

SCN2Vit : Separable CNN-vision transformer with optimized transfer learning and YOLOv7 for bean leaf disease detection

N. Kumar *

Department of Mathematical and Physical Sciences

College of Arts and Sciences

University of Nizwa

P.O. Box 33

PC 616 Birkat Al Mouz

Nizwa

Sultanate of Oman

Mahendra Swain [†]

Peerless Securities Limited

Kolkata 700069

West Bengal

India

Aftab Ahmed [§]

Natural and Medical Sciences Research Center

University of Nizwa

P.O. Box 33

PC 616 Birkat Al Mouz

Nizwa

Sultanate of Oman

Abstract

Ubiquitously promoting digital system emulates the human visual capabilities to get forward every manual step for fully automated system. Agriculture is one of the most necessary and fruitful domains for computer scientists to meet the satisfactory requirement of global and sustainable development. The meaning of sustainable development is pasteurized due to natural resources. The pollutant air, global warming badly impact of plants leaves. This article attempts to present a machine learning based leaf disease modeling framework.

* E-mail: naresh@unizwa.edu.om (Corresponding Author)

[†] E-mail: mr.mahendraswain@gmail.com

[§] E-mail: aftab.ahmad@unizwa.edu.om

Plant diseases have contiguous impact and spread sharply; therefore, it is important to have an early detection of the disease. Considering the importance and badly impacting issues in leaf disease detection, we introduce Separable CNN Vision Transformer (SCN2Vit) model which utilizes a multi-headed vision transformer learnt from mobileNetv2 in transfer learning mode. The learnt feature to enhance the performance is trained from YOLOv7, a faster and real-time most accurate object detector. For optimizing the network Nesterov-Accelerated Adaptive Momentum (NADM) is used and ensured that SCN2Vit-YOLOv7 meets the challenges satisfactorily. The experiments were performed to test the model on publicly available collected from Makerere AI laboratory NaCRR institute.

Subject Classification: 94A08.

Keywords: Leaf disease detection, Transfer learning, YOLOv7, Transformer, MobileNet(v2).

1. Introduction

Growth of human population on the earth is obviously raised the issues of food production. Lacking proper enough food is compromised with quality of food. This is because 90% of food is received from plants. The very common and regular problems for the farmers are noticed while farming is affected from diseases to plants and leaves. Once the disease is detected veraciously, actual control measures can be taken. The main source of food production comes from crops in India is almost 70% which is the main source of income. Agriculture is the keystone of the developing nations. Loizou et al. [1] and Diao, Hazzell, and Thurlow [2]. The output shows that the correctness of methodology can be higher than the standard method, therefore minimizing the influence of illness on agricultural production and being commendatory to sustainable development of agriculture is possible. As per the recent research surveys by India Brand Equity Foundation (IBEF), in India, agriculture is the main subsistence for about 58% in Selvam and Kavitha [3]. In many rural areas it is tough for farmers to approach the experts who can estimate the diseases in plants in Abade, Ferreira, and Vidal [4]. Recently, some technologies based on convolution neural network performed experiments of real-time disease detection of the plants in Gajjar et al. [5] and Kashyap et al. [57]. Most of the verification is performed manually even today, which may not be easy for the disease and its form to be identified. There is a need for automatic disease detection in plants because wrong identification gives the loss in production of agriculture in Massi et al. [6]. The importance of automatic infection detection has been posed by this issue and has forced the development of certain methods or systems that can detect the disease more reliably. The crucial problem of detecting and diagnosing diseases more efficiently depends on obtaining correct knowledge about the symptoms of diseases.

Some farmers do not have accurate knowledge about the diseases, and they cannot identify the diseases which lead to high agriculture production loss. For manual observation, every farmer is considered to have minimum knowledge

about the plant's growth and the diseases of the plants, so it will be easy if they find spots on leaves. In different plants, there are different diseases of different forms, and they have their own remedy to get cured. Manual method of identification of disease is hard task because naked eye is not capable to identify all the disease, however it is deadly error-prone and time-consuming process. The lack of awareness leads them to incorrect recommendations. The root causes of diseases in plants is due to genetic disorder, living biotic agents, microorganisms, and pathogen agents like bacteria, viruses, fungi, protozoa and helminths. These pathogens spread the disease from one infected plant to another healthy plant easily. The change in the weather is also a cause of diseases in plants. All these reasons may lead to the agriculture production in Haung et al. [7], and in Bora, Parasar, and Charhate [8].

