Plant Disease Identification using CNN

Suhas Kadu, Milind Kulkarni, Khushi Bhartiya, Vidhi Mehta, Siddhesh Memane,

Abstract —Plant diseases result in significant economic losses, as well as a reduction in agricultural product quality and quantity. In addition, due to a lack of infrastructure, their identification becomes a difficult task. If sufficient care is not taken in this region, it can have major consequences for plants, affecting product quality, quantity, and productivity. Smartphone-based disease diagnosis is now achievable thanks to a combination of rising global smartphone usage and recent breakthroughs in computer vision enabled by deep learning. A deep convolutional neural network model has been trained using a public dataset of almost 87k rgb images of healthy and diseased crop leaves, categorized into 38 classes containing 28 different diseases. The trained model achieves an accuracy of 95.84 on the test dataset. The model is converted to a TFLite model. Using the model, a flutter app is developed. Through this app, an image of a leaf of a plant can be uploaded or a photo of a leaf of a plant can be taken via camera, then the disease is identified.

Keywords — — convolutional neural network, tensorflow, image data generator, tflite, overfitting.

I. INTRODUCTION

Diseases are changes in a plant's natural state that affect or stop its vital operations like photosynthesis, transpiration, pollination, fertilization, germination, and so on. Pathogens, such as fungus, bacteria, and viruses, as well as adverse environmental circumstances, cause these diseases. As a result, detecting plant disease at an early stage is critical. Variations in symptoms indicated by diseased plants may lead to an incorrect diagnosis, since untrained gardeners and hobbyists may have more difficulty determining it than a skilled plant pathologist. An automatic method that can help amateur gardeners identify plant diseases based on the plant's appearance and visual symptoms could be very useful.
There are some natural methods to protect plants but they are not used on a large scale. Identifying the affected plants is very important to treat and protect the plants.

Farmers experience great difficulties in switching from one disease control policy to another. The naked eye observation of experts is the traditional approach adopted in practice for the detection and identification of plant diseases.But it is quite time-consuming and not always accurate. As a result, finding a quick, low-cost, and accurate way to automatically diagnose diseases from symptoms on plant leaves is critical.

With advancements in AI, its role has increased in the field of agriculture. AI is used for a lot of applications, one of them which involves the diagnosis of affected plants. This diagnosis can be done using digital image processing but they are not very accurate. Deep learning has many algorithms specifically for image processing using Artificial Neural Networks. Artificial Neural Networks perform well but Convolutional Neural Networks are widely used for these image processing tasks. ConvoNets have the ability to develop an internal representation of a two-dimensional image and to learn position and scale variants in images.

In this paper, a simple approach has been discussed in which a public dataset of plants is used. For training and building the model, GPU-enabled kaggle and google colab is used. A deep convolutional neural network is used to build a model with 94.64% accuracy. This model is converted into a TFLite model using tflite converter. Then an app is developed using this TFLite model and flutter framework. Using this app an image of a leaf of a plant can be uploaded or a photo of a leaf of a plant can be taken via camera, then the disease is identified and remedy is also suggested. This app will work on both android and ios, and identification of disease is very easy and fast.

II. LITERATURE REVIEW

Prakanshu[1] discussed a replacement approach of naked eye detection using image processing. By the utilization of deep convolutional networks, a disease recognition model, supported leaf image classification was developed. The model is able to distinguish between diseased leaves and healthy ones or from the environment by using CNN. All essential steps required for implementing the disease recognition model have been fully described throughout the paper, from collecting the pictures used for training and validation to image pre-processing and augmentation and eventually the procedure of coaching the deep

CNN and fine-tuning. After fine-tuning the parameters of the network, an overall accuracy of 88% was achieved. Furthermore, the trained model was tested on each class individually. It was mentioned that the system can be integrated in web applications or desktop applications to automate this process and can also be installed on drones in the future.

