Efficient deep learning architecture for the classification of diseased plant leaves

Muniyandi Sadhasivam¹, Manoharan Kalaiselvi Geetha¹, James Gladson Maria Britto²
¹Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar, India
²Department of Computer Science and Engineering, Malla Reddy College of Engineering, Hyderabad, India

ABSTRACT

The classification of plant leaf diseases via machine learning and deep learning algorithms has a great deal of potential for enhancing agricultural operations by allowing the early and accurate diagnosis of diseases. These systems can potentially develop into useful instruments for environmentally responsible farming and increased food safety as technological advancements continue. In this work, an efficient deep learning architecture has been developed to classify the diseased plant leaves. A ten-layer architecture is designed, which includes 5-convolutional layers using different numbers of filters (32, 64, 128, 256, and 512) and for dimension reduction, five max-pooling layers are used. The PlantVillage dataset which consists of more than 50,000 plant leaf samples is used to analyze the proposed architecture's performance. The performances are evaluated across different training and testing configurations and different dropout configurations. When compared to well-known transfer learning methods using visual geometric group (VGG16), AlexNet, and GoogleNet architectures, the proposed architecture obtains a higher level of performance with 98.18% classification accuracy.

Keywords:
Classification
Deep learning
Feature extraction
Pattern recognition
Plant leaf

1. INTRODUCTION

Plant diseases cause a serious risk to agricultural production, which may have a severe impact in terms of both the quality and quantity of crops produced. The diagnosis and classification of these diseases at an early stage are very necessary for the efficient management of diseases and the preservation of crops. Algorithms that learn from machine data have recently emerged as strong tools for automating the process of recognizing and categorizing plant leaf diseases. This has made it possible to respond to outbreaks in a manner that is both faster and more accurate. With the help of traditional classification systems like k-nearest neighbor [1], decision trees [2], support vector machine (SVM) [1], and neural networks [3], a broad range of automated approaches have been developed. The accuracy of traditional classification systems is reliant on the features that are extracted (colour and textures) as well as the classifiers that are chosen. The weak classifier may cause the dominant characteristics to fail, and vice versa. In addition, the creation of an accurate categorization system necessitates the use of two distinct modules.

Deep learning strategies have, in recent years, produced some amazing achievements in various fields, including natural language processing, object identification, medical diagnosis, and others. Convolutional neural networks (CNNs) are among the most essential and well-known models used in deep learning techniques. There are still theoretical and practical barriers to overcome in this field, even though...
Deep learning techniques have advanced quickly and shown the specific benefits they provide. This paper presents a deep learning architecture for computer-assisted plant leaf disease classification, with the goal of enhancing the accuracy of classification.

Ant colony optimization with CNN (ACO-CNN), a novel deep learning approach for illness detection and classification. ACO was used to study how well it could detect diseases in plant leaves [4]. The proposed rice plant disease recognition system uses computer vision, image processing, machine learning, and deep learning to protect paddy crops from five major diseases that plague Indian rice fields: bacterial leaf blight, false smut, brown leaf spot, rice blast, and sheath rot is described in [5].

The most significant challenges and issues in identifying leaf diseases are described in [6]. Methods are compared in terms of the agricultural product they were designed to work with, as well as their efficiency, effectiveness, benefits, and downsides. SVM has been widely utilized for disease classification, as shown by the review study's analysis of the algorithms used most frequently in machine learning. The modified InceptionResNet-V2, a variant of the CNN model, is presented with a transfer learning strategy for disease recognition in tomato leaf images [7]. Seven diseases of tomato leaves are classified by the modified algorithm.

The various methods presented for detecting plant diseases are discussed in [8]. 160 diverse research works are considered in this study, and the methods included range from single-network to hybrid to real-time detection. About 57 researchers considered multiple plants, whereas 103 focused on only one. Fifty datasets on plant leaf diseases are discussed, including those that are freely accessible and those that are not. A customized version of the ResNet model for the classification and detection of plant leaf diseases is described in [9]. Initially, ResNet-50 extracts various features from plant leaf images that include color and texture properties. Additionally, the modified red deer optimization algorithm (MRDOA) is implemented as an optimum feature selection method in order to acquire optimized and salient features while reducing the size of the MRDOA. Further, a deep learning CNN classifier model is used in order to achieve improved classification performance.

A new model for precision agriculture that uses deep learning for automated plant disease detection and classification is described in [10]. U2Net-based background removal is used in the initial step of the extraction process, during which the leaf and fruit regions are extracted. After that, the Adam optimizer is used along with the SqueezeNet model as a feature extractor, and the Adam optimizer is responsible for tuning the hyperparameters. The classification of plant diseases is carried out using the extreme gradient boosting classification. AlexNet, a simple sequential model, MobileNet, and Inception-v3 are all representations of deep learning models that may be used to identify disease in leaf is discussed in [11].

