Research Article

Fungi Classification using Convolution Neural Network

Sukanya S. Gaikwad[a], Shivanand S. Rumma[b], Mallikarjun Hangarge[c]

[a] Research Scholar, Dept. of Computer Science, Gulbarga University, Kalaburagi, Karnataka, India.
[b]Chairman, Dept. of Computer Science, Gulbarga University, Kalaburagi, Karnataka, India.
[c]Associate Professor, Dept. of Computer Science, Karnatak Arts, Science, and Commerce College, Bidar, Karnataka, India.
[a]gsukanya116@gmail.com, [b]shivanand_sr@yahoo.co.in, [c]mhangarge@yahoo.co.in

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: This paper presents a model based on Convolution Neural Network (CNN) to identify and classify the fungi those causes disease to apple plant leaf. In this paper, apple scab, rust, black rot, and healthy leaf are studied and classified. The plant pathology dataset (publically available) consists of 9164 images are used for experimentation. The proposed CNN model identifies and classifies the apple leaves into these four categories. This model can successfully detect and classify diseases with an accuracy of 88.9%.

Keywords: Convolution Neural Network (CNN), Fungi diseases, Apple plant.

1. Introduction

The primary classification of living organisms is into five groups. Monera, Protista, Fungi, Plantae and Animalia. Among these powerful kingdoms of living organisms on earth, fungi have about 100,000 known species on the planet. The study of fungi is called Mycology, and the current research in this field is at the molecular level. For humans, the fungi are both beneficial and harmful. In this research, we are focusing on the harmful fungi which cause damage to crops and plants.

It is a known fact that India is an agricultural country, where most of the population rural in specific is dependent on agriculture for their living, thereby significantly contributing to the Indian economy. Plant diseases being the most significant reasons that lead to the destruction of plants and crops, require attention and care at many stages. Among most plant diseases, the primary cause can be attributed to the virus, bacteria, soil bacteria, airborne fungi, fungi, etc. Of these, fungi are accountable for many diseases in plants. Plant diseases affect the development of harvest yield of the plants and have social, biological and prudent effects on horticulture. The foods are grown on the ground, half lost due to fungi, as noticed in several studies. To monitor and make sure about hygiene and foodstuff, we need to keep a constant eye on the growth of fungus and its spore's essence in the current habitat. Utilization of scientific techniques becomes imperative as a part of the complex process (*i.e.*, recognizing visible symptoms and signs); natural judgment comes in handy in the process.

The manual method of inspection carried out by farmer, and agriculture experts require examining crops visually. This evaluation process is tedious, time-consuming and very subjective. Most of the methods used in identification are traditional or manual method. These require expensive equipment, expert labour and huge processing time. Hence, obtaining accurate, unbiased and spontaneous results has inspired us to propose a generic algorithm for automated, cost-effective systems for identifying and classifying the fungi.

This paper aims to identify and classify three types of fungi affected apple leaf diseases, which includes apple scab due to Venturia inaequalis, apple rust due to Gymnosporangium juniper-Virginiana, apple black rot due to botryosphaeria obtuse and a healthy leaf of apple. The proposed method uses a deep learning-based CNN model to identify and classify these different apple leaves.

The arrangement of the paper is as follows: Section II describes the literature survey. Section III narrates the dataset's preparation, and in Section IV, we describe the proposed CNN model used to identify and classify apple leaf diseases. Section V gives the experimental results and conclusions are summarized in Section VI.

2. Literature Survey

Yusuke Kawasaki *et al.*, [1] proposed a leaf-based disease identification system based on convolution neural networks (CNN). The CNN model is trained on 800 cucumber leaf images, giving good results in classifying cucumbers leaf into two disease classes and a non-diseased one. Dheeb Al Bashish *et al.*, [2] developed a software solution on image processing for automatic detection and classification of leaf diseases. The collected dataset is from Al-Ghor area in Jordan. They have tested five diseases: early scorch, cottony mould, ashen mould, late scorch, and tiny whiteness. Classification carried out using a neural network based on the backpropagation algorithm and achieved a precision of around 93%. Sa'ed Abed *et al.*, [3] focused on identifying two types of fungi affected bean diseases: bacterial brown spot and powdery mildew. The dataset included 40 testing images, which correctly classified with an accuracy of 100%. Muhammad Waseem Tahir *et.al.,* [4] proposed a CNN model for detecting different fungi types. Forty thousand eight hundred labelled images of 6 classes used to develop the fungus dataset and obtained an accuracy of 94.8%. Pujari JD *et al.,* [5] developed an automated system to detect fungi on crops like sugarcane, chilli and cotton. A total of 2616 samples used for classification; Cotton Alternaria leaf spot gave an accuracy of 94% and the sugarcane leaf redroot gave the lowest of 72%. Kuldeep S *et al.,* [6] proposed work on rust disease of pea plants to identify and classify them at the microscopic level. Five hundred images for testing; among these, correctly classified are 448. Halil Durmu *et al.,* [7] compared the accuracy of two pre-trained networks, i.e. AlexNet and SqueezeNet, from the dataset collected from plant Village. They concluded that SqueezeNet architecture gave good accuracy compared to AlexNet architecture. H. Park *et al.,* [8] developed a mechanism to diagnose and predict disease of strawberry leaf, fruit, or stem image taken by a smart phone. 1000 images were used for the model training and obtained an excellent accuracy of 89.7%. Meilani Wulandari *et al.*, [9] proposed work to identify the most poisonous fungi, Basidiomycota, which is the most common cause of death in humans from mushroom poisoning.1020 images were used for training and were able to detect fungus found to be 89.71%. H. Al-H*airy et al.,* [10] proposed an automatic detection and classification of five different diseases of plants. The experiment was conducted on 33 samples and has got an accurate detection of leaf diseases.

