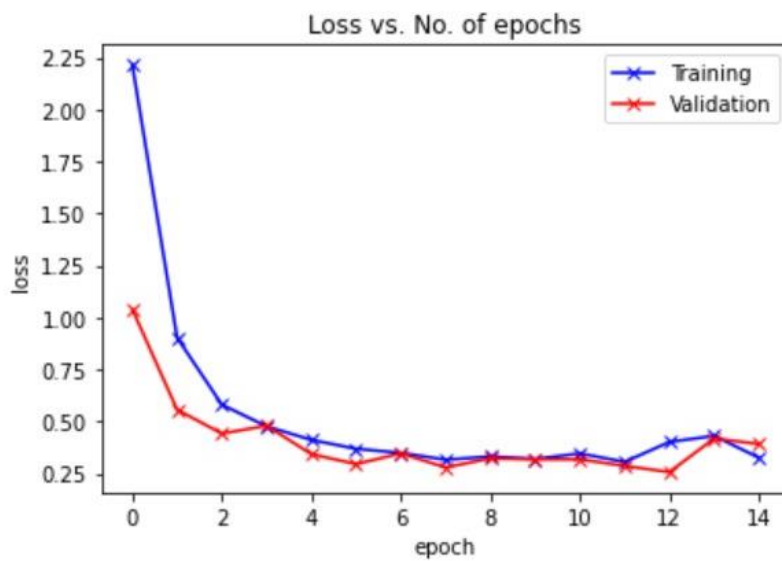
**Figure 4.** Resnet34 – Error rate of training and test data

The accuracy of the test data is shown in figure 4. The accuracy is increased after 10 epochs.



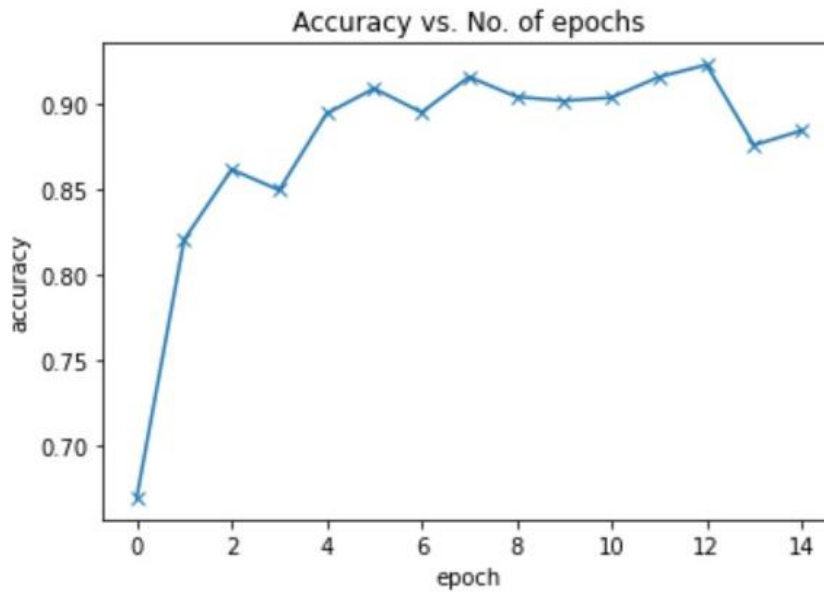
**Figure 5.** Resnet34 – Test data accuracy

The error rate of the VGG16 Model is shown in figure 5 . The error rate of the training data and testing is reduced gradually when the number epochs increased. The accuracy of test data of VGG16 model is shown figure 6. In this model after 12 epochs, the accuracy is decreasing.



**Figure 6.** VGG16- Error rate of training and validation data





**Figure 7.** VGG16- Test data accuracy

The performance comparison of Resnet34 and VGG16 is shown in table 2. Both model performance is good. The Resnet34 performs better than VGG16 for plant disease prediction application because Resnet34 optimize the residual and it deals with vanishing gradient problem.