The development of a disease detection model using images of diseased-healthy plant pairs and a CNN algorithm consisting of five pre-trained models is discussed in [12]. The crop classification phase is the first in the disease detection model's three-stage classification process, including the disease classification step. An artificial neural network (ANN) is optimized for plant leaf analysis [13]. The data is initially included for preprocessing, then the important features are retrieved, and finally the whale optimization algorithm is used to choose the required features. Following that, ANN is used to classify the data. The ANN classification strategy uses the feed-forward neural network. ANN is a highly adaptable technology used extensively to solve various problems. In this research, classification is used throughout each step to exclude certain options, which results in improved prediction accuracy.

The diseases that affect eleven (11) different plants and the methods that may be used to identify the diseases from images of plant leaves using CNN-based deep learning models are described in [14]. It provides a summary of the research that has used hyperspectral images for the diagnosis of plant diseases as well as the various data sources that have been used in different studies. A hybrid deep learning-based method for diagnosing diseases in tomato plants based on images of their leaves is discussed in [15]. Two pretrained models, namely EfficientNetB3 and MobileNet are combined to correctly identify tomato leaf diseases. In addition, the issue of model overfitting was handled using a variety of approaches, including regularization, dropout, and batch normalization.

A hierarchical deep learning CNN (HDLCNN) to detect the diseases in the leaf is presented in [16]. The technique of median filtering is used in the first phase of the pre-processing procedure that is carried out. After the image has been processed, an intuitionistic fuzzy local binary pattern is added; this pattern is responsible for determining the features of the leaf. After that, the HDLCNN is used to detect and classify the disease. Different deep learning architectures are designed for a particular plant diseases classification such as tomato plant [17]–[19], anthracnose and red-rust leaf disease [20], coffee leaf disease [21], peanut leaf disease [22], and citrus leaf disease [23]. The main objective of this study is to design an efficient deep learning architecture for accurate and robust classification of plant leaf diseases using leaf images in order to overcome the challenges that have been presented.
2. METHOD AND MATERIALS

The proposed system for plant leaf disease detection is a machine learning system designed to classify the plant leaf into healthy or diseased. The leaf images are fed to the deep learning architecture, where they are processed, and the deep features are retrieved with the help of two simple components: convolution layers (for the extraction of features) and pooling layers (for the reduction of features). When applied to a specific challenge, the efficacy of deep learning highly depends on the specific configuration of the layers described above. The classification process begins with the extraction of deep features, followed by the use of fully linked layers and an output layer. Figure 1 shows the arrangement of convolution and max-pooling layers in the proposed architecture. It consists of interleaved convolutional and pooling layers. At the end of these, the feature map is flattened to a one-dimensional array and the rest of the network has the structure of a neural network. Table 1 shows the proposed architecture structural details.

![Arrangement of layers in the proposed architecture](image)

<table>
<thead>
<tr>
<th>#Layer</th>
<th>Input size</th>
<th>Output size</th>
<th>#Filters</th>
<th>Kernel size</th>
<th>Operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>256x256</td>
<td>256x256</td>
<td>32</td>
<td>3x3</td>
<td>Convolution with zero padding</td>
</tr>
<tr>
<td>2</td>
<td>256x256</td>
<td>128x128</td>
<td>64</td>
<td>2x2</td>
<td>Pooling</td>
</tr>
<tr>
<td>3</td>
<td>128x128</td>
<td>128x128</td>
<td>64</td>
<td>3x3</td>
<td>Convolution with zero padding</td>
</tr>
<tr>
<td>4</td>
<td>128x128</td>
<td>64x64</td>
<td>-</td>
<td>2x2</td>
<td>Pooling</td>
</tr>
<tr>
<td>5</td>
<td>64x64</td>
<td>64x64</td>
<td>128</td>
<td>3x3</td>
<td>Convolution with zero padding</td>
</tr>
<tr>
<td>6</td>
<td>64x64</td>
<td>32x32</td>
<td>-</td>
<td>2x2</td>
<td>Pooling</td>
</tr>
<tr>
<td>7</td>
<td>32x32</td>
<td>32x32</td>
<td>256</td>
<td>3x3</td>
<td>Convolution with zero padding</td>
</tr>
<tr>
<td>8</td>
<td>32x32</td>
<td>16x16</td>
<td>-</td>
<td>2x2</td>
<td>Pooling</td>
</tr>
<tr>
<td>9</td>
<td>16x16</td>
<td>16x16</td>
<td>512</td>
<td>3x3</td>
<td>Convolution with zero padding</td>
</tr>
<tr>
<td>10</td>
<td>16x16</td>
<td>8x8</td>
<td>-</td>
<td>2x2</td>
<td>Pooling</td>
</tr>
<tr>
<td></td>
<td>Fully connected layer</td>
<td>32,768</td>
<td>2</td>
<td>-</td>
<td>Classification</td>
</tr>
</tbody>
</table>

