ISSN: 2349 – 4891

International

Journal of Recent Research and Applied Studies

(Multidisciplinary Open Access Refereed e-Journal)

Early Detection and Categorization of Corn Leaf Diseases using Deep Learning Model

R. Jayanthi[1], A.Ahamed Noorea Fayaz[2], S.Poovarasan[3], A.Praveen[4]

[1]Professor and Head, [2,3,4]UG Students – Final Year, Department of Electronics and Communication Engineering, Nandha College of Technology, Perundurai, Tamilnadu, India

Abstract

Agriculture is a very significant field for increasing population over the world to meet the basic needs of food. Many farmers are cultivating in remote areas of the world with the lack of accurate knowledge in disease detection, however, they rely on manual observation on grains and vegetables, as a result, they are suffering from a great loss. Accurate detection of corn leaf diseases is a complex challenge faced by farmers during the growth and production stages of corn. Digital farming practices can be an interesting solution for easily and quickly detecting plant diseases. To address such issues, this paper proposes a method based on an improved deep learning Convolutional Neural Network (CNN) model for accurately detecting three common diseases of corn leaves: gray spot, leaf spot and rust. First, this technique is applied on Kaggle datasets of corn leaves to investigate the symptoms of unhealthy leaf. Then, the feature extraction and classification process are performed in dataset images to detect leaf diseases using CNN model with applying image processing. For three corn leaf diseases, the approach citied in this paper has an average accuracy of 96%. It has a better accuracy than the other pretrained model. The deep learning algorithm proposed in this paper is of great significance in intelligent agriculture, ecological protection and agricultural production.

Keywords: Cornleafdiseasedetection,Deeplearning,Convolutional Neural Network(CNN), Image processing.

1. Introduction

Corn is currently the highest-yielding food crop around the world, an important food, and industrial raw material. The stable and healthy development of corn production plays a pivotal role in food security, farmer’s income growth and the national economy. Corn diseases directly affect its yield and quality. There are more than a dozen common diseases in corn, most of which occur in leaves, ears, and roots. Among them, leaf spots and rust are typical [1]. Leaf spot, there are oval or rectangular, spindle-shaped lesions on the leaves, with yellow-brown halos around them, 5-10cmlongand1.2-.5cmwide.Inseverecases, several lesions are connected, and the leaves die early. Rust disease mainly occurs in the middle and upper leaves of the plant. At first, small light-yellow dots scattered or clustered on the front of the leaf, then protruded and expanded to round to oblong, yellowish-brown, or brown, and the surrounding epiderm is turned up. Gray leaf spot, also known as corn Cercosporaleaf spot and corn mildew, is a more severe disease. The initial stage of the disease is light brown spots in the shape of water stains, which extend parallel to the veins and are often rectangular and spread all over the leaves.

It may be more difficult for inexperienced farmers to detect diseases than for professional plant pathologists [2]. As a verification system in disease diagnostics, an automatic system that is designed to identify plant diseases by the plant's appearance and visual symptoms could be of great help to farmers. Many efforts have been applied to the quick and accurate detection of leaf diseases which will be very helpful for the farmers. By using the digital image processing techniques, support vector machine (SVM), neural networks and other methods, we can detect and classify leaf diseases. An SVM - based multi - classifier was proposed by Song et al. [3] and was applied to identify a variety of maize leaf diseases. The best recognition accuracy was 89.6%. The method of classification using SVM is only applicable to small samples, for a large number of samples, it cannot achieve high recognition accuracy. Chen and Wang [4] proposed a method for the identification of maize leaf diseases based on image processing technology and a probabilistic neural network (PNN). The best recognition accuracy of this method was 90.4%. However, for the PNN classifier, the identification accuracy and speed of this method decreases as number of training sample increases. Deep learning has made tremendous advances in the past few years. It is now able to extract useful feature representations from a large number of input images. Deep learning provides an opportunity for detectors to identify crop diseases in a timely and accurate manner, which will not only improve the accuracy of plant protection but also expand the scope of computer vision in the field of precision agriculture. This paper proposes a corn leaf disease detection method based on the deep learning to improve corn leaf disease detection accuracy, to train the deep learning model and improve the results. The proposed CNN model is compared with classic deep learning models such as VGG Net and ResNet to compare the results. The purpose of this work can effectively detect three common corn leaf diseases, which can be applied to the agricultural sector for crop protection. To obtain a highly corn leaf disease identification accuracy, it is highly significant to design a recognition model with fewer parameters and higher recognition accuracy. The rest of this paper is organized as follows: Section II gave the related work. Section III introduced corn data and preprocessing methods and proposed a deep learning model for corn disease detection. Section IV explained the proposed model description and in Section V analyzed the experimental results. Finally the paper is concluded in Section VI.

