Multi-Class Plant Leaf Disease Detection Using a Deep Convolutional Neural Network

Shriya Jadhav, Vellore Institute of Technology, India
Anisha M. Lal, Vellore Institute of Technology, India*

ABSTRACT

Traditional machine learning methods of plant leaf disease detection lack successful performances due to poor feature representation and correlation. This paper presents a novel methodology for automatic plant leaf disease detection using cascaded deep convolutional neural network (CDCNN) which focuses on increasing the feature representation and correlation factors. It provides distinctive features that give low intra-class variability and higher inter-class variability. CDCNN were performed on a plant-village leaf disease database which consists of 13 classes of tomato, potato, and pepper bell plant diseases; DCNN model performs better with an overall accuracy, recall, and precision of 98.50%, 0.98, and 0.97 respectively. Additionally, performance of the proposed algorithm is evaluated on real time cotton leaf database for bacterial blight, leaf miner, and spider mite diseases detection and provides 99.00% accuracy. The proposed DCNN outperforms well compared to traditional machine learning and deep learning models and is able to detect the diseases present in the leaves of the plant.

KEYWORDS

Convolutional Neural Network, Deep Learning, Plant Leaf Disease, Precision Agriculture

1. INTRODUCTION

Plants and fruits are primary energy sources for humans and animals. The plant leaves play an essential role in the growth of plants through photosynthesis. The leaves of the plants are also beneficial for humanity because of their medicinal properties (Dhingra et. at 2018). In the Asian and African countries, more than 50% population directly depends upon agriculture for food, shelter, medicine, and employment (Patel et al., 2017). The economy and living of the small farm holders entirely depends upon the production of agriculture products. However, any type of disease degrades the quality of crops and agriculture products and hence leads to huge economical loss of framers. Various crops are damaged because of diseases which declines the quantity and quality of agricultural production. Plant diseases are broadly categorized into parasitic diseases and non-parasitic diseases. The parasitic diseases may occur due to various pathogens such as viruses, bacteria, fungus, chromista, etc.; pests such as millet, mammals, slugs, rodents, etc; weeds such as monocots and dicots. In contrast, non-parasitic plant diseases may occur due to excess or shortage of water, temperature, irradiation, minerals, and nutrients (Kaur et al., 2019) (Poojary et al. 2019).

Crop diseases are chief hazard to food security, but disease detection is challenging in many parts of the world as a result of the unavailability of infrastructure. Recent growth in smartphone technologies paved the way for computer vision-based disease detection. Plant diseases establish financial disasters to the smallholder farmers whose livelihood depends on vigorous crops (Harvey et al., 2015).
The manual leaf disease detection technique is time-consuming, tedious, and depends upon expert knowledge. The manual leaf disease detection is often subjected to less accuracy and inefficient because of fatigues, tiredness, and lack of interpretation of disease. Thus, image processing-based automatic machine learning and deep learning methods are used for plant leaf disease detection techniques (Chapaneri et al., 2020)( Joshi et.al 2020)(Kaur et al 2019).

Machine learning models can be used for thing identification, review, prediction, categorization, and grouping of object. The learning task in ML categorized as “Supervised”, “unsupervised” and “semi-supervised”. Depending upon the problem statement different models is used.

Deep neural networks are advances of the neural network that have been successfully applied recently for many computer vision-based applications. Deep neural networks are constructed by stacking the series of layers of nodes. The deep learning algorithms’ performance can be improved by tuning the parameters of the deep learning layers. The accuracy of deep learning models depends on the size of the database (LeCun et al., 2015)( Schmidhuber et al., 2015).

This paper presents multi-class plant leaf disease detection based on a deep convolutional neural network (DCNN). Three-layered DCNN architecture that consists of convolution, rectified linear unit layer, maximum pooling layer (MP), fully connected layer, and a classification layer is used. The database contains healthy and disease samples of tomato, potato, and pepper plant leave . Various diseases consider for the implementation are bacterial spot, early blight, late blight, target spot, mites, leaf mold, and septoria leaf spot.

