International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-10 Issue-3, September 2021

# Plant Disease Detection and Classification using CNN

**Rinu R, Manjula S H**

***Abstract:** Agriculture is one field which has a high impact on life and economic status of human beings. Improper management leads to loss in agricultural products. Diseases are detrimental to the plant's health which in turn affects its growth. To ensure minimal loss to the cultivated crop, it is crucial to supervise its growth. Convolutional Neural Network is a class of Deep learning used majorly for image classification, other mainstream tasks such as image segmentation and signal processing. The main aim of the proposed work is to find a solution to the problem of 38 different classes of plant diseases detection using the simplest approach while making use of minimal computing resources to achieve better results compared to the traditional models. VGG16 training model is deployed for detection and classification of plant diseases. Neural network models employ automatic feature extraction to aid in the classification of the input image into respective disease classes. This proposed system has achieved an average accuracy of 94.8% indicating the feasibility of the neural network approach even under unfavorable conditions.*

***Keywords**: Deep Learning, VGG16 model, Convolutional Neural Network.*

## I. INTRODUCTION

Agriculture is a sector which has a huge impact on life and economic stature of humans. Agriculture proves to be a primary source of livelihood, for about 58% of India's population. India ranks second globally in terms of farm yields. It was reported in the year 2018 that agriculture opened the doors of employment for more than 50% of the employees, hence contributing to 18–20% to country's GDP. India has thus proven to be one of the leading nations in term of agricultural yield and productivity. With this majority of population rely on agriculture, it is very crucial to recognize the problems faced in this sector. There are a numerous of problems that the agriculture field faces such as inefficient farming strategies and techniques, inadequate use of compost, manures and fertilizers, insufficient water supply, various diseases attack on plants and so on. Diseases are exceedingly harmful to the well-being of plants which in turn influence its growth. The attack of these numerous types of diseases on plants results in a huge loss in the yield performance in terms of quality as well as quantity. Plants affected by diseases add up to for about 20-30% of the entire crop deprivation. Thus, recognizing diseases in plant becomes very crucial in order to avoid any massive losses in production, performance and in the amount of the agricultural outcome. Since manual recognition is extremely time consuming and more prone to inaccuracy, leading to wrong treatment. The recent development in technology, and this evolution has made it feasible and possible thus made its way for plant disease detection and identification and contribute to provide better treatment for plants in case of any plant having diseased conditions. The proposed system of recognition of plant leaf diseases focuses on 14 varieties of plants which include apple, blueberry, cherry, corn, grape, orange, peach, pepper, potato, raspberry, soybean, squash, strawberry and tomato. This system is built on concepts of Deep Learning- Convolutional neural networks (CNN) is posted for formation of a statistical model that is executed on the input image and transforms the input to classify output tags.

## II. RELATED WORK

Eftekhar Hossain et al., [1] proposed a system for recognizing the plant leaf diseases with the appropriate classifier K-nearest neighbor (KNN). The features that were extracted through the images of diseased image were used to execute the classification. In the paper, the system KNN classifier classified the diseases commonly found in plants like bacterial blight, early blight, bacterial spot, leaf spot of various plant species. This method exhibited an accuracy of 96.76%. Sammy et al., [2] proposed a CNN for the classifying the disease types and in this paper the author used 9 different varieties of leaf diseases of tomato, grape, corn, apple and sugarcane. In this paper the training is conducted on the system for nearly about 50 epochs and they used 22 size of batch. In this model with the help of categorical cross-entropy, Adam optimizer is conducted. Accuracy obtained is 96.5%. Ch Usha Kumari et al., [3] developed a system that deploys the methods of K- Means clustering and Artificial Neural Network and performs computation of various features like Contrast, Correlation, Energy, Mean, Standard Deviation and Variance were performed. The major limitation was that accuracy of four different diseases was analyzed and the average accuracy is comparatively low.

Merecelin et al., [4] put forward a detailed study of identification of disease in plant (appleand tomato leaf) using the concepts of CNN. The model was trained on leaf image dataset containing 3663 images of apple and tomato plant leaf achieving an accuracy of 87%.

Manuscript received on September 09, 2021.
Revised Manuscript received on September 16, 2021.
Manuscript published on September 30, 2021.

