EasyChair Preprint

№ 9760

ORPN Algorithm to Diagnosis and Detect Plant Diseases Based on Image Segmentation Using Machine Learning Techniques

Rajesh Kanna Rajendran and Mohana Priya Thiruvengadam

February 22, 2023

ORPN ALGORITHM TO DIAGNOSIS AND DETECT PLANT DISEASES BASED ON IMAGE SEGMENTATION USING MACHINE LEARNING TECHNIQUES

Abstract

The Plant disease modernization in agricultural land is the main concern for every country, as the food demand is increasing at a fast rate due to an increase in population. Indian economy is extremely dependent of farming productivity. Therefore, in field of cultivation, detection of disease in floras plays an important role. Also, the increased use of expertise today has increased the efficacy and accuracy of noticing diseases in plants and animals. The thought of classification in machine learning techniques deals with the problem of identifying to which set of categories a new population belongs. Our attention is to illuminate the facts about the diseases and how to perceive them promptly with artificial intelligence. It proposes to deliberate the use of Ai techniques to detect diseases in plants robotically. In this research article, the O-RPN(Optimized Region Proposal Network) is utilized to identify and localize the leaves in complex surroundings. O-RPN Algorithm comprises the feature of indications through Chan–Vese (CV) techniques. The CV algorithm based on region shows promising results for segmenting images free of noise and weak edge. Furthermore, different data sets related to plant diseases are compared with CNN and SVM.

Keywords: Image Segmentation, Feature Extraction, RPN,CNN,SVM

INTRODUCTION

The usage of skill in the finding and search process to increase the accuracy and reliability of these processes. For example, the people who use the modern technology to analyse the diseases that arise unexpectedly are at a higher chance of controlling them than those that do not. Crop illnesses are a important hazard to human existence since they are likely to lead to droughts and starvations. The modern techniques like Computer vision is a form of artificial intelligence (AI) that performs computers to understand and identify objects. It is mainly useful in testing drivers, automobile industries, avoid parling space, and driverless car and now in medical processes to detect and analyse objects [1]. In specific, machine learning and deep learning techniques have achieved remarkable results in biomedical image classification process.

Classical approaches rely on image pre-processing and the extraction of features which are then fed into one of the AI techniques. Popular procedure selections are Support Vector Machines (SVM), k-Nearest Neighbours (k-NN), Fully Connected Neural Networks (FCNN), Decision Trees, Random Forests etc. In the last years, the researchers shifted almost completely to the AI methods for image classification tasks.

EXISTING SYSTEM

Leaf image outline is that the key downside in leaf image identification. Till now, several form choices are extracted to explain the leaf image form. Though, there's no accurate application to categorize the leaf once capturing its image and identifying its characteristics, however. In plant leaf image classification leaf is categorized and maintained its completely different morphological options.

RELATED WORK

Sanga et al. [1] advanced a disease finding tool for banana plants with five different CNN architectures. These architectures were VGG-16, ResNet-152, ResNet-50, ResNet-18, and InceptionV3. They originate that ResNet-152 outstripped others with an accuracy of 99.2%. They also established a mobile application so that farmers could easily detect diseases in banana plants by uploading leaf images of their banana plants with their smartphones. This mobile claim used the InceptionV3 model for disease prediction with 99% confidence.

Mohameth et al.[2] used different modern CNN architectures and different classifiers for automatic plant disease detection on the Plant Village dataset. They employed VGG-16, ResNet-50, and GoogLeNet CNN architectures for feature extraction, and for classification, they used k-Nearest Neighbour and Support Vector Machine (SVM) classifiers. They observed that SVM with ResNet-50 outperformed others with an accuracy of 98%.

Khamparia et al.[3]proposed a Deep Convolutional Encoder Network system for seasonal crops disease identification. They considered 900 leaf images of three crops: potato, tomato, and maize, distributed in six classes. They achieved 100% training accuracy while the testing accuracy of their model was 86.78%. Since the training accuracy was much higher as compared to the testing accuracy, so there was a chance that the trained model overfitted on the training data. In their paper, they have also mentioned that their system used approximately 3.3 million training parameters, which is much higher than the number of training parameters 9,914 used

in the proposed work. They have used Autoencoder and CNN for seasonal crop disease identification. On the other hand, the proposed novel hybrid model is based on CAE and CNN.