1.1 Motivation

Leaf infection is the infiltration of disease-causing organisms such as bacteria, viruses, and fungi into leaf tissues, resulting in leaf and plant deterioration. Goyal, Kumar, and Gupta [9] performed research experiments recognize leaves of various plants which helps to find the intrinsic features of the plants leaves. Considering the fact by Turkoglu, Yanikoglu, and Hanbay [10] had extensive set experiments on various leaves and developed PlantDiseaseNet model. This is also noticeable that about two-thirds of the world's three billion rural people live on small farms, of which approximately 500 million exist as per record in Massi et al. [6]. Also, pests and diseases have been blamed for yield losses of more than 50% in many cases. For the next generation of agriculture, it must include all the possible solutions in any situation by Beikmohammadi and Faez [11]. Beans are the rich source of niacin B3, a coenzyme which plays a very important role in metabolic process. The hormones secreted in metabolic process are responsible for appetite satiation and digestion which indirectly impact on human body fitness. We mainly focus on the experiments of detecting disease in the leaf green bean by Mullins and Arjmandi [12] ensured the importance of bean cure in controlling many serious diseases like blood and breast cancer, melanoma, lymphoma etc. The leaves of beans work as an antiviral for HIV, leukemia and coronaviruses. In the study of Horticulture, oil seeds and other products are considered cash crops. Bean plants are also considered as Horticulture crops which are fundamentally part of the legume family, and they are seeds that grow in pods, making them fruit. Beans are botanically categorized as a legume, which is a type of plant food.

Deep learning has grown in prominence over the previous decade because of its accuracy when taught with large amounts of data utilized in Sachar and Kumar [58]. Traditional machine learning methods, on the other hand, require less time to train and more time to test Chakraborty et al. [13]. The texture and color of the leaves are the most prominent visual properties in Dubey and Jalal [14] used texture and shape feature to classify the leaf disease in apple. Therefore, classifying leaf disease is essential for assessing agricultural products, improving market value, and satisfying quality criteria. It is also

beneficial to identify and take further actions to prevent the spread of illnesses. The two most important qualities of plant disease detection machine-learning algorithms to accomplish are speed and accuracy. It is necessary to create techniques such as automated plant disease detection and categorization using leaf images processing methods. This approach will be beneficial to farmers since it will notify them at the appropriate time before the illness spreads over a broad region.

Due to various high-resolution screens, enormous computational power, large, fixed sets of attachments, like sophisticated cameras, smart phones offer extremely new methods to assist in illness identification. It is commonly predicted that by 2020, there may be 5-6 billion cell phones on the planet. By 2015 ending, 69 percent of global populace will have access to digital broadband indemnity, with mobile wireless service prevalence exceeding 47 percent; a 12-fold growth since 2007. A combination of broad smart phone use, high-definition cameras, and speed CPUs in movable devices has resulted in circumstances in which illness diagnosis due to automatic picture recognition, if scientifically viable, can be given access on a previously unheard-of scale by Mathew and Mahesh [29].

To correctly diagnose plant disease, a plant pathologist must have strong observation skills and be able to detect typical signs by Mullins and Arjmandi [12]. Promoted AI technology had uncovered the path to creating an automated framework capable of producing speedy, more accurate illness detection findings in Sethy et al. [15] and Thai, Le, and Nguyen [16]. Artificial intelligence-based technologies that autonomously detect a wide range of illnesses are being employed nowadays. The main contributions of bean leaf disease detection model can be highlighted as follows-

- We used separable convolution specifically MobileNetV3 in end-to-end fashion. The features learnt from the separable convolution follow transfer learning to get classified by the vision transformer. MobileNetV2-CNN based transfer learning ensured satisfactory results in many computers vision re- search in Kumar and Bansal [17], Alkhawaldeh et al. [18].
- The proposed Separable CNN Vision Transformer (SCN2Vit) adopted multi-head attention layer which targets disease sensitive parts of the leaves.
- The learnt SCN2Vit features are smoothly detected with YOLOv7 in Wang, Bochkovskiy, and Liao [19] which can easily surpass the feature of YOLOv4 in Bochkovskiy, Wang, and Liao [20] in real time detection and accuracy.
- Leaf disease detection for plants using MobileNet(V2) method is delineated to collect leaf images from dataset and to perform satisfactorily while trained YOLOv7.

The main motive of this project is to prevent crops from diseases. In India, Agriculture plays a major role. But there are problems while cultivating, the

main problem is diseases to the leaf/plant which results in loss of whole field. To overcome these problems, we proposed a method which classifies the diseases in leaf. Classification in leaf whether it is affected or not results in early prevention of crops from drowning of fields from diseases.

