A Deep Learning Approach for Plant Material Disease Identification

R K Tripathi
Dept. of CEA, IET, GLA University, Mathura, 281406, UP, INDIA
rajesh.tripathi@gla.ac.in

Abstract. Plant Material Disease Identification is essential for the food safety. To increase the crop production for the growing population of the world, the proper treatment is required on proper time to save the plant. Therefore, disease diagnosing on time is very important. This paper uses a deep learning convolutional neural network model to identify the plant disease. The pre-existing deep learning model Alexnet has been employed for plant disease identification in which an external feature of segmented plant material (leaves) is passed to the deepest fully connected layer. This combination of extracted feature by Alexnet and external feature of segmented plant material helps in plant disease identification. Experimental analysis has been done on a standard dataset Plant Village which has total 54,306 leaf images of 15 distinct plants having 38 diseases. The presented CNN approach worked well and outperformed to the existing approach.

Keywords: Plant material disease; feature fusion; deep learning; CNN

1. Introduction

Plant material disease identification and providing treatment on time highly essential to improve the food production and security for the large population of the world. As per the growth of the population of the world, agriculture production must be increased by a good number of percentage in coming decade to fulfill the requirements of the food. Generally, farmers use different chemicals such as fungicides, bactericides and nematicides for the treatment of plant diseases without knowing the actual disease due to less knowledge of plant diseases [1]- [2]. These chemicals cause for human cancerous diseases and also has an adverse effect on agro-ecosystem. Therefore, a plant disease recognition system is required that can detect and recognize the disease in early stage so that proper treatment can be given on time without using harmful chemicals and also sustain to the agro-ecosystem.

The disease in the plants material affects quality of vegetables, grains, fruits, and legumes which causes heavy loss in production of crop. In current scenario, advanced technology of molecular biology and biotechnology evolved revolution in the area of plant disease recognition [2]- [6], [15], [25]- [27].

Future prospects of the plant disease identification is to help to farmers of the world to diagnose the disease at early stage so that treatment can be done to increase the productivity. The proposed deep learning approach is robust to recognize the plant disease; the contribution of the proposed approach is as follows:

- The proposed deep learning approach has improved discriminative power after combining the external features of segmented plant material (leaves) to the activations of the deepest fully connected layer.
- Colored plant leaves and segmented plant leaves performed well and improved the recognition
accuracy of the plant diseases than existing segmentation based approaches and other algorithms.

- This approach has been evaluated on a large standard dataset PlantVillage and outperformed to the state-of-the-art approaches.

Remaing part of this paper is arranged as follows: Section II discusses the plant disease recognition state-of-the-art, Section III discusses the proposed approach, Section IV presents the result analysis, and section V presents the conclusion.

2. Related Work
This section presents the progress of the state-of-the-art in the field of plant material disease identification. In the recent years, many plant material disease recognition methods have been developed. Historical approaches of past decade have been supplemented by integrated pest management (IPM) approaches [1]. In the beginning of this decade, many works [2]-[5] have been presented by using the shape, color and texture features to detect the plant diseases after analyzing diseased leaf images. Mohanty et al. [6] proposed a deep learning approach for plant disease recognition. Aravind et al. [7] utilized two existing deep learning models to detect the diseases in the tomato crop of PlantVillage dataset. There are a few image segmentation approaches and classifiers like decision trees, support vector machines (SVM), and neural networks, etc., are used in both the diagnosis process of plant and human skin disorders [8]. There are various approaches in visible light including neural networks, real-time monitoring, and classification have discussed in [9]. Kulkarni et al. [10] presented a segmentation technique using CIE L*a*b color space and utilized a Gabor filter to generate input for neural network, that yields 91% recognition accuracy of the plant disease. Camargo et al. [11] used shape, lacunarity, texture fractal dimension, gray levels, dispersion, Fourier descriptors and gray histogram discrimination for the disease identification. There are several researchers have presented the different solutions for specific diseases such as corn diseases by [12] and citrus pathogens by [13] etc.

