International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Review on Convolutional Neural Network to Plant Leaf Disease Classification

* Mr Dheeraj Kumar

* Mr Jayesh Patil

Prof.Gunjan Behl & Prof. Rasika Patil.

Abstract

Abstract: Crop productivity can be drastically reduced as a result of different diseases, putting food security in jeopardy. Thus, plant diseases detecting to be accurately is necessary and urgent. Traditional classification methods, such as naked-eye inspection and laboratory tests, have a number of drawbacks, including the fact that they are time-consuming and subjective. Deep learning (DL) methods, particularly those based on convolutional neural networks (CNN), are now widely used in the classification of plant diseases. They have partially solved the problems of traditional classification methods. In this paper, we looked at the most recent CNN networks that are relevant to plant leaf disease classification. We summarized the DL principles that are used to classify plant diseases. Additionally, we summarized the challenges and their respective solutions of CNN used in classification of plant disease.

Keywords: Classification of plant disease, DL, CNN.

Author correspondence:
Dheeraj Kumar
Email:-dheerajkymar0411@gmail.com
Mr.Jayesh patil
Email:-jayesp14698@gmail.com

1 Introduction

Plant diseases have a huge impact on the production of food crops. An iconic example is the Irish potato famine of 1845–1849, which resulted in 1.2 million deaths (2). The diseases that affect a variety of common plants are shown in Table 1. Plant diseases can be divided into fungal, oomycete, bacterial, and viral types. Researchers and farmers have never stop trying to figure out how to create an intelligent and effective approach or method for classifying plant diseases. Laboratory test approaches to plant samples, such as polymerase chain reaction, enzyme-linked immunosorbent assay, and loop-mediated isothermal amplification, are highly specific and sensitive in identifying diseases (2).

TABLE 1

1. Common diseases of several common plants.

Plant	Fungal	Bacterial	Viral	
Cucumber	Downy mildew, powdery mildew, gray mold, black spot, anthracnose	Angular spot, brown spot, target spot	Mosaic virus, yellow spot virus	(3)
Rice	Rice stripe blight, false smut, rice blast	Bacterial leaf blight, bacterial leaf streak	Rice leaf smut, rice black-streaked dwarf virus	(3; 4)
Maize	Leaf spot disease, rust disease, gray leaf spot	Bacterial stalk rot, bacterial leaf streak	Rough dwarf disease, crimson leaf disease	(4)

However, conventional field scouting for diseases in crops still relies primarily on visual inspection of the leaf color patterns and structures. People observe the symptoms of diseases on plant leaves with the naked eye and diagnose plant diseases based on experience, which is time and labor consuming and requires specialized skills (6).

At the same time, due to the diversity of plants, the disease features of different crops differ, this situation adds to the complexity of plant disease classification. Meanwhile, numerous stud- ies have focused on using machine learning to classify plant diseases. The following three steps are involved in using machine learning technologies to detect plant diseases: first, using prepro- cessing techniques to remove the segment that are infected part; second, extracting the main features for further analysis; third, using supervised classification or unsupervised clustering algorithms to classify the features (7; 8; 9). Most machine learning studies have focused on the classification of plant diseases by using features, such as the texture (7), type (7), and color (8) of plant leaf images. The main classification methods include support vector machines (10), K-nearest neighbor (11), and random forest (12).

2 Related Work

Many work were done related to the plant disease classification using CNN. They are sum- marized below.

Mohanty et al. tested with a public database of 54,306 images collected under controlled conditions, researchers used the traditional network models AlexNet and GoogLeNet to detect

14 crop species and 26 diseases. They achieved a high accuracy of 99.35%, demonstrating the method's viability. However, because of the insufficient diversity in the training set, the model accuracy was considerably lowered when it was evaluated on a set of images collected under settings other than those used for training. (14).

Ferentinos used CNN models (i.e., AlexNet, GoogLeNet, and VGG) to detect and recognize plant diseases with a public dataset PlantVillage. When the model was trained and tested with PlantVillage, the best success was 99.53% with the VGG model. However, when they trained the VGG model with laboratory images and tested it with field images, the success rate was only up to 33.27% (5).

Fuentes et al. was to develop a reliable DL-based detector for real-time tomato disease and pest detection. All images of plant pest and diseases were captured, including background differences, varying lighting situations, and objects of various sizes. Because of the insufficient amount of samples, the precision would be lower in practical application (16).

Mukti et al. used a transfer learning model based on ResNet50 to identify the plant diseases. Their dataset contains 87,867 images. A total of 80% of the dataset was used for training and 20% for validating. They achieved a high accuracy of 99.80% (1).