* Correspondence Author

**Rinu R***, Department of Computer Science and Engineering, UVCE, Bangalore University, Bengaluru, India. Email: rinuraghu97@gmail.com

**Manjula S H**, Department of Computer Science and Engineering, UVCE, Bangalore University, Bengaluru, India. Email: manjulash@bub.ernet.in

Jiayue et al., [5] performed the recognition of tomato fruits with disease, the technique called YOLOv2 CNN was used. YOLOv2 is based on regression model and uses a target detection algorithm, which exhibits fast detection speed and good accuracy. The MAP (mean Average Precision) was estimated to be around 97%. The major limitationof the paper was the need to perform different tuning if the images. Robert G et al., [6] proposed a system using CNN to detect the type of tomato leaf diseases. This paper reported that the F-RCNN trained model obtained 80% confidence score, while accuracy of 95.75% was obtained by the Transfer Learning model. The automated image seizing method registered 91.67 % accuracy. Halil et al., [7] proposed a deep learning model was deployed with two different deep learning network architectures, Alex Net and then Squeeze Net. The training and validation of these deep learning networks were performed on the Nvidia Jetson TX1. The Alex Net achieved an accuracy of 95.6% and on the other hand Squeeze Net model achieved an accuracy of 94.3%. Sabrol et al., [8], the authors used an easy and uncomplicated mechanism is utilized for doing the process of classification of the different kind of diseases that occur in tomato leaves namely Early blight, Yellow curl virus, late blight, Mosaic virus, Bacterial spot and Healthy. The dataset contained 400 images clicked using a digital camera. Supervised learning method have been used for classification, where in the accuracy achieved was high, but decision tree has certain disadvantages – if instance of noisy data overfitting happens.

## III. KEY TECHNOLOGIES

### A. Deep Learning

Deep learning, a category of machine learning algorithms which uses various layers to do the extraction of higher level from the raw input. Deep learning is a machine learning method that instruct a computer to do filtration of inputs across the layers Deep learning illustrates the way human brain does the filtration of information. Many deep learning techniques utilizes the neural network architectures. The term “deep” cite to the various hidden layers present inside neural network. In contrast to this conventional neural network that consists of 2-3 hidden layers, the deep neural networks can have as much as one hundred and fifty.

### B. Convolutional Neural Network

One variant of deep neural networks is called as convolutional neural networks (CNN). A CNN combines well-read features with input data, and then it uses 2D convolutional layers, and hence makes this architecture more suitable for processing 2D data, like images. CNNs abolish the demand for manual feature removal and extraction for the classification of the images. The CNN model of its own extracts features straight from images. The features that are extracted aren’t pre-trained; they are well-read while the network is trained on few groups of images. The Convolutional Neural Network (CNN) model has numerous of layers which execute the processing of image in convolutional layers include- Input layer, Output Layer, Convo Layer, , Fully, Soft-max layer, Connected layer, Pooling Layer.

### C. VGG 16 Model

VGG16 is a CNN model used for Large-Scale Image. There are two tasks to be performed for best recognition of plant diseases. The first is to detect objects within an image coming from *several* classes, which is called object localization. The second is to classify images, each labelled with one of *several* categories, which is called image classification. The CNN model has seven different layers. Each layer has certain information processed in them. Those seven layers are as follows: Input layer, Output Layer, Convolutional Layer, Fully, Soft-max layer, connected layer, Pooling Layer.

**Input layer**: It contains data in the form of image. The parameters include height, width, depth and color information of the image (RGB). Input size is fixed to 224 X 224 RGB image.

**Convo layer**: Convolutional layer is also called as feature extraction layer. This layer extracts the prominent features from the given collection of images using dot products of the image dimensions.

**Pooling Layer:** The pooling layer helps to reduce the computational power in order to process the data by decreasing (or) reducing the dimensions of the featured matrix obtained by using the dot products.

**Fully connected layer:** It comprises of loads, neurons and biases. It connects neurons from one convolutional layer to another.

**Softmax Layer/ Logistic Layer:** Softmax executes multi-classification. Logistic layer executes the binary classification. It determines the probability of the presence of a given object in the image. If the object is present in the image, then the probability is ‘1’otherwise it is ‘0’.

**Activation Function- ReLU:** It transforms the total weighted input through the node and puts it into the operation, activates the node. Rectified Linear Unit (ReLU) is an activation function used in the neural networks for convolutional operations.

