Apple Foliar Disease and detection using Machine learning

Babli Pal

*[1](Department of CSE, Global Nature Care Sangathan's Group of Institutions, Jabalpur)

Abstract

Apple (Malus pumila) is commercially the most important temperate fruit and is fourth among the most widely produced fruits in the world after banana, orange and grape. China is the largest apple producing country in the world. **Himachal Pradesh** covering the districts of Shimla, Siramour, Kullu, Mandi, Chamba and Kinnaur considering the vast potential for increasing exports. The diseases like apple scab, apple rust, and brown rot affect apple trees. Diseases significantly affect the apple, so their identification is very important. Correct identification of these diseases is crucial for establishing a good disease control strategy to avoid time and financial losses. In general, machines can greatly reduce the possibility of human error. In particular, computer vision techniques developed through deep learning have paved a way to detect and diagnose these plant diseases on the leaf. In this work, the model DF-Tiny-YOLO was developed to detect and identify various leaf diseases in apple trees. The dataset is from Kaggle 2020 and 2021 and was financially supported by the Cornell Initiative for Digital Agriculture. DF-Tiny-YOLO was proposed for leaf disease classification in apple trees and the results of the efficiency of the model are compared with other state-of-the-art deep learning approaches. The results of the experiments in the validation dataset show that the proposed DF-Tiny-YOLO model achieves the highest values compared to other deep learning models in the original and extended datasets with 98.7% accuracy for Plant Pathology 2020 and 92.6% for Plant Pathology 2021.

Keywords: Computer vision, Deep learning, Kaggle, Foliar disease, YOLO.

Introduction

Leaf diseases of apples refer to the common diseases found on apple leaves, namely scab, rust, and powdery mildew, among others [1]. Apple scab, caused by a fungal pathogen, is one of the most economically important fungal diseases of apple in the world [2,3]. The symptoms of apple scab are clearly visible fungal structures on the surface of the leaf. Rust disease also causes severe losses when environmental conditions are favorable for disease growth. For example, in a plant affected by Rust [4], small yellow spots appear on the leaf surface.

Farmers spend a lot of money on disease control and inadequate technical support, but the results always lead to poor disease control. Foliar diseases spread rapidly and can destroy large portions of the yield in a very short time. In some cases, these diseases destroy the entire crop if the disease is not controlled quickly and accurately. Foliar diseases are a challenge to crop production in most countries. They reduce crop yields, fruit quality, and nutritional value, resulting in lower returns for the farmers [4].

Hundreds of apple diseases disseminate in fruits, leaves, branches, roots, and other areas, but often initially appear in leaves, which are easily observed, collected, and managed. Therefore, they are an important reference for disease identification and effective automated detection of diseases is essential. However, judging differences among diseases is difficult due to the complexity of blade veins [5], resulting in unsatisfactory outcomes of experimental detection methods [6].

Machine learning models learn, recognize patterns, and make decisions with minimal human intervention. Ideally, machines increase accuracy and efficiency and eliminate the possibility of human error [5]. The use of AI in agriculture helps farmers gain insights into their crops and use the data to increase their overall production. Various computer vision techniques can be applied to gain the desired insights [6,7].

Recent advances in computer vision enabled by deep leaning have paved the way for more accurate disease diagnosis. Using large public datasets of diseased and healthy plants and leaves, a CNN can be trained to identify various leaf diseases [8,2]. With the increasing availability of smartphones, the approach of training deep-learning models on a large scale has emerged as a clear way to diagnose crops on a large scale [5]. Every year, farmers worldwide are affected by foliar diseases. Our research could contribute greatly to the automation of disease detection worldwide and potentially help millions of people [1].

The article deals with various aspects of diseases and classifies them based on the characteristics of the condition of the leaves. It is important to identify the cause at the root, which is beneficial and time-saving for both the agricultural sector as well as the farmers.

Literature Review

In recent years, ML and DL are widely used to detect plant diseases, which helps farmers identify the right foliar disease and apply the appropriate treatments [9]. Digital images are widely used in computer vision to identify the diseases for further classification based on their symptoms [10,11,12,13]. However, it is challenging to accurately identify disease from leaves due in part to the resolution, background light, and shadows of the leaves, among other [14]. Machine learning (ML) and deep learning (DL) approaches are well suited for processing image data, especially in agriculture, and can be used to detect and classify plant diseases from the collected images [15].

Agarwal et al. proposed a model consisting of 3 maximal pooling layers followed by two densely connected layers. After testing with different numbers of convolutional layers from 2 to 6, it was found that 3 layers provide the best accuracy [2]. The proposed model achieves a very impressive accuracy, i.e., 96%. The database used for the developed framework consists of nearly 50,000 images of 171 diseases, including 21 plant species. The original samples were divided into smaller images containing individual lesions or localized symptom regions. This was done to increase the size of the dataset and to test how the CNN would perform with more localized information.