Deep learning approaches has been described based on the layering architecture of convolutional neural network which helps in identification of diseases of plants. Convolutional Neural Network (CNN) works better for large dataset and gives better accuracy. In the recent years, many deep learning approaches have been proposed [14]-[22]. Amara et al. [15] used LeNet architecture over the PlantVillage dataset to detect the diseases of Banana leaf i.e. black sigatoka and black speckle. DeChant et al. [16] proposed a CNN pipeline to detect the diseases of Corn images and achieved 97%. Fuentes et al. [18] proposed a CNN approach to detect tomato diseases and achieved 83% accuracy of own dataset. Lu et al. [20] presented a deep learning approach to detect the diseases of rice plants over own dataset and achieved 95% accuracy. Brahimi et al. [21] employed CNN architecture using Alexnet and GoogleNet to detect the plant diseases. Oppenheim and Shani [22] used VGG model for Potato disease detection using own dataset and achieved 96% accuracy. Recently, Barbedo [14] used a pre-trained GoogleNet and applied transfer learning to detect the 10 different plant diseases. Most of the deep learning approaches performed best in case of colored plant leaves for disease recognition. A few deep learning approach used segmented plant leaves for disease recognition which provided better results than the grayscale plant leaves. Yosuke et al. [23] discussed analysis of a CNN architecture based on Inception-V3 in terms of accuracy, recall, precision, and f-measure. Sibiya et al. [24] presented a CNN model to detect the disease on maize plant only of Plant Village [6] dataset and achieved accuracy 92.85%. Recently, Lobna et al. [25] employed a VGG-16 pre-trained model on PlantVillage [6] dataset to improve the result and achieved 95.87% accuracy and achieved 98.67 after applying Gaussian filtering.

3. Proposed Deep Learning Approach
In this section, we described the layering architecture of convolutional neural network which helps in identification of diseases of plants. Convolutional Neural Network works better for large dataset and gives better accuracy. The proposed deep learning approach consists of 3 basic steps: plant material (leaves) pre-processing and coulered segmentation, Alexnet architecture is applied over coloured input images, and then segmented feature is fused at fully connected layer to improve the accuracy of the system.
3.1. Plant Material (Leave) Image Pre-processing
Images of the plants are resized to 200×200 dimension using as training and testing of deep convolutional neural network. Colored plant material (leaves) are converted into grayscale dataset. A hybrid method of Otsu and median filter [27] have been applied over plant material (leaves) for color segmentation as shown in Fig. 1.

Fig.1. presents the segmented colored image using Otsu and Median Filtering [27].

3.2. Deep feature representation of Plant Material Disease
A deep CNN is the best approach for the plant disease recognition in terms of accuracy. In this, deep Alexnet CNN architecture has been used and an external feature is passed to the last layer of this architecture to increase the strength of discriminative features so that recognition of diseases of plants can be improved. In this deep learning architecture, 5 convolutional, 3 pooling and 3 fully connected layers are arranged where every convolutional layers are followed by rectified linear unit (ReLU). Then, illumination variations are handled by adding two local layers for normalization. Then, a regularization process is applied using dropout over all fully connected layers due to the large parameters i.e. fc_6 = 4096, fc_7 = 4097, and fc_8 = 10575 over Plant Village dataset. The extracted features from fully connected layer are used for plant disease representation. Figure 2 presents the architecture of Alexnet model in which segmented image feature is fused at last layer that helps in the improvement of disease identification accuracy.

Fig. 2. presents the proposed CNN model Alexnet using Segmented Feature Fusion.
3.3. External Feature Fusion
Features of segmented images of plant diseases are given as set of n input of each category to the fully connected layer -normalized fc7 features. In addition to the Deep CNN features, an externally computed feature of coloured segmented images is fused with the activations of the deepest layers. Features are inserted in the fully connected layer fc6. Here, feature fusion means the concatenation of the external features with the features of fully connected layer 6. At test time, the activation of fc7 with inserted external features are computed. A normalization and distance measure Euclidean is applied to classify correctly to the plant diseases. A deep neural network is the best approach for the plant disease recognition while dataset is of large size.

4. Experimental Results and Discussion
Evaluation of the proposed deep learning approach has been done on standard dataset Plant Village [6]. This dataset consists of total 54,306 plant material (leaves). There are 15 different plants and 38 distinct plant material diseases. Figure 3 shows the normal and diseases plant leaves of different plants. Two versions such as the coloured and segmented images are created for Plant Village [6] dataset for using deep CNN and as an external feature fusion respectively.

![Fig. 3. Images of 38 different normal and plant diseases leaves of standard dataset PlantVillage.](image-url)
This approach has been implemented on a machine having configuration-8th Gen i7, 2.20 GHz processor, 8 GB RAM Computer with 4GB NVIDIA GEFORCE GTX. To measure the performance of the presented approach, comparison has been done with best performer of the plant disease identification on Plant Village Dataset Deep CNN [6], recent performances- Inception V3 [28], ResNet-50 [28], DenseNet169 [28], VGG [16], VGG-16 with Gaussian approach [16].

Table I presents the comparison of precision, recall rate and F-measure while considering train-test data ratio 20%-80%. Overall accuracy with the specified train-test ratio is also comparable with Deep CNN [6] and outperformed by small margin. Table 2 presents the performance in terms of precision, recall rate, F-measure while considering data divided into 80%-20% train-test ratio. Overall accuracy is better than Deep CNN [6], Inception V3 [28], ResNet-50 [28], DenseNet169 [28] approaches and outperformed with small margin.