2. Related Work

Accurate modeling and finding the most critical factors in the analysis is one of the required steps in preprocessing stage. However, the convolutional neural network is critical for feature extraction. It can automatically extract image features and has good adaptability to image displacement, scaling, and distortion [5]. Therefore, deep learning models are applied in the current research because of their excellent efficacy [6]. Deep learning is rapidly becoming the standard technology for image classification [7]. It has been applied to many fields such as medical image recognition, remote sensing image recognition, autonomous vehicle drivand face recognition, text clustering, lunar impact crater identification, and age estimation, epidemic prevention and control [8]. In the field of agriculture, many studies have

ISSN: 2349 – 4891
International
Journal of Recent Research and Applied Studies
(Multidisciplinary Open Access Refereed e-Journal)

been conducted on the classification of plant pests and diseases [9], such as tea [10], apple [11], [12], rice [13], mango [14], cucumber [15], etc.

Saeed et al [16] proposed an automated crop disease identification system that was evaluated algorithmically on tomato, corn and potato crops. They used partial least square (PLS) regression, fusion and selection of features extracted by the CNN model, which were then passed to multiple classifiers to obtain the final recognition. Multiple classifiers can be used to improve accuracy of the model. The average accuracy achieved by the PLS-based fusion and selection method is about 90.1%, which not only improves the recognition accuracy but also reduces the computation time.

3. Materials and Methods

3.1 Dataset

The corn data set used in this study is from the kaggle dataset for our research performance analysis which contains healthy or unhealthy leaf images. Three types of corn leaves (gray spot, rust and leaf spot) were selected for detection. Thousand images of each disease are chosen along with thousand healthy images resulting in a total of four thousand disease images. Part of the disease images is illustrated in Fig. 1.

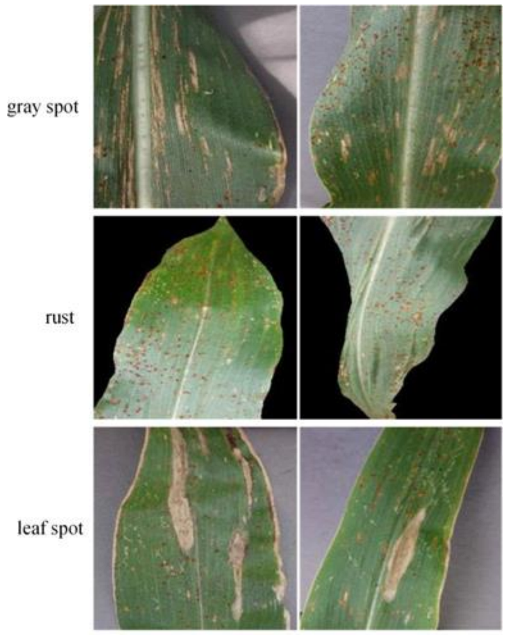

Figure 1 Corn Leaf Disease Images

3.2 Image Preprocessing

The process of pre-processing technique transforms raw input leaf image data sets into desirable process datasets format to develop the quality of leaf images and to eliminate the undesired portions from the leaf images. These processes occur in various phases such as data cleaning, integration, reduction and transformation which are shown in Fig. 2.

