

Automated Quality Assessment of Crops Using CNN - Keras

Meghashree, Alwyn Edison Mendonca, Ashika S Shetty

Department of Computer Science, Srinivas Institute of Technology Valachil, Mangaluru, Karnataka, India

ABSTRACT

Article Info

Volume 7, Issue 3

Page Number: 641-645

Publication Issue :

May-June-2021

Article History

Accepted : 20 June 2021

Published : 30 June 2021

Plant disease is an on-going challenge for the farmers and it has been one of the major threats to the income and the food security. This project is used to classify plant leaf into diseased and healthy leaf, to improve the quality and quantity of agricultural production in the country. The innovative technology that helps in improve the quality and quantity in the agricultural field is the smart farming system. It represented the modern method that provides cost-effective disease detection and deep learning with convolutional neural networks (CNNs) has achieved large successfulness in the categorisation of different plant leaf diseases. CNN reads a really very larger picture in a simple way. CNN nearly utilised to examine visual imagery and are frequently working behind the scenes in image classification. To extract the general features and then classify them under multiple based upon the features detected. This project will help the farmers financially in increasing the production of the crop yield as well as the overall agricultural production. The paper reviews the expected methods of plant leaf disease detection system that facilitates the advancement in agriculture. It includes various phases such as image preprocessing, image classification, feature extraction and detecting healthy or diseased.

Keywords : Image Preprocessing, Classification, Convolutional Neural Network, Keras, Deep Learning.

I. INTRODUCTION

Agriculture is an important source in the economic development of India. Approximately 70% of Indian economy relies on farming. Thus, harm to the crops would lead to heavy loss in productivity and would ultimately infect the economy. Leaves being the most sensitive part of plants show disease symptoms at the earliest. The crops need to be monitored against diseases from the very first stage of their life-cycle to the time they are ready to be harvested. At first, the

method used to proctor the plants from diseases was the conventional open eye observance that is a time-consuming method which requires experts to manually monitor the crop fields to develop automatic and semi-automatic plant disease detection structures. These system of rules have so far resulted to be fast, low-cost and more accurate than the conventional method of automatic observance by farmers. Thus, this invokes researchers to make use of more intelligent technological structures for plant disease detection which do not require human

intervention. In this project we are going to build a convolution neural network to solve the problem of identifying diseased and healthy leaf. Keras would be used for model building and all the code would be implemented on Jupyter notebook.

CNN(convolutional neural network)

The name “convolutional neural network” indicates that the network employs a mathematical (convolution) operation. Convolution is a specialised kind of linear activity. Convolutional networks are only system networks that usage convolution In place of general-purpose matrix multiplication In at least one of their layers. A convolutional network consists of an input and AN output part, as well as multiple hidden layers. The hidden layers of a CNN typically consist of a series of convolutional layers that convolve with a multiplication or other dot product. The stimulation part is normally a RELU layer, and is later followed aside additional convolutions such As pooling layers, to the full connected layers and normalisation layers, relate to as hidden layers because their input and output are masked aside the stimulation function and lastconvolution.The terminal convolutional In activity, oftentimes relate back propagation In state to more accurately weight the last product.

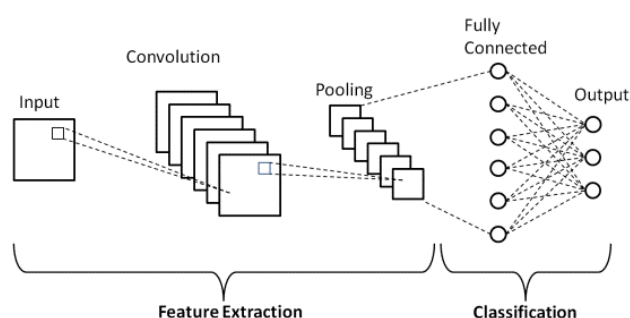


Fig 1. Basic CNN architecture

II. LITERATURE REVIEW

A web based tool was proposed by M.Bhange who has developed to identify fruit diseases by uploading fruit image to the system . Features extraction uses the following parameters such as colour, morphology and color coherence vector. Cluster has been done using the k means rule. SVM is utilized for arrangement of infective or non-infected. This work achieved an accuracy of 82% to identify pomegranate disease[1].