Bin Lu[2] proposed an accurate identifying approach for apple leaf diseases based on deep convolutional neural networks in contrast to earlier research that used complex image preprocessing and could not guarantee high recognition rates for apple leaf diseases. Here sufficient pathological images were generated and a novel architecture of a deep convolutional neural network based on AlexNet was designed to detect apple leaf diseases. A dataset of 13689 images was generated using direction disturbance, light disturbance and PCA jittering. and the proposed deep convolutional neural network model was trained to identify the four common apple leaf diseases. To optimize network parameters NAG algorithm was used to accurately identify the diseases. Whole model was implemented on a GPU platform, Caffe framework. The experimental results showed that the proposed approach achieved an overall accuracy of 97.62%. Other deep neural network models, such as Faster RCNN, YOLO, and SSD have been planned to be applied.

Vijay Singh[3] presented an algorithm for image segmentation technique that is used for automatic detection and classification of plant leaf diseases. In this paper a combination of image processing, clustering, Genetic algorithm and SVM is used. Images are taken through digital cameras, and then image processing techniques like filtering, smoothening, clipping are applied. Then masking is done on green color pixel components, if the pixel's value is less than threshold then zero is assigned to that pixel, after that masked cells are removed. Then the remaining segments are classified using Genetic algorithm and clustering. For classification, extraction and comparison of the co-occurrence features for the leaves with the corresponding feature values are stored in the feature dataset. First, the Minimum Distance Criterion and then SVM classifier are used to do the classification. The paper also covers surveys on different disease classification techniques that can be used for plant leaf disease detection. With very less computational efforts the optimum results were obtained, which showed the efficiency of the proposed algorithm in recognition and classification of the leaf diseases. It was mentioned that To improve recognition rate in classification process Artificial Neural Network, Bayes classifier, Fuzzy Logic and hybrid algorithms can also be used.

Hassan[4] implemented deep convolutional-neural-network (CNN) models to identify and diagnose diseases in plants from their leaves. In this paper, standard convolution were replaced with depth=separable convolution, which reduces the parameter number and computation cost. The implemented models were trained with an open dataset of 53,407 images consisting of 14 different plant species, and 38 different categorical disease classes and healthy plant leaves. To evaluate the performance of the models, different parameters such as batch size, dropout, and different numbers of epochs were incorporated. The implemented models achieved a disease-classification accuracy rates of 98.42%, 99.11%, 97.02%, and 99.56% using InceptionV3, InceptionResNetV2, MobileNetV2, and EfficientNetB0, respectively, which were greater than that of traditional handcrafted-feature-based approaches. In comparison with other deep-learning models, the implemented model achieved better performance in terms of accuracy and it required less training time. Here, only images of only leaves facing upward were considered. In the future, performance on noisy images would be evaluated and improved, and the dataset would be expanded with more varieties of disease images.

In this paper [5], digital Image processing is combined with machine learning algorithms to build a machine learning model. Markov Random Field, KNN, ANN and SVM are the algorithms used. Source of dataset and accuracy are not mentioned. Limitation is no. of models can be reduced and CNN can be used. In this paper [6], the PlantVillage dataset is used. AlexNet and GoogleNet models are used to build an efficient model. AlexNet and GoogleNet models give accuracy 85.53% and 99.34% respectively. Although less diverse images were considered and images of leaves facing upwards are considered. In this paper [7] the author automated the method of classification diseases on potato plants from a publicly available plant image database called ‘Plant Village’. An author uses segmentation approach and utilization of support vector machines to demonstrate disease classification over 300 images with an accuracy of 95%.

III. Methodology/Experimental

A. Information about modules

- **Tensor flow:** Tensor Flow is an open-source library primarily for deep learning applications developed by Google. Traditional machine learning is also supported. Without keeping deep learning in mind, Tensor Flow was originally developed for large numerical computations. However it also proved to be very useful for the growth of deep learning, and therefore Google open-sourced it. This is written in C++, CUDA, and Python.
- **Convolutional Neural Networks:** A convolutional neural network (ConvoNet / CNN) is a deep learning algorithm that can record input images, assign meanings to different aspects / objects (learnable weighting and distortion) in the image, and distinguish them from each other. ConvoNet requires much less pre-processing compared to other classification algorithms. Filters are developed manually in a basic way, but with sufficient training, ConvoNet can learn these filters / properties.
- **TFLite:** TensorFlow Lite is an open source, product-ready, cross-platform deep learning framework that transforms pre-trained models in TensorFlow into a special format that can optimize speed and memory. A special format model can be provided on edge devices such as mobile phones with Android or IOS, or Linux-based embedded devices such as Raspberry Pi and microcontrollers to perform inferences at the edge.
- **ReLu Activation functions** The Relu or Rectified Linear Unit is an activation function which returns 0 if it receives negative value as input. In case of positive value it returns it as it is. It is widely used in deep learning.