2. Related Work

Many existing techniques are declared by researchers for plant disease detection. To reduce processing time, different researchers consider different datasets based on resolutions while conducting these experiments. They stated in this re- search, A Review on Machine Learning Classification Strategies for Plant Disease Detection in Li, Zhang and Wang [21], how they conducted a comparative study on five different types of machine learning classification techniques for plant disease detection. Plant illnesses can be recognized by looking at the leaf of the plants, stems, and roots. This work offers a method for detecting and classifying tomato leaf disease in Liu and Wang [22]. They designed a CNN that can automatically extract and classify features. Plant leaf disease research makes extensive use of color information. The filters are tested on three channels based on RGB components in our model. For training the network, the output feature vector of the convolution component was input into the LVQ. The SVM was educated and checked in the classification using the kernel function. The images are segmented to assess the affected portion of the leaf to be compared using the CNN Classifier after being processed using a set of feature extraction mechanisms. The findings are then used to diagnose and identify tomato plant disease, as well as propose a suitable solution to farmers as a preventative measure. Pre-processing and feature extraction for both the input image and the dataset images were performed to improve prediction accuracy in Karthik et al. [23]. The pictures are split to identify the damaged area of the leaf to be contrasted using the CNN Classifier after being processed with a variety of feature extraction methods.

The use of transfer learning for rice disease classification was explored in Krishnamoorthy et al. [24]. We used AlexNet for feature extraction and SVM for classification. The suggested model can accurately diagnose rice illnesses with an accurate rate of 91.37% for a training- testing division of 80 percent - 20%. They initially suggest a Densely Connected Convolutional Networks (DenseNet) based transfer learning technique to identify plant illnesses in Deep Learning based Plant Disease Detection for Smart Agriculture in Nie et al. [25], which wants to strengthen on edge servers with increased computational capacity. To detect strawberry verticillium wilt, a disease detection network focused on Faster R-CNN and multi-task learning was suggested in Nie et al. [25]. The authors presented Rider-CSA- based DBN for effective image processing based on plant disease identification technique in the publication Deep neural network-based Rider-Cuckoo Search Algorithm for plant disease detection in Cristin et al. [26] which is detailed in this part. To accomplish disease discovery in plants, the suggested model goes through four steps.

The results of the experiments demonstrate that our approach is more successful than the original Faster RCNN at detecting and recognizing tomato illnesses. However, because the dataset comes from a laboratory, only a single leaf disease can be recognized in the picture in Bi et al. [27]. Two CNN architecture models, Inception v3 and MobileNets, were tested to see how they compared in terms of performance, training speed, and model size, among other things. To supplement restricted databases on Indian flora and illnesses, the system employs Generative Adversarial Networks (GANs) in Zhang, Song, and Zhang [28]. An Inception v3 model was pre-trained on the ImageNet database for the developed framework, and the outdoor unit was retrenched to deal with the dataset of plant illnesses. All the layers work together to extract the most relevant characteristics from the dataset's pictures. Several objectors developed in Mathew and Mahesh [29], YOLO-LITE in Huang, Pedoeem, and Chen [30], YOLOv3 for apple leaf in Tian et al. [31], Improved TinyUOLOv3 Kumar et al. [32], Improved YOLOV3-CNN, and DDMA-YOLO in Bao et al. [33] ensured that detection plays a very important role in identification of objects. Mathew and Mahesh [29] used Yolov5 for detecting leaf disease in ball pepper plant. Recently developed fastest real-time object detector YOLOv7 in Wang, Bochkovski, and Liao [19] beats the state of arts in multimedia domain. YOLOv7 introduced different trainable bag-of-freebies (BoF) techniques which enable detector performance without braining any computational overhead.

3. Proposed Methodology

The work presented in this paper provides the proposed system of detecting the leaf disease classification. The proposed model is depicted in Figure 1. The initially raw features are learned by separable convolution MobileNetv2 with transfer learning. The second module includes multi-headed vision transformer. The weight in convolution layers is updated according to Nesterov-accelerated adaptive momentum (RowReuse) in Zhao et al. [34] as represented in Eqn. 1. To compute the moments of varying information, Adam utilizes exponentially moving averages which enables to improve the performance of disease detection in the leaf image

$$w_{t+1} = w_t - \frac{\alpha}{\sqrt{v}+c} \left(\beta_t \times m_t + \frac{(1-\beta_1) \frac{\partial E(w_t - m_i)}{\partial x}}{1-\beta^t} \right) \quad (1)$$

Where $E(w_t, M_i)$ defines entropy of M input images of class i with weight w_t . The value of β_1 is chosen 0.999 and α is set at 0.002. v is infinity norm constraint appreciated by fixed constant c . As MobileNet (v1) is a good feature extractor for image detection and segmentation, it gives a turbo speed of 35%.

