Automation, Control and Intelligent Systems
2023; 11(1): 1-7
http://www.sciencepublishinggroup.com/j/acis
doi: 10.11648/j.acis.20231101.11
ISSN: 2328-5583 (Print); ISSN: 2328-5591 (Online)

Classification and Detection of Cabbage Leaf Diseases from Images Using Deep Learning Methods

Myna A. N., Manasvi K., Pavan J. K., Rakshith H. S., Yuktha D. Jain

Department of Computer Science and Engineering, Navkis College of Engineering, Hassan, India

Email address:
mynaan@gmail.com (Myna A. N.)

To cite this article:
Myna A. N., Manasvi K., Pavan J. K., Rakshith H. S., Yuktha D. Jain. Classification and Detection of Cabbage Leaf Diseases from Images Using Deep Learning Methods. *Automation, Control and Intelligent Systems*. Vol. 11, No. 1, 2023, pp. 1-7. doi: 10.11648/j.acis.20231101.11

Received: December 26, 2022; **Accepted**: January 30, 2023; **Published**: March 4, 2023

Abstract: The presented work uses Deep learning methods to detect diseases in cabbage leaves. The plant disease detection is constrained by human visual capabilities. Because, most of the early symptoms detected are microscopic. This process is tedious, time consuming and prediction is a challenging task. Hence, there is a need for developing a methodology that automatically recognizes, classifies and detects plant infection symptoms. Five major types of diseases namely Leaf miner, Diamond backmoth, Blackrot, Maggvots and Downy mildew are considered. Initially, the input images are classified into two types as healthy and diseased. Further, the diseased images are categorized into five different varieties. Around 3000 images of cabbage leaves are used containing healthy and infected leaves. Different phases namely preprocessing, feature extraction, training, testing and classification are used the proposed methodology. The accuracies of 93.5% and 90.5% are achieved for healthy and diseased leaf images. Classification accuracies for different types of diseased images are 89.9, 89.5, 91.8, 90.5 and 90.8 for Leaf miner, Diamond backmoth, Blackrot, Maggvots and Downy mildew respectively. The overall classification accuracy of 92% is attained. The developed methodology is found to provide good classification accuracy. The developed model finds its applications in APMCs, online purchase, Agricultural departments etc.

Keywords: Deep Learning, CNN, VGG-16, Image Preprocessing, Feature Extraction

1. Introduction

India is well-known in the field of agriculture where, about 70% of the population relies on agriculture. Indian farmers select appropriate crops for their farm. The cultivation of these crops for best yield and quality produce is not technical. It can be enhanced by the aid of technological support.

The management of persistent crops needs close monitoring and controlling of diseases which can affect production considerably and the post-harvest life afterword. Image processing techniques are employed in agricultural applications for a variety of purposes. The plant disease detection is constrained by human visual capability. Because, most of the early symptoms observed are microscopic. This practice is monotonous, time consuming and prediction is a difficult talk. Hence, there is a need for developing a methodology that can automatically recognize and detects infection symptoms. A symptom is a deformation, observed as proof of its existence.

In this work, the diseases which occur on cabbage leaves are considered. Cabbage is a green leafy vegetable which ranges from 500 grams to 1,000 grams and can be grown in three months. Although it has good yield, it has some major diseases which are much affective on the production rate of cabbage. So, the producers are unable to find out the accurate disease and accurate pesticide for these diseases. Figure 1 shows the cross section of cabbage image.

Figure 1. *Cross Section of Cabbage image.*

2. Related Work

Preparation of Dataset: Dataset used in the approach contains dataset with numerous diseases in various plants [1] thru [14]. In the work carried out, few of the business-related crops, cereal crops, vegetable crops and fruit crops such as sugarcane, cotton, potato, chilly, banana etc., are considered. The images are resized to 256*256 size during pre-processing of images. Plant dataset used has fifty-four thousand three hundred and six images of various plant leaves which are separated into eighteen categories [5, 11]. The work is implemented using K-Means Clustering and SVM Algorithm that detects and distinguishes different types of leaf and skin diseases [2]. The dataset used in the methodology contains the images of plant diseases to test the performance that are frequent to the leaves [1] thru [12]. The images are acquired with a Nikon 7200d camera in the proposed approach [9]. The technique proposed comprises of 54,300 plant leaf images [10]. The additional class in the dataset through background images is helpful to obtain better accuracy [15]. It is stated that Deep learning model is found to be an efficient model that can differentiate the leaves from its surrounding.