The objectives of this article are to are as follows:

- To design and implement simple three layered Cascaded Deep Convolutional Neural Network (CDCNN) model.
- To improve the feature distinctiveness of the leaf image.
- Effectively detection of diseases not only for a particular plant but for multiclass plant.
- Testing the designed model in the real itme surroundings.

The remaining paper is arranged as: Section II gives a brief description of previous work carried out on plant leaf disease detection. Section III elaborated the CDCNN methodology in detail. Section IV depicts the experimental results and discussion. Finally, section V concludes the paper and suggests the direction for future work.

2. RELATED WORK

Machine learning is becoming one of the prime technology in extracting knowledge from available data also it has successfully helped in delivering dynamic solution for real-time problems (Alzubi J et.al 2018). Various computer vision-based techniques have been presented for leaf disease detection based on machine learning and deep learning techniques. In the local binary pattern, features are extracted to describe the diseased, regular vine plant leaves texture features and classified using one-class support vector machine(SVM). It achieved 95% accuracy for leaf disease detection (Pantazi et al., 2019). Traditional features are irrelevant and inessential, which degrades the classification accuracy. Sandip Kumar et al. 2018 Presented exponential spider monkey optimization (ESMO) algorithm to select salient features obtained using the spatial domain subtractive pixel adjacency model (SPAM) approach. It has shown 92.12% accuracy for the plant leaf detection using the SVM classifier.

Kaur et al. (2018) have suggested semi-automatic using various color and texture features of soybean leaf disease detection. In their approach, the k-mean segmentation technique is used to segment the leaf region. There is a need for the selection of proper clusters for feature extraction. It achieved 90% accuracy on the samples obtained from the plant-village database. Ramesh et al. 2018
presented leaf disease detection using a histogram of oriented gradients (HOG) and a random forest classifier. It is observed that the shape of the leaf plays a vital role in disease detection.

Waghmare et al. (2016) investigated multi-class SVM for grape leaf disease detection, which has given accuracy of 96%. They mentioned that testing accuracy of leaf disease detection could be increased by increasing the training samples.

Recently various deep learning-based techniques have been implemented for leaf disease detection. Deep learning-based methods have shown improved performance over machine learning-based techniques. Mohanty et al. (2016) presented a deep convolutional neural network for the disease detection of 14 crop species. They have developed a plant-village leaf disease database that includes 54306 images of 26 diseases. Their model has given an accuracy of 99.35% for the controlled database. Ramcharan et al. (2017) proposed cassava disease detection using DCNN trained using transfer learning. It has shown overall accuracy of 93% for cassava disease detection, which is affordable and easily deployable on the digital platform. Ferentinos et al. (2018) presented a convolutional neural network for leaf disease detection, resulting in 99.53% accuracy for the 25 plants dataset. Deep learning-based methods often result in over-fitting for the smaller database. Goncharov et al. 2018 proposed deep siamese CNN to avoid the problem of smaller data. They achieved 90% accuracy for detecting three grape plant diseases as Esca, Black rot, and Chlorosis. Abbas et al. (2021) by using Conditional Generative Adversarial Network (C-GAN) created synthetic images and used DenseNet121 model for training. It improved network generalizability and accomplished accuracy of 99.51% for five class classification for tomato plant but there is need for testing the model for various parts of the plant other than leaves.