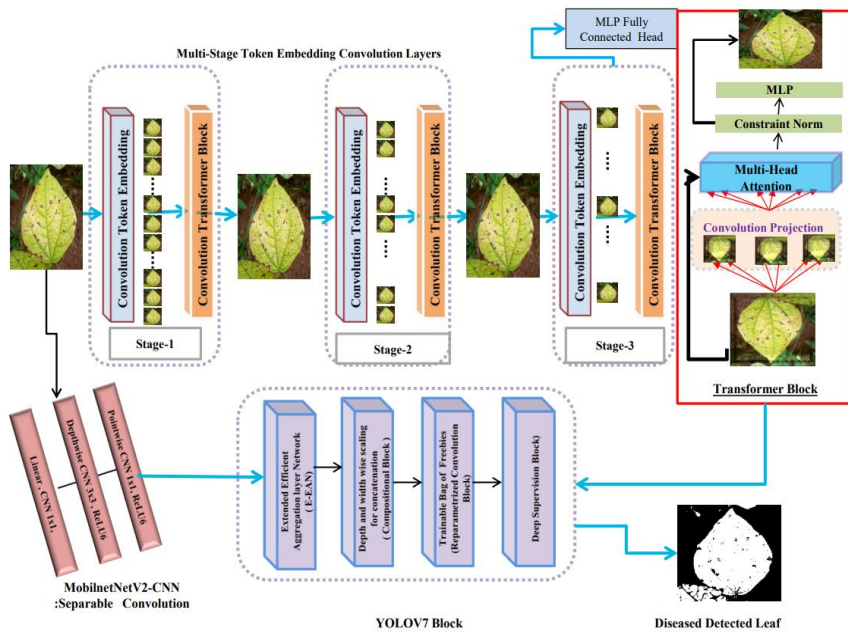


Figure 1
Transfer Learning Vision Transformer-YOLOv7: bean leaf disease detection model
 containing: a separable convolution MobileNetv2 for transfer learning to be adopted Vision Transformer and SCN2Vit features are learnt by YOLOv7

YOLOV7:

The game changing evolution in object detection field was led by 'You only look once (YOLO) in Redmon et al. [35]' an object detection technology which superseded all the detectors which raise issue with finer objects in Liu et al. [36], Kumar and Srivastava [37]. YOLOv7 is recently developed a real-time object detection technique. Till date, it is fastest and produces most accurate results in detecting objects in real-time mode. In the decade 2010 to 2020 several important deep network-based detectors were developed such as ResNet) on He et al. [38], (DenseNet) in Haung et al. [39], YOLO9000 in Redmon and Farhadi [40] and again YOLOv3 in Redmon and Farhadi [41]. All the raised computational issues with these detectors are addressed by developing comparatively light networks such as MobileNetv1 in Weyand, Andreotto, and Adam [42], MobileNetv2 in Sandler et al. [43], YOLO3-4-5 in Sozzi et al. [44], and YOLO-LITE in Huang, Pedoeem, and Chen [30]. Deep learning-based detectors learned end-to-end fashion using regression model which becomes the foundation of YOLO-series detectors in Bochkovskiy, Wang, and Liao in [20], in Wong et al. [45], in Zhu et al. [46], and in Kumar et al. [47]. All these detectors follow GPU based computations. Mehta and Rastegari [56] presented a lightweight version of vision transformer. Being GPU based computation hard to afford, transfer learning in Paymode and Malode [48] is introduced. The

advancements in YOLOv7 represent the extension of ag- aggregation layer and concatenation with scaling using Bag of Freebies (BoF) which reduces the detector into a pre-trained and parameterized CNN.

3.1 *Convolution Vision Transformer:*

Transformers learning with vision features are inbuilt with multi-headed self-attention modules. Due to its long-range dependencies, the Vision Transformer (Vit) in Arnab et al. [49] can learn the diverse features smoothly. Vision transformers alone cannot learn sufficient features from the limited amount of data. We initially introduce separable convolution specifying MobileNetv2 for associating the transfer learning with vision transformers. The hyperparameter tuning of pre-trained model MobileNetv2 was trained over RowReuse optimizer. All the pre-trained models were trained over different optimizers with a fixed learning rate of 0.001, batch size 32, epochs 25, and dropout 0.3 as shown in Table 1.

3.2 *Data Collection Resources*

iBean dataset: This dataset contains 1296 images of leaf bean which are collected from different locality in Uganda under the project Makerere AI lab NaCRR Institute. In the dataset (total 1295 images) three variations are included: healthy (428 images), angular leaf spot disease (432 images), and rust (436 images). The infection of angular spot disease is caused by *Pseudomonas syringae* and rust disease is caused by the fungus of *Uromyces appendiculatus*. The training, testing, and validating samples of the dataset are licensed by MIT Cambridge. The dataset for performing the experiments is divided into 80/20 ratio of training and validation sets.

4. Results and Discussion

In this paper, as per the dataset images shown in Figure 2, the three defined classes of bean leaf are scaled to take the input obligation. Later the leaf images are trained in MobileNetv2 using transfer learning which includes the optimizer selection and activation function.

4.1 *Optimization method*

The algorithm used to optimize separable convolution is supported by RowReuse approach in Kandemir et al. [50] which is carried out in iterative fashion by comparing numerous solutions until the most satisfactory one is achieved. To tune with the parameters which will result in low-cost function, RowReuse optimizer is more useful which ensures better accuracy in classifying the plant disease. The actual disease that leaf has is angular leaf spot and the result which shows the prediction output is also same as the actual one.

