Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 2 (2024)

Plant Disease Identification Using Convolution Neural Network

Riya Singhal[a], Manan Gaur[b], Ritika Rathi[c], Rashi Garg[d] , Mr. Md. Shahid

Department of CSE,Meerut Institute of Engineering & Technology, Meerut 250001, India

Abstract:In the present day, the prevalence of plant diseases is emerging as a significant factor contributing to reduced food production and increased losses for farmers. Hence, it is imperative to employ a method capable of delivering swift and precise results. The recent expansion of deep learning has proven instrumental in addressing both conventional and unconventional challenges more effectively. The Convolutional Neural Network (CNN) has emerged as a cutting-edge approach for state-of-the-art identification and detection. To tackle the issue of plant diseases, we have established a comprehensive dataset featuring 37 different plants and crops for training and validating our model. Our implementation involves Resnet (Residual Neural Network), a specific CNN architecture. We capture images of diseased plant leaves and employ CNN-based classification for disease detection. Our model demonstrates superior accuracy compared to numerous previously utilized techniques for disease detection.

Keywords**:** Convolution Neural Network, Resnet, EfficientNet, Neural architecture search, Focal loss function.

1. Introduction:

In India, agriculture stands as the predominant economic sector, supporting approximately 50% of the country's population [10]. A diverse range of vegetables, fruits, and pulses are cultivated, contributing around 17% to 18% to the nation's GDP [10]. Despite its significance, a substantial portion of crops faces destruction each year due to diseases and pests. Plant diseases are inevitable, often detected by farmers only when they have already affected 10% of the plant. Additionally, existing techniques for disease identification are often inadequate, potentially yielding incorrect results based on observed symptoms. The repercussions of crop failure have a profound impact on the economic well-being of farmers, contributing to adverse conditions. Notably, there has been a notable increase in the suicide rate among farmers in recent years. Therefore, the timely detection of diseases plays a crucial role in the agricultural sector to mitigate significant losses. Achieving this necessitates meticulous diagnosis and surveillance. Leaves, being a vital component of plants, offer valuable insights into the presence of diseases, making their careful examination imperative.

1.1. Deep Learning:

Globally, deep learning architectures, as outlined in reference [11], are widely applied for image identification and disease detection in both humans and plants. Issues like image recognition and classification have been significantly simplified with the adoption of deep learning techniques, including but not limited to Artificial Neural Networks (ANN) [12], Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). Extensive research has already been conducted utilizing these technologies.

1.2. Convolution Neural Network:

CNN is fundamentally a neural network designed for the identification of shapes, colors, and patterns in images with minimal processing, making it particularly effective for diverse identification and detection tasks. Each image is treated as a matrix of pixels by the computer. Through convolution with a filter matrix, a new matrix representing a different function is obtained, enabling tasks such as edge detection.

This concept allows the detection of specific patterns or spots in an image. Khatri utilized deep learning, specifically linear vector quantization ANN, for handwritten digit identification, achieving a maximum accuracy of 94.9% [9]. Our approach focuses on diagnosing diseases in brinjal leaves, employing image processing and artificial neural network techniques.

The brinjal plant leaf data is processed using the K-means clustering algorithm for segmentation, and an artificial neural network is applied for image classification [1].Pujari proposed image processing methods for the identification and classification of fungal diseases in various crops, comparing SVM and ANN classifiers, with SVM demonstrating superior accuracy [2].

Prakash developed a leaf disease detection and classification system using K-means clustering for segmentation, GLCM for feature extraction, and SVM for classification, achieving a 90% accuracy.

In our work, we cast plant disease detection as a classification problem and employ CNN to address it. The proposed CNN model demonstrates superior performance compared to other methods used for disease identification and detection. The subsequent sections outline our CNN model, discuss results, and present conclusions.

2. Method Used:

The flowchart in Figure 1 provides an overview of the steps involved in constructing the model and evaluating its performance. The initial step involves acquiring the image dataset, followed by

preprocessing tasks such as image shaping, resizing, and conversion to an array format of fixed size. Similar preprocessing steps are applied during image testing. The processed image is then augmented and loaded into the CNN model constructor.