Bedi et al. (2021) presented a hybrid model and used convolutional autoencoder (CAE) network and convolutional neural network (CNN) which gained accuracy of 99.35% for training and 98.38% for testing, however model was designed for identification of only one type of disease for peach plant i.e. Bacterial Spot disease. Ashwinkumar et al. (2020) applied optimal mobile network-based convolutional neural network (OMN-CNN) on tomato leaf images which accomplished accuracy of 0.987, precision of 0.985, recall of 0.9892, F-score of 0.985, and kappa of 0.985 respectively. Hernández et al. (2020) suggested probabilistic deep learning-based plant disease detection model. The Stochastic gradient descent (SGD) and Monte Carlo dropout likely to fabricate overconfident incorrect predictions whose maximum probabilities is larger than their correct counterparts whereas stochastic gradient langevin dynamics (SGLD) generates less confident outcomes for correctly and incorrectly classified samples. When testing was performed in real-time environment image quality, covariate shift and little other aspect showed decrease in the performance. Tiwari et al. (2021) utilized dense convolutional neural network architecture which demonstrated the average cross-validation accuracy of 99.58% and average testing accuracy of 99.19% with overall processing time taken for one leaf image was 0.016 seconds. They mentioned that performance for larger size dataset was better than that for small size image dataset. Shah et al. (2021) proposed ResTS (Residual Teacher/Student) for visualization and a classification technique for finding the plant disease. The performance of ResTS outshined the original Teacher/Student architecture with F1 score of 0.991 and 0.972 respectively. The experiment was performed on 54306 images and on 14 crop varieties. Wang et al. (2021) suggested that instead of working on crop-disease pair, identification and classification on separating crop and its disease achieved better performance. Accuracies accomplished on crop and disease in detachment for laboratory environment was 99.99% and 99.7% and real-time environment was 84.11% and 75.58%. They proposed Trilinear convolutional neural networks (T-CNN) model that make use of 3 CNNs, VGG-16, Inception v3, and ResNeXt-101, as the support networks for the model. Tahir et al. (2021) recommended a innovative deep classification model for apple diseases based on transfer learning and Inception v3. Variance control method is used to down-sample the features extracted to eliminate unneeded features from the feature vector. Finally, the most discriminative features are classified by a classifier with a maximum accuracy of 97%. It improved the classification performance of the model by effectively removing redundant features; however, the model is weak in generality as it...
is limited to only one plant category of apple. Adeel et al. (2020) projected deep learning approach for grapevine leaf disease detection, which initiated the concept of transfer learning and uses pre-trained models AlexNet and ResNet-101 for feature extraction. Yager entropy along with the kurtosis technique are used for feature selection, and to spot the most evident and robust feature set, tag along by least squares support vector machine for classification and gained accuracy of 99%. However, there are limitations to these studies like sole crop or a tiny number of diseases, and crop disease images are derived from datasets collected in a controlled laboratory environment that may be difficult to deal with complex real-world growing environments. Sharma et al (2020) in paper give you an idea about the feasibility of training a convolutional neural network (CNN) model using segmented and annotated images as an alternative of full images. Model performance increases from 42.3% to 98.6% when the same CNN model is trained using the segmented images (S-CNN) compared to training using full images (F-CNN). Also quantitative analysis of self-classification confidence confirmed improvement with 82% of test dataset showing increase in confidence. Unlike other deep learning models for automatic disease detection whose performance reduces when applied to real world images the proposed models proves to overcome this shortcoming too. Sachdeva et al. (2021) presented a deep convolutional neural network (DCNN) model using Bayesian learning process for classification of the diseases in potato, tomato and pepper bell plants by using 20,639 images having 15 distinct classes from plantVillage database. Bayesian process is used at the top of residual network for efficient feature learning which resulted increase in the accuracy rate of 98.9% with no overfitting problem, however the computation time needed to process each image is high. Uday Pratap et al. (2019) for classification used multilayer convolutional neural network (MCNN) and obtained accuracy of 97.13%. Testing was performed on 1070 mango leaves. Arib et al. (2018) proposed CaffeNet model which was built on a Caffe framework to label paddy pest and paddy diseases. The abovementioned model was fine-tuned over 30,000 iterations and obtained an accuracy score of 87%. Lu et al. (2017) made use of supervised DL model for wheat disease identification. It is tested using a 50k labeled wheat images. The model contains 4 diverse CNN models for recognizing the seven types of wheat disease classes. The results stated that the VGG-16 model has attained a higher detection rate of 97.95%.