Instead of taking pictures in the natural condition, Zhong and Zhao took pictures with a solid background [5]. Images of all symptoms were resized to 128 by 128. The dataset was split 8:2 for the training and test datasets by randomly selecting images from the dataset. After deduplication, the dataset contained 2,462 images, with 85% of the images used for training and 15% for validation. The accuracy of this method for the test dataset was 93.71%.

Militante et al. proposed a model, i.e., a combination of a convolutional layer, an activation layer, a pooling layer, and a fully connected layer [6]. The images used in this study were in color and were reduced to 96 by 96 for further processing. An accuracy of 96.5% was achieved with 75 epochs while the model was well trained. A maximum accuracy of 100% was achieved when random images of plant varieties and diseases were tested.

Arsenovic et al. presented a dataset of 79,265 images. Traditional augmentation methods and generative adversarial networks are used for image augmentation. Moreover, a 2-stage NN architecture was proposed for classification and a test accuracy of 93.67% was achieved with the trained model. The DCNN model [16] was presented for detecting the leaf disease of apple tree by combining DenseNet and Xception. The results show that the developed model achieved 98% accuracy.

Yu et al. designed two subnetworks; the first is used for segmentation to identify features, and the second model is used for classification. In the experiments, the proposed model provided an accuracy of 89.4% [17]. The use of AI in agriculture helps farmers gain insights about their crops and use this data to increase their production wisely. Using the proposed methods, Thapa et al. captured 3,650 of high-resolution images of several apple leaf diseases and annotated the dataset with the help of an expert in the field of pathology to confirm the annotations for the images, which were difficult to distinguish based on symptoms [1]. The overall test accuracy achieved by a ResNet50 network pre-trained on ImageNet was 97%.

Raschka et al. proposed a CNN model and achieved 97.62% accuracy in identifying four different types of apple leaf blight, detecting infected parts on the leaf, and classifying between healthy and infected fruits, e.g., apples [7]. However, there are very few studies dealing with apple foliar disease and most of them are limited to a specific type of disease, either biotic or abiotic.

Priyadharshini et al. proposed a deep CNN-based architecture (modified LeNet) for four maize leaf disease classifications, and the trained model achieved 97.89% accuracy [25].

Xue et al proposed the improved Tiny-YOLO mango detection network combining dense connections, with an accuracy rate of 97.02% [26].

Jan and Ahmad [27] developed an apple pest and disease diagnostic system to predict apple scabs and leaf spots. Entropy, energy, inverse difference moment (IDM), mean, standard deviation (SD), and perimeter, etc.,
were extracted from apple leaf images and a multi-layer perceptron (MLP) pattern classifier and 11 apple leaf image features to train the model, which had 99.1% diagnostic accuracy.

Sun et al [28] proposed a lightweight CNN model that can be deployed on mobile devices to detect apple leaf diseases in real time to test five common apple leaf diseases, with a mAP value of 83.12%.

Proposed Methodology

A framework called DF-Tiny-YOLO is presented in this paper. Basically, our work is based on the transfer learning approach, where the first and most important step is to collect the dataset [1]. The second step is to project and clean the database using image processing steps [18] to find outliers and class imbalances. This was followed by an exploratory data analysis of the dataset with all graphs and class distribution of the foliar diseases. Extensions such as rotation, transformation, and flips were applied to increase the diversity/learning capability of the model [19,20]. After pre-processing all the data, the data were fed into the training pipeline with 5-fold cross-validation, properly validating the training data. These iterations are repeated continuously until we find a stable cross-validation value for our training data that matches the test data.

Conclusion

The world around us relies heavily on the agricultural sector to provide food. Early detection of plant diseases is critical to the industry. In this article, DF-Tiny-YOLO model is proposed to identify leaf disease in apple trees. The proposed model is applied to two data sets: Plant Pathlogy 2020 and Plant Pathalogy 2021. The proposed DF-Tiny-YOLO model achieves 98.7% accuracy, which is higher than other TL models (B3, B4, Inception V3, VGG16, ResNet50, and 101). In addition, the performance of the proposed model outperforms the other models for both datasets.

The above results show the efficiency of the proposed model in identifying leaf diseases on apple trees for major and minor classes, i.e., multiclassification.