In the data cleaning phase, it eliminates the undesired distortion, manages the missing data and rectifies then consistent data. At the integration stage, multiple and the heterogeneous data as well as data redundancy in leaf image datasets is an ordinarily encounter situation of data retrieval strategies which resolves multiple data conflicts and arranges a unified representation of data. A large size datasets increases the storing space size and computational difficulty due to the different feature dimensions. In the process of data reduction, a large volume of data is reduced to increase the performances and efficiency of image processing. The operations of data transformation perform the data smoothing, aggregation, feature construction, data normalization and discretization to inhibit the dependability of the attributes in the data assessment structures and units for data images conversion. These leaf image datasets are resized and converted into 256×256 dimension for training datasets and testing datasets analysis. So, pre-processing technique can provide preparing datasets to identify leaf diseases through the leaf image datasets.

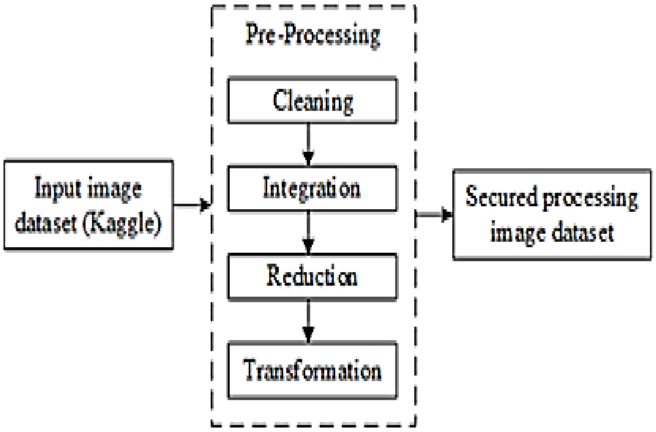

Figure 2 Phases of Leaf Image Preprocessing

3.3 Image augmentation

Image augmentation is involved for changing and facilitating of the leaf image representation to accurately identify leaf disease. Thus, the training and testing leaf image datasets are augmented to diminish the chance of over-fitting and to enrich the simplification of the model. The process of augmentation is applied to resize the original leaf image dataset using flipping, cropping and rotation techniques as well as to convert the leaf images into RGB using color transformation technique. However, the augmented leaf images are created to maintain the balanced quality and size of images in the healthy and unhealthy leaf datasets.

3.4 Feature Extraction

ISSN: 2349 – 4891
International
Journal of Recent Research and Applied Studies
(Multidisciplinary Open Access Refereed e-Journal)

Feature extraction is the very important phase of the image processing technique to provide a suitable platform and optimal constraints. The feature extractor of the CNN based detection framework can extract the image feature vectors of the leaf disease.

The feature extraction technique analyzes the properties of a leaf image such as color, shape and texture in a convenient way. So, this extraction technique is able to assist in proper classification of different leaf disease classes. For the leaf diseases, the feature extraction mechanism extracts the features of various lesion shapes and colors.

3.5 Convolutional Neural Network

Convolutional neural networks are a category of neural networks designed for image recognition and classification and have achieved excellent results. Different from the traditional approaches, CNNs can learn high- level robust features directly from the original image instead of extracting the specific features manually. In the identification of plant species and diseases, it is demonstrated that CNNs can provide better performance than the traditional feature extraction methods. A typical CNN architecture mainly consists of convolution layers, pooling layers, and full connection layers, which are described as follows.

A. Convolutional Layers

The convolutional layer is the crucial component of CNN, which extracts the specific features of the image by the different sizes of the convolution kernel. A set of feature maps of input images can be extracted after applying the convolutional layers several times. Let Hi represent the feature map of the i^{th} layer in CNN, then the Hi can be generated as follows:

$$H_i = \varphi\ (H_{i-1}W_i+b_i)$$

Where H_i denotes the feature map of the current network layer, H_i-1is the convolution feature of the previous layer (H0is the original image.), W_i is the weight of the i-th layer, b_i is the of set vector of the i-th layer, and φ (·) represents the rectified linear unit (ReLU) function.