Ferentinos[4] employed five different modern CNN architectures named Alex Net, AlexNet OWTBn, GoogLeNet, Over feat, and VGG for plant disease detection using the Plant Village dataset. In his paper, he found that VGG outperformed other CNN architectures with an accuracy of 99.5% using 138 million trainable parameters.

Tiwari et al. [5]proposed an automatic disease detection system for potato plants. This system used different CNN architectures such as VGG-19, VGG-16, and InceptionV3 for feature extraction and different classifiers such as Logistic Regression, k-Nearest Neighbour classifier, Support Vector Machine (SVM), and Neural Network for disease detection. They concluded that VGG-19 with Logistic Regression outperformed others with an accuracy of 97.8%.

MATERIAL AND METHODS

In this research work, introduced a novel method to deal with the problems raised by such multifaceted images, for simple and plant leaves. A first segmentation step based on graph-cut approach is first performed and later used to guide the evolution of leaf boundaries, and implement classification algorithm to classify the diseases and endorse the nourishments to affected leaves.

Image Classification Steps

The projected image classification method is separated into the following stages:

- Image acquisition
- Pre-processing
- Segmentation
- Disease Prediction
- Fertilizer Recommendation

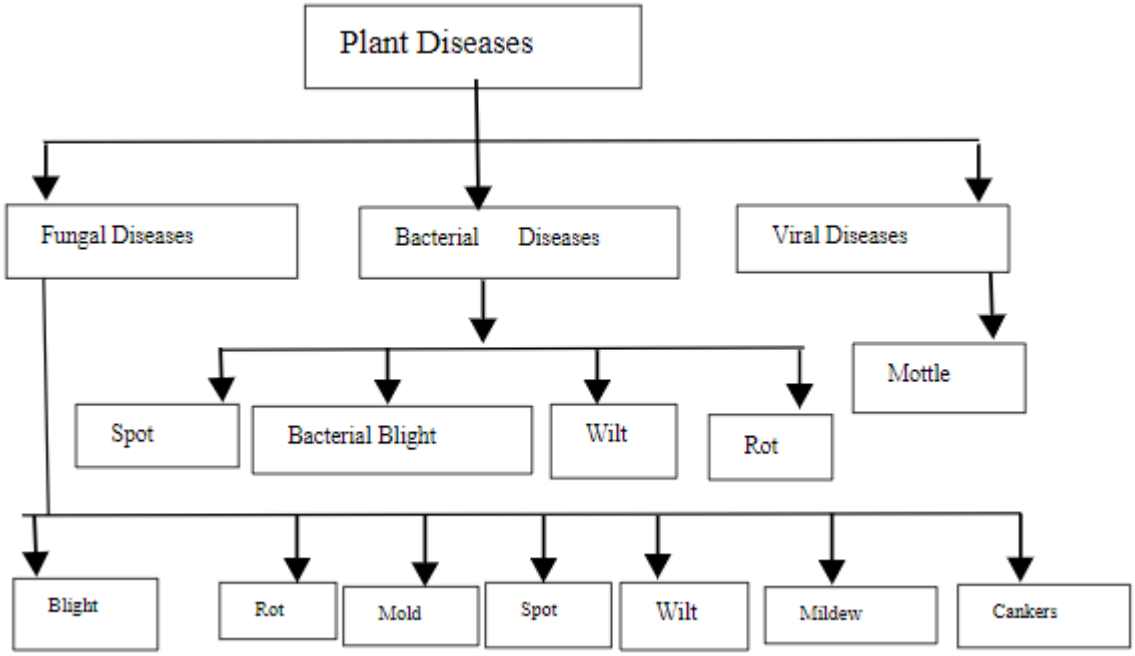


Figure 1: Classification of Plant Diseases

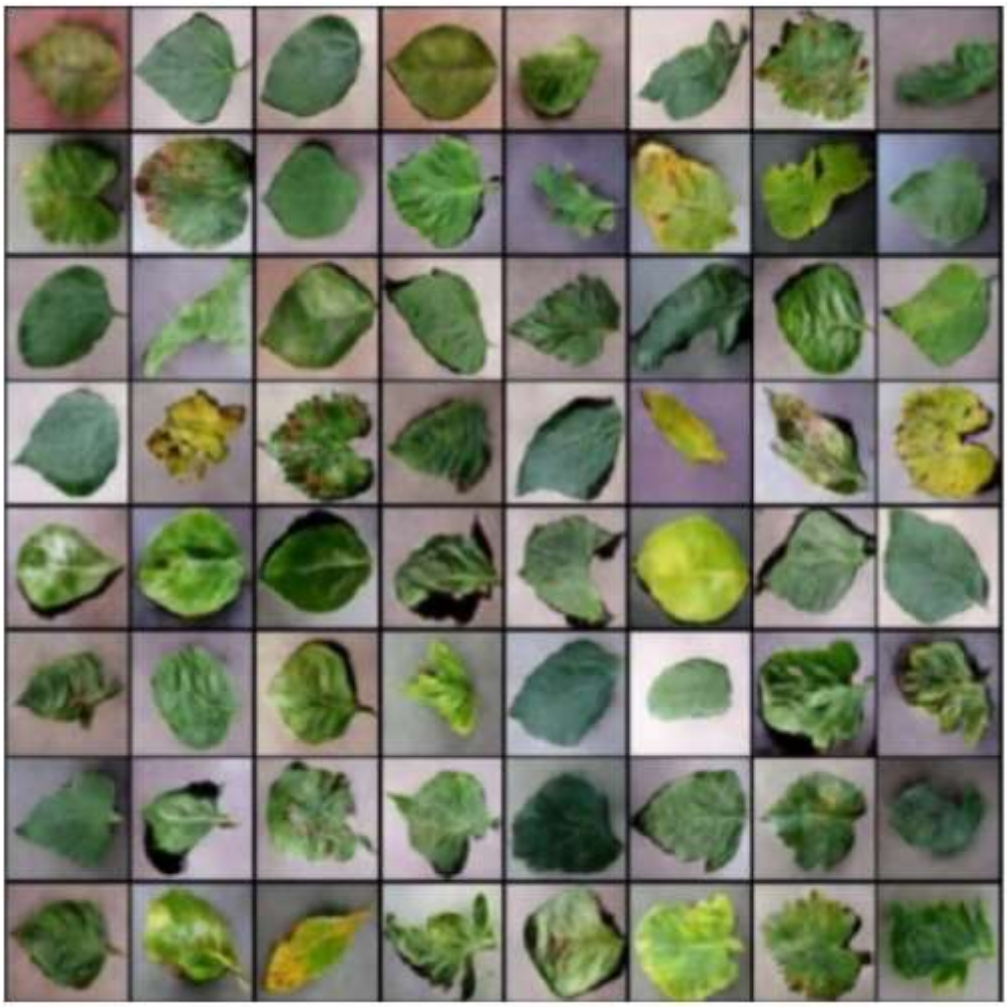


Figure 2: deafferent phenotypes of various plants

PROPOSED METHODOLOGY

Image Processing

It is a development of performing about actions on an image, in order to get an improved image or to extract some useful information from it. Gaining, Segmentation, Augmentation, Feature Extraction. Image acquisition is the primary key step of digital image processing. Image acquisition is modest process when given an