References

1. Ranjita Thapa, Kai Zhang, Noah Snavely, Serge Belongie, and Awais Khan. The plant pathology challenge 2020 data set to classify foliar disease of apples. *Applications in Plant Sciences*, 8(9):e11390, 2020.
2. Mohit Agarwal, Rohit Kumar Kaliyar, Gaurav Singal, and Suneet Kr Gupta. Fcnn-lda: A faster convolution neural network model for leaf disease identification on apple's leaf dataset. In *The International Conference on Information & Communication Technology and System (ICTS)*, pages 246–251, 2019.
3. Jayme Garcia Arnal Barbedo. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Computers and Electronics in Agriculture*, 153:46–53, 2018.
4. Alieu MOININA, Rachid LAHLALI, and Mohammed BOULIF. Important pests, diseases and weather conditions affecting apple production in morocco: Current state and perspectives. *Revue Marocaine des Sciences Agronomiques et V´et´erinaires*,7(1), 2019.
5. Yong Zhong and Ming Zhao. Research on deep learning in apple leaf disease recognition. *Computers and Electronics in Agriculture*, 168:105146, 2020.
6. Sammy V Militante, Bobby D Gerardo, and Nanette V Dionisio. Plant leaf detection and disease recognition using deep learning. In *IEEE Eurasia Conference on IOT, Communication and Engineering*, pages 579–582, 2019.
7. Sebastian Raschka. Model evaluation, model selection, and algorithm selection in machine learning. *arXiv preprint arXiv:1811.12808*, 2018.
8. Alieu Moinina, Rachid Lahlali, Dustin MacLean, and Mohammed Boulif. Farmers' knowledge, perception and practices in apple pest management and climate change in the fes-meknes region, morocco. *Horticulturae*, 4(4):42, 2018.
9. Anne-Katrin Mahlein. Plant disease detection by imaging sensors–parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease*,100(2):241–251, 2016.
10. Jayme Garcia Arnal Barbedo. An automatic method to detect and measure leaf disease symptoms using digital image processing. *Plant Disease*, 98(12):1709–1716, 2014.
11. Pengbo Dai, Xiaofei Liang, Yajing Wang, Mark L Gleason, Rong Zhang, and Guangyu Sun. High humidity and age-dependent fruit susceptibility promote development of trichothecium black spot on apple. *Plant Disease*, 103(2):259–267, 2019.
12. Thomas W¨ohner and Ofere Francis Emeriewen. Apple blotch disease (marssonina coronaria (ellis & davis) davis)–review and research prospects. *European journal of plant pathology*, 153(3):657–669, 2019.
13. Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. Deep neural networks based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience*, 2016, 2016.
14. Sachin B Jadhav, Vishwanath R Udupi, and Sanjay B Patil. Identification of plant diseases using convolutional neural networks. *International Journal of Information Technology*, pages 1–10, 2020.
15. Jihen Amara, Bassem Bouaziz, and Alsayed Algergawy. A deep learning-based approach for banana leaf diseases classification. *Datenbanksysteme f¨ur Business, Technologie und Web (BTW 2017)-Workshopband*, 2017.
16. Xiaofei Chao, Guoying Sun, Hongke Zhao, Min Li, and Dongjian He. Identification of apple tree leaf diseases based on deep learning models. *Symmetry*, 12(7):1065, 2020.
17. Hee-Jin Yu, Chang-Hwan Son, and Dong Hyuk Lee. Apple leaf disease identification through region-of-interest-aware deep convolutional neural network. *Journal of Imaging Science and Technology*, 64(2):20507–1, 2020.
18. Xinjian Guo, Yilong Yin, Cailing Dong, Gongping Yang, and Guangtong Zhou. On the class imbalance problem. In *The International Conference on Natural Computation*, volume 4, pages 192–201, 2008.
19. Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, and Youngjoon Yoo. Cutmix: Regularization strategy to train strong classifiers with localizable features. In *The IEEE/CVF International Conference on Computer Vision*, pages 6023–6032, 2019.
20. Hongyi Zhang, Moustapha Cisse, Yann N Dauphin, and David Lopez-Paz. mixup: Beyond empirical risk minimization. *arXiv preprint arXiv:1710.09412*, 2017.
21. Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning*, pages 6105–6114, 2019.
22. Hamail Ayaz, Erick Rodr´ıguez-Esparza, Muhammad Ahmad, Diego Oliva, Marco P´erez-Cisneros, and Ram Sarkar. Classification of apple disease based on non-linear deep features. *Applied Sciences*, 11(14):6422, 2021.
23. Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.
24. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
25. Priyadharshini RA, Arivazhagan S, Arun M, Mirnalini A. Maize leaf disease classification using deep convolutional neural networks. Neural Computing and Applications. 2019; 31(12).
26. Xue Y, Huang N, Tu S, Mao L, Yang A, Zhu X, et al. Improved YOLOv2 Identification Method for Immatureb Mangoes. Transactions of the Chinese Society of Agricultural Engineering. 2018; 34(07):173–9.
27. Jan M, Ahmad H. Image Features Based Intelligent Apple Disease Prediction System: Machine Learning Based Apple Disease Prediction System. International Journal of Agricultural and Environmental Information Systems (IJAEIS). 2020; 11(3).
28. Sun H, Xu H, Liu B, He D, He J, Zhang H, et al. MEAN-SSD: A novel real-time detector for apple leaf diseases using improved light-weight convolutional neural networks. Computers and Electronics in Agriculture 2021;189.