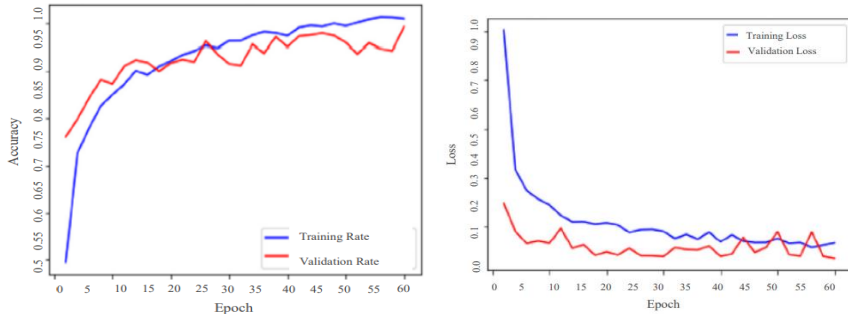


Figure 2
Training and validation score for accuracy (left) and loss (right)
of the proposed SCN2Vit

From the graph in Figure 2 (left) and Figure 2 (right), it is clearly observed that if the val-loss increases and val-acc decreases, it means the method is overfilling but not learning. In the proposed SCN2Vit-YOLOv7, val-loss is reduced and val-acc is increased which means our method is built exactly and learning and working of model is validated. In Figure 2 (right) and Figure 2 (left), the loss and accuracy respectively for both training and testing of the proposed method is shown.

4.2 Training and Testing parameters:

From Figure 3 and Figure 4, we can see the accuracy and loss of the proposed model in different epochs. When the accuracy increases, loss decreases which means the model is built and working properly. From Table 1, the comparative state of the arts is observed that, when the training images of 1034 and testing 128 images are taken for the proposed model, the final accuracy of the model is 98.8% and eventually the loss is very low as 0.11%.

Table 1
Comparative Results of State-of-the-arts (SOA) general leaf disease Detection

S-O-A	Loss	Acc.	Leaf
MCLD-CNN(2022) [48]	0.32	95.71%	Multiple-crop
ViT+VGG16 [51]	-	96%	Olive
SNDPN(2022) [7]	0.25	97.59%	Tomato
UNet-Resnet34(2021) [53]	0.5	91.07%	Bean
Con-Vit(2022) [52]	0.2	96.85%	Apple
Citrus-YOLOV7 [54]	0.92	97.29%	Citrus Orchard
VGG16+MSVM(2020) [55]	-	99.4%	Egg Plants
Our	0.085	98.80%	Bean

4.3 Experiments for ablation studies

In extended state of the arts, we preformed Mobile Vision Transformer (Mobilenetv2) Sandler et al. [43] presented in Figure 1. ON the basis of computation, the confusion matrix is presented in Figure 3, and Figure 4 on bean leaf disease dataset as ablation part to justify the validity of our model. Figure 4 (right part) shows validation loss (0.25) immediately goes down for achieving high performance of the model and high accuracy 99.2%. Figure 5 shows ROC curve ensuring the performance of classifier on various categories of bean leaf diseases.

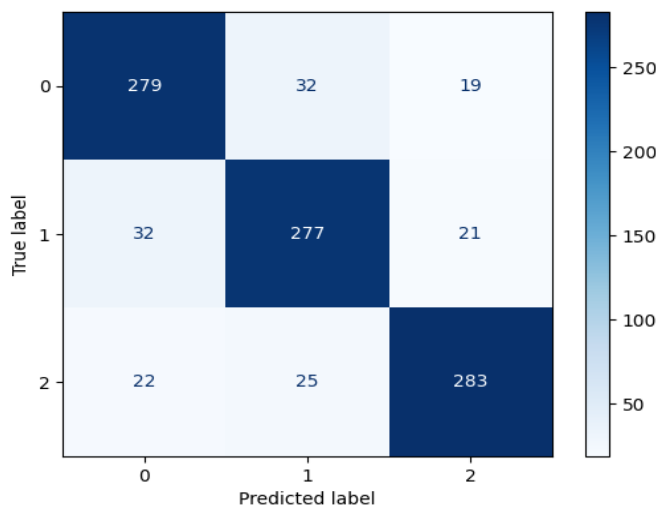


Figure 3
Confusion matrix on bean leaf disease dataset while performing with MobileNetv2