Chit Su Hlaing, SaiMaungMaungZaw presented Plant Diseases Recognition for Smart Farming Using bModel based on the Statistical Features. It has shown the advantages of GP distribution model for SIFT descriptor and successfully applied in plant disease classification. Moreover, it projected feature achieves A great trade-off between execution and categorisation accuracy. Although it's proposed characteristic can be with success model the SIFT feature and applied in recognizing the plant diseases, it is needed to put effort to modify our proposed feature aside considering and co-operation with other image process methods[2].

G. Saradhambal, proposed an approach to produce a system for automatic plant disease detection.To predict the infected area of the leaves a research was carried out by applying k-means algorithm and the Otsu's classifier. Both the shape and texture features were extracted in the proposed work. In this work few shape oriented features were extracted which includes area, colour axis length, solidity and perimeter also texture oriented features were comparatively different, correlation, energy and mean [3]. And lastly, classification in this research was done using a neural network based classifier.

H. V. Singh has worked on automating the detection and classification of plant diseases by implementing genetic algorithm as the image segmentation technique. For the training and testing sets a small leaf images are taken namely banana, beans, lemon

and rose. Colour co-occurrence method has been used for feature extraction considering both color and texture features. The Least Distance Criterion using k-mean clustering and the SVM classifier have been used to classify the diseases showing an accuracy of 86.54% and 95.71% respectively [4].

Combining the Minimum Distance Criterion classifier with the genetic algorithm increases the accuracy to 93.63%. J.D. Pujari has taken a number of crop types namely, fruit crops, vegetable crops, cereal crops and commercial crops to detect fungal diseases on plant leaves. Different techniques have been adopted for each type of crop [5].

- k-means clustering is also used as the segmentation method used for the fruit crops, texture features were classified using ANN and nearest neighbor algorithm achieving an total average aggregate of 90.723%.

- chan-vase method is used for segmentating vegetable crops, local binary patterns for texture feature extraction and k-nearest neighbor algorithm for 87.825%.

- The commercial crops have been segmented using grab- cut algorithm. PNN as classifiers with an overall average accuracy of 84.825%.

- Using k-means Clustering cereal crops has been segmented and canny edge detector. Colour, form, texture, colour texture and random change in features have been selected. SVM and closest neighbor classifiers utilised in getting an overall average accuracy of 83.72%. Plant Disease segmenting image sample was presented by XiaoyanGuo, MingZhang, Yongqiang Dai supported on pulse united Neural Network with shuffle frog leaping algorithmic method. A novel image segmentation sample SFLA-PCNN for plant diseases based on hybrid frog- hopping algorithm is presented. Using the weighted sum of cross entropy and image segmentation compactness as the fitness function of SFLA, the picture of potato late blight disease is taken as a trial segmentation picture to find the optimal

configuration parameters of PCNN neural. Image segmentation is a key step in feature extraction and disease recognition of plant diseases images[6].

III. PROBLEM STATEMENT

One of the important sectors of Indian Economy is Agriculture. Employment to almost 50% of the countries workforce is provided by Indian agriculture sector. India is known to be the world's largest producer of pulses, rice, wheat, spices and spice products. Farmer's economic growth depends on the quality of the products that they produce, which relies on the plant's growth and the yield they get. Therefore, in field of agriculture, detection of disease in plants plays an instrumental role. Plants are highly prone to diseases that affect the growth of the plant which in turn affects the ecology of the farmer. In ordering to observe a plant disease at very initial stage, usage of automated disease detection method is beneficial. The symptoms of plant diseases are very thin in different parts of a plant such as leaves, etc. Manual spotting of plant disease using leaf images is a tiresome job. Hence, it is required to develop computational methods which will make the process of disease detection and classification using leaf images automatic.