image that is previously in digital form. Commonly, the image acquisition stage includes pre-processing, such as scaling etc. Image Segmentation actions partition an image into its constituent parts or objects. Usually, autonomous segmentation is the most problematic tasks in digital image processing. A uneven separation procedure that carries a long way near successful resolution of imaging problems that require objects to be identified individually. Image Augmentation is a method that is used to affectedly enlarge the dataset, parameters that are generally used to increase the data sample count are zoom, shear, rotation, pre-processing function and so on. Image feature extraction is one of the most important segments in this project. Feature selection works by selecting the best features based on unilabiate statistical tests. The features are carefully selected based on their unique differences between the different types of leaves.

Optimized RPN Algorithm

In this research paper, plant leaf disease recognition progressions an appropriate and faultless dataset is required. First step of the proposed system is training the system. Machine learning algorithm-based dataset is used to perform the disease identification and feature extraction. Basically, as a training data healthy and unhealthy leaves are required. For this training data in this research used an commercial agricultural plants and particularly we take tomato and Beans plants. 500 images are collected from different places for train system. 750 images are composed for test the system for plant disease identification. we used Leaf Classification dataset from Kaggle repository. For more accurate data, need go to some farming place and take some decent and clear digital photo. After gathering, next processes are cleaning those data. Using python script compared the image from collected dataset. It is possible that data can be repeatable.