Various factors affect the performance of deep learning, such as image background, limited annotated dataset, image capturing conditions, symptom variations, disorders with similar symptoms, etc. Again, the deep learning technique’s performance is limited because of data imbalance problem, over-fitting, and quality of leaf images (Barbedo et al., 2019). The data obtained from various source may contain plentiful features, not all of which would be relevant to the learning process. These features need to be removed and a subset of the most important features needs to be obtained, making feature selection one of the important factor. CDCNN would help in efficient feature selection and accurate classification (Alzubi et al. 2018).

3. PROPOSED METHODOLOGY: CDCNN

The CDCNN system consists of a three-layered deep convolutional neural network (DCNN). The DCNN architecture consists of a convolution layer, ReLU layer, pooling layer, fully connected layer, and a classification layer. The flow diagram of the CDCNN system is shown in figure 1. It consists of the training and testing phase. During the training phase DCNN is trained for various healthy and diseased images of plants. The CDCNN architecture consists of three cascaded layers of CNN followed by a classification layer. The architecture of the CDCNN is given in figure 2.

The convolution layer gives the correlation and connectivity between local regions of the plant leaf surface. The convolution layer accepts the original leaf images in color format. It can describe discriminative attributes of the texture, cracks, consistency, and shape of a plant leaf (Francis er. Al 2019). In this layer, the input image is convolved with various convolutional filters (Albawi et al., 2017). Feature maps generated due to each filter may represent distinctive characteristics of the leaf.
For the convolution, $3 \times 3$ filter $h(w, w)$ is utilized. The first, second and third layer consists of 96, 256 and 384 filters respectively. The convolution operation can be given by equation 1:

$\text{Conv}(m, n) = im(m, n) * h(w, w) = \sum_{i=1}^{R} \sum_{j=1}^{C} im(i, j) \cdot h(i - m, j - n)$

(1)

Convolution operation reduces the dimension of the original image $im$, therefore to maintain the size of the original image is zero-padded from the outer border. Convolution filter is stride over the input image by one pixel, representing the higher correlation between local regions of the plant leaf image.

All negative values it presents in convolutional layer output are rounded to zero in this layer, and all non-negative values are kept and it is as given in equation 2. ReLU activation function minimizes...
the vanishing gradient problem and makes CNN features faster and efficient for training. It brings
the linearity in the data, which is simple to optimize:

\[
ReLU(i, j) = f(x) = \begin{cases} 
0, & \text{conv}(i, j) < 0 \\
\text{conv}(i, j), & \text{conv}(i, j) \geq 0
\end{cases}
\] (2)

In the maximum pooling layer, the maximum of the pooling window is selected to grab the
salient information from the crop leaf and reduce the feature maps. For the pooling operation, a 2x2
window is selected, which strides over the image for down-sampling. The pooling window strides
by 2 pixels over the rows and columns. The dimensions are reduced to half of the ReLU layer. The
MP layer operation is given by equation 3:

\[
MP(i, j) = \max_{1 \leq i \leq R, 1 \leq j \leq C} \left( \max_{i : M, j : M} ReLU(i, M, j, M) \right)
\] (3)

Each neuron of one layer is connected with all other neurons of all different layers in a fully
connected layer. A fully connected layer helps to increase the connectivity of the features of leaf
images. The softmax layer gives the probabilities of each output class. The sum of all the probabilities
of the soft-max layer is equal to one. The class with a higher probability will be considered as the
result class (Liu et al., 2015).