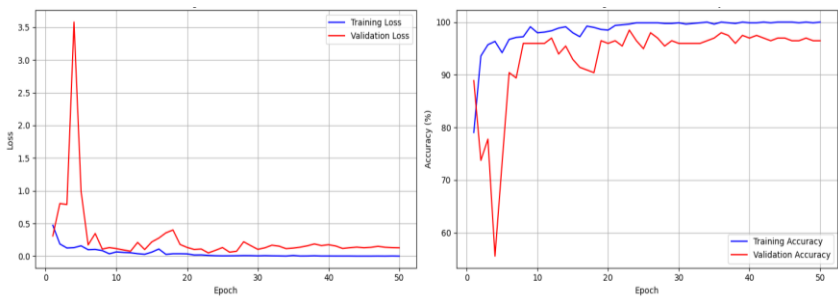


Figure 4
Training and validation score for accuracy (left) and loss (right) of the proposed MobileNetv2

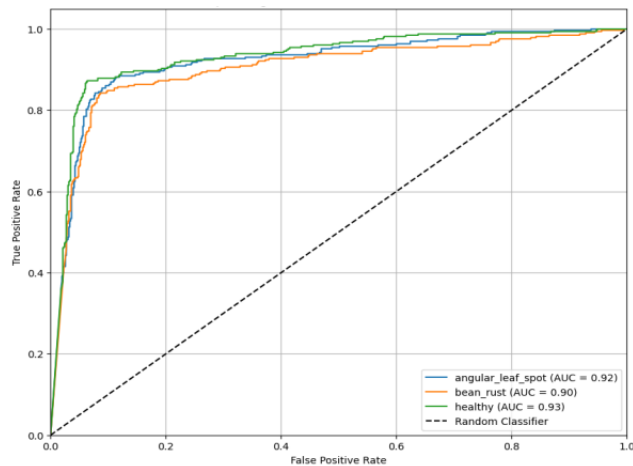


Figure 5
Receiver Operating Characteristic (ROC) for all the classes of leaf disease

5. Conclusion and Future Recommendation

Due to the huge number of parameters utilized, deep learning approaches also require a long time to train, but they probably be cheaper to test. Traditional machine learning methods, on the other hand, require short time for training and large amount of time for testing. As a result, early and precise identification might substantially reduce global economic losses while also diminishing the negative environmental effect of fertilizers and pesticides. To address this issue, an automated framework that identifies illnesses on plants and leaves is required. The present techniques have numerous limits and advantages, according to the literature review. So here we need to propose a method where the diseases on leaf are classified. For this work was, MobileNetv2 which is a deep learning method, has a triumphant result which we get for classification of different bean leaf diseases using the bean leaf captured images.

References

- [1] E. Loizou, C. Karelakis, K. Galanopoulos, and K. Mattas, "The role of agriculture as a development tool for a regional economy," *Agricultural Systems*, vol. 173, pp. 482–490 (2019).
- [2] X. Diao, P. Hazell, and J. Thurlow, "The role of agriculture in african development," *World development*, vol. 38, no. 10, pp. 1375– 1383 (2010).

- [3] L. Selvam and P. Kavitha, "Classification of ladies finger plant leaf using deep learning," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–9 (2020).
- [4] A. Abade, P. A. Ferreira, and F. de Barros Vidal, "Plant diseases recognition on images using convolutional neural networks: A systematic review," *Computers and Electronics in Agriculture*, vol. 185, p. 106125 (2021).
- [5] R. Gajjar, N. Gajjar, V. J. Thakor, N. P. Patel, and S. Ruparelia, "Real-time detection and identification of plant leaf diseases using convolutional neural networks on an embedded platform," *The Visual Computer*, vol. 38, no. 8, pp. 2923–2938 (2022).
- [6] I. El Massi, Y. Es-saady, M. El Yassa, and D. Mammass, "Combination of multiple classifiers for automatic recognition of diseases and damages on plant leaves," *Signal, Image and Video Processing*, vol. 15, pp. 789–796 (2021).
- [7] X. Huang, A. Chen, G. Zhou, X. Zhang, J. Wang, N. Peng, N. Yan, and C. Jiang, "Tomato leaf disease detection system based on fc-sndpn," *Multimedia Tools and Applications*, pp. 1–24 (2022).
- [8] R. Bora, D. Parasar, and S. Charhate, "A detection of tomato plant diseases using deep learning mndlnn classifier," *Signal, Image and Video Processing*, pp. 1–9 (2023).
- [9] N. Goyal, N. Kumar, and K. Gupta, "Lower-dimensional intrinsic structural representation of leaf images and plant recognition," *Signal, Image and Video Processing*, pp. 1–8 (2022).
- [10] M. Turkoglu, B. A. Yanikoglu, and D. Hanbay, "Plantdiseasenet: Convolutional neural network ensemble for plant disease and pest detection," *Signal, Image and Video Processing*, vol. 16, no. 2, pp. 301–309 (2022).
- [11] A. Beikmohammadi and K. Faez, "Leaf classification for plant recognition with deep transfer learning," in *2018 4th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*. IEEE, pp. 21–26 (2018).
- [12] A. P. Mullins and B. H. Arjmandi, "Health benefits of plant-based nutrition: focus on beans in cardiometabolic diseases," *Nutrients*, vol. 13, no. 2, p. 519 (2021).
- [13] K. K. Chakraborty, R. Mukherjee, C. Chakraborty, and K. Bora, "Automated recognition of optical image based potato leaf blight diseases-