Coulibaly et al. used using transfer learning to recognize mildew diseases in pearl millet. This approach was based on CNN model VGG16 with public dataset ImageNet. They achieved a high accuracy of 94.50% (17).

TABLE 2

Pertained Model	**Dataset**	**Best Accuracy**	
ResNet50	PlantVillage (extended)	99.80%	(1)
VGG16	Millet crop images (own)	95.00%	(9)
VGG16	Plant images	93.00%	(17)
VGGNet	ImageNet	91.83%	(15)

3 Methodology

DL is a branch of machine learning and image categorization, object recognition, and natural language processing are some of the applications of deep learning.

DL is a neural network-based approach for automatically selecting data features. For discovering dispersed features and properties of sample data, it combines low-level features toproduce abstract high-level features. Its accuracy and generalization ability in image recognition and target detection are superior to traditional methods. The most common types of networks today are multilayer perceptron, CNN, and recurrent neural network (RNN).

CNN is the most widely used for plant leaf disease classification. CNN consists of convolutional, pooling, and fully connected layers. The convolutional layer uses the local correlation of the information in the image to extract features.

The process of convolution operation and pooling operation is shown in Figure 1 and Figure 2.

Input(6x6x1)

3	4	6	5	1	3
5	3	2	4	3	2
5	4	3	3	2	6
1	1	2	5	3	3
2	3	3	4	1	2
3	3	2	4	2	4

Filter(3x31)

1	0	1
1	0	1
1	0	1

Stride=1

Output(4x4x1)

24	23	17	23
18	20	15	24
16	20	14	24
13	20	13	23

Figure 1. The process of convolution operation.

Input(4x4x1)

24	23	17	23
18	20	15	24
16	20	14	24
13	20	13	23

Stride=2
Max Polling

Output(2x2x1)

24	24
20	24

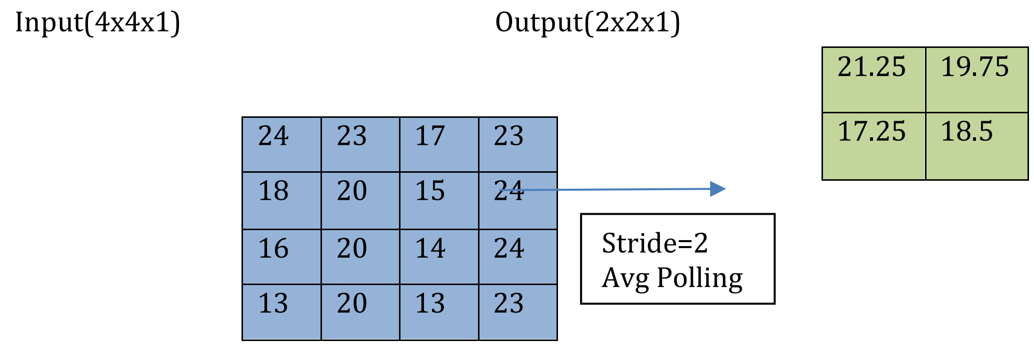

Figure 2. The process of pooling operation.

For classification tasks, various CNN-based classification models have been developed in DL- related research, including AlexNet, VGGNet, GoogLeNet, ResNet, MobileNet, and EfficientNet. AlexNet was proposed in 2012 and was the champion network in the ILSVRC-2012 competi- tion (13).

Until 2015, Microsoft lab proposed the ResNet network and won the first place in the clas- sification task of the ImageNet competition (13). The classification of images is a fundamental task in computer vision. It is also the foundation of technologies such as object detection, image segmentation, image retrieval, and others.

In CNN-based architecture to extract features, which mainly include convolutional, max- pooling, and full connection layers. The convolutional layer is mainly used to extract features of plant leaf images. The convolutional layer extracts some edge and texture information, the medium layer extracts complicated texture and part of the semantic data, and the deep layer extracts high-level semantic elements.

The convolutional layer is followed by a max-pooling layer, which is used to extract the necessary information of the image. At the last step this classifier is used to classify the high- level semantic elements extracted by the feature extractor.

4 Comparative Analysis

CNN-based models are the most used for identifying plant leaf diseases. This section intro-duces and summaries the issues and solutions that have arisen throughout the development of CNN-based DL techniques for plant disease detection and classification.

The problems are caused by external and internal factors are mentioned below: -

4.1 Inadequate Datasets

The most significant issue with CNN-based DL's application of plant disease classification is a lack of large and diverse datasets. This situation is partially responsible for all of the other issues that have been presented.