Table 1. Performance Analysis of the proposed and the existing deep CNN [6] approach on colored, grayscale and segmented datasets 20%-80% train-test ratio.

| Approach          | Input Images | Precision | Recall | F-measure | Overall Accuracy |
|-------------------|--------------|-----------|--------|-----------|------------------|
| Deep CNN [6]      | Colored      | 0.9742    | 0.9737 | 0.9736    | 0.9738           |
|                   | Grayscale    | 0.9368    | 0.9369 | 0.9361    | 0.9371           |
|                   | Segmented    | 0.9727    | 0.9727 | 0.9724    | 0.9726           |
| Proposed Approach | Colored and  | 0.9748    | 0.9739 | 0.9737    | 0.9741           |
|                   | Segmented    |           |        |           |                  |

Table 2. Performance Analysis of the proposed and an existing deep CNN [6] approach on colored, grayscale and segmented datasets on transfer learning with 80%-20% train-test ratio.

| Approach          | Input Images | Precision | Recall | F-measure | Overall Accuracy |
|-------------------|--------------|-----------|--------|-----------|------------------|
| Deep CNN [6]      | Colored      | 0.9928    | 0.9927 | 0.9927    | 0.9928           |
|                   | Grayscale    | 0.9728    | 0.9727 | 0.9726    | 0.9725           |
|                   | Segmented    | 0.9893    | 0.9891 | 0.9891    | 0.9892           |
| Inception V3 [28] | Colored      | 0.9200    | 0.9400 | 0.9300    | 0.9710           |
| ResNet-50 [28]    | Colored      | 0.9200    | 0.9400 | 0.9400    | 0.9820           |
| Dense Net 169 [28]| Colored      | 0.9200    | 0.9300 | 0.9300    | 0.9740           |
| Proposed Approach | Colored and  | 0.9937    | 0.9936 | 0.9936    | 0.9936           |
|                   | Segmented    |           |        |           |                  |

Most of the researchers of this filed have presented their results on specific type of plants, a few have presented their performance on Plant Village dataset [6]. Table 3 presents overall accuracy comparison with best and a few recent works. Table 3 proves the robustness of the proposed CNN model which presents comparable and better result with small margin. The presented approach has been compared with Deep CNN [6], and recent works-VGG-16 [25], VGG-16 with Gaussian [25], Inception V3 [28], ResNet-50 [28] and DenseNet169 [28] and proved that overall accuracy outperformed with the help of feature fusion with the use of Alexnet model.

The standard dataset has been divided into the ratio of 20%-80 %, and 80%-20 % for training, and testing to evaluate the performance of the proposed deep CNN approach. Table I and II presents the performance of the proposed approach and an existing approach Deep CNN (Alexnet) [6] that has computed precision, recall, F-measure and overall accuracy on colored, grayscale and segmented PlantVillage dataset [6]. Results of existing Deep CNN (Alexnet) [6] approach is the best, and also performs better with segmented images and grayscale images respectively.
Table 3. Performance comparison of existing approaches and proposed approach

| Approach                        | Input Images       | Accuracy  |
|---------------------------------|--------------------|-----------|
| Deep CNN [6] (2016)             | Colored            | 0.9928    |
| VGG-16 [25] (2020)              | Colored            | 0.9587    |
| VGG-16 with Gaussian [25] (2020)| Colored            | 0.9867    |
| Inception V3 [28] (2020)        | Colored            | 0.9710    |
| ResNet50 [28] (2020)            | Colored            | 0.9820    |
| DenseNet169 [28] (2020)         | Colored            | 0.9740    |
| Proposed Approach               | Colored and Segmented | 0.9936   |

The proposed approach used the colored image as direct input to the deep CNN and features of segmented images as an external feature to the fully connected layer of the Alexnet. Recall measures how many of the sample photographs displaying the disease have been recognized correctly. Specificity measures how many of the samples that do not display the specific disease, have been recognized correctly. Results have been analyzed in terms of precision, recall, F-measure and overall accuracy.

\[
\text{Recall} = \left( \frac{TP}{TP + FN} \right) \times 100
\]

(1)

\[
\text{Precision} = \left( \frac{TP}{TP + FP} \right) \times 100
\]

(2)

\[
\text{F-measure} = \left( \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \right) \times 100
\]

(3)

\[
\text{Accuracy} = \left( \frac{TP + TN}{TP + TN + FP + FN} \right) \times 100
\]

(4)

Fig. 4. Result Images (a)-(c) shows Cherry leaf with no disease, Tomato Leaf with Late_blight and Squash Leaf with a disease Powdery_mildew.
Table I clearly shows the performance of proposed deep CNN approach over the Plant Village [6] dataset. It has been also proved that the proposed approach works better in each ratio of training and testing ratio than existing approach. Fig.4 presents the examples of three images - cherry leaf with no disease, tomato leaf with late_blight disease and squash leaf with a disease powdery_mildew identification by proposed approach.