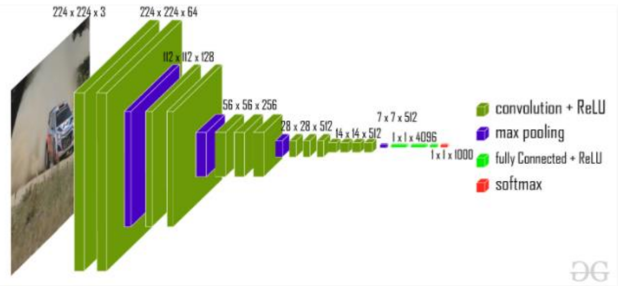


**Figure 1: VGG16 Model**

## IV. PROPOSED MODEL

On the yearly basis the agriculture field witnesses’ major reduction in productivity performance and crop losses since farmers are unable to identify the disease type in the initial stages. The farmer’s naked eye observations of leaves are often not capable of identifying the disease type and often need an expert to make predictions.

Retrieval Number: 100.1/ijrte.C64580910321
DOI: 10.35940/ijrte.C6458.0910321
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

These losses have a major impact on the production and thus on the lives of farmers. The proposed model designed and developed an automated system which is used for identification of plant diseases that helps to determine if the plant is infected by a disease or not.The following are the steps in which the process is carried out:

1. Acquiring of the plant image dataset 14 different varieties of plant images containing 38 different classes of plant diseases.
2. Pre-processing of the image in different convolutional layers.
3. Classification of plant diseases stating if the given plant leaf image is diseased or healthy.

## V. DESIGN

### A. System Architecture

The proposed System architecture comprises of data acquisition from a huge dataset, processing at different convolutional layers and then the classification of plant diseases which declares if the plant image is of a healthy class or diseased class.

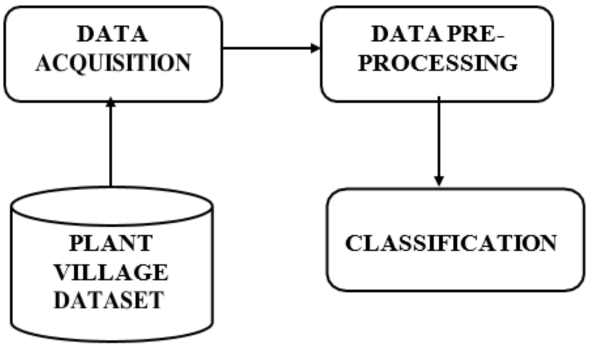


**Figure 2: System Architecture**

### B. Data Flow Diagram

Data Flow Diagrams (DFDs) describe the processes of how the transfer of data takes place from the input till prediction of the corresponding output.

#### 1. Data Flow Diagram – Level 0

The DFD Level 0 depicts the users to input the image of the plant leaves. The system in turn detects and recognizes the plant leaf disease.

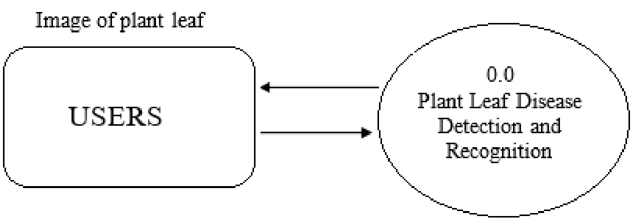


**Figure 3: Level 0 Data Flow Diagram**

#### 2. Data Flow Diagram – Level 1:

The Figure 4 displays the DFD Level 1, where the CNN model takes the image from the training dataset and then CNN model predicts the type of disease of the leaf.

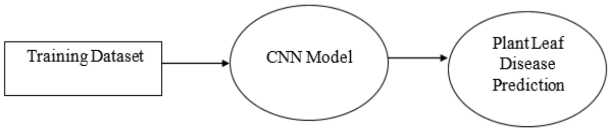


**Figure 4: Level 1 Data Flow Diagram**

#### 3. Data Flow Diagram – Level 2:

DFD Level 2 goes one step deeper into parts of 1-level DFD. It can be used to plan or record the specific/necessary detail about the system's functioning.