IV. METHODOLOGY

A. Proposed Model

The CNN model consists of various steps involved during the training process. The model's first layer is a Conv2D layer. Since, it is the first layer of the model, input shape of the images that are going to be supplied to the model is being mentioned. Following layer is a batch normalisation layer. And then one stimulation layer corresponding to the conv2d layer. Further here there is some other set of conv2d, batch normalisation and stimulation layer with various number of kernels in the conv2d layer. Later on a

Max Pooling layer is present and so a drop out layer is exist.

The same set of layers is again repeated with different number of kernel's and drop out rate. Also, the convolution layers point with this set. Following are the full connected layer.

The Sequential model API is used to create a model in layer- by-layer. In this 'add()' function is used to add layers to our model. This model must know what is the input shape it should expect. So therefore, the first layer in a Sequential model necessarily used to receive information just about its input shape.

Dropout layer is used to avoid over fitting on the dataset and dense is the output layer contains only one neuron which decide to which category image belongs.

Activation layer are used to apply the activation function to the output of that layer. The use of the activation function is to initiate non-linearity into the output of a neuron. Activation function such as Relu and sigmoid are used in the model.

We use batch normalization method to normalize the inputs of each layer and achieve faster convergence and also improve the speed performance and steadiness of ANN.

Flatten Layer is used to flatten the dimensions of the image obtained after convolving it and the epochs tells us the number of times model will be trained in forward and backward pass.

In Conv2D layer is used for blurring, sharpening, embossing, edge detection. This is realised by doing a convolution between a kernel and an image appearance. And then comes the fully connected layers which consists only 2 layers. The very Initial one is the global average pooling layer. This layer is used to minimize over fitting by reducing the overall performance of factor in the model. Second layer and the final layer is the Dense layer with sigmoid activation. Global Average Pooling 2D layer is

utilised to minimise over fitting by decreasing the total number of parameters in the model. GAP layers perform a more extreme type of dimensionality reduction, where a tensor with dimensions $h \times w \times d$ is reduced in size to have dimensions $1 \times 1 \times d$. This layer will reduce each $h \times w$ attribute represents to a single performance by simply taking the average of all hw values.

When the model is compiled, the optimizer algorithm, loss function and the list of metrics of parameters are to be assumed. The model trained for 5 epochs. The table below is showing the end result of training per epoch I.e train accuracy, test accuracy, train loss, test loss and epochs.

TABLE I. Training Result Per Epoch

```
Epoch 1/5
50/50 [=====] - ETA: 0s - loss: 2.3263 - accuracy: 0.4531WARNING:tensorflow:Can save best model only w
1th val_acc available, skipping.
WARNING:tensorflow:Early stopping conditioned on metric `val_acc` which is not available. Available metrics are: loss,accuracy,
val_loss,val_accuracy
50/50 [=====] - 391s 8s/step - loss: 2.3263 - accuracy: 0.4531 - val_loss: 1.4887 - val_accuracy: 0.69
25
Epoch 2/5
50/50 [=====] - ETA: 0s - loss: 1.1476 - accuracy: 0.7563WARNING:tensorflow:Can save best model only w
1th val_acc available, skipping.
WARNING:tensorflow:Early stopping conditioned on metric `val_acc` which is not available. Available metrics are: loss,accuracy,
val_loss,val_accuracy
50/50 [=====] - 368s 7s/step - loss: 1.1476 - accuracy: 0.7563 - val_loss: 0.9369 - val_accuracy: 0.80
00
Epoch 3/5
50/50 [=====] - ETA: 0s - loss: 0.7320 - accuracy: 0.8369WARNING:tensorflow:Can save best model only w
1th val_acc available, skipping.
WARNING:tensorflow:Early stopping conditioned on metric `val_acc` which is not available. Available metrics are: loss,accuracy,
val_loss,val_accuracy
50/50 [=====] - 368s 7s/step - loss: 0.7320 - accuracy: 0.8369 - val_loss: 0.6286 - val_accuracy: 0.85
37
Epoch 4/5
50/50 [=====] - ETA: 0s - loss: 0.5205 - accuracy: 0.8838WARNING:tensorflow:Can save best model only w
1th val_acc available, skipping.
WARNING:tensorflow:Early stopping conditioned on metric `val_acc` which is not available. Available metrics are: loss,accuracy,
val_loss,val_accuracy
50/50 [=====] - 349s 7s/step - loss: 0.5205 - accuracy: 0.8838 - val_loss: 0.4368 - val_accuracy: 0.91
75
Epoch 5/5
50/50 [=====] - ETA: 0s - loss: 0.4241 - accuracy: 0.9044WARNING:tensorflow:Can save best model only w
1th val_acc available, skipping.
WARNING:tensorflow:Early stopping conditioned on metric `val_acc` which is not available. Available metrics are: loss,accuracy,
val_loss,val_accuracy
50/50 [=====] - 313s 6s/step - loss: 0.4241 - accuracy: 0.9044 - val_loss: 0.3882 - val_accuracy: 0.92
25
```