2.1. Convolution layers

The process of convolution consists of two stages: the first stage involves multiplying the original image’s pixel values with the filter’s coefficients, and the second stage involves adding the results obtained from the first stage’s multiplications. The filters of the convolutional layer may be thought of as having the purpose of identifying edges and curves (low-level features). The input data are passed through successive convolutional layers, which results in the acquisition of more complex features [24]. It can be defined as;
Efficient deep learning architecture for the classification ... (Muntyandi Sadhasivam)

\[ FM^q_p = \sum_i x_i^{q-1} * k_{ip} \]  

where \( FM^q_p \) is the \( p^{th} \) feature map of \( q^{th} \) layer and \( k \) is the kernel (filter coefficients). Figure 2 shows the outputs from the convolution layer. Figure 2(a) shows the input image and Figure 2(b) shows the obtained feature map from the 3\(^{rd}\) convolution layer. As the 3\(^{rd}\) convolution layer uses 64 filters to extract the feature map, Figure 2(b) shows feature maps obtained from all the convolution filters in that layer.

![Figure 2](image)

**Figure 2.** Convolution layer outputs (a) plant leaf image and (b) feature map from the 3\(^{rd}\) convolution layer

### 2.2. Pooling layers

When using a CNN, the computation cost tends to become very high due to the large number of hyperparameters, which may lead to the issue of overfitting. Thus, after the convolutional layer, it is common practice to include pooling layers. Max-pooling and average pooling are the two types of pooling that are used the most often [25]. Max-pooling applies a kernel to the input and generates an output that is the greatest number in the convolved area, while average pooling generates an output which is the average for each sub-region. In this work, max-pooling is employed as the specific features with high values are significantly more relevant than the location of these features. Figure 3 shows the working of the max-pooling layer. Figure 3(a) shows the reduced feature map (2\times2) for a 4\times4 feature map shown in Figure 3(b).
2.3. Fully connected layer

It is the last layer of the proposed architecture where the classification of plant leaves takes place. In general, a back propagation neural network is employed in this layer. In the field of machine learning, feature vectors are used as a means of mathematically representing numerical or symbolic qualities of an object/image, which are referred to as features. “Feature space” refers to the vector space that is associated with these vectors. Thus, the feature space is the input to the fully connected layer inputs. These input signals are weighted and summed (along with biases) to provide the neuron’s activation [26]. Figure 4 shows the feed-forward neural network.

The backpropagation method is often used during the training process for deep learning networks such as CNN. Because it needs a known output and intended output for each input value in order to compute the gradient of the loss function, this training approach is classified as a supervised training method. The backpropagation method contains the following two phases:

- The first phase of training a neural network involves the forward propagation of input data across all of the network's nodes in order to create output values for the network. The generation of the new weight value for each network neuron requires the calculation of the error function as well as the backward propagation of the output values through the network.
- In the second phase, the gradient for each network is computed using the neuron’s weight and threshold. Then, the gradient’s weight and threshold are subtracted from an amount equal to the learning rate of the algorithm.

The activation function, also known as the transfer function, produces the neuron's output [27]. The ReLU activation function is shown in Figure 5. It demonstrates that the ReLU function always returns the same +ve integer values, and when given –ve integer values, it always returns zero. As a result, this function is less prone to the problem of vanishing gradients.

![Figure 3. The working of the max-pooling layer (a) feature map values and (b) reduced feature map](image)

![Figure 4. Feed forward neural network](image)

![Figure 5. ReLU function](image)
To combat over-fitting, dropout ignores a certain number of randomly chosen nodes in the weight update for each epoch. The features are element-wise multiplied by a mask of the same size to carry out this operation. The amounts of the two values that the elements of the mask which have values of 0 and 1, may have been determined by the dropout rate. The percentage of neurons that are lost is represented by this rate. Convolutional layers may further include dropout. Some features, which correspond to lost nodes, are replaced with zeroes by the element-wise multiplication of the features and the dropout mask.