The full plant illness identification model based on Machine learning algorithms including three steps, the localization of plant leaves, the segmentation of plant images, the extraction of plant disease, and the identification of plant disease. The model used in this research paper mainly consists of the following three steps. The first phase is to locate the diseased leaves. The O-RPN algorithm is used to find and recover the diseased leaves in the complex situation. The second step is the segmentation of diseased plant leaves. CV algorithm is used to segment the plant image of diseased leaves. The third step is the identification of leaf disease species. The existing model is trained to understand the plant disease recognition in the simple background. This proposed research model compared with the well-known existing techniques in machine learning algorithm like CNN and SVM algorithm.

Feature extraction

The objective of feature extraction is to extract individual variances that can differentiate different identification features. On the other hand, the resolution is to extract leaf features from the high dimensional data to obtain useful information to define and illustrate the key plant leaf information

RESULTS AND DISCUSSION

On the Plant leaf classification dataset, the application of the proposed methodology was carried out. It comprises of approximately 3230 pictures from 2 distinct types of tomato leaf diseases. a Python was used to implement the model. The filter size is fixed as 5x5, while the amount of filters is gradually expanded as we move from block to block. To reduce the size of the feature maps, the max-pooling layer is used to speed up the training process770 pictures were set aside for testing out of the 3230 images and 2460 images were used for training. The optimization was performed in the proposed techniques using Adam optimizer as the loss function with categorical cross entropy to improve the image based plant disease prediction and diagnosis in agricultural land.

Performance Metrics

Precision: The proportion of positive identification is actually correct.

$$\text{Precision} = TP / (TP+FP) \quad \ldots..1$$

Recall: The proportion of actual positives is identified correctly.

$$\text{Recall} = TP / (TP+FN) \quad \ldots..2$$

F-Measure: Defined as the weighted harmonic mean of precision and recall.

F-Measure = 2TP / (2TP + FP + FN) …..3

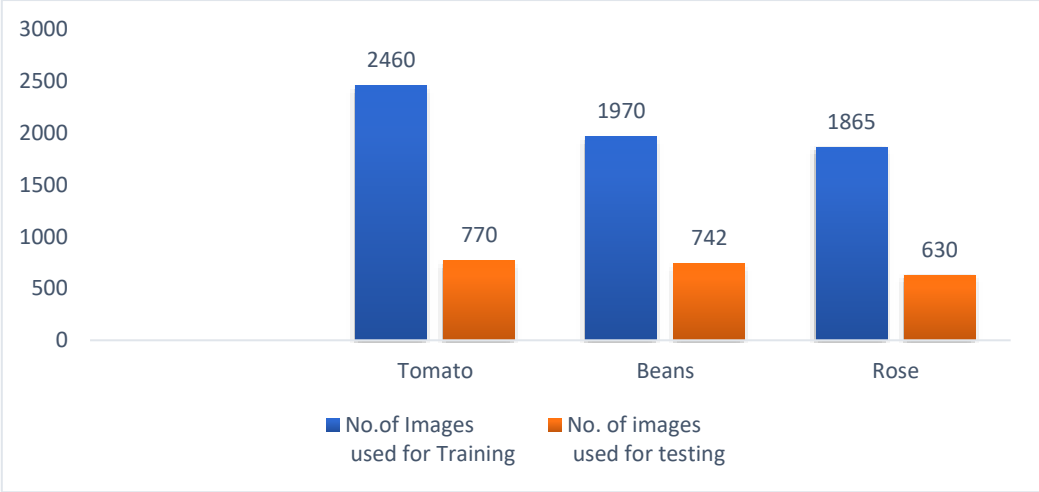
Accuracy: It refers to the closeness of a measured value to a standard or known value.

Accuracy = (TP + TN) / (FP + TP + FN + TN) …..4

The Precision, Recall, F-Measure and Accuracy for the O-RPN, CNN and SVM are calculated and given in table 1 the corresponding graph is given in Figure 3.