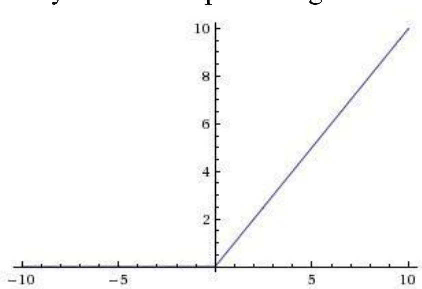

Fig. 1: ReLu function Graph

Mathematically Relu is defined as

$$f(x) = \max(0, x) \qquad (1)$$

where x is the input.

- **Softmax:** Softmax is an activation function which calculates the probabilities and assigns them to classes in the output vector. The class having maximum probability is the result. Mathematically it is represented as,

$$soft\max(z_i) = \frac{\exp(z_i)}{\sum_{j=1}^{n} \exp(z_j)} \qquad (2)$$

where z represents input vector consisting n classes

i represents i-th element of input vector

$\exp(z_i)$ represents exponential function applied to z_i

$\sum_{j=1}^{n} \exp(z_j)$ ensures the sum of probabilities is one

n represents the of classes

B. Flow Chart

Initially necessary modules and dataset are imported. TensorFlow and keras are used for this project. The dataset is splitted into 80:20 for training and validation and the classes of different diseases are extracted. The numbers of samples are increased using ImageDataGenerator; the samples are zoomed and rotated using this. The model is built using CNN. Image size 256x256 and batch size is 256. Images of one leaf are plotted after each layer of CNN to identify the change in the image and features. ReLu is used as an activation function in hidden layers and Softmax is used at the output layer. Adam optimizer is used. The model is trained for 10 epochs and graphs are plotted for accuracy and loss. Testing and prediction is done and the model is converted into a TFLite model. Lastly the model is integrated with flutter to make a working app.

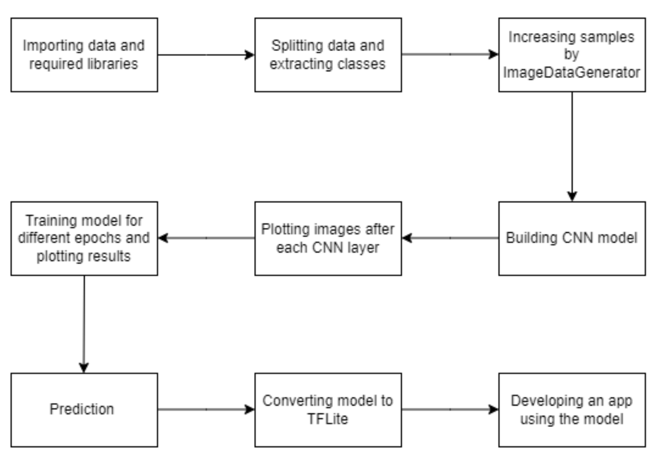

Fig 2: Flowchart of project

IV. Results and Discussions

A. Model summary and results

The table below shows summary of models and its parameters

Table no. 1: Model Summary

Input Shape	256x256x3
Batch size	256
Trainable parameters	11,930,502
Conv2Dlayer	3
MaxPooling layer	2
Dense layer	3
Flatten layer	1
Dropout layer	1
Epochs	20
Testing Accuracy	94.64%