2.1. Image Acquisition

Masking and removal the green pixels and segmenting the objects using Otsu's method are performed [10]. The images collected are unprocessed and is the effect of hardware used to capture it.

2.2. Preprocessing

In preprocessing phase, Image annotation and augmentation are carried out and is shown in Figure 2.

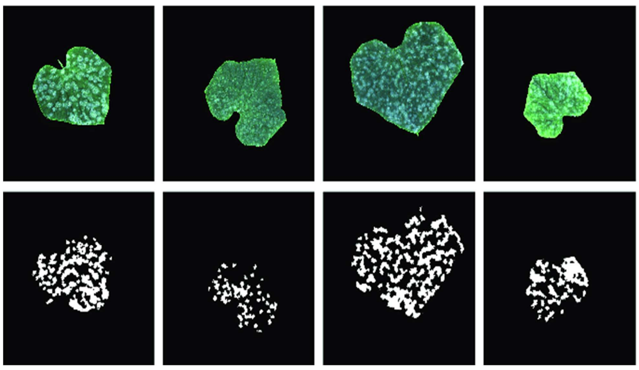

Figure 2. *Image annotation and augmentation.*

The manual annotation is employed on the areas of whole image having disease [2, 4]. To identify a transformation matrix, three points are needed in the corresponding locations in the output image. a 3×3 transformation matrix is used in perspective transformation [6, 7].

Noise removal from images

The Cumulative Distribution Function (CDF) is applied to share the intensity values. Various pre-processing techniques are applied to eliminate noise from the images [10]. For an example, image clipping is used to get the specific region of leaf. Image smoothing can be carried out using the smoothing filter. Also, image enhancement is employed to increase the contrast.

2.3. Image Segmentation

Segmentation is the process, where, the images are partitioned into different segments. Segmentation is applied to identify the region of interest, which is having an important effect and easier to analyze the image. Image segmentation is used to trace objects and boundaries in the images. Each pixel in a region, with respect to some property, such as color, texture and pixel intensity is identified [3, 6, 13]. The performance of proposed methodology using segmented images is always better, when, compared with the model using grey-scaled images [14]. But, it is slightly lesser than the model that uses colored version of the images.

2.4. Feature Extraction

Feature extraction is based on extracting unique features learnt from a pre-trained Convolutional Neural Network [4, 6, 8]. A low-cost technique of apple disease diagnosis using a neural network and fruit classification into four classes of scab, bitter rot, black rot, and healthy fruits is developed [1]. This method uses color and texture features. In this work, unique features are extracted using deep learning models, namely ResNet50, Google Net InceptionResNetV2, Squeeze-Net, pool5-drop_7x7_s1 etc.

2.5. Training the Model

The work proposes a smart and efficient method for detection of crop diseases using vision and machine learning techniques [4, 5]. The Methodology for disease detection in the early stage using ANN and image processing techniques are proposed in [4, 10]. The approach is mainly based on ANN model for classification. The recognition rate of up to 91% is achieved. The performance of conventional multiple regression, ANN and SVM are compared in [11].

2.6. Existing System

In plant leaf classification, leaves are classified and diseases are identified using following techniques

1) Fuzzy Logic
2) Principal Component Analysis (PCA)
3) K-Nearest Neighbour (KNN)
4) Support Vector Machine (SVM)
5) Artificial Neural Network (ANN)
6) Single Shot Detector (SSD)

2.7. Disadvantages

1) The deployed model has not focused on identifying the particular plant or leaf diseases and the dataset used for the training and implementation is not sufficient.
2) The accuracy rate of the model is not satisfactory.
3) Feature extraction of the model is not accurately

defined.

4) It has not focused on increasing the classification accuracy and severity of leaf infection.