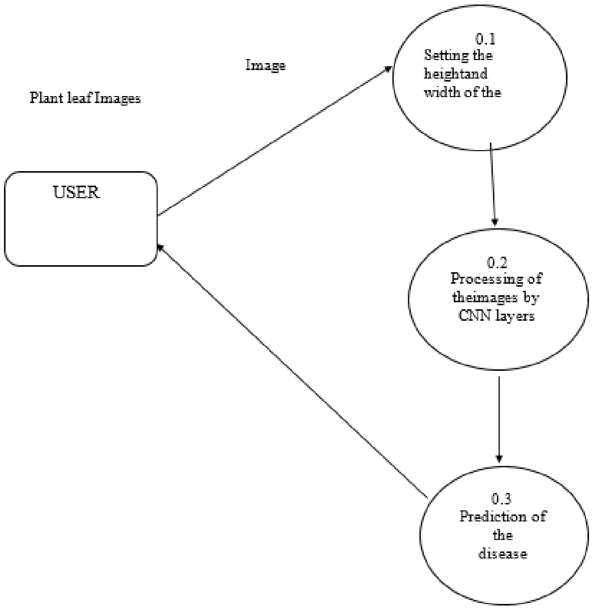


**Figure 5: Level 1 Data Flow Diagram**

## VI. IMPLEMENTATION DETAILS

### A. Platform

"Python" is used as the development platform. It includes libraries which is necessary to run the code. The libraries used are- Tensor Flow, NumPy ,Pandas, Keras .The software platform used for the implementation purpose :

**Google Collaboratory**

Google provides a no-cost cloud service application relying on Python Jupyter Notebook. Training code is written on google collaboratory.

**Sublime Text**

Sublime is a Text editor and it more convenient to work with code base. The testing code is written on this platform since it provides a responsive GUI.

## VII. EXPERIMENTAL ANALYSIS AND RESULTS

### A. Data Analysis

The dataset that is used in this proposed system project is the Plant Village dataset and it was downloaded from the Kaggle website; the dataset consisted of images of diseased and healthy plant leaf images.

Upon exploration, we found that the dataset did not have any missing values. The dataset was further explored to understand the various spices and diseases of plant leaf. The dataset consisted of 14 different plant varieties. The training dataset has a total of 54305 images in total.