B. Results of Training, Testing and Prediction

- **Training:**

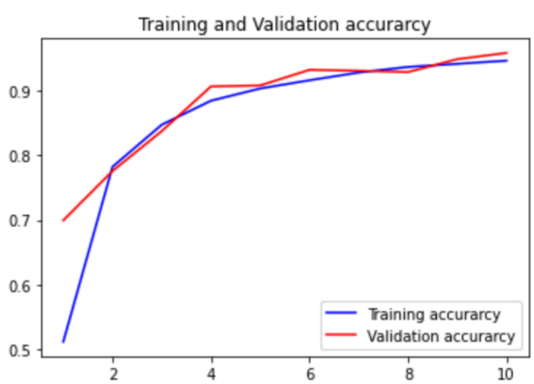

Fig. 3: Training & Validation accuracy graph

Fig. 4: Training & Validation graph

- **Prediction:**

```
img_path = "../input/new-plant-diseases-dataset/test/test/AppleCedarRust4.JPG"

predict_x=model.predict([prepare(img_path)])
classes_x=np.argmax(predict_x,axis=1)

disease=image.load_img(img_path)
plt.imshow(disease)
print (Classes[int(classes_x)])
```

```
Apple___Cedar_apple_rust
```

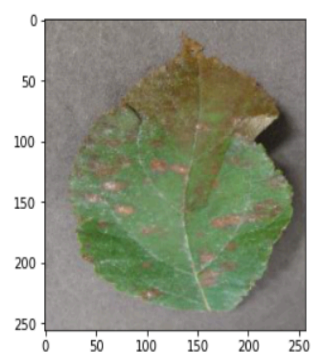

Fig. 5: Prediction of image

In the above figure an image of Apple leaf is passed. The leaf is infected with Apple Cedar Rust disease, which is correctly predicted by the model. The model archives 94.64% on the training set, and 95.84% accuracy on the validation set. Results of Prediction can also be observed.

- **Converting model to TFLite:**

```
import tensorflow as tf

keras_model = tf.keras.models.load_model("./new_my_h5_mode
converter = tf.lite.TFLiteConverter.from_keras_model(keras
tflite_model = converter.convert()

with open('my_model.tflite', 'wb') as f:
  f.write(tflite_model)
```

Fig. 6: Converting model to TFLite

Saving the model in local storage and then loading the model and converting it to TFLite using TFLite Converter.