2.8. Motivation

Cabbage plant leaves are prone to various disorders and attacks that are caused by different diseases. There are many reasons namely humidity, temperature, nutritional surplus or losses, intensity etc. The most common diseases like bacterial, viral and fungal, affect the plants. The infections along with the cabbage leaf may show different physical characteristics such as change in shape, color etc. Due to similar characteristics, the changes are difficult to be recognized, which makes their detection a challenge. The earlier detection and treatment can avoid many losses in the cabbage plant. Hence, the work on "*Classification and detection of Cabbage leaf diseases From Images using Deep Learning Methods*" is proposed.

The Convolutional Neural Network (CNN) is used to detect cabbage leaf disease so that people with lesser expertise in software should also use it easily. Sample images of different diseases are collected. A total of five hundred images of each type are collected that will be classified into training images and testing images. The image database is collected from tensor flow and keras library.

3. Proposed Methodology

Different phases used in the proposed methodology are shown in Figure 3.

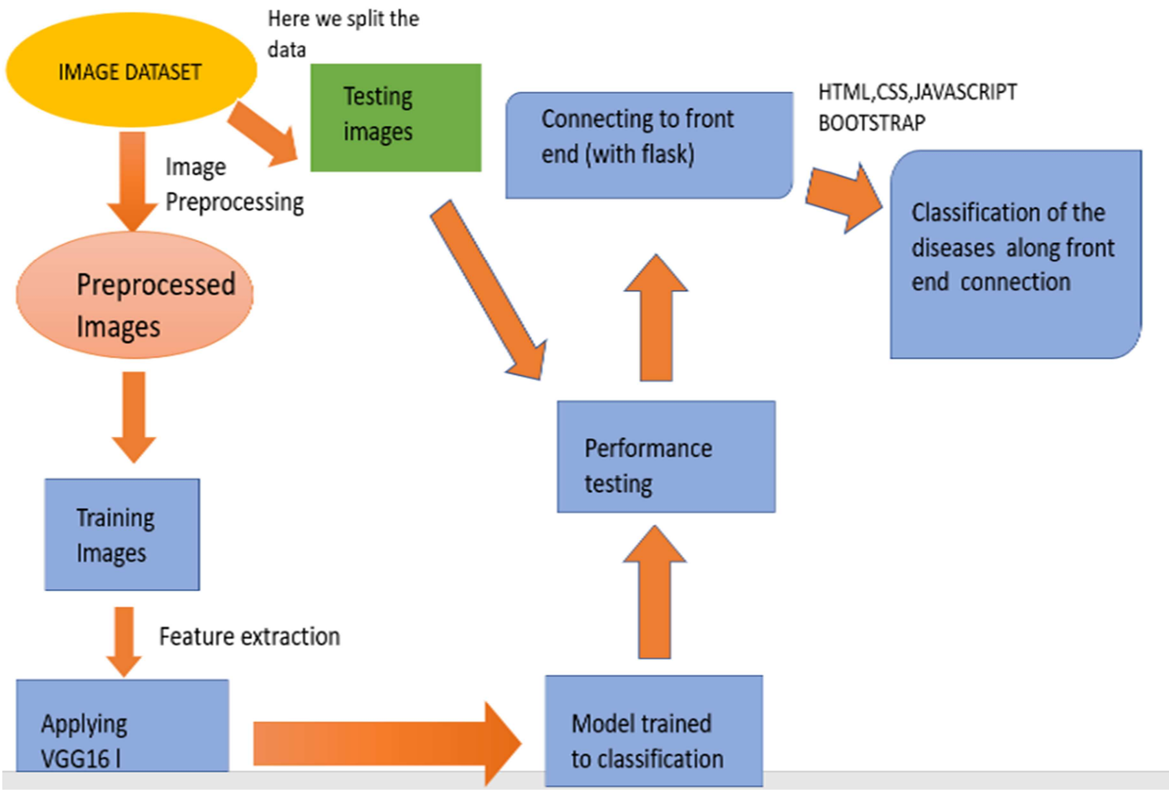

Figure 3. *Methodology used in the proposed work.*

1) *Image dataset*: A dataset is a curated set of digital photographs that is used to test, train and evaluate the model. Around three thousand images of cabbage leaves are used containing healthy and infected leaves.
2) *Data Annotation*: Data annotation is the process of categorizing and labeling of data in deep learning applications. Training has to be properly categorized and annotated for precise use case. Cabbage leaves suffer from five types of diseases, namely Leaf miner, Diamond backmoth, Blackrot, Maggvots and Downy mildew.
3) *Data Augmentation*: Data augmentation in data analysis is a method used to increase more images by adding up slightly modified images of already existing data or newly created data using existing data. In the proposed project work, images are classified based on the color, size, infected area and other expertise knowledge given by farmers.
4) *Deep Learning Techniques*: It is a machine learning technique. We have used Deep Convolutional Neural networks and compared its performance with other neural networks.
5) *Training Data*: 70% of the data will be used for training using deep learning method.