Mohanty et al. tested the classic network models AlexNet and GoogLeNet with a public database of 54,306 images collected under controlled conditions to identify 14 crop species and 26 diseases. They obtained a top accuracy of 99.35%, which demonstrates the feasibility of this method. However, the accuracy of the model was greatly reduced when

it was tested on a set of images taken under conditions different from the images used for training because of the insufficient diversity of the training set. (14)

Fuentes et al. aimed to introduce a robust DL-based detector for real-time tomato disease and pest recognition. All images of plant diseases and pests were taken in-place, including background variations, different illumination conditions, and multiple sizes of objects. The precision would be lower in practical application due to the insufficient number of samples. (19)

Ample datasets have a significant impact on the practical application. However, external fac- tors such as season and climate can readily influence data collection, and image classification is a time-consuming and hard operation. Because of these considerations, creating a useful dataset is incredibly challenging. Transfer learning, data augmentation approach, citizen research, and data sharing are currently methods for resolving dataset issues.

TABLE 3

Studies on transfer learning technology applied to the identification task.

Pertained Model	Dataset	Best Accuracy	
ResNet50	PlantVillage (extended)	99.80%	(1)
VGG16	Millet crop images (own)	95.00%	(9)
VGG16	Plant images	93.00%	(17)
VGGNet	ImageNet	91.83%	(15)

The data augmentation technologies can efficiently increase the number of datasets. We show some traditional image data augmentation methods, such as rotation, mirror symmetry, and adjusting saturation.

Liu et al. used data augmentation technologies to solve the problem of insufficient apple pathological images for the identification of four apple leaf diseases. The researchers used direction disturbance (rotation transformation and mirror symmetry), light disturbance, and principal component analysis jittering to disturb natural images. With the application of these image processing technologies, the dataset expanded from 1053 images to 13,689 images, and the accuracy with the expanded database improved 10.83%. (20)

4.2 Symptom Variations

When detecting plant illnesses, we normally presume that the disease's symptoms would remain constant.

The symptoms of plant diseases are the results of the interaction of diseases, plants, and the environment (23). Changes in any of the three can cause changes in disease symptoms, as explained further below.

In general, plant disease has the following three variations: (1) at different development stages of the disease, the symptoms shown may be different (23; 24); (2) in the same period, multiple diseases may be observed on the same plant leaves. If multiple diseases are clustered together, then the symptoms may change drastically, which brings difficulty in identifying the types of diseases (24); (3) similar symptoms may appear among different diseases, which increases the difficulty of disease classification.

The combination of diseases, plants, and the environment can cause a wide range of symp- tom changes, which makes image capture and annotation difficult. This problem can be solved using two methods: -

1. Collecting images of specific diseases that contain the entire range of variation. (24)
2. Gradually enriching the diversity of the database in practical applications (23)

The first method is unrealistic since gathering photos of the entire range of variation is atime- and money-intensive undertaking, and it's uncertain whether researchers have captured all variations. The alternative way is far more realistic, and it is currently being utilized frequently by academics to efficiently boost data diversity.

TABLE 4

Studies on using data augmentation technologies to expand the dataset.

Expanded dataset	Methods	Best Accuracy	
From 1053 to 13,689 images	Direction disturbance and light disturbance and PCA jittering.	97.62%	(20)
From 10,820 to 32,460 images	Noise addition, color jittering, and radial blur	96.17%(improved 3.15%)	(21)
From 54,309 to 87,848 image	Cropping, resizing	99.53%	(5)
From 1567 to 46,409 images	Segmentation, resizing	94.00% (improved 12%)	(22)

5 Discussion

summaries and clarifies all pertinent information to assist readers in selecting one or more criteria and comparing various DL models at a glance. Most writers adopt comparable network designs, as indicated in Table 5, and hence achieve similar experiment results.(5)

TABLE 5

Studies on different CNN methods applied to plant leaf disease identification.

Task	Dataset	Best Accuracy	
AlexNet, GoogLeNet	Identify 14 crop species and 26 diseases	54,306 images from PlantVillage	99.35% (14)
VGGNet	Detect diseases and pests in tomato plants using images captured in-place by camera devices	5000 images	83.00% (27)
VGGNet	Identify rice and maize leaf diseases	500 & 466 images	92.00% (25)
AlexNet, GoogLeNet	Classify nine diseases of tomato leaves	14,828 images	99.18% (26)
AlexNetOWTBn, VGG	Plant disease detection and diagnosis	87,848 images (PlantVillage)	99.53% (best)

6 Conclusion

In plant disease detection and categorization, DL approaches have become widely used. Traditional machine learning approaches have been solved (or partially solved) by it. Picture classification, target identification, and image segmentation are all applications of deep learning (DL), which is a subset of machine learning. In this paper, we reviewed the latest CNN networks relevant to plant leaf disease classification.