Dropout is used because each node is comparatively independent of the other nodes and the updating of the weights in the network no longer relies on the joint activity of hidden nodes with fixed relations [28]. The idea behind bagging and dropout is quite similar. Both strategies aim to create numerous sub-classifiers based on diverse subsets of training data in order to avoid a scenario in which certain properties pre-dominate the training process. Figure 6 shows the dropout regularization in deep learning. Figure 6(a) shows no dropout in the architecture and Figure 6(b) shows the inclusion of dropout in the architecture. Dropout avoids the co-adaptation of neurons, and the network is encouraged to learn properties that are more robust and generalizable.

3. RESULTS AND DISCUSSION

The performances of the proposed architecture are analyzed using the PlantVillage dataset [29]. Within the realms of plant disease categorization and agricultural research, the dataset known as PlantVillage is quite popular. It comprises photographs of plant leaves affected by various illnesses and conditions. The dataset is open to the public and is often used in the process of developing and evaluating machine learning models for the diagnosis and categorization of plant diseases. The dataset includes 54,303 leaf images, healthy (15,084 images), and unhealthy or diseased (39,221 images), organized into 38 categories according to species and illness. Sample images from the PlantVillage dataset are shown in Figure 7.

Performance metrics can evaluate how effectively a system works for plant leaf disease classification. The commonly used popular performance indicators such as accuracy, precision, and recall are evaluated. When conducting an analysis of a system for the classification of diseases, it is essential to take into account all of these metrics together in order to get an in-depth comprehension of the system's performance. This is particularly important in real-world situations that include unbalanced datasets or different levels of disease prevalence. In addition, the selection of metrics could be determined by the particular objectives and prerequisites of the application. Table 2 shows the performance metrics used in this study.
Table 2. Performance metrics used by the proposed system

<table>
<thead>
<tr>
<th>Performance metrics</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formula</td>
<td>$\frac{TP + TN}{TP + FN + TN + FP}$</td>
<td>$\frac{TP}{TP + FP}$</td>
<td>$\frac{TP}{TP + FN}$</td>
</tr>
</tbody>
</table>

Table 1 true positive (TP), false positive (FP), false negative (FN), true negative (TN) represents TP (number of diseased leaves correctly identified as diseased), FP (number of healthy leaves identified as diseased), FN (number of diseased leaves identified as healthy), and TN (number of healthy leaves correctly identified as healthy). The training and testing split of the database is an essential stage in developing and evaluating machine learning models. The proposed system uses random split approach. It includes randomly dividing the data into two groups, one of which is used for training the model, and the other of which is used for assessing the performance of the model. Table 3 shows the obtained performances by the proposed system.

Table 3. Accuracy of the proposed leaf disease classification system across different training and testing configurations

<table>
<thead>
<tr>
<th>Training</th>
<th>Testing</th>
<th>Deep learning architectures</th>
<th>Proposed system</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>50%</td>
<td>VGG16 [30]</td>
<td>71.91</td>
</tr>
<tr>
<td>60%</td>
<td>40%</td>
<td>AlexNet [31]</td>
<td>75.66</td>
</tr>
<tr>
<td>70%</td>
<td>30%</td>
<td>GoogleNet [32]</td>
<td>80.96</td>
</tr>
<tr>
<td>80%</td>
<td>20%</td>
<td>Proposed system</td>
<td>88.82</td>
</tr>
<tr>
<td>50%</td>
<td>50%</td>
<td>VGG16 [30]</td>
<td>75.93</td>
</tr>
<tr>
<td>60%</td>
<td>40%</td>
<td>AlexNet [31]</td>
<td>80.50</td>
</tr>
<tr>
<td>70%</td>
<td>30%</td>
<td>GoogleNet [32]</td>
<td>84.68</td>
</tr>
<tr>
<td>80%</td>
<td>20%</td>
<td>Proposed system</td>
<td>91.27</td>
</tr>
<tr>
<td>50%</td>
<td>50%</td>
<td>VGG16 [30]</td>
<td>81.17</td>
</tr>
<tr>
<td>60%</td>
<td>40%</td>
<td>AlexNet [31]</td>
<td>84.92</td>
</tr>
<tr>
<td>70%</td>
<td>30%</td>
<td>GoogleNet [32]</td>
<td>88.03</td>
</tr>
<tr>
<td>80%</td>
<td>20%</td>
<td>Proposed system</td>
<td>95.59</td>
</tr>
<tr>
<td>50%</td>
<td>50%</td>
<td>VGG16 [30]</td>
<td>87.62</td>
</tr>
<tr>
<td>60%</td>
<td>40%</td>
<td>AlexNet [31]</td>
<td>92.55</td>
</tr>
<tr>
<td>70%</td>
<td>30%</td>
<td>GoogleNet [32]</td>
<td>94.92</td>
</tr>
<tr>
<td>80%</td>
<td>20%</td>
<td>Proposed system</td>
<td>98.18</td>
</tr>
</tbody>
</table>