6) *Testing Data*: Remaining 30% of the data is used for testing.
7) *Validation of Data*: This will be used to know how much model will get trained.

3.1. Preprocessing of Cabbage Leaves

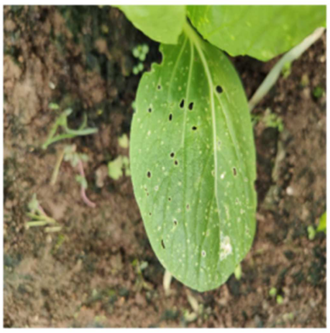

Figure 4. *Before preprocessing.*

The cabbage leaf before preprocessing is shown in Figure 4. In preprocessing phase, the raw images is inputted from the databases undergo preprocessing before it is fed into CNN classifier. Images are resized to create a base size and remove noise. Each image is established into three dimensional vectors namely, P, Q and R. The vectors P, Q, and R represent width, height and RGB channel respectively. Firstly, the images are converted into gray scale. The segmentation is carried out using edge based segmentation and is shown in Figure 5.

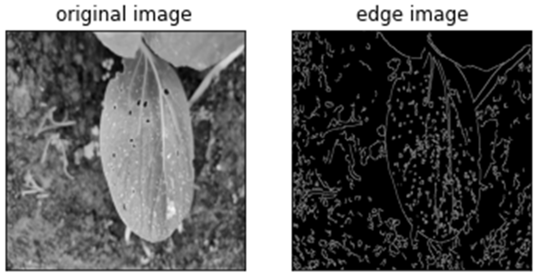

Figure 5. *Gray-scale of image and Edge based segmentation.*

After segmentation, k-means clustering is applied to form the clusters of images. The five labels are assigned to the clusters. Clustering of images is shown in Figure 6.

Figure 6. *Clustering of images using k-means.*

3.2. CNN Model

Few decades ago, Deep Learning had proved to be a very influential tool because of its capability to handle huge amount of data. Convolutional Neural Networks (CNNs) are one among most popular Deep Neural Networks (DNNs) used.

Convolutional Neural Network (CNN): Convolution is applied on two functions, which generates third function. It specifies the shape of one function customized by another function. It comprises of three layers namely, Convolutional layer, pooling layer and fully connected layer.

3.3. VGG-16 Architecture

VGG-16 is a CNN model proposed by K. Simonyan and A. Zisserman, University of Oxford. The VGG-16 architecture used in the work is shown in Figure 7.

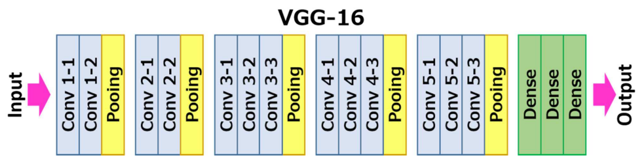

Figure 7. *VGG-16 architecture.*

The inputs given to conv1 layer are of size 224 x 224 RGB images. The images are given through a heap of convolutional layers. The filter size of 3×3 used. It also uses 1×1 convolution filters that are viewed as a linear transformation of the input channels.

The convolution tramp is positioned to one pixel and the spatial padding of the layer is carried out. The spatial resolution is preserved after convolution. All the layers are not followed by max-pooling. Max-pooling is carried out over 2×2 filter with stride 2.