The CNN model is trained using the database of the training set, enabling it to identify test images and the diseases they depict. Results are observed after this successful training and preprocessing phase. Following the training, the model undergoes a comparison with the test image, and the accuracy of the test within the model is noted and verified. This comprehensive process ensures the model's capability to accurately identify diseases in test images.

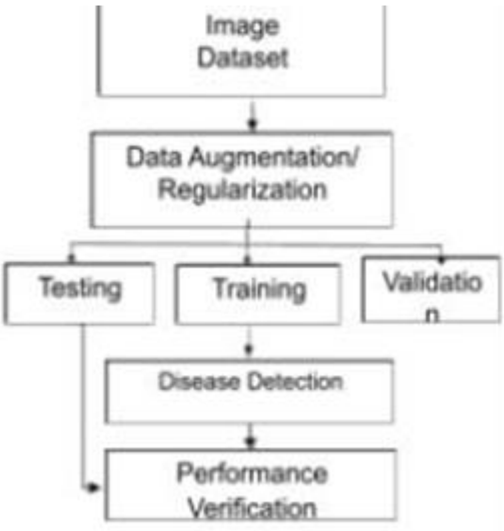

Figure 1– Flowchart of the steps involved inbuilding the model used.

2.1. CNN classification model:

In summary, the model architecture described employs the ResNet architecture in PyTorch, consisting of four main layers: convolution, pooling, dropout, and fully connected. These layers are further divided into six convolution layers, four pooling layers, two dropout layers, and four fully connected layers. The use of bottleneck layers aims to achieve the best loss. Key details include:

- Convolution layers use a stride of (1,1) and (2,2) layers.
- Padding is applied in convolution layers to prevent information loss and maintain the original image size as output.
- Test and train data are resized to a height and width of 224 pixels and loaded in batches of 128 images.
- Max Pooling is employed, and a dropout layer is used to prevent overfitting and enhance generalization.
- Instance normalization is applied for better performance, normalizing each data separately.
- ReLU activation function is used for all layers except the last one.
- Softmax activation function is applied to the last layer to obtain probabilities between 0 and 1.

Unfortunately, the detailed structure of convolution and pooling layers, as well as the overall model architecture, is not provided in the summary text but is available in a figure.