B. Pooling Layers

The function of pooling layers is reducing the spatial dimension, which can reduce computational complexity and effectively control the risk of over-fitting. In the l-th pooling layer, the output feature on the j^{th} local receptive field can be calculated by

$$x_j^l= down\ (x_{jl}\text{-}1,\ s)$$

Where, down (·) representsthedown-samplingfunction,$x_j^{l-1}$is the feature vector in the previous layer, and sis the pooling size.

C. Fully connected layers

ISSN: 2349 – 4891
International
Journal of Recent Research and Applied Studies
(Multidisciplinary Open Access Refereed e-Journal)

After convolutional and pooling layers, there are one or several fully-connected (FC) layers, and the purpose of FC layers is using the extracted features for image classification. The Softmax function is usually employed to conduct the class prediction with the features extracted from the previous layers. Mathematically, the Softmax function is written by

$$\text{Softmax } (z)_j = e^{zj} / \sum\nolimits^{k}_{k=1} e^{zk} (for j = 1,...,K)$$

Where K represents the dimension of the z vector.

3.6 Equipment

The Keras framework, Google colab development environment and Python language are used to train and test the complete model on a computer. The relevant parameters are shown in Table 1.

Table 1 Hardware and software parameters

Name	**Parameters**
Memory	Above 4 GB
Processor	Intel core i5
Graphics	NVIDIA
Operating system	Windows 10 64 bits
Development environment	Google colab
Language	Python

4. Proposed Model Description

Dataset is divided into three parts one for training, one for validating and other for testing which are shown in Fig. 4.1. Splitting of dataset is 70/20/10 ratio randomly. 70% for the training dataset, 10% for validating dataset and 10% for testing dataset. Training dataset consists of 2698 images and testing, validating data consists of 400 images. Training of model is done using 2698 images were kept unseen by model so that accuracy of model can be checked. First some pre-processing is applied on dataset in form of augmentation to increase size of dataset in order to achieve better accuracy. Then images sizes are reduced by 256x256pixels. After that a convolution neural network based model will be created with multiple pooling and convolution layers and a dense layer for prediction. Seven convolution layers with 3x3 filter are used and seven Max Pooling 2D layers with 2x2. Batch normalization is also used in this mode. Batch normalization is used to scale data on particular scale but the difference is that it not just does it on input layer but it also does it at other hidden layers. At last model is trained on Plant Village dataset.