C. Results of App

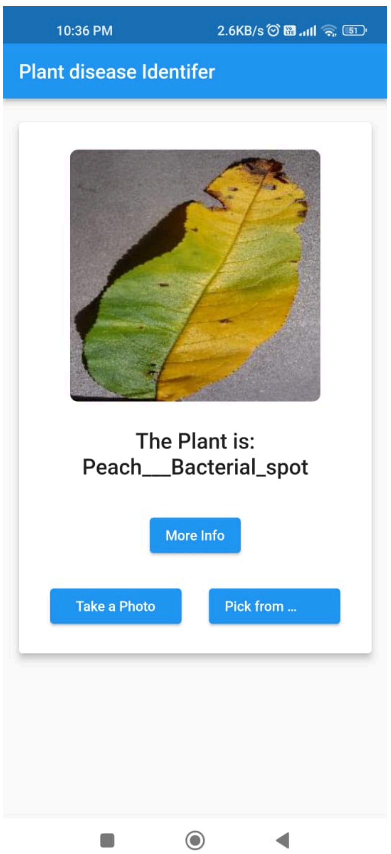

Fig 7.1: Results screen

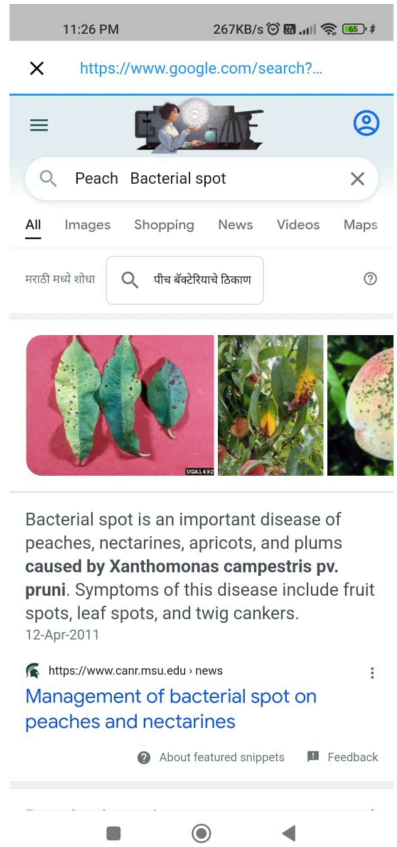

Fig 7.2: Remedies screen

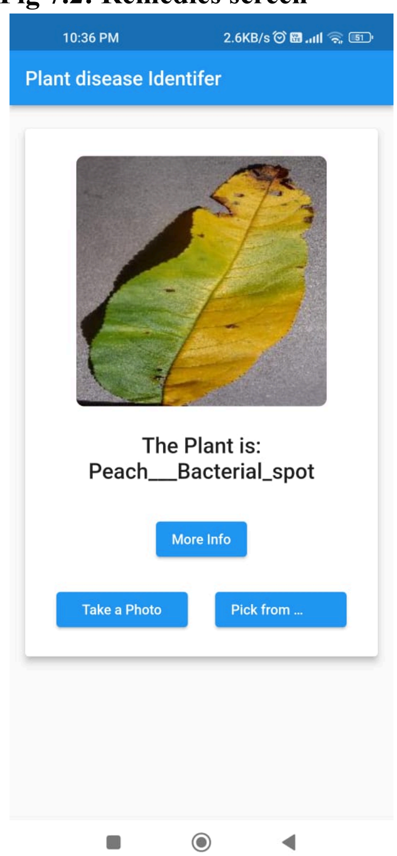

Fig 7.3: Results screen

From above screenshots it can be observed how the model is identifying images. Two options are available for uploading images i.e. take a photo from camera or pick image from gallery. On clicking More Info button information about the disease can be obtained inside the app.

V. Limitations

As every system has some limitations, there are also some limitations in this project. The testing accuracy is less as compared to the training set, because of overfitting, thus the model gives incorrect results in some cases. The TFLite model gives poor performance as compared to the trained model and also it has a very large size.

VI. Future Scope

In this project only sequential keras models are used, a model like GoogleNet, InceptionV3 can be used to increase the accuracy. The TFLite model can be optimized to reduce its size. The model is currently in the local storage; it can be deployed on a cloud service and automated. More no features like instant remedies, demo videos, tips for plant care can be added to the app.The recommendation system of pesticides for the detected disease can be implemented.

VII. Conclusion

In this paper a plant disease identifier model is developed using Deep learning and Convolutional Neural Networks. This model identifies 28 different diseases and gives an accuracy of 94.64% on test images. The model is converted into a TFLite model and integrated with Flutter to make an app. This app can be used for the identification of plant diseases by uploading the image of a leaf of a plant. This app also gives some information on the disease identified and it works on both Android and IOS. Due to Deep learning, CNN, and TFLite, any plant with a single disease can be identified and instant remedies can also be found in the app.

Acknowledgment

We would like to thank our honorable Director Prof. (Dr.) R.M Jalnekar, Vishwakarma Institute Of Technology, Pune and HOD Prof. Sandip Shinde for extending strong moral support and encouragement. We would also like to express our gratitude to our project guide Prof. Milind Kulkarni for his valuable guidance to complete the project successfully.

References

[1] Plant Disease Detection Using Convolutional Neural Network January 2021, International Journal of Advanced Research.

[2] Identification of Apple Leaf Diseases Based on Deep Convolutional Neural Networks Bin Liu 1, Yun Zhang, DongJian He and Yuxiang Li, 29 December 2017.

[3] Vijai Singh, A.K. Misra, Detection of plant leaf diseases using image segmentation and soft computing techniques, Information Processing in Agriculture, Volume 4, Issue 1, 2017.

[4] Hassan, Sk & Maji, Arnab & Jasiński, Michał & Leonowicz, Zbigniew & Jasińska, Elżbieta. (2021). Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach. Electronics. 10. 1388. 10.3390/electronics10121388.

[5] Gavhale, Ms & Gawande, Ujjwala. (2014). An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques. IOSR Journal of Computer Engineering. 16. 10-16. 10.9790/0661-16151016

[6] Mohanty Sharada P., Hughes David P. Salathé Marcel, Deep Learning for Image-Based Plant Disease Detection, Frontiers in Plant Science, Vol. 7, 2016.

[7] Detection of potato diseases using image segmentation and multiclass support vector machine Islam, Monzurul; Dinh, Anh; Wahid, Khan; Bhowmik, Pankaj, 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), 2017-06-15.