The first two layers have 4096 channels each. The third layer performs 1000-way ILSVRC classification. It has 1000 channels for each class and the final layer is known as soft-max layer.

The loss function plays significant role as it must be done with the correct modelling problem. The equation (1) is used to compute loss function.

$$LE = -\frac{1}{N}\sum_{i=1}^{N} log\left(\frac{p}{yi}\right) \quad (1)$$

The equation (2) is applied to calculate Cross-Entropy.

$$Cross\ Entropy\ Loss = \sum_{i=1}^{2} ti\log(Pi) \quad (2)$$

The phases involved in the classification process are training the model and testing for the accuracy. For each epoch, the training batches are repeated to compute the loss and adjust the network weights. At the end of each epoch, the loss and performance accuracy are calculated. The accuracy will convey the number of predictions that are true. For testing, the images are present in any of the training and validation dataset are used as the inputs. Using a complete different image dataset would test the efficiency of the model. It will also verify, if the model can predict correctly, even if a completely different dataset is used as an input.

4. Results and Discussion

The main objective of the research work carried out is to recognize and find out the whether a cabbage leaf is unhealthy or healthy. Also, it identifies the type of disease occurred. There are 3000 images among them 450 are validation images. All the images are labelled. The proposed model is able to classify the diseases with 92% accuracy using VGG16. The accuracy of classification can still be increased when trained with larger dataset and by using pre-trained models.

The output of VGG16 is shown in Figure 10. The model predicts the image as healthy, when, healthy cabbage leaf image is inputted.