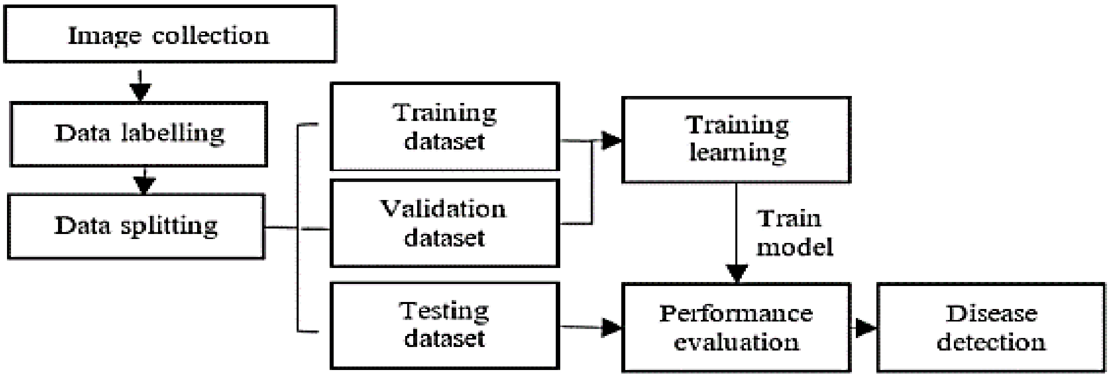

Figure 3 Working flow of proposed CNN model

Table 2 Parameters of proposed CNN model

Layer Name	Output Shape	Parameters
Convolutional Layer C1	3 x 3	896
Pooling layer P1	2 x 2	0
Convolutional Layer C2	3 x 3	9248
Pooling Layer P2	2 x 2	0
Batch Normalization B1	___	128
Convolutional Layer C3	3 x 3	9248
Pooling Layer P3	2 x 2	0
Batch Normalization B2	___	128
Convolutional Layer C4	3 x 3	9248
Pooling Layer P4	2 x 2	0
Batch Normalization B3	___	128
Convolutional Layer C5	3 x 3	9248
Pooling Layer P5	2 x 2	0
Batch Normalization B4	___	128
Convolutional Layer C6	3 x 3	9248
Pooling Layer P6	2 x 2	0
Batch Normalization B5	___	128
Convolutional Layer C7	3 x 3	9248
Pooling Layer P7	2 x 2	0
Batch Normalization B6	___	128
Flatten	___	0
Dense	___	33024
Dense 1	___	1028

Total Parameters: 91204
Trainable Parameters: 90820
Non-trainable Parameters: 384

IJRRAS
ISSN: 2349 – 4891
International
Journal of Recent Research and Applied Studies
(Multidisciplinary Open Access Refereed e-Journal)

5. Results and Discussions

This study shows the importance of plant disease detection in these days. This model was developed using Deep Learning in python.10% images from Plant Village dataset were used to test the accuracy of this model. These images are from 4 different classes. 10% of each class randomly selected for testing. Some real time images were also used. Those images were captured from local environment. They do not belong to any class which are present in dataset. But model give us more than 96% accuracy on those images as well by telling either leaf is healthy of unhealthy. Total 100 images were used and 96 were classified correctly. Some images were captures at night with the help of flash light and some images have dirt upon it so that they were misclassified. Some of the images which we captured from local environment. Testing dataset gives accuracy more than 96%. Below is the Training and Validation accuracy graph generated by our model on testing dataset.

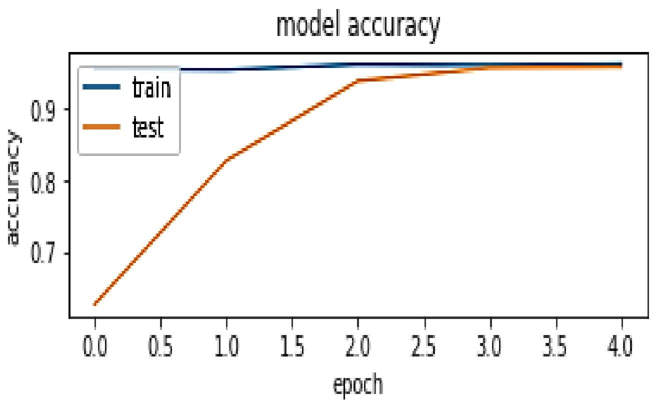

Figure 4 Training and Testing Accuracy of Proposed Model

Table 3 Classification of proposed CNN model

Class	Total	Correctly Classified	Misclassified
Corn_(maize)_healthy	186	186	0
Corn_(maize)_Grey_spot	182	175	7
Corn_(maize)_Common_rust	191	188	3
Corn_(maize)_Northern_Leaf_Blight	191	175	16

Table 4 Final model performance

Model	Dataset for Training	Dataset for Testing	Training Accuracy	Testing Accuracy
CNN	Plant Village (80%)	Plant Village (20%)	97%	96%+
CNN	Plant Village (80%)	Actual Environment (100Images)	97%	95%+

6. Conclusion and Future Work

This study has utilized deep learning capabilities to achieve automatic corn leaf disease detection system. This system is based on a simple classification mechanism which exploits the feature extraction functionalities of CNN. For prediction finally, the model utilizes the fully connected layers. The research was carried out using the publically accessible collection of 4000 images, and 100 images from experimental conditions and actual environment. The system has achieved an overall 96% testing accuracy on publically accessible dataset. It is concluded from accuracy that CNN is highly suitable for automatic detection and diagnosis of plants. This system can be integrated into mini-drones to live detection of diseases from plants in cultivated areas. Though this system is trained on Plant Village dataset with only 4 classes it could tell if the plant has a disease or not as somehow symptoms are same in all kinds of plants. In addition, more actual environment images can be added to the dataset to improve the accuracy on real-condition images of leaves and classify more plant types as well as disease types. In the future, this system can also adopt 3 layer approach where the first layer detects if there's any plant in an image or not, second layer tells the plant type and the third layer tells if there is any disease or not and what type of disease is there if any.