Mini-batch gradient descent (MBGD) algorithm is used for learning of DCNN where the
optimization is performed by splitting the \( n \) training samples into small batches \( b \). In this average
or sum of the gradient is selected to diminish the changes in gradient. MBGD unites the robustness
of the stochastic gradient descent algorithm and the efficiency of batch gradient descent to reduce
computational complexities. Total \( T = n/b \) number of iteration are considered per training epoch (Danner
et al., 2015) (Konečný et al. 2015) (Ruder et al. 2016). The weights (\( w \)) of CNN filters are tuned using
the error function given in equation 4:

\[
E_i[f(w)] = \frac{1}{b} \sum_{i=(i-1)b+1}^{ib} f(w, x_i)
\] (4)

where \( x_i \) is \( i \)th training data sample. The weights are updated after every iteration using learning rate
\( \mu \) as given in equation 5:

\[
w^{i+1} = w^i - \mu \nabla_w E_i[f(w^i)]
\] (5)

Table 1 gives details about the initial parameters used for DCNN implementation.

Table 2 describes the configuration of DCNN layers, sub-layers, output size of the respective
layer, striding value, and output feature size. Increasing the number of CNN layers increases the
connectivity and distinctiveness of the features. The feature map of the convolution layer and ReLU
layer is the same, and the only difference is that the convolution layer may consist of negative values.
In contrast, the ReLU layer feature map consists of only positive values. After every pooling layer,
the size of the ReLU layer feature map gets halved.
3.1 The Proposed Algorithm for CDCNN

The cascaded-CNN (C-CNN) is a novel deep learning architecture comprised of multiple CNNs, each predicting a specific aspect. The cascaded CNN includes three parts: the shared hidden layers part, the global atmospheric light estimation subnetwork, and the medium transmission estimation subnetwork. In the network diagram, different color blocks represent the different operations (Li et al. 2018). Based on the similar logic CDCNN is designed.

1. Input: Single image of size 256 x 256
   Plant database is divided into two parts 70% for training and 30% for testing
2. Output: Class probability given by softmax classifier
3. Initialization Phase:
Network hyperparameter initialization
Learning Rate
Number of Iterations
Number of epoch
Number of Layers
Number of Filters
Activation Function
Dropout rate
Hidden Neurons in Fully connected Layer, etc.

4. Training Phase:
im_train: training sample
L: labels of training sample
For i=1:total_train_samples
    Model= Train (DCNN (im_train(i):Sample, L: label)
DCNN=
    {Convolution Layer 1, ReLU1, Max pooling 1},
    {Convolution Layer 1, ReLU1, Max pooling 1},
    {Convolution Layer 1, ReLU1, Max pooling 1},
    {Fully Connected layer}
    {Softmax Classifier layer}
end

5. Testing phase:
im_test: testing sample
result = Test(im_test)

4. EXPERIMENTAL RESULTS AND DISCUSSION

Due to digitalization vast amount of data is generated, machine learning takes the advantage of this generated data for knowledge processing. Size of the database has a greater impact on the results of machine learning and deep learning accuracy and performance. Database and deep learning needs to be integrated to support complex analytics and predictions. The PlantVillage dataset consists of 54303 healthy and unhealthy leaf images divided into 38 categories by species and disease. The performance of the CDCNN approach is evaluated on a plant-village leaf disease dataset (Xu et.al 2018). The sample images from the PlantVillage dataset are shown in Figure 3.

The performance of the training, testing, and validation of the CDCNN is shown in Figure 4. It is observed that the network can be trained in the first 20 epochs. It has demonstrated overall training, validation, and testing accuracy of 99.80%, 98.50%, and 97.00%, respectively.

The effectiveness of the DCNN is evaluated using accuracy, recall, and precision using equation 6-8. TP represents the number of correctly classified healthy samples. TN represents a number of correctly classified diseased samples. FN represents the number of wrongly classified healthy samples, and FP represents a number of improperly classified diseased images:

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

(6)
Recall \( TP \) FN \( = \times 100 \) (7)

Precision \( TN \) FP \( = + (8) \)

The performance of the C- DCNN is compared with the implementation of the existing local binary pattern texture (LBP) descriptor and histogram of oriented gradient (HOG) based shape descriptor using K-nearest neighbor (KNN) and support vector machine (SVM) classifiers. The C- DCNN performs better than traditional machine learning-based approaches because of its higher feature