- es using deep learning," *Physiological and Molecular Plant Pathology*, vol. 117, p. 101781 (2022).
- [14] S. R. Dubey and A. S. Jalal, "Apple disease classification using color, texture and shape features from images," *Signal, Image and Video Processing*, vol. 10, pp. 819–826 (2016).
 - [15] P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, "Deep feature based rice leaf disease identification using support vector machine," *Computers and Electronics in Agriculture*, vol. 175, p. 105527 (2020).
 - [16] H.-T. Thai, K.-H. Le, and N. L.-T. Nguyen, "Formerleaf: An efficient vision transformer for cassava leaf disease detection," *Computers and Electronics in Agriculture*, vol. 204, p. 107518 (2023).
 - [17] B. A. Kumar and M. Bansal, "Face mask detection on photo and real-time video images using caffe-mobilenetv2 transfer learning," *Applied Sciences*, vol. 13, no. 2, p. 935 (2023).
 - [18] R. S. Alkhawaldeh, M. Alawida, N. F. F. Alshdaifat, W. Alma'aitah, and A. Almasri, "Ensemble deep transfer learning model for arabic (indian) handwritten digit recognition," *Neural Computing and Applications*, vol. 34, no. 1, pp. 705–719 (2022).
 - [19] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Yolov7: Train-able bag-of-freebies sets new state-of-the-art for real-time object detectors," *arXiv preprint arXiv:2207.02696* (2022).
 - [20] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934* (2020).
 - [21] L. Li, S. Zhang, and B. Wang, "Plant disease detection and classification by deep learning—a review," *IEEE Access*, vol. 9, pp. 56683–56698 (2021).
 - [22] J. Liu and X. Wang, "Tomato diseases and pests detection based on improved yolo v3 convolutional neural network," *Frontiers in plant science*, vol. 11, p. 898 (2020).
 - [23] R. Karthik, M. Hariharan, S. Anand, P. Mathikshara, A. Johnson, and R. Menaka, "Attention embedded residual cnn for disease detection in tomato leaves," *Applied Soft Computing*, vol. 86, p. 105933 (2020).
 - [24] N. Krishnamoorthy, L. N. Prasad, C. P. Kumar, B. Subedi, H. B. Abrahah, and V. Sathishkumar, "Rice leaf diseases prediction using deep

- neural networks with transfer learning," *Environmental Research*, vol. 198, p. 111275 (2021).
- [25] X. Nie, L. Wang, H. Ding, and M. Xu, "Strawberry verticillium wilt detection network based on multi-task learning and attention," *IEEE Access*, vol. 7, pp. 170 003–170 011 (2019).
 - [26] R. Cristin, B. S. Kumar, C. Priya, and K. Karthick, "Deep neural network based rider-cuckoo search algorithm for plant disease detection," *Artificial intelligence review*, vol. 53, no. 7, pp. 4993–5018 (2020).
 - [27] C. Bi, J. Wang, Y. Duan, B. Fu, J.-R. Kang, and Y. Shi, "Mobilenet based apple leaf diseases identification," *Mobile Networks and Applications*, pp. 1–9 (2020).
 - [28] Y. Zhang, C. Song, and D. Zhang, "Deep learning-based object detection improvement for tomato disease," *IEEE Access*, vol. 8, pp. 56 607–56 614 (2020).
 - [29] M. P. Mathew and T. Y. Mahesh, "Leaf-based disease detection in bell pepper plant using yolo v5," *Signal, Image and Video Processing*, vol. 16, no. 3, pp. 841–847 (2022).
 - [30] R. Huang, J. Pedoeem, and C. Chen, "Yolo-lite: a real-time object detection algorithm optimized for non-gpu computers," in *2018. IEEE International Conference on Big Data (Big Data)*. IEEE, pp. 2503–2510 (2018).
 - [31] Y. Tian, G. Yang, Z. Wang, H. Wang, E. Li, and Z. Liang, "Apple detection during different growth stages in orchards using the improved yolo-v3 model," *Computers and electronics in agriculture*, vol. 157, pp. 417–426 (2019).
 - [32] N. Kumar, A. K. Jilani, P. Kumar, and A. Nikiforova, "Improved yolo-v3-tiny object detector with dilated cnn for drone-captured images," in *2022 International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*. IEEE, pp. 89–94 (2022).
 - [33] W. Bao, Z. Zhu, G. Hu, X. Zhou, D. Zhang, and X. Yang, "Uav remote sensing detection of tea leaf blight based on ddma-yolo," *Computers and Electronics in Agriculture*, vol. 205, p. 107637 (2023).
 - [34] H. Zhao, J. An, M. Yu, D. Lv, K. Kuang, and T. Zhang, "Nesterov-accelerated adaptive momentum estimation-based wavefront distortion correction algorithm," *Applied Optics*, vol. 60, no. 24, pp. 7177–7185 (2021).