References

1. X. Zhang, Y. Qiao, F. Meng, C. Fan, and M. Zhang, ''Identification ofmaize leaf diseases using improved deep convolutional neural networks,''*IEEE Access*, vol.6, pp.30370–30377, 2018.

2. S. A. Miller, F. D. Beed, and C. L. Harmon, ''Plant disease diagnos- tic capabilities and networks,'' Annu. Rev. Phytopathol., vol. 47, no. 1, pp. 15–38, 2009.

3. K. Song, X. Y. Sun, and J. W. Ji, ''Corn leaf disease recognition based on support vector machine method,'' Trans. Chin. Soc. Agricult. Eng., vol. 23, no. 1, pp. 155-157, Jan. 2007.

4. L. Chen and L. Y. Wang, ''Research on application of probability neural network in maize leaf disease identification,'' J. Agricult. Mech. Res., vol. 33, no. 6, pp. 145–148, Jun. 2011.

5. H. Chen, W. Li, and X. Yang, ''A whale optimization algorithm with chaos mechanism

based on quasi-opposition for global optimization problems,'' Expert Syst. Appl., vol. 158, Nov. 2020, Art. no. 113612.

6. Z. Lv, L. Qiao, A. K. Singh, and Q. Wang, ''Fine-grained visual computing based on deep learning,'' ACM Trans. Multimedia Comput., Commun., Appl., vol. 17, no. 1, pp. 1–19, Apr. 2021.

7. J. G. A. Barbedo, ''Plant disease identification from individual lesions and spots using deep learning,'' Biosyst. Eng., vol. 180, pp. 96–107, Apr. 2019.

8. R. Guan, H. Zhang, Y. Liang, F. Giunchiglia, L. Huang, and X. Feng, ''Deep feature-based text clustering and its explanation,'' IEEE Trans. Knowl. Data Eng., early access, Oct. 6, 2020, doi: 10.1109/TKDE.2020.3028943.

9. E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, ''A comparative study of fine-tuning deep learning models for plant disease identification,'' Comput. Electron. Agric., vol. 161, pp. 272–279, Jun. 2019.

10. G. Hu, X. Yang, Y. Zhang, and M. Wan, ''Identification of tea leaf dis- eases by using an improved deep convolutional neural network,'' Sustain. Comput., Informat. Syst., vol. 24, Dec. 2019, Art. no. 100353.

11. J. Zhang, L. He, M. Karkee, Q. Zhang, X. Zhang, and Z. Gao, ''Branch detection for apple trees trained in fruiting wall archi- tecture using depth features and regions-convolutional neural net- work (R-CNN),'' Comput. Electron. Agricult., vol. 155, pp. 386–393, Dec. 2018.

12. B. Liu, Y. Zhang, D. J. He, and Y. Li, ''Identification of apple leaf diseases based on deep convolutional neural networks,'' Symmetry, vol. 10, no. 1, p. 16, Jan. 2018.

13. Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, ''Identification of rice diseases using deep convolutional neural networks,'' Neurocomputing, vol. 267, pp. 378–384, Dec. 2017.

14. S. S. Chouhan, U. P. Singh, and S. Jain, ''Web facilitated anthracnose disease segmentation from the leaf of mango tree using radial basis func- tion (RBF) neural network,'' Wireless Pers. Commun., vol. 113, no. 2, pp. 1279–1296, Jul. 2020.

15. J. Ma, K. Du, F. Zheng, L. Zhang, Z. Gong, and Z. Sun, ''A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network,'' Comput. Electron. Agricult., vol. 154, pp. 18–24, Nov. 2018.

16. F. Saeed, M. A. Khan, M. Sharif, M. Mittal, L. M. Goyal, and S. Roy, ''Deep neural network features fusion and selection based on PLS regres- sion with an application for crops diseases classification,'' Appl. Soft Comput., vol. 103, May 2021, Art. no. 107164.