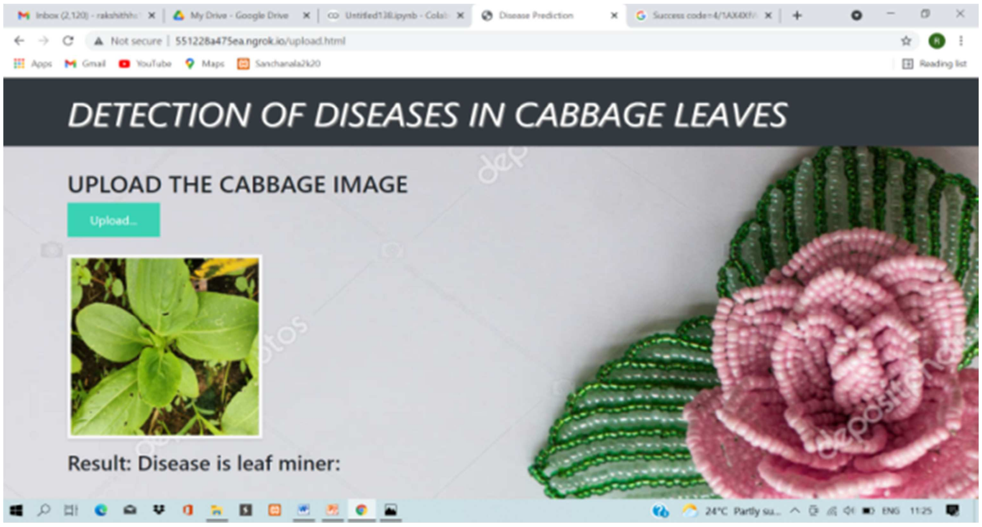

Figure 8. *Detection of Diseases.*

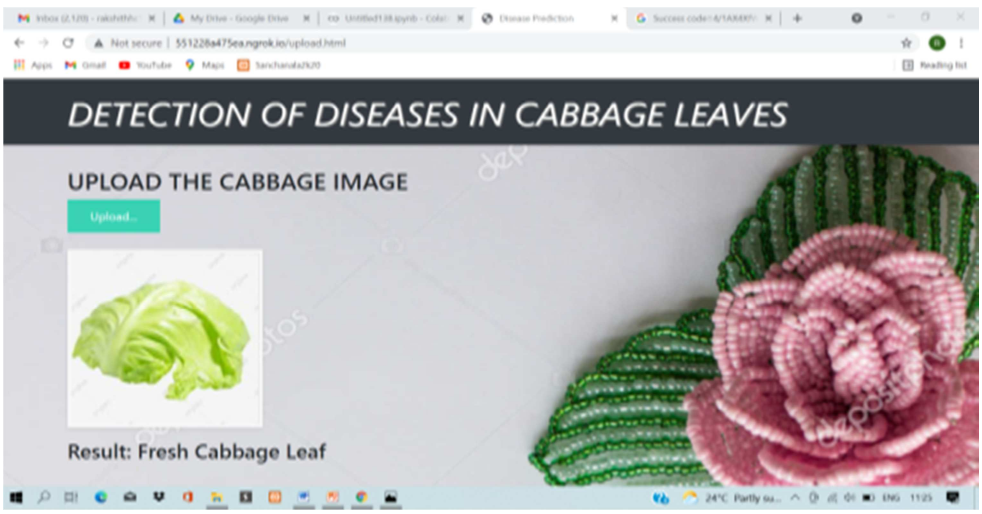

Figure 9. *Detection of Fresh Cabbage Leaves.*

Figure 10 shows the accuracy during training and testing loss during training and testing phase.

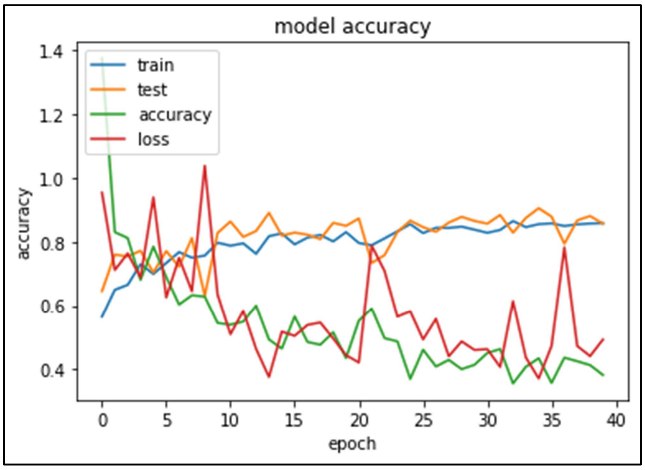

Figure 10. *Accuracy during Training and Testing loss.*

The classification accuracies of healthy and infected leaf images are shown in Figure 11. The accuracies are 93.5 and 90.5 respectively. The overall accuracy of 92% is obtained.

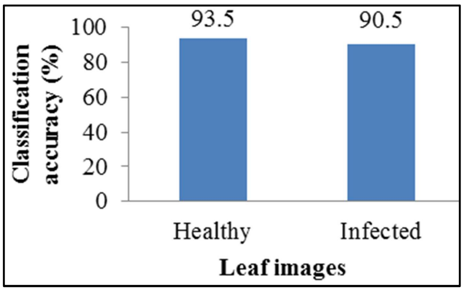

Figure 11. *Classification accuracies for healthy and infected images.*

The classification accuracies of different diseased leaf images are shown in Figure 12. The classification accuracies of Leaf miner, Diamond backmoth, Blackrot, Maggvots and Downy mildeware are 89.5%, 89.5%, 91.8%, 90.5% and 90.8% respectively. The overall recognition accuracy of 90.5% is obtained.

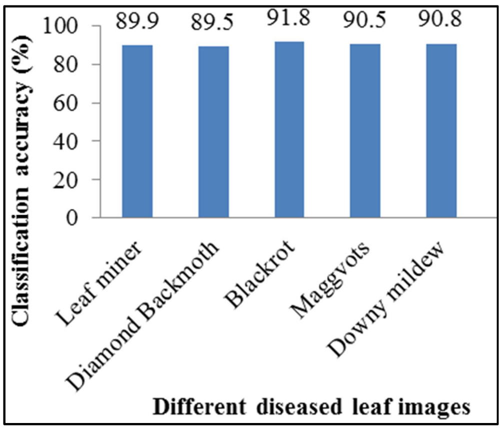

Figure 12. *Classification accuracies for different types of diseased images.*

5. Conclusion

The work carried out has discussed the development model of the deep learning model that identifies, whether the cabbage leaf is infected or not. Further, the type of disease occurred on the particular leaf is identified. In future, the developed methodology can be used to capture the real-time images using drones from the farm land and then predict the diseases. The optimal solution for the predicted diseases can be developed. The developed model finds its applications in APMCs, online purchase, Agricultural departments etc.