It can be seen from Table 3 that the proposed system maintains a performance advantage over the other three architectural designs during each testing and training split. It obtains the best accuracy in each scenario, showing that it is the most effective model for this dataset since it is designed specifically for it. The accuracy of all models, in general, improves as a result of increasing the data for training. It seems that having more data available for training is favorable to the performance of the model. When compared to the other known designs (VGG16, AlexNet, and GoogleNet), the performance of GoogleNet is often superior to that of VGG16 and AlexNet in the majority of cases. This suggests that GoogleNet, in comparison to VGG16 and AlexNet, might be an option that works better for plant leaf classification. When using an 80%-20% split, all models reach their best level of accuracy because having a bigger training set often results in improved model generalization. To avoid overfitting issues, dropout regularization is introduced. The dropout ratio is increased in multiples of 0.2, and their performances are evaluated. Figure 8 shows the performance s of the proposed system using different dropout ratios for an 80:20 split set.

Figure 8. Performance of the proposed leaf disease classification system across different dropout configurations for 80:20 split set

It can be seen from Figure 8 that the selection of the dropout ratio has a considerable influence on the level of accuracy achieved by the model. In this particular scenario, the maximum accuracy is achieved by using a dropout ratio of 0.2, which suggests that it is an appropriate option for limiting overfitting while also retaining model performance. When the dropout ratio is set to 0.8, there is a considerable decline in
accuracy that brings the overall percentage down to 47.15%. This indicates that an excessive amount of dropout regularization might be detrimental to model performance in this setting, which can result in underfitting. The findings highlight how important it is to strike a balance between minimizing overfitting and retaining the capacity of the model. A dropout ratio of 0.2 represents a decent balance, resulting in highest accuracy.

4. CONCLUSION

The classification of diseases in plant leaf images is an essential application of computer vision and machine learning in agriculture. If successful, this application has the potential to enhance crop output dramatically, minimize environmental impact, and promote sustainable farming practices. This paper presents an efficient plant leaf classification system using a ten-layer deep learning architecture. Different convolution filters are used in each convolution layer, along with a max-pooling layer. For classification, a fully connected layer is employed with different configurations, such as dropout ratio and random split ratios. Results show that the system gives 98.18% classification accuracy to classify the 38 different plant leaf diseases into either healthy or diseased using an 80%-20% split ratio for training and testing along with a 0.2 dropout ratio. In the future, this work can be extended to identify the plant leaf diseases and improve the proposed system by including the inception concept.

REFERENCES

Biographies of Authors

Muniyandi Sadhasivam is Assistant Professor at Computer Science and Engineering, SRM Institute of Technology, India. He is pursuing PhD degree in Computer Science and Engineering at Annamalai University. His research area is image processing. He has completed his Master of Engineering at Annamalai university, Bachelor of Engineering at Anna University. He received state first award in Tamil Nadu Government in India. He has been working as Assistant Professor at computer science and Engineering since 2012. He published a book ‘Computer Graphics’. He published a patent and also granted a patent. He has attended national and international conferences. He can be contacted at email: msadhasivam156@gmail.com.

Dr. Manoharan Kalaiselvi Geetha is Working as Professor at Annamalai University in Computer Science and Engineering Department in India. She Holds a PhD degree in Computer Science and Engineering with specialization in Artificial Intelligence, Machine Learning, Image and Video Processing. She guided and Awarded PhD for Six Students and she guiding Five research scholar currently. She completed two research Project wroth of twenty-four lakhs and above. She attended more than sixteen National conferences and fifteen international conferences. She has teaching Experience More than 29 years. She is life time member of ISTE and CSI. She published more than 75 international journals and five National Journal. She published three Books in well reputed Publication. She can be contacted at email: geesiv@gmail.com.

Dr. James Gladson Maria Britto is working as a professor Computer science and Engineering Department in Malla Reddy College of Engineering, India. He completed his PhD in Manonmaniam sundaranar university India. He has more than 15 years Teaching Experience. He Published a book Named Computer Graphics. He attended More than 15 international conferences and ten national conferences. He published 10 International Journals and two patents. He attended many FDP and seminar in various reputed institutions. He also a journal editor for BEIESP-Blue Eyes Intelligence Engineering and Sciences Publication Pvt Ltd from January, 2018. He can be contacted at email: cse.gladson@mrce.in.