We describe how CNN methods are used to classify plant diseases and highlight the DL prin- ciples that are involved in this process. We also summarize some problems and corresponding solutions of DL used for plant disease classification with external and internal factors as men- tioned below: -1. Inadequate Datasets and 2. Symptom Variations.

Furthermore, we discussed the future development direction in plant disease classification, for example, plant electrophysiology and the combination of the mobile phone client and the server-side program would be good future research directions (28).

References

[1] I Z Mukti and D Biswas. Transfer Learning Based Plant Diseases Detection Using ResNet50. *Proceedings of the 2019 4th Interna- tional Conference on Electrical Information and Com- munication Technology (EICT)*.

[2] D P Hughes and M Salathe, An open access repository of images on plant health to enable the development of mobile disease diagnostics.

[3] J Kianat, M A Khan, M Sharif, T Akram, A Rehman, and T Saba, A joint framework of feature reduction and robust feature selection for cucumber leaf diseases recognition.

[4] S Zhang, S Zhang, C Zhang, X Wang, and Y Shi, Cucumber leaf disease identification with global pooling dilated convolutional neural network.

[5] K Ferentinos, Deep learning models for plant disease detection and diagnosis.

[6] S Sankaran, A Mishra, R Ehsani, and C Davis, A review of advanced techniques for detect- ing plant diseases.

[7] J G Barbedo Arnal, An Automatic Method to Detect and Measure Leaf Disease Symptoms Using Digital Image Processing.

[8] Q Feng, L Dongxia, S Bingda, R Liu, M Zhanhong, and W Haiguang, Identification of Alfalfa Leaf Diseases Using Image Recognition Technology.

[9] E Omrani, B Khoshnevisan, S Shamshirband, H Saboohi, N B Anuar, and M H N Nasir, Potential of radial basis function- based support vector regression for apple disease detec- tion.

[10] T Rumpf, A K Mahlein, U Steiner, E C Oerke, H W Dehne, and L Plümer, Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance.

[11] E Hossain, M F Hossain, and M A Rahaman, A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease Using KNN Classifier.

[12] R M Mohana, C K Reddy, K Anisha, P R Murthy, and B R, Random forest algorithms for the classification of tree-based ensemble.

[13] K He, X Zhang, S Ren, and J Sun, Deep Residual Learning for Image Recognition.

[14] S P Mohanty, D P Hughes, and S Marcel, Using Deep Learning for Image-Based Plant Disease Detection.

[15] J Chen, J Chen, D Zhang, Y Sun, and Y A Nanehkaran, Using deep transfer learning for image-based plant disease identification.

[16] A Fuentes, S Yoon, S C Kim, and D S Park, A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition.

[17] S Coulibaly, B Kamsu-Foguem, D Kamissoko, and D Traore, Deep neural networks with transfer learning in millet crop images.

[18] A Abdalla, H Cen, L Wan, R Rashid, and Y He, Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field.

[19] A Fuentes, S Yoon, S C Kim, and D S Park, A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition.

[20] L Bin, Z Yun, H Dongjian, and L Yuxiang, Identification of Apple Leaf Diseases Based on Deep Convolutional Neural Networks.

[21] Z Lin, S Mu, A Shi, C Pang, G Student, X Sun, and G Student, A Novel Method of Maize Leaf Disease Image Identification Based on a Multichannel Convolutional Neural Network.

[22] Arnal Barbedo and J G, Plant disease identification from individual lesions and spots using deep learning.

[23] S Coulibaly, B Kamsu-Foguem, D Kamissoko, and D Traore, Deep neural networks with transfer learning in millet crop images.

[24] J G Barbedo Arnal, A review on the main challenges in automatic plant disease identifica- tion based on visible range images.

[25] J Chen, J Chen, D Zhang, Y Sun, and Y A Nanehkaran, Using deep transfer learning for image-based plant disease identification.

[26] M Brahimi, K Boukhalfa, and A Moussaoui, Deep Learning for Tomato Diseases: Classifi- cation and Symptoms Visualization.

[27] A Fuentes, S Yoon, S C Kim, and D S Park, A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition.

[28] S K Chatterjee, O Malik, and S Gupta, Chemical Sensing Employing Plant Electrical Signal Response-Classification of Stimuli Using Curve Fitting Coefficients as Features.