Table 1: Performance analysis of Proposed O-RPN vs Existing techniques among the image based plant disease prediction

Disease Samples	No. of Images used for Training	No. of images used for testing	Accuracy		
			O-RPN	CNN	SVM
Tomato	2460	770	98.75	96.89	97.86
Beans	1970	742	98.67	97.13	97.32
Rose	1865	630	98.32	97.89	96.54



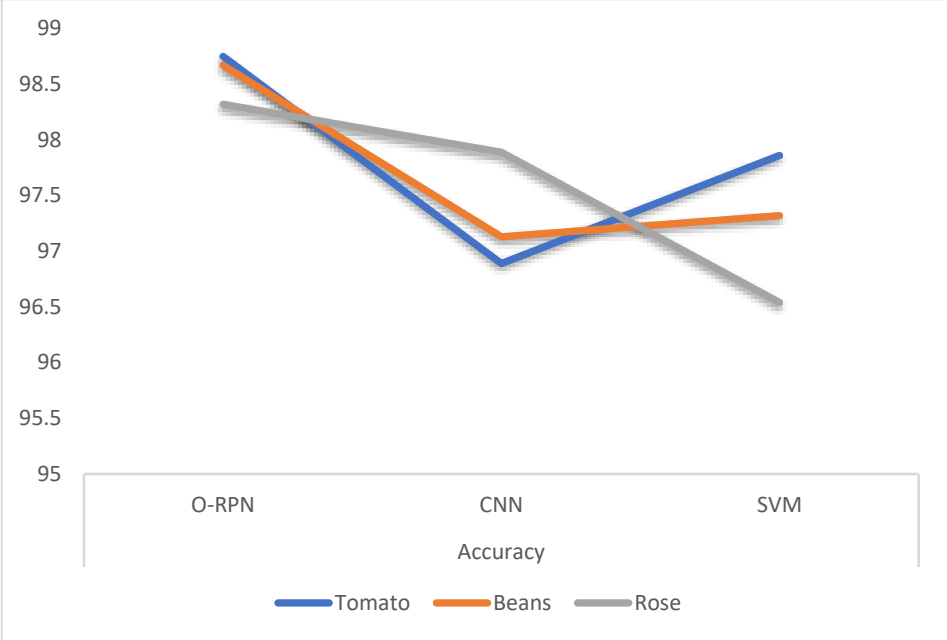


Figure 3: Comparison of O-RPN model with Existing techniques

The proposed research model O-RPN performs better and provides high accuracy when compared to existing techniques in this plant disease identification and extraction.

CONCLUSION

This research paper shows that the plant disease identification and recovery model based on Machine learning has the features of unsupervised, high accuracy, good universality, and high training efficiency. However, there are many challenges in accuracy practicability of plant disease detection in the complex environment. The O-RPN model not only adjusts to complex environments, but also rises the accuracy of identification. Compared with the traditional model, the model proposed in this paper not only promises the strength of the CNN and SVM, but also decreases the number and quality necessities of the model on the data set and obtains better results. Then, the proposed O-RPN model could help farming production personnel to prevent and cure the plant disease speedily.

REFERENCES

1. S.L. Sanga, D. Machuve, K. Jomanga,Mobile-based deep learning models for Banana disease detection Technol. Appl. Sci. Res., 10 (3) (2020), pp. 5674-5677
2. F. Mohameth, C. Bingcai, K.A. Sada,Plant disease detection with deep learning and feature extraction using Plant Village,J. Comp. Commun., 8 (6) (2020), pp. 10-22

3. Khamparia, G. Saini, D. Gupta, A. Khanna, S. Tiwari, V.H.C. de Albuquerque,Seasonal crops disease prediction and classification using deep convolutional encoder network,Circ. Syst. Sign. Proc., 39 (2020)
4. K.P. Ferentinos,Deep learning models for plant disease detection and diagnosis,Comput. Electron. Agric., 145 (2018), pp. 311-318
5. TD. Tiwari, M. Ashish, N. Gangwar, A. Sharma, S. Patel, S. Bhardwaj,Potato leaf diseases detection using deep learning,2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE, Madurai, India, India (2020), pp. 461-466.
6. C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich Going deeper with convolutions 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Boston, MA, USA (2015), pp. 1-9
7. P. Sharma, Y.P.S. Berwal, W. Ghai, Performance analysis of deep learning CNN models for disease detection in plants using image segmentation,Inform. Proc. Agric. (2019), pp. 1-9
8. L.C. Ngugi, M. Abelwahab, M. Abo-Zahhad,Recent advances in image processing techniques for automated leaf pest and disease recognition – a review, Inform. Proc. Agric. (2020), pp. 1-25
9. A. Karlekar, A. Seal,SoyNet: soybean leaf diseases classification,Comput. Electron. Agric., 172 (2020), p. 105342
10. K. Golhani, S.K. Balasundram, G. Vadamalai, B. Pradhan,A review of neural networks in plant disease detection using hyperspectral data,Inform. Proc. Agric., 5 (3)