**Table 2.** Resnet34 and VGG16 comparison

| S.No.       | Resnet34 | VGG16  |
|-------------|----------|--------|
| Accuracy    | 97.77    | 97.58  |
| Loss        | 0.0857   | 0.0875 |
| Sensitivity | 1.0000   | 0.9741 |
| Specificity | 0.9921   | 0.9561 |

## 6. CONCLUSION

The prediction of plant disease using convolutional neural network was proposed and implemented. The Resnet34 and VGG16 CNN architectures were used for prediction of plant diseases. The convolutional layer was used for extraction of features, pooling layer was used for sample downsizing and dense layer was used for prediction. The ReLU activation function was used in convolutional layer. The softmax activation function was used in dense layer. Nearly 150000 plant leaves were taken from 14 different plants for experimentation. The proposed architectures were able recognize 38 differing types of plant diseases.

**REFERENCES**

- [1] Gokulnath, B. V., & Usha Devi, G. **A survey on plant disease prediction using machine learning and deep learning techniques**. *Inteligencia Artificial*, 22(63), 0-19, 2017.
- [2] Ashok, S., Kishore, G., Rajesh, V., Suchitra, S., Sophia, S. G., & Pavithra, B. **Tomato Leaf Disease Detection Using Deep Learning Techniques**. In 2020 5th International Conference on Communication and Electronics Systems (ICCES) (pp. 979-983). IEEE, 2020.
- [3] Ferentinos, K. P. **Deep learning models for plant disease detection and diagnosis**. *Computers and Electronics in Agriculture*, 145, 311-318, 2018.
- [4] Reddy, K. N., Polaiiah, B., & Madhu, N. **A literature survey: plant leaf diseases detection using image processing techniques**. *IOSR J. Electron. Commun. Eng.*, 2017.
- [5] Lowe, A., Harrison, N., & French, A. P. **Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress**. *Plant methods*, 13(1), 1-12, 2017.
- [6] Kanagaraj, S., Hema, M. S., & Gupta, M. N. **Environmental Risk Factors and Parkinson" s Disease–A Study Report**. *International Journal of Recent Technology and Engineering (IJRTE)* ISSN, 2277-3878 2019.
- [7] Yang, X., & Guo, T. **Machine learning in plant disease research**. *European Journal of BioMedical Research*, 3(1), 6-9, 2017.
- [8] Nagasubramanian, K., Jones, S., Singh, A. K., Sarkar, S., Singh, A., & Ganapathysubramanian, B. **Plant disease identification using explainable 3D deep learning on hyperspectral images**. *Plant methods*, 15(1), 1-10., 2019.
- [9] Khamparia, A., Saini, G., Gupta, D., Khanna, A., Tiwari, S., & de Albuquerque, V. H. C. **Seasonal crops disease prediction and classification using deep convolutional encoder network**. *Circuits, Systems, and Signal Processing*, 39(2), 818-836., 2020
- [10] Hema, M. S., & Chandramathi, S. **Federated query processing service in service oriented business intelligence**. In *International Conference on Advances in Communication, Network, and Computing* (pp. 337-340). Springer, Berlin, Heidelberg, 2011.
- [11] Saleem, M. H., Potgieter, J., & Arif, K. M. **Plant disease detection and classification by deep learning**. *Plants*, 8(11), 468, 2019
- [12] Arsenovic, M., Karanovic, M., Sladojevic, S., Anderla, A., & Stefanovic, D. (2019). **Solving current limitations of deep learning based approaches for plant disease detection**. *Symmetry*, 11(7), 939, 2019.
- [13] Ramesh, S., Hebbar, R., Niveditha, M., Pooja, R., Shashank, N., & Vinod, P. V. **Plant disease detection using machine learning**. In 2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C) (pp. 41-45). IEEE, 2018

- [14] Dhingra, G., Kumar, V., & Joshi, H. D. **Study of digital image processing techniques for leaf disease detection and classification.** Multimedia Tools and Applications, 77(15), 19951-20000, 2018.
- [15] Wang, G., Sun, Y., & Wang, J. **Automatic image-based plant disease severity estimation using deep learning.** Computational intelligence and neuroscience, 2017.
- [16] Prashanthi, V., & Srinivas, K. **Plant disease detection using Convolutional neural networks.** International Journal of Advanced Trends in Computer Science and Engineering, 9(3), 2020.
- [17] Ghosh, G., & Chakravarty, S. **Grapes Leaf Disease Detection Using Convolutional Neural Network.** International Journal of Modern Agriculture, 9(3), 1058-1068, 2020.
- [18] Vetal, S., & Khule, R. S. **Tomato plant disease detection using image processing.** International Journal of Advanced Research in Computer and Communication Engineering, 6(6), 293-297,2017.