- [35] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788 (2016).
- [36] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*. Springer, pp. 21–37 (2016).
- [37] A. Kumar and S. Srivastava, "Object detection system based on convolution neural networks using single shot multi-box detector," *Procedia Computer Science*, vol. 171, pp. 2610–2617 (2020).
- [38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778 (2016).
- [39] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708 (2017).
- [40] J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7263–7271 (2017).
- [41] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," *arXiv preprint arXiv:1804.02767* (2018).
- [42] T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861* (2017).
- [43] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4510–4520 (2018).
- [44] M. Sozzi, S. Cantalamessa, A. Cogato, A. Kayad, and F. Marinello, "Automatic bunch detection in white grape varieties using yolov3, yolov4, and yolov5 deep learning algorithms," *Agronomy*, vol. 12, no. 2, p. 319 (2022).
- [45] A. Wong, M. Famuori, M. J. Shafiee, F. Li, B. Chwyl, and J. Chung, "Yolo nano: a highly compact you only look once convolutional neural network for object detection," in *2019 Fifth Workshop on Energy Efficient Machine Learning and Cognitive Computing-NeurIPS Edition (EMC2-NIPS)*. IEEE, pp. 22–25 (2019).

- [46] X. Zhu, S. Lyu, X. Wang, and Q. Zhao, "Tph-yolov5: Improved yolov5 based on transformer prediction head for object detection on drone-captured scenarios," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2778–2788 (2021).
- [47] A. Kumar, A. Kalia, A. Sharma, and M. Kaushal, "A hybrid tiny yolo v4-spp module based improved face mask detection vision system," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–14 (2021).
- [48] A. S. Paymode and V. B. Malode, "Transfer learning for multi-crop leaf disease image classification using convolutional neural network vgg," *Artificial Intelligence in Agriculture*, vol. 6, pp. 23–33 (2022).
- [49] A. Arnab, M. Dehghani, G. Heigold, C. Sun, M. Luc'ic, and C. Schmid, "Vivit: A video vision transformer," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6836–6846 (2021).
- [50] M. Kandemir, H. Zhao, X. Tang, and M. Karakoy, "Memory row reuse distance and its role in optimizing application performance," in *Proceedings of the 2015 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, pp. 137–149 (2015).
- [51] H. Alshammari, K. Gasmi, I. Ben Ltaifa, M. Krichen, L. Ben Ammar, and M. A. Mahmood, "Olive disease classification based on vision transformer and cnn models," *Computational Intelligence and Neuroscience*, vol. 2022 (2022).
- [52] X. Li and S. Li, "Transformer help cnn see better: A lightweight hybrid apple disease identification model based on transformers," *Agriculture*, vol. 12, no. 6, p. 884 (2022).
- [53] S. H. Abed, A. S. Al-Waisy, H. J. Mohammed, and S. Al-Fahdawi, "A modern deep learning framework in robot vision for automated bean leaves diseases detection," *International Journal of Intelligent Robotics and Applications*, vol. 5, no. 2, pp. 235–251 (2021).
- [54] J. Chen, H. Liu, Y. Zhang, D. Zhang, H. Ouyang, and X. Chen, "A multiscale lightweight and efficient model based on yolov7: Applied to citrus orchard," *Plants*, vol. 11, no. 23, p. 3260 (2022).
- [55] A. Krishnaswamy Rangarajan and R. Purushothaman, "Disease classification in eggplant using pre-trained vgg16 and msvm," *Scientific reports*, vol. 10, no. 1, pp. 1–11 (2020).

- [56] S. Mehta and M. Rastegari, "MobileViT: Light-weight, general-purpose, and mobile-friendly vision transformer," in *Proc. 2022 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, pp. 2120–2130 (2022).
- [57] S. S. Kashyap, B. Unhelkar, S. S. Shankar, T. Chakrabarti, P. Chakrabarti, B. Sivaneasan, and M. Margala, "Optimized decision science approach for accurate detection of rice plant diseases," *J. Inf. Optim. Sci.*, vol. 45, no. 8, pp. 2081–2090 (2024), doi: 10.47974/JIOS-1652.
- [58] S. Sachar and A. Kumar, "Deep learning for automatic identification of plants through leaf," *J. Inf. Optim. Sci.*, vol. 44, no. 4, pp. 709–716 (2023), doi: 10.47974/JIOS-1269.

Received December, 2025