References

[1] Azgomi, Hossein, Fatemeh Roshannia Haredasht, and Mohammad Reza Safari Motlagh. "Diagnosis of some apple fruit diseases by using image processing and artificial neural network." *Food Control* 145 (2023): 109484.

[2] Badiger, Manjunatha, Varuna Kumara, Sachin CN Shetty, and Sudhir Poojary. "Leaf and skin disease detection using image processing." *Global Transitions Proceedings* 3, no. 1 (2022): 272-278.

[3] Aiswarya, K., and N. R. Sreekumar. "Plant Disease Detection Using Quantum Image Processing." In *2022 International Conference on Industry 4.0 Technology (I4Tech)*, pp. 1-6. IEEE, 2022.

[4] Kulkarni, Pranesh, Atharva Karwande, Tejas Kolhe, Soham Kamble, Akshay Joshi, and Medha Wyawahare. "Plant disease detection using image processing and machine learning." *arXiv preprint arXiv: 2106. 10698* (2021).

[5] Vishnoi, Vibhor Kumar, Krishan Kumar, and Brajesh Kumar. "Plant disease detection using computational intelligence and image processing." *Journal of Plant Diseases and Protection* 128, no. 1 (2021): 19-53.

[6] Sharma, Parul, Yash Paul Singh Berwal, and Wiqas Ghai. "Performance analysis of deep learning CNN models for disease detection in plants using image segmentation." *Information Processing in Agriculture* 7, no. 4 (2020): 566-574.

[7] Annabel, L. Sherly Puspha, T. Annapoorani, and P. Deepalakshmi. "Machine Learning for Plant Leaf Disease Detection and Classification–A Review." In *2019 International Conference on Communication and Signal Processing (ICCSP)*, pp. 0538-0542. IEEE, 2019.

[8] Türkoğlu, Muammer, and Davut Hanbay. "Plant disease and pest detection using deep learning-based features." *Turkish Journal of Electrical Engineering and Computer Sciences* 27, no. 3 (2019): 1636-1651.

[9] Monishanker Halder, Ananya Sarkar, Habibullah Bahar, "Plant Disease Detection by Image Processing: A Literature Review", *SDRP Journal of Food Science & Technology*. Vol-3, Issue-6, pp. 534-538, February 2019.

[10] M. Deepan, P., and M. Akila. "Detection and Classification of Plant Leaf Diseases by using Deep Learning Algorithm." *International Journal of Engineering Research & Technology (IJERT)* (2018).

[11] Jayswal, Hardikkumar S., and Jitendra P. Chaudhari. "Plant Leaf Disease Detection and Classification using Conventional Machine Learning and Deep Learning." *International Journal on Emerging Technologies 11 (3)*: 1094-1102 (2020).

[12] Sladojevic, Srdjan, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. "Deep neural networks based recognition of plant diseases by leaf image classification." *Computational intelligence and neuroscience* 2016.

[13] Prasanna Mohanty, Sharada, David Hughes, and Marcel Salathe. "Using Deep Learning for Image-Based Plant Disease Detection." *arXiv e-prints* arXiv-1604, 2016. https://www.frontiersin.org/articles/10.3389/fpls.2016.01419/full

[14] Alka Dixit, Erande Rani, LokhandeYogita, NighotRutuja, Kote S. V, "Plant Disease Detection", International Journal on Emerging Technologies. Volume-9, Issue 03, pp. 66-70, February 2016.

[15] Mukesh Kumar Tripathi, Dhananjay D. Maktedar, Machine Learning Based Approaches for Disease Detection and Classification of Agricultural Products, 2016 International Conference on Computing Communication Control and automation (ICCUBEA), Pune, 2016, pp. 1-6, doi: 10.1109/ICCUBEA, 2016.