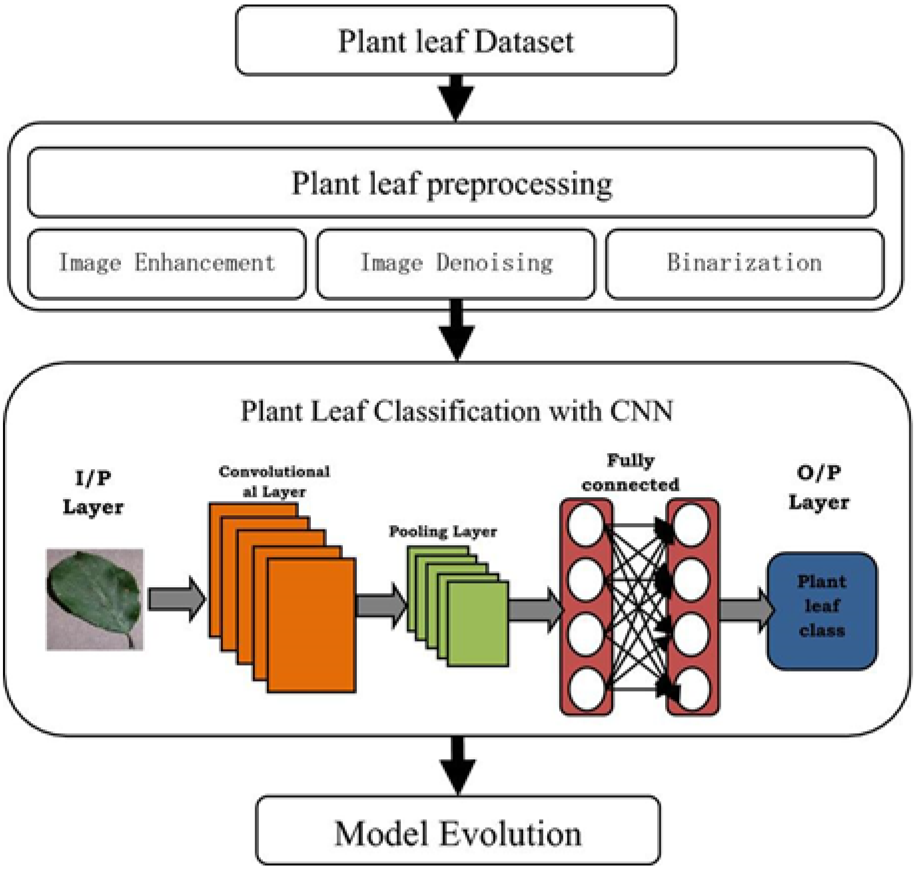

Figure 2- Detailed diagram of CNN model

In short, the described neural network model consists of various layers and processes:

- Conv2D Layer: Convolves the input image into multiple images using a specified activation function.
- MaxPooling2D Layer: Utilizes max-pooling to extract the maximum values from the dataset images, aiding in feature map calculation.
- Flatten Layer: Flattens the dimensions of the images into a single column after convolution.
- Dense Layer: Creates a fully connected layer with a specified number of nodes.
- Dropout Layer: Reduces overfitting by randomly dropping a proportion of nodes during training.

- Image Data Generator: Augments the dataset by applying transformations like resizing, shearing, zooming, and horizontal flipping to improve model robustness.
- Training Process: Involves loading images from the specified directory, using the Flow from directory function, and fitting the data into the CNN model with the fit generator. Parameters like steps per epoch indicate the number of batches in a single epoch.
- Epochs: Represents the number of cycles during the training of the dataset.
- Validation Process: Incorporates validation data to assess model performance, and the validation dataset overrides the training dataset during this phase.

2.2. Training & testing model:

It seems like your message got cut off. Could you please provide more details or complete your question about the dataset preprocessing and the 38 different diseases in plant leaves? I'd be happy to help with any information or classification you need.

Figure 3- Training Model

The dataset has been preprocessed, including actions such as resizing, shaping, and conversion to an array format at a fixed size. The dataset comprises 38 different diseases in plant leaves. Any image from this dataset can be selected and used as a test image within the software. The model is trained using a dataset containing various layers in a Convolutional Neural Network (CNN), such as Dense, Dropout, Activation, Flatten, Convolution2D, and MaxPooling2D. After successful training and preprocessing, the model can analyze and identify diseases in a test image based on its learning from the training dataset. The prediction is made by comparing the features of the test data with the patterns learned during training.

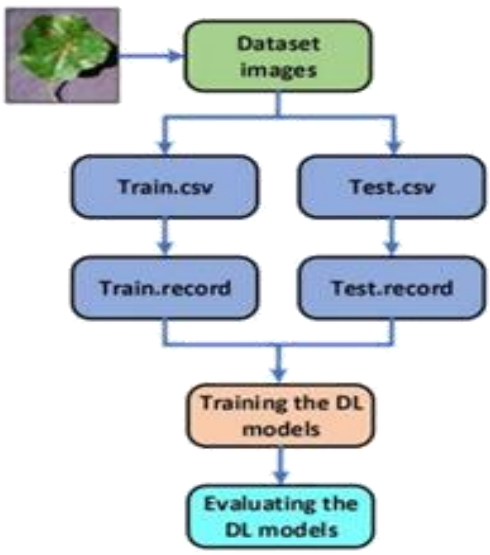

Figure 4 -Testing Model

3. Experimental Results:

The results are derived from training a Convolutional Neural Network (CNN) on a dataset comprising both original and augmented images. CNNs are effective at learning features for visual imagery, especially when trained on larger datasets, leading to improved results. The specific dataset used for this purpose consists of 30 plant diseases, encompassing thirty thousand five hundred colored leaf images.We have trained the model disease dataset. The Cross Function has been calculated by finding the Cross Entropy. The model has beenoptimized with the help of Stochastic GradientDescent technique. We have taken 0.005 as the learning rate. We have trained the network for 100 epochs. We have taken a dataset of 30 plant diseaseswhich is about thirty thousand five hundred leaf images. These images are colored Figure-5 Graph depicting the cost function with the number of epochs.