3. Preparation of Dataset

We are working on fungi affected apple leaf diseases. The most commonly caused leaf diseases of fungi are apple scab caused by Venturia inequality, apple rust caused by Gymnosporangium juniper-Virginiana, apple black rot caused by botryosphaeria obtuse and a healthy leaf of apple. Four different classes are involved in the study, as described below.

Apple Scab: Apple scab caused by the fungus Venturia inequalis. It typically shows up in mid-spring and often familiar through rainy weather. In the spring, during wet weather, the spores are moved away by the wind on the newly emerging leaves, which are vulnerable.

Apple scab initially appears as olive-coloured lesions on the backside of the leaves. As the fungus grows, the leaves' top sides produce similar olive- coloured lesions, and this well turns and becomes black with noted edges.

Apple Rust: Gymnosporangium juniper-Virginiana causes Apple rust. The exciting part of this disease is that it initially holds off a new plant already infected, such as cedar, to grow further. The hyphae grow into large galls. It shows yellow or orange spots on the leaves and distorted or spotted fruit..

Apple Black Rot: The black rot and frog eye leaf spot are the same disease at different disease cycles. The condition is caused by botryosphaeria obtuse. The leaves are covered with holes or have small brown spots on them. In due course, it may spread to different parts of the tree, eventually killing the tree and destroying it.

Below are instance images of diseased leaves and a healthy leaf shown in Figure 1 to 4.

Fig (1): Apple Scab

Fig (2): Apple Black rot

Fig (3): Apple Rust

Fig (4): Apple Healthy

The experimental study carried out with a publically available Plant Pathology dataset. The dataset consists of 9164 images, and they are categorized into apple scab, black rot, rust and healthy leaf.

4. Proposed Model

The proposed CNN model has sequential layers, and each layer uses the previous layer as input to the model. CNN requires minimal pre-processing and is very good at analyzing images. With its multilayered structure, CNN is good at separating the desired features.

The CNN model's basic building blocks are the Convolution layer, Pooling layer, Activation Function, and the Fully Connected layer. Below figure 5 shows the general architecture of the CNN model.

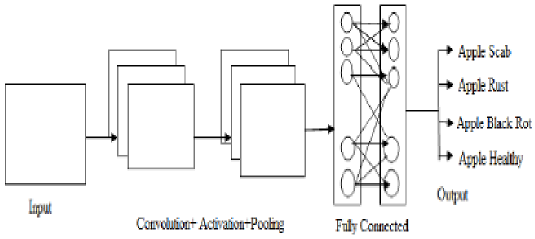


Fig 5: A general CNN architecture.

Convolution layer: The primary function of this layer is to convolve a filter (for example, 3x3 or 5x5) on the original image, i.e., the dot product of kernel (filter) and the image matrix. The obtained result summed up into a final matrix representing all the pixels obtained after the dot product. Below fig 6 shows the convolution operation.

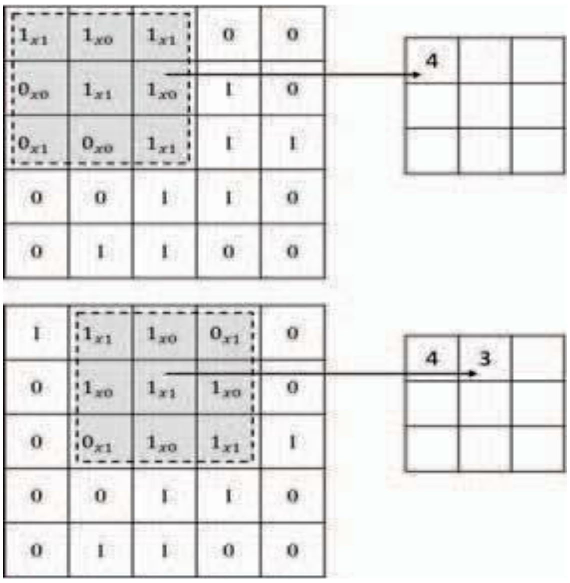


Fig 6: Convolution operation of an image with 3x3 kernel.