## B. Result Analysis

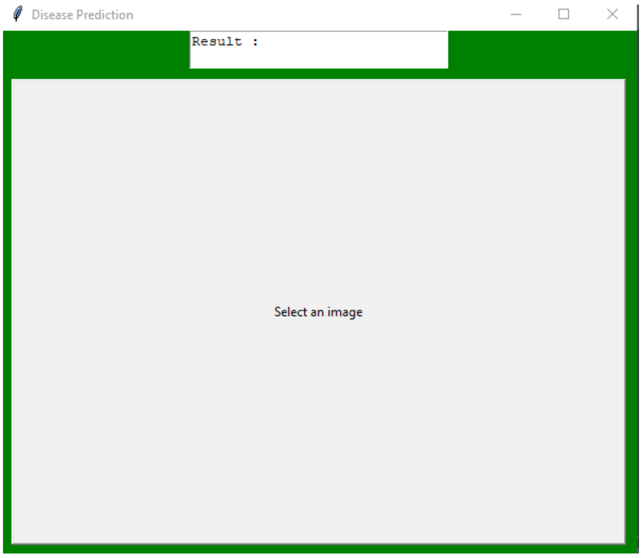


**Figure 6: GUI of the plant disease detection application**

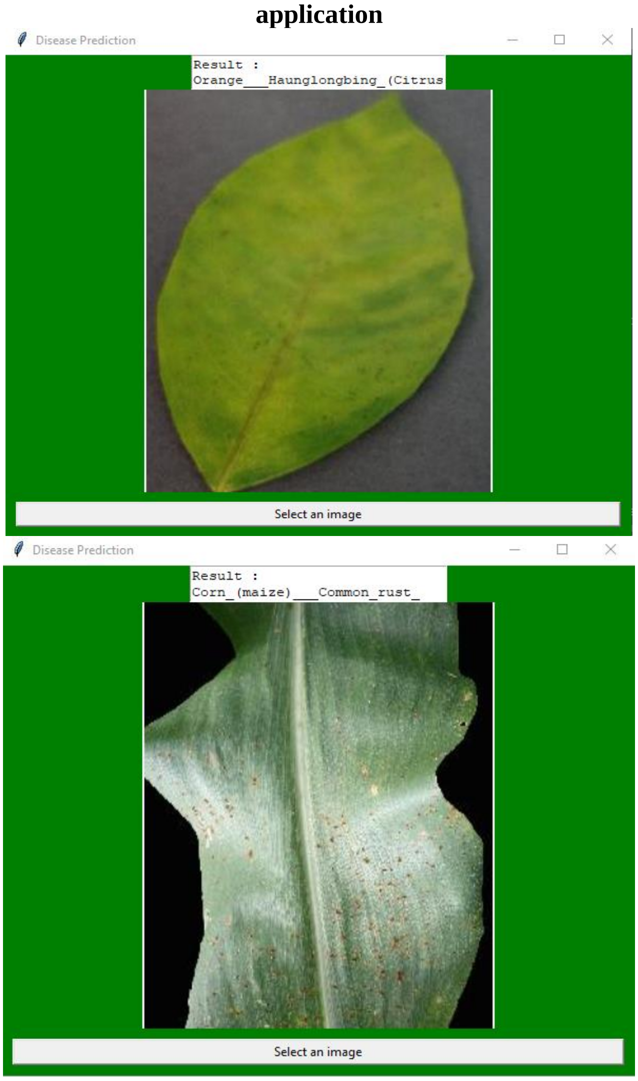


**Figure 7: The application displaying the output disease of orange and corn plant on accepting the input image**

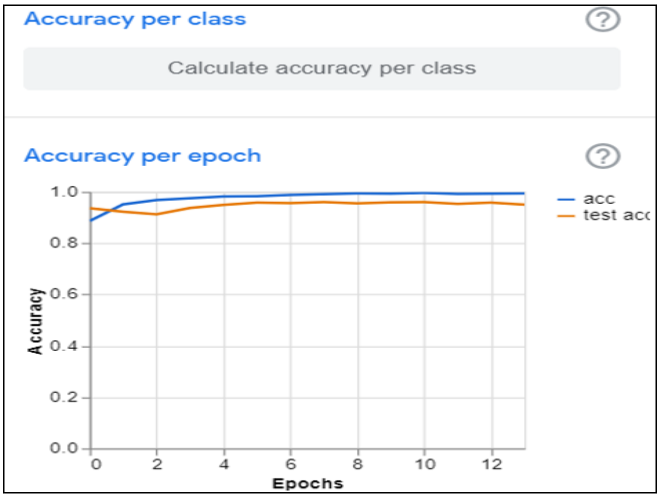


**Figure 8: Accuracy graph- accuracy vs epochs**

The graph shown here depicts the two curves- Blue curve showing training accuracy.Orange curve showing testing accuracy. The accuracy achieved is around 94.8%.

## VIII. CONCLUSION

A large part of the Indian population relies on agriculture, hence it becomes very essential to detect and recognize the leaf diseases that results in losses, since agriculture is critical to the growth of the economy. This project based on deep learning approach called CNN is utilized to build 13 different plant leaf disease identification, detection and recognition system. This approach utilized a minimum set of layers to identify the diseases of seven classes. The neural network is trained with Plant Village dataset. A Graphical User Interface is designed for this system. This GUI permits the user to choose the images from the dataset. User can select any image from the dataset and the image gets loaded, following which the prediction of the disease will be shown on the User Interface. Convolutional neural network, trained for identifying and recognizing the plant leaf disease, could classify and predict the diseases correctly for almost all the images with few anomalies thus and obtained 94.8% accuracy.

## ACKNOWLEDGMENT

This study and research is carried out in Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bengaluru, Karnataka, India, under the guidance of Dr. Manjula S H. I also thank our Head of the Department Dr. Dilip Kumar S M and The Principal of the institution, Dr. H N Ramesh for providing us all the required facilities and suitable environment to successfully complete this research.

## REFERENCES

1. Eftekhar Hossain, Md. Farhad Hossain, Mohammad Anisur Rahaman, "A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease Using KNN Classifier", Proceeding of the International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox's Bazaar, Bangladesh, 2019.