It is clearly depicted by the figure that when the epoch increased, the cost function was decreased and had become constant at the end.

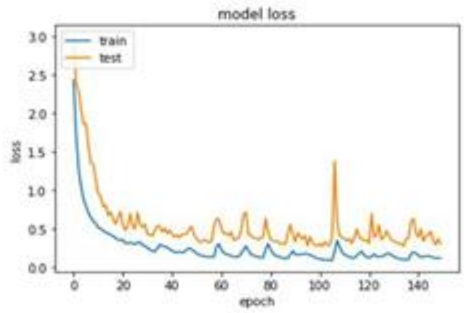

Figure-5 Depicts a graph showing the cost function's variation with the number of epochs during training.

The figure clearly shows that as the number of epochs increased, the cost function initially decreased and eventually stabilized or became constant towards the end of training.

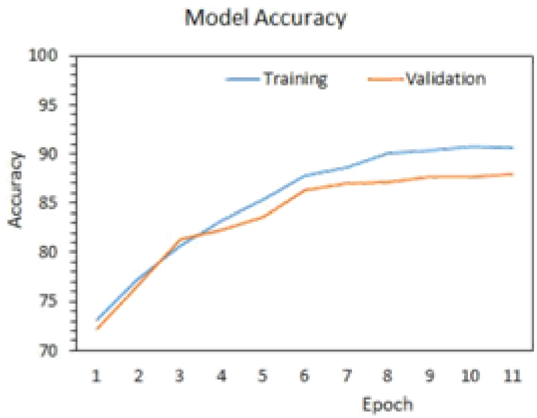

Figure-6 Illustrates a graph displaying the test accuracy of the model.

We achieved a notable accuracy in our model. After comparing it with different models and presenting the results in a table, our model outperformed the others, demonstrating superior accuracy.

4. Conclusion:

Automated plant disease detection systems leverage convolutional neural networks to identify and classify plant diseases, combining the expertise of plant pathology specialists with the ability to extract symptomatic traits. Our results indicate significant progress in predicting plant diseases using CNNs. The trend towards developing new CNN architectures for plant disease recognition is evident, demonstrating improved accuracy compared to traditional methods.

References:

[1] Anand R, V. S. J., "An Application of image processing techniques for Detection of Diseases on Brinjal Leaves using K-Means Clustering Algorithm," Fifth International Conference On Recent Trends In Information Technology, Coimbatore, 2016.

[2] Pujari, Jagadeesh D., et al., "Identification and classification of fungal disease affected on agriculture/horticulture crops using image processing techniques," 2014 IEEE International Conference on Computational Intelligence and Computing Research.

[3] Prakash, R. Meena, et al., "Detection of leaf diseases and classification using digital image processing," 2017 international conference on innovations in information, embedded and communication systems (ICIIECS).

[4] Phadikar, Santanu, and Jaya Sil, "Rice disease identification using pattern recognition techniques," 2008 11th International Conference on Computer and Information Technology. Narmadha, R. P., and G. Arulvadivu, "Detection and measurement of paddy leaf disease

[5] symptoms using image processing," 2017 International Conference on Computer Communication and Informatics (ICCCI).

[6] Orillo, John William, et al., "Identification of diseases in rice plant (oryza sativa) using back propagation Artificial Neural Network," 2014 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM).

[7] Şengür, Abdulkadir, et al., "A retinal vessel detection approach using convolution neural network," 2017 International Artificial Intelligence and Data Processing Symposium (IDAP).

[8] Pooja, V., et al., "Identification of plant leaf diseases using image processing techniques," 2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR).Khatri, Sunil Kumar, et al., "Recognizing images of handwritten digits using learning vector quantization artificial neural network," 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRTO).

[9] Financial Express (2018) " India economic survey 2018: Farmers gain as agriculture mechanisation speeds up, but more R&D needed." [Online] Accessed on 8 March, 2019 https://www.financialexpress.com/budget/india-economic-survey-2018-for-farmers-agriculture-gdp-msp/1034266/.

[10] Bengio, Yoshua, et al., "Deep learning." Nature 521.7553 (2015): 436-444.Zurada, Jack

[11] M.,"Introduction to artificial neural systems," St. Paul: West publishing company, 1992.