Figure 3. Sample of PlantVillage a) potato early blight b) Potato late blight c) Tomato mosaic virus d) Pepper bacterial spots e) Tomato spider mites f) Tomato leaf mold

Figure 4. Performance of training, validation, and testing of DCNN

\[ \text{Recall} = \frac{TP}{TP + FN} \times 100 \] (7)

\[ \text{Precision} = \frac{TN}{TN + FP} \] (8)
representation capability. The performance comparison of the CDCNN technique with conventional methods is shown in Table 3. The framework for overall system is shown in Figure 5 where the main database is divided into two sections: one for training and the other for testing, then the image passes through the CDCNN where identification and classification are performed.

Table 3. Performance comparison with implementation of machine learning techniques

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Accuracy (%)</th>
<th>Recall</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP+KNN</td>
<td>86.50</td>
<td>0.88</td>
<td>0.85</td>
</tr>
<tr>
<td>HOG+KNN</td>
<td>87.00</td>
<td>0.89</td>
<td>0.86</td>
</tr>
<tr>
<td>LBP+SVM</td>
<td>88.20</td>
<td>0.91</td>
<td>0.89</td>
</tr>
<tr>
<td>HOG+SVM</td>
<td>88.00</td>
<td>0.90</td>
<td>0.89</td>
</tr>
<tr>
<td>CDCNN</td>
<td>98.50</td>
<td>0.98</td>
<td>0.97</td>
</tr>
</tbody>
</table>

The performance of the CDCNN approach is compared with previous techniques of plant leaf disease detection as given in Table 4. It is observed that the CDCNN approach gives satisfactory improvement in the plant leaf disease detection accuracy.

The effectiveness of the CDCNN approach is also validated for the real-time cotton plant leaf disease detection. Total 600 images of bacterial blight, leaf miner, and spider mite diseases, and 400 images of normal cotton leaf are collected for experimentation using mobile phone camera with resolution of 256x256. Out of total leaf image samples, 70% and 30% samples are selected for training and testing. The performance of CDCNN approach for real-time database shows impressive results like standard database as given in Table 5.
5. CONCLUSION AND FUTURE SCOPE

Thus, this paper presents a deep convolutional neural network for plant leaf disease detection such as early blight, late blight, target spot, bacterial spot, mosaic virus, and yellow leaf. DCNN has given better feature representation and connectivity of the basic features. The CDCNN can provide better discriminative features for multi-class leaf disease detection for various plants. The CDCNN results in accuracy, recall, and precision of 98.50%, 0.98, 0.97, respectively. The performance of the CDCNN is compared with the implementation of LBP and HOG features along with KNN and SVM classifiers. It has shown significant improvement over the traditional approaches. Thus compared to traditional machine learning based approaches CDCNN have better results in terms of accuracy. When judge against models like CaffeNet, VGG16, CAE-CNN, MCNN; CDCNN gained accuracy of 98.50%. In real-time environment too CDCNN outperformed other models with accuracy of 99%. It was tested on local cotton leaves database consisting of 1000 images. In the future, the performance of the deep learning technique can be optimized using optimization techniques. Further, the Cascaded- algorithm is intended to investigate real-time plant data and images captured at different environmental conditions in the future.
REFERENCES


Anisha M. Lal is an Associate Professor in the School of Computer Science and Engineering at VIT University, Vellore, India. She received her B.E and M.E degrees in Computer Science and Engineering from Anna University in 2003 and 2005, respectively. She has been awarded her Doctorate in Computer Science and Engineering at VIT University in 2016. She has teaching and research experience of about 16 years. She has published more than 20 papers in international journals and presented around 3 papers in national and international conferences. Her research interests include image processing, remote sensing, biomedical, deep learning, and soft computing. She has served as a reviewer in many academic conferences and journals in the related areas.