5. Conclusion and Future Scope
The proposed deep CNN approach are fused by external features which makes improvement in the accuracy, recall, precision, and f-measure of plant disease recognition. Coloured input image is passed to Alexnet model and coloured segmented image is fused at the fully connected layer which helps in the improvement of the overall accuracy. This approach provides better accuracy and outperformed to the state-of-the-art approach by small margin. The combined features of coloured and segmented images with Alexnet model improved the accuracy.

Future prospects of the plant disease identification is to help to farmers of the world to diagnose the disease at early stage so that the treatment can be done to increase the productivity. In future, improvement can be achieved using different CNN models on Plant Village dataset and other challenging datasets with complex background and unconstrained environment. The treatment information related to diagnosed disease would be added in the proposed system which will help to illiterate farmers.

References
[1] Ehler L E 2006 Pest management science 62 787.
[2] Hillnhtrer C and Mahlein A K 2008 Gesunde Pflanzen - GESUNDE PFLANZ 60 143.
[3] Camargo A and Smith J 2009 Biosystems Engineering - BIOSYST ENG 102 9.
[4] Rumpf T, Mahlein A K, Steiner U, Oerke E C, Dehne H W and Plmer L 2010 Computers and Electronics in Agriculture 74 91.
[5] Al Bashish D, Braik M and Bani-Ahmad S 2011 Information Technology Journal 10 267.
[6] Mohanty S P, Hughes D P and Salath’e M 2016 Using deep learning for image-based plant disease detection. Frontiers in plant science 7 1419.
[7] Krishnaswamy Rangarajan A, Purushothaman R and Ramesh A 2018 Procedia Computer Science 133 1040.
[8] Petrellis N 2018 Symmetry 10 270.
[9] Barbedo J 2013 SpringerPlus 2 660.
[10] Kulkarni A 2012 Applying image processing technique to detect plant diseases.
[11] Camargo A and Smith J S 2009 Comput. Electron. Agric. 66 12 1125.
[12] Lai J C, Ming B, Li S, Wang K, Xie R Z and Gao S J 2010 An image-based diagnostic expert system for corn diseases.
[13] Schaad N W and Frederick R D 2002 Canadian Journal of Plant Pathology 24 250.
[14] Barbedo J 2019 Biosystems Engineering 180 96.
[15] Amara J, Bouaziz B and Algargawy A 2017 A deep learning-based approach for banana leaf diseases classification pp. 79–88.
[16] DeChant C, Wiesner-Hanks T, Chen S, Stewart E, Yosinski J, Gore M, Nelson R and Lipson H 2017 Phytopathology 107 1426.
[17] Ferentinos K 2018 Computers and Electronics in Agriculture 145 311.
[18] Fuentes A, Yoon S, Kim S C and Park D S 2017 A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition Sensors.
[19] Bin L, Zhang Y, He D and Li Y 2018 Symmetry 10 11.
[20] Lu Y, Yi S, Zeng N, Liu Y and Zhang Y Neurocomputing 267.
[21] Brahimi M, Kamel B and Moussaoui A 2017 Applied Artificial Intelligence 1.
[22] Oppenheim D and Shani G 2017 Advances in Animal Biosciences 8 244249.
[23] Yosuke T and Fumio O 2019 Plant Phenomics.
[24] Sibiya M, Sumbwanyambe M 2019 A Computational Procedure for the Recognition and Classification of Maize Leaf Diseases Out of Healthy Leaves Using Convolutional Neural Networks AgriEngineering 1 pp.119–131

[25] El-Magd L A, Darwish A, Hassanien A E 2020 Artificial Intelligence-Based Plant’s Diseases Classification Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020) DOI: 10.1007/978-3-030-44289-7_1

[26] Sun G , Jia X , and Geng T 2018 Plant Diseases Recognition Based on Image Processing Technology Hindawi Journal of Electrical and Computer Engineering Vol. 2018 https://doi.org/10.1155/2018/6070129

[27] Firas Ajil Jassim, Fawzi H. Altaani 2013 Hybridization of Otsu Method and Median Filter for Color Image Segmentation Computer Vision and Pattern Recognition

[28] Sagar A and Jacob D 2020 On Using Transfer Learning For Plant Disease Detection Journal of bioRXiv Cold Spring Harbor Laboratory