2. Sammy V. Militante, Bobby D. Gerardo, Nanette V. Dionisio, "Plant Leaf Detection and Disease Recognition using Deep Learning" Proceeding of the IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), Yunlin, Taiwan, 2019, pp 579-582
3. Ch. Usha Kumari, S. Jeevan Prasad, G. Mounika, "Leaf Disease Detection: Feature Extraction with K-means clustering and Classification with ANN "Proceedings of the 3rd International Conference on Computing Methodologies and Communication (ICCMC)", Erode, India, 2019, pp 1095-1098
4. Mercelin Francis, C. Deisy, "Disease Detection and Classification in Agricultural Plants Using Convolutional Neural Networks — A Visual Understanding", Proceeding of the 6th International Conference on Signal processing and Integrated network, Noida, India, 2019, pp 1063-1068.
5. Jiayue Zhao, Jianhua Qu, "A Detection Method for Tomato Fruit Common Physiological Diseases Based on YOLOv2 "Proceeding of the 10th International Conference on Information Technology in Medicine and Education", Qingdao, China, China, 2019
6. Robert G. de Luna, Elmer P. Dadios, Argel A. Bandala, "Automated Image Capturing System for Deep Learning-based Tomato Plant Leaf Disease Detection and Recognition" Proceeding of the IEEE Region 10 Conference, Jeju, South Korea, 2018, pp 1414-1419.
7. Halil Durmus, Ece Olcay Gunes, "Disease detection on the leaves of the tomatoplants by using deep learning" Proceeding of the 6th International Conference of theAgro-Geoinformatics, 2016, Fairfax, VA, USA
8. H. Sabrol , K. Satish, "Tomato plant disease classification in digital images using classification tree" Proceeding of the International conference on communication and signal processing, Melmaruvathur, India, 2016 pp 1242-1246
9. Usama Mokhtar, Mona A.S. Ali, Hesham Henfy, "Tomato leaves diseases detection approach based on support vector machines" Proceeding of the 11th International Computer Engineering Conference, Cairo, Egypt, 2015 pp 246-250
10. Suma V.R Amog Shetty, Rishab F Tated, Sunku Rohan, Triveni S Pujar, "CNNbased Leaf Disease Identification and Remedy Recommendation System ", Proceedings of the Third International Conference on Electronics Communication and Aerospace Technology, Coimbatore, India, 2019 pp 395-399.
11. Akshay K, Vani M, "Image Based Tomato Leaf Disease Detection", Proceedings of the 10th International Conference on Computing, Communication andNetworking Technology, Kanpur, India, 2019.
12. Sharada P Mohanty, David P Hughes, and Marcel Salath, "Using deep learning for image-based plant disease detection". In: Frontiers in plant science 7 (2016), p. 1419.
13. S. D. Khirade and A. B. Patil. "Plant Disease Detection Using Image Processing", Proceeding of the International Conference on Computing Communication Control and Automation. Feb. 2015, pp. 768–771.
14. M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep Learning for Tomato Diseases: Classification and Symptoms Visualization", vol. 31, no. 4, pp. 299–315, 2017.
15. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks". Proceeding of the Advances in neural information processing systems, 2012, pp. 1097–1105.
16. K. Elangovan and S. Nalini, "Plant disease classification using image segmentation and SVM techniques", International Journal of Computational Intelligence Research, vol. 13(7), pp. 1821-1828, 2017.
17. Y. Dandawate and R. Kokare, "An automated approach for classification of plant diseases towards development of futuristic decision support system in indian perspective," Proceedings of the International Conference Advances in Computing, Communications and Informatics (ICACCI), 2015, pp. 794-799.
18. Mohanaiah, P., P. Sathyanarayana, and L. GuruKumar. "Image texture feature extraction using GLCM approach." International journal of scientific and research publications 3, no. 5 (2013)
19. Monika Jhuria, Ashwani Kumar, and Rushikesh Borse. "Image processingfor smart farming: Detection of disease and fruit grading." In 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013), pp. 521-526. IEEE, 2013.
20. Husin, Zulkifli Bin, Ali Yeon Bin Md Shakaff, Abdul Hallis Bin Abdul Aziz, and Rohani Binti S. Mohamed Farook. "Feasibility study on plant chili disease detection using image processing techniques." In 2012 Third International Conference on Intelligent Systems Modelling and Simulation, pp. 291-296. IEEE, 2012.
21. Mrunalini R. Badnakhe, Prashant R. Deshmukh, Infected Leaf Analysis and Comparison by Otsu Threshold and k-Means Clustering, International Journal of Advanced Research in Computer Science and Software Engineering.


## AUTHORS PROFILE

**Rinu R** is a student at Department of Computer Science and engineering, UVCE, Bangalore University, Bengaluru, Karnataka, India.

**Dr. Manjula S H** is currently working as a Professor, Department of Computer Science and Engineering, UVCE, Bangalore University, Bengaluru. She is a member of Academic Senate, VTU Belgavi. She obtained BE, and ME. in Department of Computer Science and Engineering from UVCE, Bengaluru. Ph.D. in Computer Science and Engineering, Chennai. She has published 120 technical articles in referred journals and international conference and authored three books. Her research interests are in the field of IoT, Cloud Computing, Wireless Sensor Networks and Data mining, AI & ML.