Activation layer: The matrix obtained at the convolution layer is small compared to the original image, and the obtained matrix run through an activation function. Some of the activation functions used in CNN are sigmoid, tanh and relu. In this model, we are using relu, and each neuron in our network activates this function. Relu function works when values more than zero are not changed, and values smaller than zero mapped to zero; as given by equation (1).

$$f(x) = 0, \quad \text{if } x < 0. \qquad (1)$$
$$x, \quad \text{otherwise.}$$

Pooling layer: The pooling function further reduces the matrix obtained from the previous layer. Here we concentrate on the essential features, which are dominant features of the image. From the matrix, we group each number which is generally the maximum (called Max pooling), which will lead the network to train it faster. Fig 7 shows the max pooling operation.

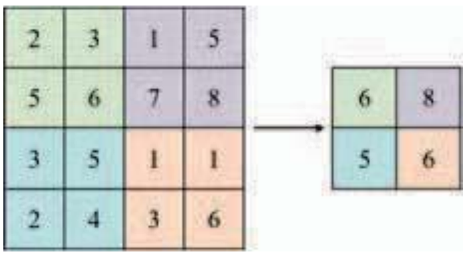


Fig 7: Max Pooling with 2x2 filter.

Fully connected layer: This layer's input is usually in a one-dimensional vector, representing the previous layer's output. This layer's output is the probabilities for different possible labels attached to the image (example: Apple scab, rot, cedar, healthy). The title that has the maximum likelihood is the classification result. Depending upon the CNN architecture, there may be multiple convolutions, activation and pooling layers. The three different input matrices, i.e. R, G and B channels for every image in the dataset, are given input to the first convolution layer. Each input image matrix is convoluted, and batch normalization is applied. Batch normalization is done to standardize the raw inputs while feeding to the next layer. After every batch normalization layer, Relu activation is used. The max pooling operation is then applied to the output matrix, which is connected to the Fully Connected (FC) layer. The FC layer's residue is connected to the softmax function, which outputs the value between 0 and 1. The adam optimizer is used to optimize the algorithm. The learning rate is at 0.001. Lastly, we train the model with 25 epochs to get accuracy. The below table 1 shows the architecture of our CNN model.

Table 1: CNN architecture of our model.

Name	**Type**	**Activations**
Image Input	Input	256*256*3
Conv_1	Convolution	256*256*8
Batch_Norm_1	Batch Normalization	256*256*8
Relu_1	RelU	256*256*8
Max_pool_1	Max Pooling	128*128*8
Conv_2	Convolution	128*128*16
Batch_Norm_2	Batch Normalization	128*128*16
Relu_2	RelU	128*128*16
Max_pool_2	Max Pooling	64*64*16
Conv_3	Convolution	64*64*32
Batch_Norm_3	Batch Normalization	64*64*32
Relu_3	RelU	64*64*32
Max_pool_3	Max Pooling	32*32*32
Conv_4	Convolution	32*32*64
Batch_Norm_4	Batch Normalization	32*32*64
Relu_4	RelU	32*32*64
Max_pool_4	Max Pooling	16*16*64
Conv_5	Convolution	16*16*128
Batch_Norm_5	Batch Normalization	16*16*128
Relu_5	RelU	16*16*128
Max_pool_5	Max Pooling	8*8*128
Conv_6	Convolution	8*8*256
Batch_Norm_6	Batch Normalization	8*8*256
Relu_6	RelU	8*8*256
Max_pool_6	Max Pooling	4*4*256
Conv_7	Convolution	4*4*256
Batch_Norm_7	Batch Normalization	4*4*256
Relu_6	RelU	4*4*256
FC	Fully Connected	1*1*4
softmax	Softmax	1*1*4
Class_output	Classification Output	--

5. Experimental Results

The dataset consists of 9164 images. The experiment conducted divides the dataset into a 60:40 split where 60% of the dataset we use for training and the remaining dataset, i.e. 40%, is used for testing purposes. The difficult task in identifying and classifying apple leaves is that the leaves with different diseases are very similar. Therefore, this similarity can lead the leaves to be mapped into the wrong classes. The CNN model trained through several iterations to classify 9164 images; we achieved an accuracy of 88.9%. The accuracy of the model is calculated as,

$$\text{Accuracy (\%)} = \frac{\text{Total number of images correctly classified}}{\text{Total number of images used for testing}} * 100$$