B. Dataset

We use Plant Village dataset for our proposed work to train and test the model. Plant Village dataset contains more than 50,000 images of classified plant diseases and healthy images. Data set consists of two sets of folder with images such as train and validation. Dataset contains two categories: Diseased and Healthy as shown in fig 2. There are 39 different classes where each class consists of at least 500 images. Since the images were chosen at random for each collection, the accuracy varies depending on which image was chosen, and we report the average accuracy and loss.



Fig. 2. Example Images from dataset of train
(a)healthy leaf (b)disease leaf of Tomato

V. RESULT

The image classification of leaf as diseased and healthy is successfully obtained. By using just small amount of dataset 92% accuracy has been obtained. If more images are used then accuracy will also increase. It is seen that image classification using CNN-Keras is very efficient and easy to implement.

VI. CONCLUSION

In this paper we build deep convolutional neural network for image classification. Despite of using only the subset of the images an accuracy of 92% was obtained. If more number of images were added to the dataset then the accuracy would have been even better. It is focused how image from given dataset in field and past data set used predict the pattern of plant diseases using CNN model. This brings some of the following insights about plant leaf disease prediction. As maximum types of plant leaves will be covered under this system, it helps the farmer in decision making of which crop to cultivate. Also, this sample takes into consideration the past production of data which will help the farmer get insight into the demand and the cost of various plants in market and improve the quality and quantity of agricultural production.

VII. FUTURE SCOPE

In this project the input is images. Further we can use sensors and some networks and convert this into live project. The farmers can directly use the sensors and

see whether the plant leaf is diseased or healthy within few seconds instead of taking pictures and loading it into the system.

VIII. REFERENCES

- [1]. Bhangе, M., Hingoliwala, H.A., 'Smart Farming: Pomegranate Disease Detection Using Image Processing', Second International Symposium on Computer Vision and the Internet, Volume58, 2015 .
- [2]. Chit Su Hlaing, SaiMaungMaungZaw, "Plant Diseases Recognition for Smart Farming based Statistical Features".
- [3]. Saradhambal, G., Dhivya, R., Latha, S., Rajesh, R, 'Plant Disease Detection and its Solution using Image Classification', International Journal of Pure and Applied Mathematics, Volume 119, Issue 14, 2018
- [4]. Singh, V., Misra, A.K., 'Detection of Plant Leaf Diseases Using Image Segmentation and Soft Computing Techniques', Information Processing in Agriculture, Volume 8, 2016 .
- [5]. Pujari, J.D, Yakkundimath, R., Byadgi, A.S, 'Image Processing Based Detection of Fungal Diseases In Plants', International Conference on Information and Communication Technologies, Volume 46, 2015
- [6]. XiaoyanGuo, MingZhang, Yongqiang Dai, "image of plant disease segmentation model based on pulse coupled neural Network with shuffle frog leap algorithm"

Cite this article as :

Meghashree, Alwyn Edison Mendonca, Ashika S Shetty, "Automated Quality Assessment of Crops Using CNN - Keras", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7 Issue 3, pp. 641-645, May-June 2021. Available at
doi : <https://doi.org/10.32628/CSEIT2173186>
Journal URL : <https://ijsrcseit.com/CSEIT2173186>