Class wise recognition accuracy is given in the table 2 below

Table 2: Class-wise recognition accuracy

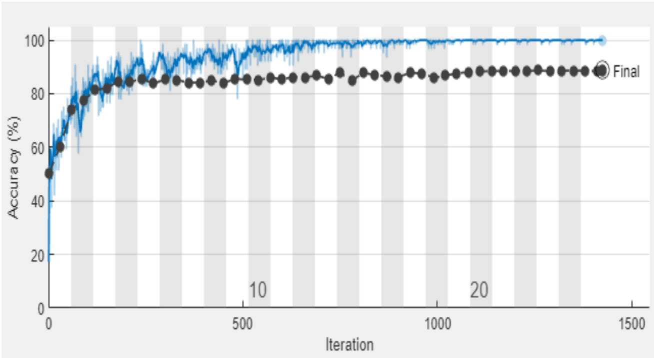
Sl. no.	Class Name	Accuracy in %
1	Apple Scab	86.6%
2	Apple Rot	86.8%
3	Apple Rust	47.6%
4	Apple Healthy	95.7%

The confusion matrix for the classified four categories is shown in the table 3 below.

Table 3: Confusion Matrix

	Apple Rot	Apple Rust	Apple Scab	Apple Healthy
Apple Rot	**931**	65	37	39
Apple Rust	118	**156**	8	46
Apple Scab	25	4	**1171**	152
Apple Healthy	13	15	90	**2628**
			Average in %	**88.9%**

The below graph fig 8 shows the accuracy obtained for training and validation data.



Fig(8) :Accuracy for training (blue line) & validation data (black line)

6. Conclusion

The proposed CNN model outperforms in the classification of fungi and yields 88.9% accuracy. The literature shows that the study of fungi classification using the Plant Pathology dataset is the first. In future, we aim to tune the network model with different architecture to achieve high classification accuracy.

7. Acknowledgement

This work is supported and funded by Karnataka Science and Technology Promotion Society (KSTePS), DST, GOVT. OF KARNATAKA.

References

1. Yusuke Kawasaki, Hiroyuki Uga, Satoshi Kagiwada and Hitoshi Iyatomi "Automated Diagnosis of Viral Plant Diseases using CNNs", Springer International Publishing Switzerland, pp. 638–645, 2015.
2. Zhang Chuanlei, Zhang Shanwen, Yang Jucheng, Shi Yancui and Chen Jia1, "Apple leaf disease identification using genetic algorithm and correlation", Int J Agric & Biol Eng, pp. 74–83, 2017.
3. Melike Sardogan, Adem Tuncer and Yunus Ozen, "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm", 3rd International conference on Computer Science and Engineering, pp. 382-385, 2018.
4. Daheeb Al Basish, Malik Braik and Sulieman Bani-Ahmad, "Detection and classification of leaf diseases using K-means based segmentation and Neural networks based classification", Information Technology Journal, pp. 267-275, 2011.
5. Zulkifli Bin Husin, Abdul Hallis Bin Abdul Aziz, Ali Yeon Bin Md Shakaff and Rohani Binti S Mohammed Farook, "Feasibility Study on Plant Chili Disease Detection Using Image Processing Techniques", Third International Conference on Intelligent Systems Modelling and Simulation, 2012.
6. Pujari JD, Yakkundimath R, Byadgi AS, "Neuro-kNN Classification System For Detecting Fungal Disease on Vegetable Crops using Local Binary Patterns", CIGR Journal, pp. 299-308, 2014.
7. Sa'ed Abed and Anwar Ali Esmaeel, "A Novel Approach to Classify and Detect Bean Diseases Based on Image Processing", IEEE Proceedings, 2018.
8. Muhammad Waseem Tahir, Nayyer Abbas Zaidi, Adeel Akhtar Rao, Roland Blank, Michael J. Vellekoop and Walter Lang, "A Fungus Spores Dataset and a Convolutional Neural Network Based Approach for Fungus Detection", IEEE Transcation on Nanobioscience, 2018.
9. Pujari JD, Yakkundimath R, Byadgi AS , "Automatic Fungal Disease Detection based on Wavelet Feature Extraction and PCA Analysis in Commercial Crops", Int J Image Graph Signal Process, pp 24–31, 2014.
10. H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and ALRahamneh, "Fast and Accurate Detection and Classification of Plant Diseases", International Journal of Computer Applications, Volume 17– No.1, 2011.